© 1989 Nature Publishing Group http://www.nature.com/naturebiotechnology

plant cells, callus cultures of tobacco were placed onto MS-agar plates containing 1 mg/l NAA, 0.1 mg/l BAP and 2,4-DCP in concentrations ranging from 1 to 10 mg/l. No toxic effects could be seen with 10 mg/l 2,4-DCP, whereas more than 2 mg/l 2,4-D severely reduced growth and turned calli brown. Perkins *et al.*[15] reported 50% growth inhibition of tobacco suspension cultures by 10 mg/l 2,4-DCP, but a generally much lower toxicity of 2,4-DCP on the development of leaves and roots when compared to 2,4-D. These results were encouraging with respect to the idea that transformation of tobacco with the DPAM gene could result in resistance to 2,4-D.

Transformation of tobacco with DPAM genes and transcript analysis. Both constructs containing the hybrid DPAM genes in combination either with the constitutive (pMCJC1007) or with the light inducible promoter (pMLJC1005) were used to transform tobacco (*Nicotiana tabacum* W38). Gene transfer was mediated by *Agrobacterium* strain C58C1 carrying the disarmed Ti-plasmid pGV3850kan[14] that is able to form cointegrates with pBR322 derived vectors. Shoots were regenerated from infected leaf discs by selection on kanamycin, screened for nopaline production and later analyzed by Northern blot hybridization for the synthesis of DPAM-RNA. Twenty out of 25 transformed plants synthesized a specifically hybridizing transcript of the expected size of 1.3 kb (data not shown).

Expression of the DPAM gene in tissue culture of tobacco. 2,4-D is known to promote callus growth and to suppress shoot regeneration from explanted tobacco leaf discs in tissue culture when applied in concentrations higher than 0.1 mg/l. To test transgenic plants for their ability to tolerate elevated levels of 2,4-D, we chose a "shoot regeneration" medium containing 1.0 mg/l BAP and 2,4-D at concentrations of 0.1, 0.2, 0.4, 1.0, 2.0, and 4.0 mg/l. Leaf discs from tobacco plants where the DPAM gene was transcribed from the constitutive 35S promoter were able to develop normal shoots on a medium containing 1.0 mg/l 2,4-D (Fig. 2A). Depending on the individual transformed plant tested, minimal shoot inhibition is observed at concentrations of 2.0 or 4.0 mg/l 2,4-D, whereas leaf discs from untransformed control plants (tobacco W38) did not show any differentiation on 0.2 mg/l 2,4-D, indicating that expression of the DPAM gene could result in up to 20-fold higher tolerance against 2,4-D.

A significantly different reaction was shown by plants where the DPAM gene is under the control of the light-inducible promoter of the ST-LS1 gene. As this promoter is only active in photosynthetically active tissues, and shoot differentiation from explanted tobacco leaf discs proceeds via a (photosynthetically not active) callus stage, as expected in the shoot suppression assay, the transgenic DPAM-plants showed the same sensitivity against 2,4-D as the nontransformed tobacco W38 (Fig. 2B).

Response to differently substituted phenoxyacetic acids. In *Alcaligenes eutrophus* JMP134 from which the DPAM gene was originally isolated, DPAM is able to degrade 2,4-D, MCPA, 4-CPA and unsubstituted phenoxyacetic acid. To investigate how transgenic tobacco expressing DPAM responds to the auxin activity of differently substituted phenoxyacetic acids, the substances listed in Figure 3 were incorporated in MS-agar together with 1.0 mg/l BAP. Leaf discs of transgenic tobacco WS7-7 and control W38 were placed on the media and shoots arising were counted after 18 days. Figure 3 shows, that transgenic plants expressed a similar tolerance to 4-CPA as to 2,4-D, a slightly reduced tolerance to MCPA and no tolerance against tri-substituted phenoxyacetic acids and all the phenoxypropionic acids we tested.

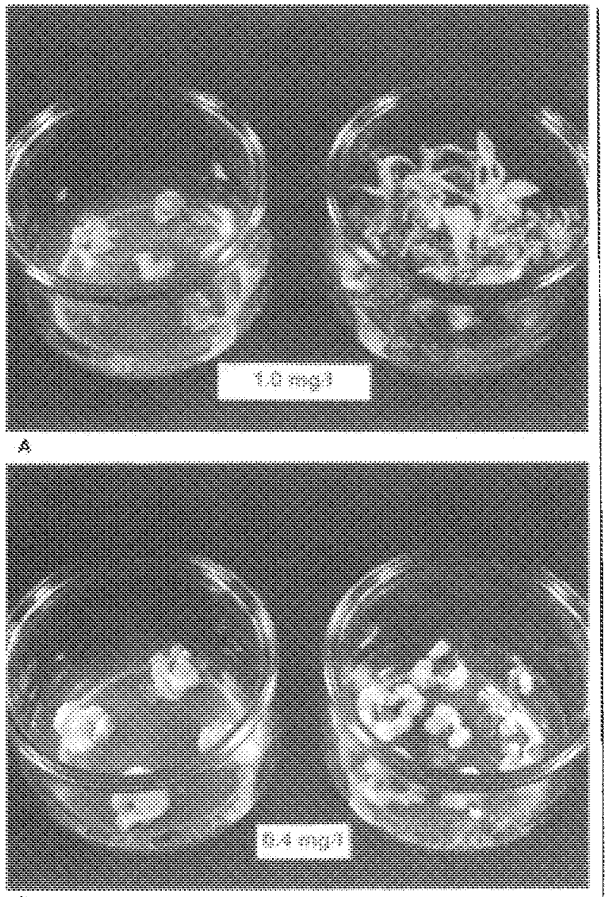


FIGURE 2 Tolerance of transgenic tobacco against 2,4-D in tissue culture. Explants from *Nicotiana tabacum* W38 control plants (left) and transgenic plants (right) were incubated on medium containing 1.0 mg/l 2,4-D. Normal shoots develop from leaf discs which constitutively express DPAM (A, right). Transgenic explants with a light-inducible DPAM gene (B, right) show the same sensitivity against 2,4-D as the nontransformed tobacco W38 (left).

Transgenic plants are resistant to the herbicidal activity of 2,4-D. Replicates of transgenic plants obtained by clonal propagation were transferred to soil and grown in the greenhouse. No difference in growth behaviour and final height could be seen between untransformed and transformed tobacco. After two weeks of growth in soil control tobacco plants and transgenic tobacco plants were sprayed with doses equivalent to 1 kg/ha, 3 kg/ha and 10 kg/ha 2,4-D-Na. Plants were further grown for 2-3 weeks in the greenhouse and then checked for resistance. Whereas 1 kg/ha totally inhibited growth of nontransformed tobacco plants, plants of the transgenic line WS7-7 which expressed the DPAM gene under the control of the constitutive 35S promoter were resistant to up to 10 kg/ha 2,4-D-Na (Fig. 4 A).

Plants of the transgenic line WS4-7 that expressed the DPAM gene under a light-inducible promoter active only in green tissue and thus did not show tolerance in a tissue culture assay, were also resistant to the herbicide when sprayed with 1 kg/ha 2,4-D-Na (Fig. 4 B). No difference in growth behaviour and final height was observed between transgenic plants treated with the herbicide and non-treated plants (Fig. 4 C).

COR-AAD1-00001131

© 1989 Nature Publishing Group http://www.nature.com/naturebiotechnology

Inheritance of 2,4-D resistance. Seeds from selfed WS7-7 plants were germinated on sterile MS medium with kanamycin. Two hundred and fifty eight seedlings were resistant to kanamycin and 83 were sensitive indicating Mendelian inheritance of a single active locus. Leaf segments from 16 two weeks old Km^r F1 plants were placed on 2,4-D medium to check for 2,4-D resistance, another 20 plants were sprayed with 3 kg/ha 2,4-D-Na. In all cases inheritance of 2,4-D resistance was observed. Southern blots showed that the parental line WS7-7 contained at least three copies of the DPAM gene, which could also be found in 14 out of 15 analysed transgenic F1 plants (data not shown). This suggests that the three T-DNA copies in line WS7-7 were linked.

DPAM as a selectable marker. The use of the constitutively expressed DPAM gene as a selectable marker for *Agrobacterium* mediated leaf disc transformation was examined. In a standard transformation procedure, kanamycin was replaced by 2,4-D at a concentration of 0.5 mg/l, which was previously shown to be appropriate to distinguish between untransformed and completely transformed tissue, but no shoot development was observed under these conditions.

In order to test if DPAM expression could be used for selection of germinating seeds, root formation from seeds was measured in response to different concentrations of 2,4-D. Inhibition curves (Fig. 5) demonstrate that no significant difference in sensitivity is seen between untransformed seedlings and seedlings containing the 35S-DPAM gene.

DISCUSSION

The *tfdA* gene from *Alcaligenes eutrophus* JMP134 encodes the enzyme 2,4-dichlorophenoxyacetate monooxygenase (DPAM) that degrades the herbicide, 2,4-D. The degradation product, 2,4-dichlorophenol, has been shown to be non-toxic to plants at least up to concentrations of 10 mg/l. We have fused the DPAM coding sequence to both a constitutive and a light-inducible promoter and transferred the constructs in tobacco using an *Agrobacterium* vector system. Transgenic tobacco plants transformed with the chimeric DPAM-genes were fully resistant to high doses of 2,4-D. In addition, leaf cells from transgenic tobacco that constitutively express DPAM are able to redifferentiate into shoots in the presence of concentrations of 2,4-D that normally suppress shoot formation.

Unlike the NPT gene[17] or the PAT gene[5], the DPAM gene could not be used as a selectable marker in leaf disc transformation. One reason may be that in a chimeric tissue transformed cells cannot develop into shoots because they are overgrown or otherwise inhibited by untransformed callus that is rapidly developing in the presence of 2,4-D. It remains to be examined whether transformation of protoplasts instead of leaf discs might overcome this problem.

It is also not possible to select transgenic DPAM plants during germination on 2,4-D. The lack of resistance of roots against 2,4-D may have several reasons, e.g. a lower expression of the 35S promoter in roots, a higher sensitivity of roots against 2,4-D or 2,4-DCP compared to shoots or an accumulation of 2,4-D or 2,4-DCP by transport.

The analysis of the DPAM-specific mRNA demonstrated a substantial variation in expression levels of around 20 to 30-fold between independent transformants. Nevertheless, the highest expression level of the DPAM gene in line WS7-7 corresponded to only three T-DNA copies in the genome (data not shown). A similar observation has been made by DeBlock et al.[5] who reported three copies of the PAT gene for the highest expressing plant. He suggested that the observed 30-fold differences in the expression of

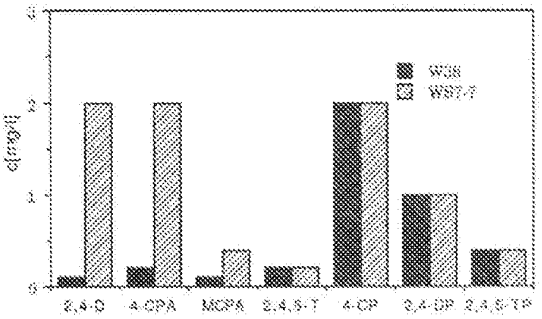


FIGURE 3 Cross tolerance of transgenic tobacco against other phenoxy herbicides. Substances listed below were incorporated into 2MS agar as described in the text. The graph demonstrates the response of control tobacco (W38, black columns) and transgenic tobacco (WS7-7, hatched columns) against different synthetic auxins. The height of the columns indicates the upper limit that allowed cells from explanted leaf discs to differentiate into normal shoots. At higher concentration only callus growth was observed. 2,4-D: 2,4-dichlorophenoxyacetic acid, 4-CPA: 4-chlorophenoxyacetic acid, MCPA: 4-chloro-2-methyl-phenoxyacetic acid, 2,4,5-T: 2,4,5-trichlorophenoxyacetic acid, 4-CP: 2-(4-chlorophenoxy)propionic acid, 2,4-DP: 2-(2,4-dichlorophenoxy)propionic acid, 2,4,5-TP: 2-(2,4,5-trichlorophenoxy)propionic acid.

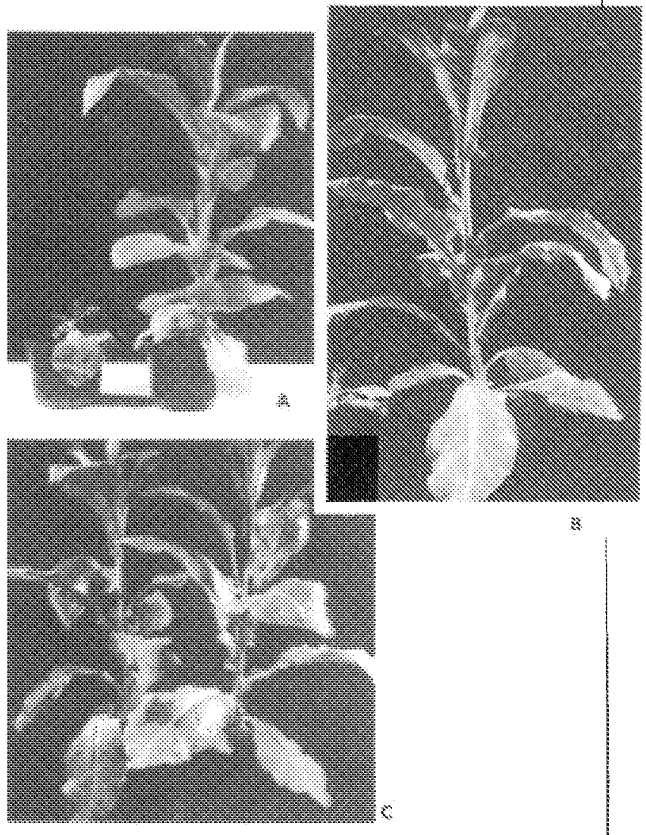


FIGURE 4 Effect of 2,4-D on transgenic tobacco plants. A. A *Nicotiana tabacum* W38 control plant (left) and a clonal replicate of the WS7-7 line that expresses the DPAM gene under the control of the CaMV 35S promoter (right) were sprayed with 10 kg/ha 2,4-D-Na. B. A *Nicotiana tabacum* W38 control plant (left) and a clonal replicate of the WS4-7 line that expresses the DPAM gene under the control of the light-inducible ST-LS1 promoter (right) were sprayed with 1 kg/ha 2,4-D-Na. C. Two replicates of the WS4-7 line were sprayed either with water (left) or with 1 kg/ha 2,4-D-Na (right). The pictures were taken three weeks after spraying.

COR-AAD1-00001132

© 1989 Nature Publishing Group http://www.nature.com/naturebiotechnology

PAT-RNA may be due to effects of the T-DNA position in the chromosome.

Expression of DPAM-RNA in transformants where the DPAM gene is under the control of the light-inducible promoter varied in the same range, but we never observed similarly high expressing variants compared to the constitutively expressing plants. The highest observed expression corresponded to a resistance against sprayed doses of 1 kg/ha 2,4-D-Na. No qualitative difference could be seen between resistance phenotypes generated by the constitutive or the light-inducible DPAM gene.

When transgenic plants expressing DPAM were sprayed with 2,4-D in the greenhouse, they were fully resistant to amounts of 2,4-D commercially applied in the field. For example, line WS4-7 tolerates 1 kg/ha, line WS7-7 tolerates 10 kg/ha 2,4-D-Na. Depending on the crop and weed species, 2,4-D is effectively applied in amounts ranging between 0.3 kg/ha and 3 kg/ha. This result is encouraging with respect to the application of the DPAM gene to create 2,4-D resistance in other commercially interesting crop plants. As 2,4-D is already a selective herbicide in monocotyledonous crop cultures, dicotyledonous crop plants will be the preferred targets for genetic engineering. Additional investigation of the fate of the DPAM reaction product, 2,4-dichlorophenol, in the plant are required. It may be worthwhile to think about introducing further genes of the bacterial 2,4-D degradative pathway into plants. In any case, toxicology and persistance of 2,4-D and its degradation products have to be newly evaluated.

Other attempts to obtain 2,4-D resistant plants have been made[7,16]. The introduction of the detoxifying enzyme DPAM by genetic engineering is up to now the only approach that results in fully resistant and phenotypically normal plants. This strategy promises to have broad application, especially with herbicides for which detoxifying pathways can be found in nature or created by genetically modifying existing pathways. Future development of new herbicidal compounds should favour substances that are easily degraded by microorganisms, thus providing environmentally safe, non-persisting herbicides together with the opportunity to introduce these selective resistances into crop plants by genetic engineering.

EXPERIMENTAL PROTOCOL

Strains. *E. coli* TB1 is a derivative of strain JM101[19] with the genotype: F'(*tra*D36, *pro*AB, *lac*I, *lac*ZΔM15) Δ(lac, pro), *supE*, *thi*, *rec*A, Sr1::Tn*10*(Tc^r) (Bart Barrel, pers. communication). *E. coli* GJ23[20] was used as a helper strain for triparental matings. *Agrobacterium tumefaciens* strain C58C1 (pGV3850kan)[14] contains a disarmed cointegrate-type Ti-plasmid.

Plasmids. pUC18 was described by Norrander[21]. pA5 (Fig. 1) contains 200 bp of the octopine synthase 3'-end[11] inserted into pUC18 (A. von Schaewen, in preparation). pA8 (Fig. 1) is a plant expression vector constructed by insertion of the CaMV 35S promoter[12] and the OCS-3' end[11] into pMPK110 (A. von Schaewen, in preparation). The multiple cloning site of pA8 was derived from pUC18. pTJSX535 contains the coding sequence of *tfdA* from *Alcaligenes eutrophus* JMP134 on a 1.5 kb XbaI-SalI fragment[9].

Media. Synthetic media for plant transformation and shoot regeneration from transformed and untransformed tissue were based on MS medium[22] containing 20 g/l sucrose. Basic chemicals and 2,4-D were from Merck, Darmstadt, FRG. All other synthetic hormones were from Sigma, St. Louis, MO, USA.

Synthetic oligonucleotides. Oligonucleotides were synthesized by an Applied Biosystems 380A equipment and purified by preparative polyacrylamide gel electrophoresis as described in the Applied Biosystems User Bulletin issue No. 13, 1984 (Applied Biosystems, Foster City, California).

Plasmid contructions. Pairs of complementary oligonucleotides, each consisting of 20 bases, were phosphorylated by polynucleotide kinase, annealed by shifting temperature from 60°C to 37°C, and then ligated with pUC18 that had been digested with BamHI and SphI. The linear product was separated from an excess of oligonucleotides by agarose gel electrophoresis using DEAE membrane NA45 (Schleicher & Schuell). pTJSX535 plasmid DNA was digested with XbaI and SphI and separated by agarose gel electrophoresis. The 1.2 kb fragment containing the DPAM coding sequence was isolated from the gel by electroelution, partially digested by FokI, and repurified by agarose gel electrophoresis. DNA fragments of around 1.1 kb isolated from the gel were ligated with the linearised pUC18 that contained the oligonucleotide. The resulting plasmid pUJC1001 was cloned in *E. coli* TB1. DNA of pUJC1001 was digested with BamHI and SphI to excise the insert, and with PvuI to avoid religation of the vector, and then ligated with pA5 that had been digested with BamHI and SphI. pA5 was used as an intermediate vector as pA8 contained a second SphI site outside the polylinker. The resulting plasmid pUJA0 was cloned in *E. coli* TB1. DNA of both pUJA01002 and pA8 was digested with BamHI and SphI, mixed and ligated. The resulting plasmid pMCJA1006 was cloned in *E. coli* TB1.

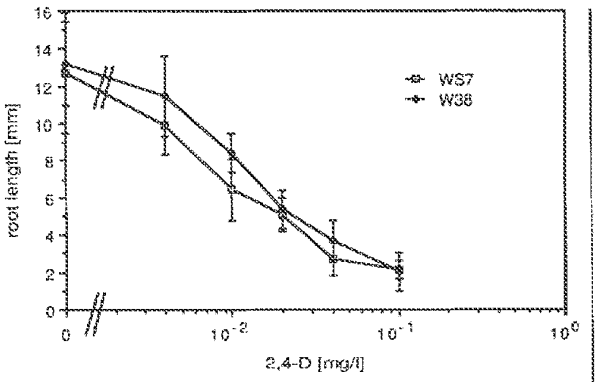


FIGURE 5 Effect of 2,4-D on rooting of seeds. Transgenic tobacco seeds containing the 35S-DPAM gene (WS7) and untransformed seeds (W38) were germinated on sterile MS agar plates containing 2,4-D at concentrations of 0, 4, 10, 20, 40, and 100 μg/l. After 14 days of incubation in the light, root length was measured. Vertical bars indicate the standard error.

Transformation of tobacco plants with chimeric genes. pMCJC1007 and pMLJC1005 were mobilized from *E. coli* TB1 into the *Agrobacterium* strain C58C1 (pGV3850kan) by a triparental mating[23] using *E. coli* GJ23 as a helper strain. Agrobacteria that by Southern hybridization were shown to contain the DPAM gene were subsequently used for leaf disc infection of *Nicotiana tabacum* W38. Transformation and regeneration of tobacco plants followed established procedures[23].

Cloning. Construction of recombinant DNA molecules, transformation into *E. coli*, and analysis of recombinant clones were performed according to established procedures[24].

DNA sequencing. DNA was sequenced by the dideoxy method[25] using plasmid DNA as a template[26].

Hybridization experiments. RNA isolation from leaves and Northern hybridization experiments were performed as described by Jones[14]. A BamHI-HindIII fragment from pUJC1001, carrying the DPAM gene, was radioactively labeled and used as a probe to detect DPAM RNA. Plant DNA was isolated essentially as described by Dellaporta[27]. For Southern hybridization experiments[28] plant DNA was cut with HindIII and probed with a cloned fragment of the nopaline synthase gene.

Herbicide application. Each plant was sprayed with 4 ml of a 1% aqueous solution of 2,4-D, Na-salt. With an estimated ground surface of 400 cm^2 per plant, this roughly corresponds to 10 kg/ha. Lower quantities were achieved by appropriate dilution of the spray solution.

Acknowledgments

We are grateful to C. Recknagel for excellent performance of plant transformation, to R. Breitfeld for propagation of plants, to U. Kutschka for analysis of plants, to A. v. Schaewen for kind gifts of plasmids, and to H.-J. Scheuermann for gas chromatography and mass spectrometry.

Received 18 January 1989; accepted 30 March 1989.

COR-AAD1-00001133

© 1989 Nature Publishing Group http://www.nature.com/naturebiotechnology

GUIDE FOR AUTHORS

Bio/Technology publishes original papers on applied research in, and on basic research with practical implications for, biotechnology. The disciplines we cover include molecular biology, microbiology, biochemistry, plant and animal biology, and chemical engineering, among others. Because of a substantial increase in the volume of high-quality manuscripts now regularly submitted to *Bio/Technology*, and in order to sustain our policy of rapid publication (usually three to four months after acceptance), we ask that research papers be no more than 3,000 words long, with a maximum of five illustrations (figures or tables). *Bio/Technology* also welcomes submission of shorter, single-topic research notes of no more than 1,500 words, with at most three illustrations. Any previously published paper may be used as an organizational and format guide. Only original research will be published; material submitted elsewhere will not be considered.

An original and two copies of the manuscript should be addressed to the journal's research editor. Please include a covering letter describing the originality and potential application of research. All manuscripts selected for possible publication will undergo peer review by two or more referees. Authors are encouraged to suggest two referees, but the final selection of reviewers is the editor's. *Bio/Technology* customarily publishes articles without page charges. In the case of articles that exceed our space constraints or contain color illustrations, however, the authors may be asked to help defray publication costs.

Reviews and book reviews are commissioned by the editor; proposals are most welcome.

—The Editors

References

1. Botterman, J. and Leemans, J. 1988. Engineering herbicide resistance in plants. Trends in Genetics **4**:219–222.
2. Cheung, A. Y., Bogorad, L., Van Montagu, M., and Schell, J. 1988. Relocating a gene for herbicide tolerance: A chloroplast gene is converted into a nuclear gene. Proc. Natl. Acad. Sci. USA **85**:391–395.
3. Comai, L., Facciotti, D., Hiatt, W. R., Thompson, G., Rose, R. E. and Stalker, D. M. 1985. Expression in plants of a mutant *aroA* gene from *Salmonella typhimurium* confers tolerance to glyphosate. Nature **317**:741–744.
4. Haughn, G. W., Smith, J., Mazur, B., and Somerville, C. 1988. Transformation with a mutant *Arabidopsis* acetolactate synthase gene renders tobacco resistant to sulfonylurea herbicides. Mol. Gen. Genet. **211**:266–271.
5. DeBlock, M., Bottermann, J., Vandewiele, M., Dockx, J., Thoen, C., Gosselé, V., Rao Movva, N., Thompson, C., Van Montagu, M., and Leemans, J. 1987. Engineering herbicide resistance in plants by expression of a detoxyfying enzyme. EMBO J. **6**:2513–2518.
6. Stalker, D. M., McBride, K. E., and Malyj, L. D. 1988. Herbicide resistance in transgenic plants expressing a bacterial detoxification gene. Science **242**:419–423.
7. Estelle, M. A., and Somerville, C. 1987. Auxin-resistant mutants of *Arabidopsis thaliana* with an altered morphology. Mol. Gen. Genet. **206**:200–206.
8. Zenk, M. H. 1974. Haploids in physiological and biochemical research, *In*: Haploids in Higher Plants. Advances and potential. Proceedings of the First International Symposium, Guelph, Ontario, Canada.
9. Streber, W. S., Timmis, K. N., and Zenk, M. H. 1987. Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. **169**:2950–2955.
10. Lütcke, H. A., Chow, K. C., Mickel, F. S. Moss, K. A., Kern, H. F., and Scheele, G. A. 1987. Selection of AUG initiation codons differs in plants and animals. EMBO J. **6**:43–48.
11. Dhaese, P., De Greve, H., Gielen, J., Seurinck, J., Van Montagu, M., Schell, J. 1983. Identification of sequences involved in the polyadenylation of higher plant nuclear transcripts using *Agrobacterium* T-DNA genes as models. EMBO J. **2**:419–426.
12. Paszkowski, J., Shillito, R. D., Saul, M., Mandák, V., Hohn, T., Hohn, B., and Potrykus, I. 1984. Direct gene transfer to plants. EMBO J. **3**:2717–2722.
13. Pietrzak, M., Shillito, R. D., Hohn T., and Potrykus, I. 1986. Expression in plants of two bacterial antibiotic resistance genes after protoplast transformation with a new plant expression vector. Nucl. Acids Res. **14**:5857–5868.
14. Jones, J. D., Dunsmuir, P., and Bedbrook, J. 1985. High level expression of introduced chimaeric genes in regenerated transformed plants. EMBO J. **4**:2411–2418.
15. Stockhaus, J., Eckes, P., Rocha-Sosa, M., Schell, J., and Willmitzer, L. 1987. Analysis of *cis*-active sequences involved in the leaf-specific expression of a potato gene in transgenic plants. Proc. Natl. Acad. Sci. USA **84**:7943–7947.
16. Perkins, E. J., Stiff, C. M., and Lurquin, P. F. 1987. Use of *Alcaligenes eutrophus* as a source of genes for 2,4-D resistance in plants. Weed Science **35**:12–18.
17. Fraley, R. T., Rogers, S. G., Horsch, R. B., Sanders, P. R., Flick, J. S., Adams, S. P., Bittner, M. L., Brand, L. A., Fink, C. L., Fry, J. S. Galluppi, G. R., Goldberg, S. B., Hoffmann, N. L., and Woo, S. C. 1983. Expression of bacterial genes in plant cells. Proc. Natl. Acad. Sci. USA **80**:4803–4807.
18. Swanson, E. B., and Tomes, D. T. 1983. Evaluation of birdsfoot trefoil regenerated plants and their progeny after *in vitro* selection for 2,4-dichlorophenoxyacetic acid tolerance. Plant Science Letters **29**:19–24.
19. Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
20. Van Haute, E., Joos, H., Maes, M., Warren, G., Van Montagu, M., and Schell, J. 1983. Intergenic transfer and exchange recombinations of restriction fragments cloned in pBR322: a novel strategy for the reversed genetics of the Ti-plasmid of *Agrobacterium tumefaciens*. EMBO J. **2**:411–417.
21. Norrander, J., Kempe, T., and Messing, J. 1983. Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene **26**:101–106.
22. Murashige, T., and Skoog, F. 1962. A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol. Plant **15**:473–497.
23. Sanchez-Serrano, J. J., Leil, M., O'Connor, A., Schell, J., and Willmitzer, L. 1987. Wound-induced expression of a potato proteinase inhibitor II gene in transgenic tobacco plants. EMBO J. **6**:303–306.
24. Maniatis, T., Fritsch, E. F., and Sambrook, J. 1982. Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
25. Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5468–5467.
26. Chen, E. U., and Seeburg, P. H. 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.
27. Dellaporta, S. L., Wood, J., and Hicks, J. B. 1983. Plant DNA mini-preparation: version II. Plant. Mol. Biol. Rep. **1**:19–21.
28. Southern, E. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.

COR-AAD1-00001134

ELSEVIER

FEMS Microbiology Letters 210 (2002) 19–23

FEMS MICROBIOLOGY Letters

www.fems-microbiology.org

Carnitine biosynthesis in *Neurospora crassa*: identification of a cDNA coding for ε-*N*-trimethyllysine hydroxylase and its functional expression in *Saccharomyces cerevisiae*

Jan H. Swiegers [a], Frédéric M. Vaz [b], Isak S. Pretorius [a], Ronald J.A. Wanders [b], Florian F. Bauer [a,*]

[a] *Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, Stellenbosch 7600, South Africa*

[b] *Laboratory for Genetic and Metabolic Diseases, Departments of Clinical Chemistry and Pediatrics, Academic Medical Center, University of Amsterdam, P.O. Box 22700, 1100 DE Amsterdam, Netherlands*

Received 25 January 2002; received in revised form 31 January 2002; accepted 12 February 2002

First published online 14 March 2002

Abstract

The biosynthesis of L-carnitine in eukaryotic organisms was first elucidated in the ascomycete *Neurospora crassa*. The first step of the pathway is catalysed by ε-*N*-trimethyllysine hydroxylase (TMLH), which converts ε-*N*-trimethyllysine into β-hydroxy-*N*-ε-trimethyllysine in a reaction dependent on α-ketoglutarate, Fe^{2+} and oxygen. Here we report on the cloning of the *N. crassa* TMLH cDNA and its functional expression in *Saccharomyces cerevisiae*. The TMLH cDNA contains an open reading frame of 1413 base pairs encoding a predicted polypeptide of 471 amino acids. The Michaelis–Menten constants of the heterologously expressed enzyme were determined for ε-*N*-trimethyllysine, α-ketoglutarate, Fe^{2+} and correspond to 0.33 mM, 133 μM and 46 μM, respectively. © 2002 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.

Keywords: Carnitine biosynthesis; ε-*N*-Trimethyllysine hydroxylase; *cbs-1*; Dioxygenase; L-Carnitine biosynthesis; Heterologous expression; *Neurospora crassa*; *Saccharomyces cerevisiae*

1. Introduction

L-Carnitine (3-hydroxy-4-*N*-trimethylaminobutyrate) is required for the transfer of activated acyl groups across intracellular membranes. In most eukaryotic systems, carnitine is an essential component of the mitochondrial carnitine cycle, which is responsible for the transfer of activated long-chain fatty acids into the mitochondria for β-oxidation [1]. In the yeast *Saccharomyces cerevisiae*, carnitine is essential for the transfer of activated acetyl groups from the peroxisome or the cytoplasm to the mitochondrion, since the peroxisomal, and presumably also the mitochondrial membranes are not permeable to acetyl-CoA [2]. Since *S. cerevisiae* does not synthesise carnitine, a strain with a non-functional glyoxylate cycle is dependent on extracellular carnitine for growth on non-fermentable carbon sources [3]. Most eukaryotes, on the other hand, are able to synthesise carnitine from ε-*N*-trimethyllysine (TML) [4–6]. TML is provided by the lysosomal hydrolysis of proteins that contain this amino acid as a result of the post-translational modification of lysine residues. In the first step of the carnitine biosynthesis, TML is hydroxylated to β-hydroxy-ε-*N*-trimethyllysine by ε-*N*-trimethyllysine hydroxylase (TMLH, EC 1.14.11.8) [5,7]. Subsequently, β-hydroxy-ε-*N*-trimethyllysine is cleaved into γ-trimethylaminobutyraldehyde and glycine by β-hydroxy-ε-*N*-trimethyllysine aldolase [5,7]. The aldehyde is then oxidised by γ-trimethylaminobutyraldehyde dehydrogenase to form γ-butyrobetaine [5,7,8]. Finally, γ-butyrobetaine is hydroxylated at the 3-position by γ-butyrobetaine hydroxylase to form L-carnitine [5,7,9].

The identity of the intermediate metabolites of the carnitine biosynthesis pathway was first elucidated in the filamentous fungus *Neurospora crassa*, using isotope-labelling experiments [4]. More recently, enzymes required for the catalysis of three of the four reactions required for

* Corresponding author. Tel.: +27 (21) 808 4346; Fax: +27 (21) 808 3771.

E-mail address: fb2@sun.ac.za (F.F. Bauer).

0378-1097/02/$22.00 © 2002 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.
PII: S0378-1097(02)00520-7

Downloaded from https://academic.oup.com/femsle/article/210/1/19/536156 by Pharmacy Hilfred Interlibrary loan user on 03 September 2020

COR-AAD1-00001135

carnitine biosynthesis have been characterised at the molecular level, either in rat, mouse or man [10–13]. However, none of the *N. crassa* carnitine biosynthesis enzymes has so far been characterised at the molecular level. Here we report the identification of the *N. crassa* cDNA encoding TMLH, the first enzyme of carnitine biosynthesis. TMLH is a non-haem ferrous–iron dioxygenase that requires α-ketoglutarate, Fe^{2+} and molecular oxygen as cofactors [14–17]. Based on homology with the recently identified rat TMLH, the putative TMLH open reading frame (ORF) was amplified by the polymerase chain reaction (PCR) from a cDNA library, and cloned into an *S. cerevisiae* expression vector. High TMLH activity could be measured in lysates of transformed yeast, allowing for the determination of the kinetic constants of the enzyme for ε-*N*-trimethyllysine, α-ketoglutarate, and Fe^{2+}.

2. Materials and methods

2.1. Cloning and expression of the N. crassa TMLH in S. cerevisiae

The predicted coding sequence of the hypothetical protein CAC18241 (GenBank™), which is derived from the *N. crassa* genome sequence of DNA linkage group V, Cosmid contig 99H12 position 13250–15449, was used to design nucleotide primers NC-TMLH-F (5′-gatcgaattc **ATG** AGA CCG CAA CGG GTA GGG GCA-3′) and NC-TMLH-R (5′-gatcgaattc **TCA** TTT TCC GCT GGT TTC TTT CCG-3′). These primers were used to amplify the putative ORF from a *N. crassa* cDNA yeast expression library (Fungal Genetic Stock Center, Kansas City, KS, USA) using the PCR technique [18]. Using the intro-

Downloaded from https://academic.oup.com/femsle/article/210/1/19/631156 by Pharmacy Hill Read International Inc user on 03 September 2020

```
Ncrassa_TMLH      1 MRPQVVGAILRSRAVVSRQPLSRTHIFAAVTVAKSSSPAQNSRRTFSSSF
Rat_TMLH          1 ---------MKRGDIAHGLRLSGFKSLFPFSLHWCHTASKSVNCTWHQHE
consensus         1           ..   .     **  .     ..    .       *.

Ncrassa_TMLH     51 RRLYEPKAEITAEGLELSPPQAVTGGKRTVLPNFWLRDNCRCTKCVNQDT
Rat_TMLH         42 DKLELQYASTVMR-----------------FDYVWLRDHCRSASCYNSKT
consensus        51  .*    *                          **** **   * *  *

Ncrassa_TMLH    101 LQRNFNTTFAIPSDIHPTKVEATKENVTVQWSDNHTSTYPWPFLSFYLTSN
Rat_TMLH         75 HQRSLDTASVDLCIKPKTIRLDESTLFFTWPDGHVTRYDLDWLVKNSYEG
consensus       101  **   *  .   *.*  .      .   * * * . *   .*

Ncrassa_TMLH    151 ARGHENDQISLWGSEAGS--RPPTVPFPRVMASDQGVADLTAMIKEPGFC
Rat_TMLH        125 QKQEVIQPRVLWNAKLYQDAQLPSVDFQGFLETKEGLKKFLQNFLLYCIA
consensus       151  .         **           *.* *   . .  *.        .*

Ncrassa_TMLH    199 FVKDTPHDDPDVTRQLLERIAFIRVTHYGGFYDFTPDLAMADTAYTNLAL
Rat_TMLH        175 FVENVPP-TQEHTEKLARRVSLIRETIYGRMWYFTSDFSRGDTAYTKLAL
consensus       201 **   * .  . *  *  *.  ** * **  . ** *   .***** ***

Ncrassa_TMLH    249 PAHTDTTYFTDPAGLQAFHLLEHKAAPSRPPPPPPPPPPPSEEKEAAGSA
Rat_TMLH        224 DRHTDTTYFQEPCGIQVFHCLKHEGT-------------------------
consensus       251   ******* .* *.* ** * * .

Ncrassa_TMLH    299 AGEAAAAAAEGCKSLLVDGFNAARILKEEDPRAYEILSSVRLPWHASGNEG
Rat_TMLH        251 ---------GGRTLLVDGFYAAQQVLQRAPEEFDLLSQVPLKHEYIENVG
consensus       301           **..****** **  .    *  ...** * *      * *

Ncrassa_TMLH    349 ITIAPDK-LYPVLELNEDTGELHRVRWNNDDRGVVPFGEKYSPSEWYEAA
Rat_TMLH        291 QCHNHMIGVGPILNIYPWNKELYLIRYNNYDRAVINTVPYDVVRRWYAAH
consensus       351           . *.* .     **  .*.** **.*.          ** *

Ncrassa_TMLH    398 RKWDGILRRKSSELWVQLEPGKPLIFDNWRVLHGRSAFSGIRRICGGYIN
Rat_TMLH        341 RTLTTELRRPENELWVKLKPGKVLFIDNWRVLHGRESFTGYRQLCGCYLT
consensus       401 *      ***    **** * *** *  *********  *.* * .** *.

Ncrassa_TMLH    448 RDDFISRWRNTNYPRSEVLPRVTG
Rat_TMLH        391 RDDVLNTARILGLHA---------
consensus       451 *** .    *
```

Fig. 1. Alignment of TMLH amino acid sequences. The deduced protein sequence encoded by the *N. crassa* TMLH-encoding cDNA (EMBL accession No. AJ421151) was aligned with the sequence of the rat TMLH protein. The *N. crassa* sequence contains an unusual, 11-residue poly(P) region (residues 278–288) in the TMLH protein.

COR-AAD1-00001136

duced *Eco*RI sites, the PCR product was cloned downstream from the constitutive *PGK1* promoter of the yeast expression vector pHVXII [19]. The correct orientation of the ORF was confirmed by restriction analysis. The construct was transformed into the yeast strain FY23Δcit2 [3] using the lithium acetate procedure [20]. Transformed yeast cells were grown on minimal glucose medium (6.7 g l^{-1} yeast nitrogen base, 20 g l^{-1} glucose and amino acids as required at 20 mg l^{-1}). Spheroplasts were prepared using zymolyase according to Franzusoff et al. [21], lysed in a 10 mM sodium phosphate buffer, pH 7.4, containing 140 mM NaCl, 1 mM DTT and 1 g l^{-1} Triton X-100 and TMLH activity was measured as described below.

2.2. TMLH assay

The amount of β-hydroxy-ε-*N*-trimethyllysine that was enzymatically produced from ε-*N*-trimethyllysine was determined by HPLC tandem MS as described by Vaz et al. [13]. For the measurement of the dependence of TMLH on ε-*N*-trimethyllysine, α-ketoglutarate and Fe^{2+}, only the concentration of these components was varied. To obtain a reliable K_m value, the enzyme sample was diluted so that the substrate consumption was below 10%.

3. Results and discussion

3.1. Cloning and sequence analysis of N. crassa cDNA

A putative *N. crassa* TMLH-encoding gene was identified by database searches based on the homology of the derived protein sequence (CAC18241) to known mammalian TMLHs. The predicted coding sequence of this hypothetical protein was used to design primers overlapping with the putative start and stop codons. Using a *N. crassa* cDNA library as template, a PCR product of 1.6 kb was amplified. Sequencing of the product revealed a smaller than expected ORF of 1413 bp, encoding a protein of 471 amino acids, which corresponds to a calculated molecular mass of 52.6 kDa (EMBL accession No. AJ421151). Alignment with the genomic sequence indicated that the coding sequence consists of eight exons, and that the coding sequence of CAC18241 had been predicted incorrectly, with several intron splice sites having been missed or miss-assigned. This also explains the smaller than expected size of the ORF encoded by the cDNA, with primer NC-TMLH-R hybridising in fact to the sequence 206–226 bp downstream from the newly identified stop codon. The protein sequence derived from the cDNA also aligns better with published TMLH protein sequences than the previously suggested sequence. The protein indeed has high homology to the three confirmed TMLH proteins identified in rat, mouse and human [13] and also to the last enzyme of carnitine biosynthesis, γ-butyrobetaine hydroxylase (BBH), which is also a dioxygenase. The positional identity to mouse and human TMLH is 33%, and 32% to rat TMLH (Fig. 1). The protein contains a unique and unusual 11 residue poly(P) region (residues 278–288). This region could separate two protein domains and should have a major effect on the tertiary structure of the enzyme. X-Ray crystallisation studies will be pursued to elucidate the effect of this region. No significant homologies to other dioxygenase enzymes were found, indicating that the TMLH and BBH carnitine biosynthesis enzymes form a separate group of highly conserved enzymes.

3.2. Expression of N. crassa TMLH in S. cerevisiae

To confirm that the identified sequence indeed encodes a TMLH, the ORF was cloned into the multiple copy ex-

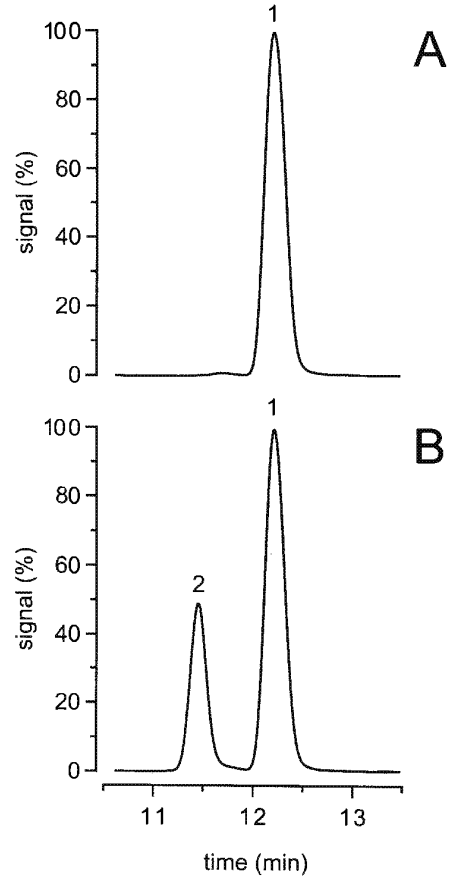


Fig. 2. Detection of β-hydroxy-ε-*N*-trimethyllysine by HPLC tandem MS. Levels of ε-*N*-trimethyllysine and β-hydroxy-ε-*N*-trimethyllysine, the substrate and product of TMLH activity, respectively, were measured after 60 min incubation in cellular extracts. Peak 1 corresponds to ε-*N*-trimethyllysine and peak 2 to β-hydroxy-ε-*N*-trimethyllysine. A: Extract from *S. cerevisiae* strain transformed with vector pHVXII without insert. B: Extract from *S. cerevisiae* strain transformed with vector pHVXII containing inserted TMLH ORF. The trace represents the total ion current of the transitions *m/z* 247 > 142 (peak 1, ε-*N*-trimethyllysine) and *m/z* 263 > 158 (peak 2, β-hydroxy-ε-*N*-trimethyllysine).

Downloaded from https://academic.oup.com/femsle/article/210/1/19/631156 by Pharmacy Hilfred International Inc user on 03 September 2020

COR-AAD1-00001137

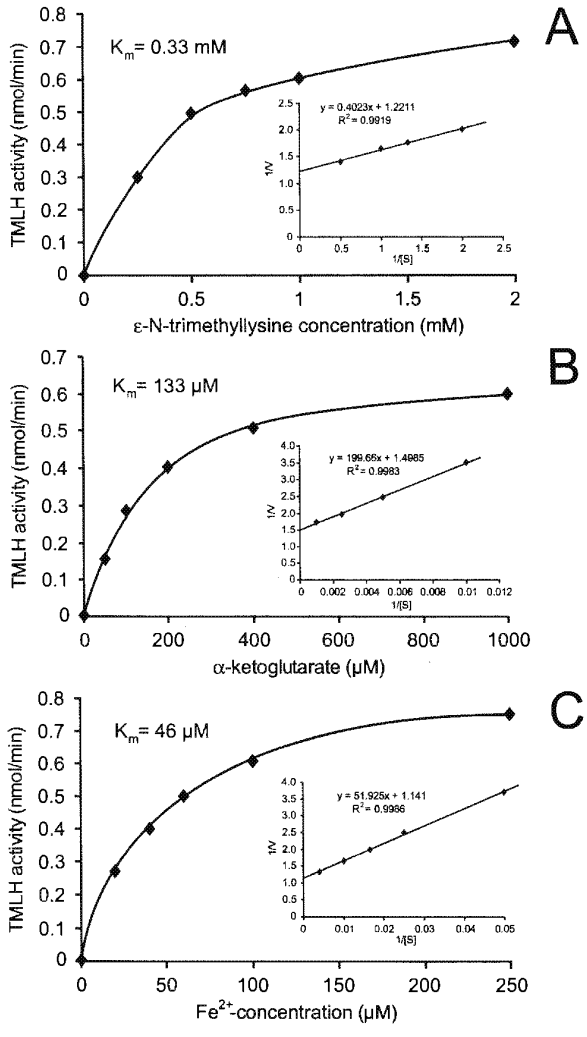


Fig. 3. K_m curve and Lineweaver–Burk plot for (A) ε-*N*-trimethyllysine, (B) α-ketoglutarate, and (C) Fe^{2+}.

pression vector pHVXII between the constitutive *PGK1* promoter and the *PGK1* terminator, and transformed into strain FY23Δcit2. TMLH activity measurement in the homogenate of a yeast transformant containing the putative TMLH cDNA revealed high TMLH activity, whereas the strain transformed with the pHVXII vector without insert showed no activity (Fig. 2). The gene was subsequently named *cbs-1* for *c*arnitine *b*iosynthesis. Contrary to rat TMLH, heterologously expressed *N. crassa* TMLH is therefore active in yeast. Rat TMLH is localised in the mitochondria and undergoes post-translational processing, probably after import into the mitochondria [13], whereas *N. crassa* TMLH is a cytosolic enzyme [4]. These differences in post-translational modification and localisation may explain the difference in activity in *S. cerevisiae*.

3.3. Characterisation of heterologously expressed *N. crassa* TMLH

The K_m values for the substrate ε-*N*-trimethyllysine, as well as the co-factors α-ketoglutarate and Fe^{2+}, were determined using Lineweaver–Burk double-reciprocal plots. The K_m values for the three compounds were found to be 0.33 mM, 133 μM and 46 μM, respectively (Fig. 3). The K_m value for ε-*N*-trimethyllysine is significantly lower than the previously reported values of 1.1 mM [13] and 1.6 mM [17] for rat TMLH. On the other hand, the K_m values for α-ketoglutarate and Fe^{2+} in rat TMLH, which were determined to be 109 μM and 54 μM, respectively, are similar to those reported here [13].

Acknowledgements

The authors thank the South African Medical Research Council and the National Research Foundation of South Africa for financial support.

References

[1] Bieber, L.L. (1988) Carnitine. Annu. Rev. Biochem. 57, 261–283.
[2] Van Roermund, C.W.T., Hettema, E.H., Van den Berg, M., Tabak, H.F. and Wanders, R.J.A. (1999) Molecular characterization of carnitine-dependent transport of acetyl-CoA from peroxisomes to mitochondria in *Saccharomyces cerevisiae* and identification of a plasma membrane carnitine transporter, Agp2p. EMBO J. 18, 5843–5852.
[3] Swiegers, J.H., Dippenaar, N., Pretorius, I.S. and Bauer, F.F. (2001) Carnitine-dependent metabolic activities in *Saccharomyces cerevisiae*: three carnitine acetyltransferases are essential in a carnitine-dependent strain. Yeast 18, 585–595.
[4] Kaufman, R.A. and Broquist, H.P. (1977) Biosynthesis of carnitine in *Neurospora crassa*. J. Biol. Chem. 252, 7437–7439.
[5] Bremer, J. (1983) Carnitine-metabolism and functions. Physiol. Rev. 63, 1420–1480.
[6] Lindstedt, G. and Lindstedt, S. (1970) γ-Butyrobetaine hydroxylase from *Pseudomonas* sp. AK 1. J. Biol. Chem. 245, 4178–4186.
[7] Rebouche, C.J. and Engel, A.G. (1980) Tissue distribution of carnitine biosynthetic enzymes in man. Biochim. Biophys. Acta 630, 22–29.
[8] Hulse, J.D. and Henderson, L.M. (1980) Carnitine biosynthesis. Purification of 4-N′-trimethylaminobutyraldehyde dehydrogenase from beef liver. J. Biol. Chem. 255, 1146–1151.
[9] Englard, S. (1979) Hydroxylation of γ-butyrobetaine to carnitine in human and monkey tissues. FEBS Lett. 102, 297–300.
[10] Galland, S., Le Borgne, F., Bouchard, F., Georges, B., Clouet, P., Grand-Jean, F. and Demarquoy, J. (1999) Molecular cloning and characterization of the cDNA encoding the rat liver γ-butyrobetaine hydroxylase. Biochim. Biophys. Acta 1441, 85–92.
[11] Vaz, F.M., van Gool, S., Ofman, R., IJlst, L. and Wanders, R.J. (1998) Carnitine biosynthesis: identification of the cDNA encoding human γ-butyrobetaine hydroxylase. Biochem. Biophys. Res. Commun. 250, 506–510.
[12] Vaz, F.M., Fouchier, S.W., Ofman, R., Sommer, M. and Wanders, R.J.A. (2000) Molecular and biochemical characterization of rat γ-trimethylaminobutyraldehyde dehydrogenase and evidence for the

Downloaded from https://academic.oup.com/femsle/article/210/1/19/536156 by guest on 03 September 2022

COR-AAD1-00001138

involvement of human aldehyde dehydrogenase 9 in carnitine biosynthesis. J. Biol. Chem. 275, 7390–7394.
[13] Vaz, F.M., Ofman, R., Westinga, K., Back, J.W. and Wanders, R.J.A. (2001) Molecular and biochemical characterization of rat ε-N-trimethyllysine hydroxylase, the first enzyme of carnitine biosynthesis. J. Biol. Chem. 276, 33512–33517.
[14] Hulse, J.D., Ellis, S.R. and Henderson, L.M. (1978) Hydroxylation of γ-butyrobetaine to carnitine in human and monkey tissues. J. Biol. Chem. 253, 1654–1659.
[15] Sachan, D.S. and Broquist, H.P. (1980) Synthesis of carnitine from ε-N-trimethyllysine in post mitochondrial fractions of *Neurospora crassa*. Biochem. Biophys. Res. Commun. 96, 870–875.
[16] Stein, R. and Englard, S. (1982) Properties of rat 6-N-trimethyl-L-lysine hydroxylases: similarities among the kidney, liver, heart, and skeletal muscle activities. Arch. Biochem. Biophys. 217, 324–331.
[17] Sachan, D.S. and Hoppel, C.L. (1980) Carnitine biosynthesis. Hydroxylation of N6-trimethyl-lysine to 3-hydroxy-N6-trimethyl-lysine. Biochem. J. 188, 529–534.
[18] Brunelli, J.P. and Pall, M.L. (1993) A series of yeast/*Escherichia coli* lambda expression vectors designed for directional cloning of cDNAs and cre/lox-mediated plasmid excision. Yeast 9, 1309–1318.
[19] Volschenk, H., Viljoen, M., Grobler, J., Petzold, B., Bauer, F.F., Subden, R.E., Young, R.A., Lonvaud, A., Denayrolles, M. and Van Vuuren, H.J.J. (1997) Engineering pathways for malate degradation in *Saccharomyces cerevisiae*. Nat. Biotechnol. 15, 253–257.
[20] Becker, D.M. and Guarente, L. (1991) In: Guide to Yeast Genetics and Molecular Biology (Guthrie, C. and Fink, G.R., Eds.), pp. 182–187. Academic Press, San Diego, CA.
[21] Franzusoff, A., Rothblatt, J. and Schekman, R. (1991) In: Guide to Yeast Genetics and Molecular Biology (Guthrie, C. and Fink, G.R., Eds.), pp. 662–674. Academic Press, San Diego, CA.

Downloaded from https://academic.oup.com/femsle/article/210/1/19/1539139 by Pharmacia Hall Library user on 03 September 2022

COR-AAD1-00001139

FEMS Microbiology Ecology 14 (1994) 191–200
© 1994 Federation of European Microbiological Societies 0168-6496/94/$07.00
Published by Elsevier

FEMSEC 00532

Enantioselective degradation of the herbicide mecoprop [2-(2-methyl-4-chlorophenoxy) propionic acid] by mixed and pure bacterial cultures

V.A. Tett, A.J. Willetts and H.M. Lappin-Scott *

Department of Biological Sciences, University of Exeter, Exeter, Devon EX4 4PS, UK

(Received 14 December 1993; revision received and accepted 28 March 1994)

Abstract: A consortium of three bacteria was isolated from top soil through their capacity to utilise the chlorinated, aromatic herbicide mecoprop as a single growth substrate. The consortium constituted a tight association of *Alcaligenes denitrificans*, *Pseudomonas glycinea* and *Pseudomonas marginalis*. The culture exclusively degraded the (*R*)-(+)-isomer of the herbicide while the (*S*)-(−)-enantiomer remained unaffected. The mecoprop-degrading community could also degrade 2,4-dichlorophenoxyacetic acid, 2-methyl-4-chlorophenoxyacetic acid and racemic 2-phenoxypropionic acid. Initially, no single member of the consortium was able to degrade mecoprop as a pure culture but after prolonged incubation, *A. denitrificans* was able to grow on the herbicide as the sole source of carbon and energy.

Key words: Biodegradation; [2-(Methyl-4-chlorophenoxy)propionic acid]; Mecoprop; Stereoisomerism; Microbial community; Bioremediation

Introduction

Chlorinated, phenoxyalkanoic compounds are xenobiotic pollutants which pose a serious threat to the environment. Consequently there has been extensive interest in the environmental persistence, biodegradability and biological activity of this group of synthetic organic compounds [1–6]. The widely used phenoxyacetic herbicides 2,4-dichlorophenoxyacetic acid (2,4-D) and 4-chloro-2-methyl-phenoxypropionic acid (MCPA) are readily degraded in soil by microorganisms. However, relatively few studies have been published on the microbial degradation of the structurally related herbicide mecoprop [2-(2-methyl-4-chlorophenoxy)propionic acid] [7–12]. Mecoprop is a systemic, post-emergence herbicide used extensively for the control of broad-leaved weeds in cereal crops, lawns and turf [13]. Alexander and Aleem [14] reported limited degradation of mecoprop in laboratory batch cultures containing soil as an inoculum. Kilpi has described degradation of mecoprop by mixed cultures isolated from soil, but only when benzoic acid was added as a co-substrate [8]. Mecoprop degradation has also been investigated by Smith [15], who demonstrated the

* Corresponding author. Tel: (0392) 263780; Fax: (0392) 263700; email: HMLappin@cen.ex.ac.uk.

SSDI 0168-6496(94)00015-O

COR-AAD1-00001140

Downloaded from https://academic.oup.com/femsec/article/14/3/191/6508684 by Oxford Academic user on 05 September 2023

breakdown of mecoprop in three different experimental soils but without any analysis of the microorganisms involved. There have been no reports of pure culture degradation with mecoprop as a sole carbon and energy source and only a limited number of studies of mixed culture degradation which for the most part, used undefined microbial communities.

Unlike 2,4-D, mecoprop possesses a methyl substituent on the aromatic ring and a propionic acid side chain which affects the susceptibility of the aromatic herbicide to microbial degradation [14]. The presence of a chiral carbon atom in the aliphatic side chain gives two stereo-isomeric forms of mecoprop with only the (*R*)-(+)-enantiomer possessing herbicidal properties [13]. For many of the herbicidal compounds which have stereogenic centres, there is generally little understanding of the relationship between stereo-isomerism and both biological activity and biodegradability [16].

The ability of microorganisms to distinguish between different isomers of chiral compounds is a reported characteristic of the degradation of some xenobiotic haloaliphatic compounds such as 3-chloro-2-methylpropionic acid (CMPA) by *Xanthobacter* species CIMW99 [17]. Despite the increasing significance of stereo-isomerism in the manufacture of herbicides there have been very few investigations undertaken on racemic haloaromatics. Ludwig and co-workers reported the stereo-selective degradation of 2-(2,4-dichlorophenoxy) propionic acid by an uncharacterised marine microbial community. However, the microbial culture was not able to degrade mecoprop [18].

The biodegradation of 2,4-D and MCPA has been well characterised, principally through the use of pure culture investigations using the herbicides as single substrates [3,5,19,20]. To date, researchers have been unable to isolate a pure culture degrader of mecoprop. This has hindered investigations of its mineralisation including characterisation of the degradative pathway of the herbicide. It is possible that in natural environments the complete degradation of mecoprop can only occur through the synergistic activities of a consortium of microorganisms [21]. This paper describes a novel laboratory study of mecoprop degradation by both pure and mixed microbial cultures. This is an indicator of the potential of microorganisms to degrade mecoprop in situ. This work also includes an investigation of the ability of a microbial community to distinguish between the (±) stereo-isomers of mecoprop. Finally, we report the ability of the consortium to degrade structurally related compounds that are often introduced into the environment as mixed formulations with mecoprop to maximise weed control [22].

Materials and Methods

Herbicides

Technical grade mecoprop, as both the free acid and as the formulation Compitox Extra, was donated by Rhone-Poulenc Agriculture Ltd. (Ongar, Essex, UK). Mecoprop, as Farmon CMPP, was donated by ICI Agrochemicals (Leadenham, Lincolnshire, UK). The chemical structure of mecoprop is depicted by Heron and Christensen [7]. MSS Optica, a solution of the potassium salt of the (*R*)-(+)-enantiomer of mecoprop, was donated by Pennine Chemical Services Ltd. (Huddersfield, West Yorkshire, UK). Documented mecoprop (99.8% pure) was obtained as the free acid from Promochem Ltd. (St. Albans, Hertfordshire, UK). MCPA, 2,4-D and 2-phenoxypropionic acid (PPA) were all obtained as free acids, MCPA and 2,4-D supplied by Sigma Chemicals Co. Ltd. (Poole, Dorset, UK), and PPA from Aldrich Chemical Co. Ltd. (Gillingham, Dorset, UK).

The (*S*)-(−)-enantiomer of mecoprop is not commercially available. It was resolved by transformation of the racemate into its respective phenylethylamine diastereoisomeric salts using a modification of the procedure of Matell [23]. Crystals of the phenylamine salt of (*S*)-(−)-mecoprop were then removed from the reaction mixture, on the basis of their relative insolubility. Equimolar quantities of the racemic acid (28.2 g, 0.13 mol) and (−)-phenylethylamine (16.0 g) were dissolved in a hot mixture of ethanol (100 ml) and water (500 ml). The solution was left to crystallise

COR-AAD1-00001141

Downloaded from https://academic.oup.com/femsec/article/13/3/191/5850694 by Cochrane Argentina user on 02 September 2022

at 5°C overnight. The phenylethylamine salt of (*S*)-(−)-2-methyl-4-chloro-phenoxypropionic acid was recrystallised until there was no further increase in optical activity. The organic acid was isolated by acidifying with 2 M HCl and extracting into ether which was then evaporated. The crude acid was then recrystallised from 75% formic acid to produce small white needles. The product had a mean optical rotation in ethanol of $-21.6° \pm 0.02°$ at 23°C which corresponds to the expected value for mecoprop [23].

Isolation and growth of microorganisms

A consortium of bacteria was isolated from garden soil by batch enrichment culture with racemic mecoprop as the sole carbon and energy source. The growth medium contained (g l^{-1} of distilled water): $MgSO_4 \cdot 7H_2O$, 0.2; $(NH_4)_2SO_4$, 0.5; KH_2PO_4, 0.5; K_2HPO_4, 1.5; NaEDTA, 0.12; NaOH, 0.02; $ZnSO_4 \cdot 7H_2O$, 0.004; $CuSO_4 \cdot 5H_2O$, 0.001; Na_2SO_4, 0.0001; $Na_2MoO_4 \cdot H_2O$, 0.001; $CoCl_2 \cdot 6H_2O$, 0.0001; $MnSO_4 \cdot 4H_2O$, 0.0004; and 0.5 ml of concentrated H_2SO_4. The pH of the medium was adjusted to 7.0. Mecoprop was added to give a concentration of 0.56 g carbon l^{-1} (1 g mecoprop l^{-1}; 4.66 mM). Solutions of chlorinated compounds were filter-sterilised to prevent thermal dechlorination. Batch cultures, set up to contain 100 ml culture in 250 ml flasks, were aerated by shaking on an orbital shaker at 28°C and at 120 rpm. Growth was measured by determining the culture absorbance at 420 nm using a Cecil 1020 spectrophotometer. For batch culture enrichments, subcultures were made when cells had reached late log phase. Appropriate dilutions of the culture were spread onto nutrient agar and minimal selective agar of the following composition: $MgSO_4 \cdot 7H_2O$, 0.2; $(NH_4)_2SO_4$, 0.5; KH_2PO_4, 0.5; K_2HPO_4, 1.5; and mecoprop at a concentration of 1 g l^{-1}. The plates were incubated at 30°C for 48 and 96 h respectively. Single colonies were isolated by repeated streaking on nutrient agar. Initially, the constituent members of the enriched communities were recognised on the basis of comparitive colony morphology. The isolates were subsequently identified using the Biolog system of identification (Biolog Inc., Hayward, USA).

Mecoprop degradation

The degradation of mecoprop by soil bacteria was measured directly by monitoring the change in the maximum absorption peak (A_{279}) in a scanning UV spectrophotometer with time. The culture supernatant was scanned from 200–500 nm to determine both the breakdown of mecoprop and the accumulation of biotransformation products. Supernatants were suitably diluted and read against a standard curve to determine their relative herbicide concentration. A decrease in A_{279} indicated breakdown of mecoprop and a shift in the λ max indicated transformation of the compound.

Chloride ion determination

Samples (6 ml) of culture removed during growth on mecoprop were assayed for chloride ion release using a Phillips PW 9413 Ion Activity Meter fitted with a Phillips chloride ion selective electrode. The meter was calibrated with a standard solution of KCl (Ciba Corning Diagnostics Ltd.). A paired samples *t*-test was conducted using Mystat: Statistical Applications (Course Technology Inc., Cambridge, UK) to determine whether the rate of chloride ion release differed from the rate of increase in culture absorbance at 420 nm. Results were considered significant when $P = < 0.005$.

Ability to degrade mecoprop after successive passages through nutrient broth

In order to determine whether mecoprop degradation was lost during non-selective growth, a 2 ml sample of bacteria grown on mecoprop was inoculated into 50 ml of nutrient broth and incubated as described above. The community was subcultured into fresh nutrient broth when culture turbidity indicated exponential growth. After twenty passages through nutrient broth, the consortium was reinoculated into mecoprop. Growth on mecoprop was assessed by determining the increase in culture absorbance at 420 nm.

Cross-adaptation

The mecoprop-enriched community was reinoculated (2 ml) into 100 ml of growth medium containing an alternative phenoxy herbicide as

Downloaded from https://academic.oup.com/femsec/article/4/3/191/5850694 by Cochrane Argentina user on 05 September 2022

COR-AAD1-00001142

the sole carbon source. MCPA and 2,4-D were added to the medium in separate batch cultures to give a final concentration of 1 g herbicide l^{-1}, corresponding to 0.54 and 0.43 g carbon l^{-1} respectively. Growth on (±) PPAwas also investigated. PPA was added to the growth medium to give a comparable concentration of total carbon to that used for growth on mecoprop. The ability of the consortium to degrade the structurally related compounds was investigated under previously described batch culture conditions.

Continuous-flow culture growth

An inoculum (2 ml) of a batch enrichment culture was grown in chemostat culture. The chemostat culture volume was 425 ml and was agitated at 200 rpm and aerated at the rate of 500 ml min^{-1}. The temperature was controlled at 30°C, and fresh sterile mecoprop based medium was transferred to the culture vessel with a Biotec LP 100–4 peristaltic pump to give a dilution rate of 0.024 h^{-1}. The experiment was run for 83 days, and at daily intervals the culture absorbance (A_{420}) and the change in colony forming units were used to assess the change in community growth and stability with time.

Results

An inoculum taken from garden soil was incubated in batch culture using mecoprop as the sole source of carbon and energy. Progressive enrichment resulted in the isolation of a consortium of three bacteria capable of growth on racemic mecoprop. The consortium comprised *Alcaligenes denitrificans*, *Pseudomonas glycinea* and *Pseudomonas marginalis*. The organisms formed a tight association and because of their similar nutritional characteristics were difficult to separate. Growth of the consortium was monitored on the racemic herbicide and both the (*R*)-(+)- and (*S*)-(−)-enantiomers in independent experiments using an equivalent amount of total carbon for comparative purposes. During growth on the racemate, onset of chloride ion release occurred after a lag phase of 8 h (Fig. 1). Degradation of the aryl moiety as monitored by the stoichiometric release of chloride ions took 37 h to complete. The final accumulated chloride ion concentration equalled half the maximum calculated value for the complete degradation of the supplied racemic mecoprop. Growth of the consortium supported by the (*R*)-(+)-enantiomer followed a similar time course (Fig. 1), although in this case the final chloride ion concentration equalled the theoretical maximum present from the added herbicide. When growth on the (*R*)-(+)-enantiomer and racemate of mecoprop had ceased, the maximum culture absorbance was 0.95 and 0.62 respectively which corresponded to 0.63 and 0.37 g l^{-1} wet weight biomass. In both types of media, there was no significant difference between the maximum specific growth rate and release of chloride ions ($P = 0.167$) and the maximum population phase was reached at the time chloride ion release was completed. There was no growth of the consortium on the (*S*)-(−)-enantiomer and no subsequent release of chloride ions (Fig. 1). Autoclaved, uninoculated growth media containing mecoprop was used as a chloride ion release control to demonstrate there was no abiotic transformation of the herbicide.

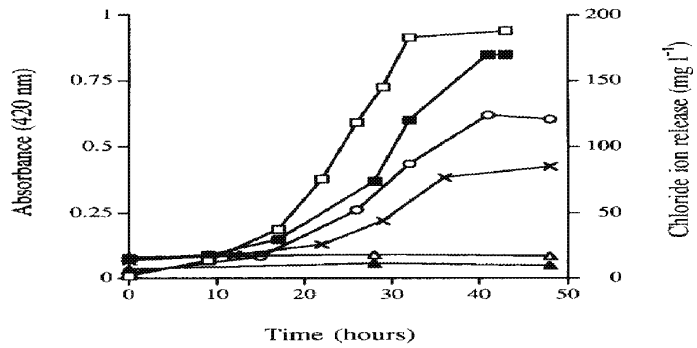


Fig. 1. Growth of the consortium on the racemate and individual (*R*)-(+)- and (*S*)-(−)-enantiomers of mecoprop (0.56 g carbon l^{-1}). Symbols: ○, culture absorbance and ×, chloride ion release from racemic mecoprop (mg chloride l^{-1}); □, culture absorbance and ■, chloride ion release from (*R*)-(+)-mecoprop (mg chloride l^{-1}); ▲, culture absorbance and △, chloride ion release from (*S*)-(−)-mecoprop (mg chloride l^{-1}).

COR-AAD1-00001143

Downloaded from https://academic.oup.com/femsec/article/4/3/191/5850884 by Cochrane Argentina user on 08 September 2020

Spectrophotometric scans of culture supernatants taken from both types of media were determined with time over the range 200–500 nm (Fig. 2). During growth and chloride ion release, there was a concomitant decrease in the size of the 279 nm peak, probably demonstrating the disruption of the conjugated aryl moiety in the primary mecoprop structure [24]. During growth on the (*R*)-(+)-enantiomer, the decrease in A_{279} corresponded to a reduction in concentration from 1 g l^{-1} to 0.03 g l^{-1}. This represents a 97% reduction in the initial herbicide concentration as compared with only a 55% decrease in mecoprop concentration during growth on the racemate. The absorption maximum of media containing the racemate and the single isomer both remained at 279 nm until after 14 h of exponential growth when it shifted to 286 nm and 291 nm respectively. After 35–40 h into the growth cycle, there was a noticeable decrease in UV absorption of the culture supernatant formulated with the (*R*)-(+)-enantiomer. The racemic carbon source showed no equivalent change and the residual mecoprop in this culture remained stable on extended incubation. The decrease in the size of the 279 nm peak during incubation of the consortium on the (*S*)-(−)-enantiomer was neglible

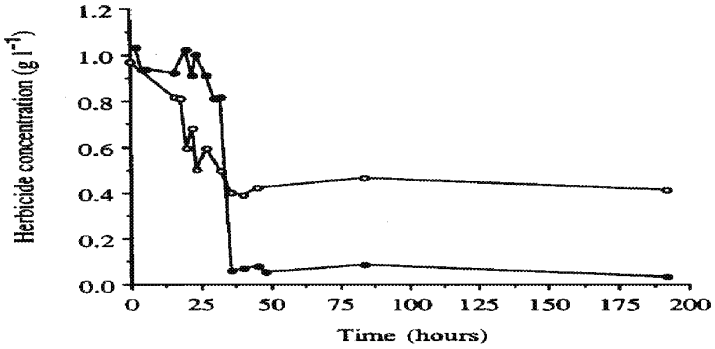


Fig. 2. Degradation of the (*R*)-(+)-enantiomer and the racemate of mecoprop by the three-member microbial community using an initial herbicide concentration of 1 g l^{-1}. Symbols: ●, (*R*)-(+)-enantiomer; ○, racemic mecoprop.

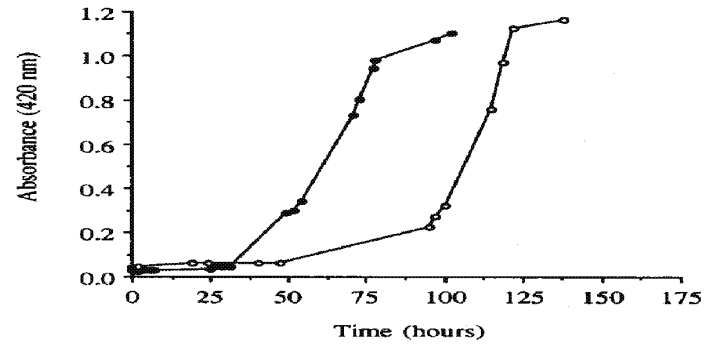


Fig. 3. Growth of the three-member microbial community on 2,4-D (○) and MCPA (●) each at 0.43 g carbon l^{-1}.

confirming chloride ion and culture absorbance analyses that it could not support growth.

The three-member community was also able to grow on MCPA and 2,4-D as sole carbon and energy sources (Fig. 3). Growth commenced after a lag period of 32 h on MCPA and 47 h on 2,4-D. There was also evidence of very slow growth on racemic PPA. For adaptation to this latter substrate, however, there was a prolonged lag phase of 10 days. When the PPA-grown culture had reached the maximum population phase, the culture absorbance (A_{420}) was 0.65. This corresponded closely to the maximum absorbance obtained from growth on racemic mecoprop (A_{420} = 0.62) at equivalent concentrations of total carbon.

In closed culture, the proportion of the community formed by each member varied with the age of the culture (Fig. 4). Typically a culture had the following composition after the completion of mecoprop degradation: *A. denitrificans* (2.2×10^9 cfu ml^{-1}), *P. marginalis* (2.4×10^7 cfu ml^{-1}) and *P. glycinea* (2.6×10^5 cfu ml^{-1}). The consortium remained able to degrade mecoprop after twenty passages through nutrient broth. After this period of non-selective culture, growth on mecoprop required 72 h for completion and there was a slight increase in the lag phase. Samples were periodically removed to test for pure culture degradation of mecoprop. Initially, none of the isolates was

COR-AAD1-00001144

Downloaded from https://academic.oup.com/femsec/article/4/3/191/5855684 by Cranfield Institute of Technology user on 03 September 2020

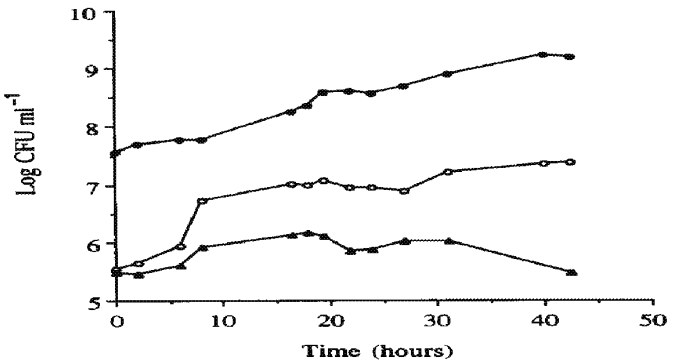


Fig. 4. Growth of community constituents on (*R*)-(+)-mecoprop (0.56 g carbon l^{-1}) in batch culture. Symbols: ●, *Alcaligenes denitrificans*; ○, *Pseudomonas marginalis*; ▲, *Pseudomonas glycinea*.

able to degrade mecoprop in pure culture or in combination with any other single member of the consortium. However, after five months of subculturing and after approximately 80 days of chemostat growth, *A. denitrificans* acquired the complete capacity to degrade mecoprop as a single substrate. The isolate from the batch culture was chosen for further study. It formed recognisable colonies on mecoprop selective agar. The ability of *A. denitrificans* to grow with mecoprop as a sole source of carbon and energy is shown in Fig. 5. Growth, expressed as an increase in culture absorbance, occurred after a lag phase of 5 h and was associated with a drop in pH from 7.1 to 6.4. Release of chloride ions was initiated 12 h into the growth cycle. The doubling time of the organism during the exponential growth phase in batch culture was 6.5 h. There was no significant difference between growth rates of the consortium and pure culture ($P = 0.184$).

In chemostat culture with (*R*)-(+)-mecoprop at a growth limiting concentration of 0.56 g of carbon l^{-1} and a dilution rate of 0.024 h^{-1}, the three-member community was maintained as a stable association for over two months of continuous growth (Fig. 6). After an initial fall in cfu, possibly attributed to wash-out of the inoculum,

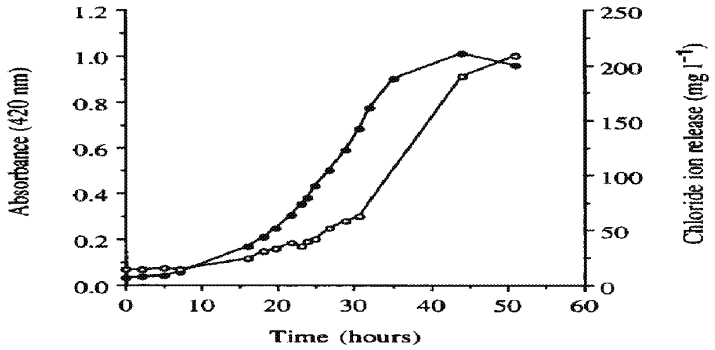


Fig. 5. Growth of *Alcaligenes denitrificans* on the (*R*)-(+)-enantiomer of mecoprop (0.56 g carbon l^{-1}). Symbols: ●, culture absorbance; ○, chloride ion release (mg chloride l^{-1}).

all three isolates stabilised with *A. denitrificans* remaining the dominant organism in the culture. The absorbance of the culture was constant at 0.74 ± 0.04, a value that corresponded to approximately 90% breakdown of the herbicide enantiomer in batch culture. After this period of growth in the chemostat, *A. denitrificans* was able to grow independently on racemic and (*R*)-(+)-mecoprop as the sole carbon and energy

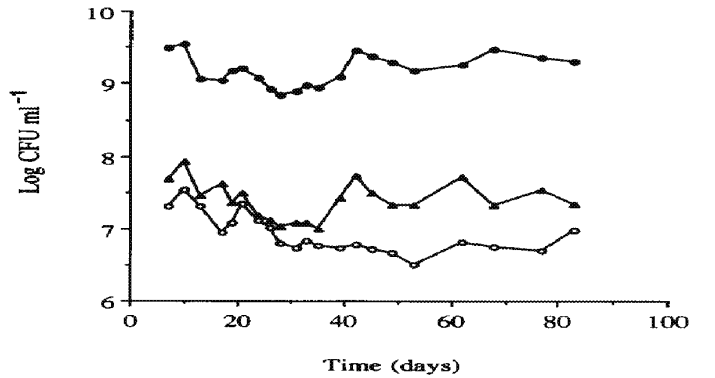


Fig. 6. Growth of the consortium on mecoprop (0.56 g carbon l^{-1}) in continuous culture. Symbols: ●, *Alcaligenes denitrificans*; ○, *Pseudomonas marginalis*; ▲, *Pseudomonas glycinea*.

COR-AAD1-00001145

Downloaded from https://academic.oup.com/femsec/article/4/3/191/6006694 by Oxford Academic user on 05 September 2023

source. The other two bacteria were not washed out of the chemostat but remained as a tight consortium.

Discussion

This study reports novel data on mecoprop degradation, concerning both the stereo-selective degradation of the (*R*)-(+)-enantiomer of the herbicide and also the first report of the degradation of mecoprop by a pure culture. The isolation of *A. denitrificans* and the demonstration of its ability to degrade mecoprop in culture solutions suggests the involvement of soil microorganisms in the dissipation of this herbicide from soil. Our results develop work by other researchers who used uncharacterised, mixed cultures by our demonstration of pure culture stereo-specific degradation of the (*R*)-(+)-enantiomer of mecoprop.

To date, studies of the biodegradation of mecoprop have been conducted indirectly with racemic or (*R*)-(+)-mecoprop [7,25]. This is the first batch culture study of microbial growth on the corresponding (*S*)-(−)-enantiomer. Several research groups have reported partial or stepwise degradation of mecoprop [7,12]. We consider that incomplete degradation of this herbicide is a consequence of stereo-selective dissimilation of (*R*)-(+)-mecoprop from the racemic mix. This interpretation is based on a comparison of growth on the (*R*)-(+)-enantiomer and the racemate where we obtained a significantly higher culture absorbance (and therefore greater cell numbers) with the isomer. Further evidence was obtained from the inability of the consortium to grow on the (*S*)-(−)-enantiomer of mecoprop. Similar culture absorbance data were obtained to suggest enantioselective degradation of the related racemate PPA. Experiments monitoring chloride ion release also showed that only half of the available racemic mecoprop was being degraded. It is proposed that the consortium exclusively degrades the (*R*)-(+)-enantiomer of mecoprop. There was no evidence of any inhibitory effect of the (*S*)-(−)-isomer, within the racemate, on the growth of the consortium.

Spectrophotometric data provided further evidence for the stereo-selective degradation of mecoprop by the consortium (Fig. 2). Since the two enantiomers are present in racemic mecoprop in the ratio 1:1 and biodegradation accounts for 55% of the available mecoprop, this supports our hypothesis that degradation of only one of the stereo-chemical forms of the herbicide has occurred. A comparison of the decrease in concentration of the herbicide in media containing (*R*)-(+)-mecoprop and the racemic herbicide is consistent with the occurrence of residual (*S*)-(−)-enantiomer. The inability of the consortium to transform this form of the herbicide was directly confirmed by a batch culture study with (*S*)-(−)-mecoprop where the herbicide concentration remained stable throughout the experimental period. The shift in the enantiomeric ratio as a result of microbial stereo-selective degradation distinguishes biodegradation from abiotic removal of the herbicide which is not able to differentiate between the two isomers. Therefore, stereo-chemistry can serve as a tool for monitoring in situ bioremediation of mecoprop [18].

The resultant displacement in peak wavelength after 21–23 h of growth with both the single enantiomer and racemic substrates may indicate transformation of (*R*)-(+)-mecoprop to form one or more primary intermediates in the degradative pathway containing a modified aryl ring. It is likely that degradation of mecoprop initially involves the elimination of one or more of the various ring substituents. A similar result has been shown for the breakdown of 2,4-D [26]. Work by Tuovinen and Oh suggested that 2-methyl-4-chlorophenol (2,4-MCP), formed by the removal of the propionate side chain, may be the first corresponding phenol intermediate formed during bacterial degradation of mecoprop [27]. The cleaved aliphatic group provides a potential source of carbon and energy for the bacteria while the aromatic ring is further degraded, ultimately into TCA cycle metabolites. Data obtained by monitoring the change in maximum absorption of the (*R*)-(+)-enantiomer with time suggests that the aromatic nucleus remains intact until approximately 35 h after inoculation (Fig. 2). We propose that at this time ring fission occurs,

COR-AAD1-00001146

Downloaded from https://academic.oup.com/femsec/article/16/3/191/6850694 by Cochrane Argentina user on 02 September 2022

producing the observed marked change in UV absorption of the growth medium. The presence of residual unmetabolised (*S*)-(−)-enantiomer probably explains why there was no analogous change for the racemic carbon source. The results confirm that stereo-configuration in addition to molecular structure determines the extent of biodegradation [14,18].

The initial inability to isolate an organism able to grow independently on mecoprop from a microbial community capable of degrading the herbicide is an indication that the genetic complement required to mineralise mecoprop was not initially present in any one organism. The maintenance of a three-member community for 83 days of chemostat growth, emphasises the importance of synergistic microbial communities in environmental decontamination [4]. The isolation of a strain of *A. denitrificans* after 5 months of subculturing that had the complete capacity to degrade mecoprop suggests that the transfer of genetic information encoding for at least one enzyme of the mecoprop degradative pathway has occurred. The consolidation of specific genetic information initially present in two or more organisms into one bacterium shows how microbial communities provide a permissive environment within which adaptation and selection can occur. The long period required for isolation has been observed previously with bacteria exposed to xenobiotic compounds [28,29]. Genes specifying biodegradation of pesticides are often borne on plasmids [30–32]. If part or all of the degradative pathway(s) of mecoprop is/are specified as plasmid-borne genes, these could act as vectors for natural gene transfer. It is considered that passage through nutrient broth may cause non-expression of mecoprop degradative enzymes which is responsible for the extended lag period of enzyme induction on reinoculating into selective media.

Many pesticides are classified as priority pollutants by the USA Environmental Protection Agency, particularly in view of the perceived environmental problems of chloroaromatics [33]. However, despite the increasing concern with the environmental persistence and residual effects of these pesticides, little information is available concerning mecoprop. The aims of the modern agrochemical industry in the manufacture of pesticides are mainly the reduction of the rate of application of a new chemical, an increase in the selectivity against the target organism and lack of toxicity or accumulation in soils [16]. Our study highlights the importance of stereo-specificity in achieving an optimal ecological profile. As only the (*R*)-(+)-enantiomer of mecoprop is active as a herbicide [13] and our results indicate that the (*S*)-(−)-isomer is not susceptible to microbial degradation, chiral preparations of (*R*)-(+)-mecoprop could be utilised to decrease the amount of this xenobiotic entering the environment via soils and crops. Furthermore, the isolation of both mixed and pure cultures capable of growth on mecoprop as the sole carbon and energy source should help future investigations to determine both the metabolic route for dissimilation and the fate in natural environments of this widely used herbicide.

Acknowledgements

The authors would like to thank Rhone-Poulenc Agriculture Ltd., ICI Agrochemicals and Pennine Chemical Services Ltd., for their kind donation of research samples of mecoprop. We also acknowledge Professor D. Ribbons for useful discussions regarding the synthesis of (*S*)-(−)-mecoprop, and Tony Davey and Rosemary McKeith for technical assistance. This work was supported by the Natural Environment Research Council and the University of Exeter.

References

1 Anderson, A.C. (1986) Environmental toxicology-biodegradation of xenobiotics. J. Environ. Health 48, 196–199.
2 Audus, L.J. (1951) The biological detoxification of hormone herbicides in soil. Plant Soil 3, 170–192.
3 Bollag, J.M., Briggs, G.G., Dawson, J.E. and Alexander, M. (1968) Enzymatic degradation of chlorocatechols. J. Agric. Food Chem. 16, 829–833.
4 Bull, A.T. (1980) Biodegradation: some attitudes and strategies of microorganisms and microbiologists. In: Con-

COR-AAD1-00001147

Downloaded from https://academic.oup.com/femsre/article/4/2/191/650809 by Oxford Academic user on 08 September 2022

temporary Microbial Ecology (Ellwood, D.C., Hedger, J.N., Latham, M.J., Lynch, J.M. and Slater, J.H., Eds.), pp. 107–136. Academic Press Inc., London.
5 Chaudhry, G.R. and Chapalamadugu, S. (1991) Biodegradation of halogenated organic compounds. Microbiol. Rev. 55, 59–79.
6 Hardman, D.J. (1991) Biotransformation of halogenated compounds. Crit. Rev. Biotechnol. 11, 1–40.
7 Heron, G. and Christensen, T.H. (1992) Degradation of the herbicide mecoprop in an aerobic aquifer determined by laboratory batch studies. Chemosphere 24, 547–557.
8 Kilpi, S. (1980) Degradation of some phenoxy acid herbicides by mixed cultures of bacteria from soil treated with 2-(2-methyl-4-chloro)phenoxypropionic acid. Microb. Ecol. 6, 261–270.
9 Kirkland, K. and Fryer, J.D. (1972) Degradation of several herbicides in a soil previously treated with MCPA. Weed Res. 12, 90–95.
10 Klint, M., Arvin, E. Jensen, B.K. and Snijders, A. (1990) Biodegradation of the pesticides atrazine and MCPP in aquifers. The Commission of the European Committees, official report.
11 Lappin, H.M., Greaves, M.P. and Slater, J.H. (1985) Degradation of the herbicide mecoprop 2-(2-methyl-4-chlorophenoxy)propionic acid by a synergistic microbial community. Appl. Environ. Microbiol. 49, 429–433.
12 Oh, K. and Tuovinen, O.H. (1991) Bacterial degradation of phenoxy herbicide mixtures 2,4-D and MCPP. Bull. Environ. Contam. Toxicol. 47, 1–8.
13 Worthing, C.R. and Walker, S.B. (1991) Pesticide manual: a world compendium, 9th Edition, pp. 539–543. BCPC publications, Bracknell, UK.
14 Alexander, M. and Aleem, M.I.H. (1961) Effect of chemical structure on microbial decomposition of aromatic herbicides. J. Agric. Food Chem. 9, 44–47.
15 Smith, A.E. (1985) Identification of 4-chloro-2-methylphenol as a soil degradation product of ring-labelled [^{14}C] mecoprop. Bull. Environ. Contam. Toxicol. 34, 656–660.
16 Tombo, G.M.R. and Bellus, D. (1991) Chirality and crop protection. Angew. Chem. Int. Engl. 30, 1193–1215.
17 Smith, M.R., van der Tweel, W.J.J. and DeBont, J.A.M. (1991) Degradation of 3-chloro-2-methylpropionic acid by *Xanthobacter* sp. CIMW 99. Appl. Microbiol. Biotechnol. 36, 246–251.
18 Ludwig, P., Gunkel, W. and Hühnerfuss, H. (1992) Chromatographic separation of the enantiomers of marine pollutants. Part 5: enantioselective degradation of phenoxycarboxylic acid herbicides by marine microorganisms. Chemosphere 24, 1423–1429.
19 Pieper, D.H., Reineke, W., Engesser, K.H. and Knackmuss, H.J. (1988) Metabolism of 2,4-dichlorophenoxyacetic acid, 4-chloro-2-methylphenoxyacetic acid and 2-methylphenoxyacetic acid by *Alcaligenes eutrophus* JMP134. Arch. Microbiol. 150, 95–102.
20 Sinton, G.L., Fan, L.T., Erickson, L.E. and Lee, S.M. (1986) Biodegradation of 2,4-D and related compounds. Enzyme Microbiol. Technol. 8, 395–403.
21 Bull, A.T. (1985) Mixed culture and mixed substrate systems. In: Comprehensive Biotechnology (Moo-Young, M., Ed.), 1st Edition, pp. 281–299. Pergammon Press, Toronto.
22 Fletcher, W.W. and Kirkwood, R.C. (1982) Herbicides and plant growth regulators, pp. 21–24. Granada, London, UK.
23 Matell, M. (1953) Optically active α-(2-methyl-4-chlorophenoxy)-propionic acid and α-(2.4-dichlorophenoxy)-*n*-butyric acid and their steric relations. Ark. Kemi 6, 365–373.
24 Evans, W.C., Smith, B.S., Moss, P. and Fernley, H.N. (1970) Bacterial metabolism of 4-chlorophenoxyacetate. Biochem. J. 122, 509–517.
25 Hallberg, K.B., Kelly, M.P. and Tuovinen, O.H. (1991) Simultaneous degradation of the herbicides 2,4-dichlorophenoxyacetic acid and 2-(2-methyl-4-chlorophenoxy)propionic acid by mixed bacterial cultures. Curr. Microbiol. 23, 65–69.
26 Loos, M.A. (1975) Phenoxyalkanoic acids. In: Herbicides: their Chemistry, Degradation and Mode of Action, (Kearney, P.C. and Kanfman, D.D., Eds.), Vol. 1, pp. 34–57. Marcel Dekker, New York.
27 Oh, K. and Tuovinen, O.H. (1991) Detection and identification of substituted phenols as intermediates of concurrent bacterial degradation of the phenoxy herbicides MCPP and 2,4-D. Microbiol. Lett. 79, 141–146.
28 Schraa, G., Boone, M.L., Jetten, M.S.M., van Neerven, A.R.W., Colberg, P.J. and Zehnder, A.J.B. (1986) Degradation of 1,4-Dichlorobenzene by *Alcaligenes* sp. Strain A175. Appl. Environ. Microbiol. 52, 1374–1381.
29 Kröckel, L. and Focht, D.D. (1987) Construction of chlorobenzene-utilizing recombinants by progenitive manifestation of a rare event. Appl. Environ. Microbiol. 53, 2470–2475.
30 Boronin, A.M. (1991) Genetics of biodegradation. Act. Biotech. 11, 177–181.
31 Kellog, S.T., Chatterjee, D.K. and Chakrabarty, A.M. (1981) Plasmid assisted molecular breeding: new techniques for enhanced biodegradation of persistent toxic chemicals. Science 214, 1133–1135.
32 Lal, R., Lal, S. and Shivaji, S. (1986) Use of microbes for detoxification of pesticides. CRC Crit. Rev. Biotechnol. 3, 1–6.
33 Keith, L.H. and Telliard, W.A. (1979) Priority pollutants I - a perspective view. Environ. Sci. Technol. 13, 416–423.

Downloaded from https://academic.oup.com/femsec/article/13/3/191/6008694 by Colorado Alliance on 03 September 2022

COR-AAD1-00001148

CURRENT OPINION
www.current-opinion.com

The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds

Eva M Top* and Dirk Springael†

Retrospective studies clearly indicate that mobile genetic elements (MGEs) play a major role in the *in situ* spread and even *de novo* construction of catabolic pathways in bacteria, allowing bacterial communities to rapidly adapt to new xenobiotics. The construction of novel pathways seems to occur by an assembly process that involves horizontal gene transfer: different appropriate genes or gene modules that encode different parts of the novel pathway are recruited from phylogenetically related or distant hosts into one single host. Direct evidence for the importance of catabolic MGEs in bacterial adaptation to xenobiotics stems from observed correlations between catabolic gene transfer and accelerated biodegradation in several habitats and from studies that monitor catabolic MGEs in polluted sites.

Addresses
*Department of Biological Sciences, 347 Life Sciences Building South, University of Idaho, Moscow ID 83844-3051, USA
e-mail: evatop@uidaho.edu
†Laboratory of Soil and Water Management, Department of Land Management, Faculty of Agricultural and Applied Biological Sciences, Catholic University of Leuven, B-3001 Heverlee, Belgium
e-mail: dirk.springael@agr.kuleuven.ac.be

Current Opinion in Biotechnology 2003, **14**:262–269

This review comes from a themed section on
Environmental biotechnology
Edited by Ian M Head and Mark J Bailey

0958-1669/03/$ – see front matter
© 2003 Elsevier Science Ltd. All rights reserved.

DOI 10.1016/S0958-1669(03)00066-1

Abbreviations

CB	chlorobenzene
2,4-D	2,4-dichlorophenoxyacetic acid
HGT	horizontal gene transfer
IS	insertion sequence
MGE	mobile genetic element
PAH	polycyclic aromatic hydrocarbon

Introduction

Xenobiotic compounds can be defined in a strict or broader sense. Xenobiotics *sensu strictu* are defined as man-made molecules, foreign to life and, as such, should have never been encountered by bacterial populations before their introduction by man. Examples are polychlorinated biphenyls (PCBs) and various pesticides. However, structurally related compounds (such as some chlorinated hydrocarbons) have been recently detected in nature, (e.g. as metabolites produced by microorganisms) [1]. Therefore, a broader definition of xenobiotics includes 'all compounds that are released in any compartment of the environment by the action of man and thereby occur in a concentration that is higher than natural' [2]. This larger group also includes heavy metals, polycyclic aromatic hydrocarbons (PAHs), and oil derivatives such as toluene. In this review, we will limit our discussion on bacterial adaptation to organic xenobiotic compounds as defined by Leisinger [2] and simply refer to them as xenobiotics or xenobiotic compounds.

There are several mechanisms, or combinations thereof, by which microbial communities can adapt to the presence of xenobiotics in their environment. Firstly, there can be an increase in population size of those organisms that tolerate or even degrade the compound by induction of appropriate genes. Secondly, the cells can adapt through mutations of various kinds, such as single nucleotide changes or DNA rearrangements that result in resistance to or degradation of the compound. Thirdly, they may acquire genetic information from either related or phylogenetically distinct populations in the community by horizontal gene transfer (HGT) — also called 'lateral gene transfer' [3]. Eventually, the individual cells best suited to resist or degrade the xenobiotic will be selected and sweep through the population until they constitute a larger fraction of the total microbial community than before the presence of the xenobiotic.

HGT between bacteria in natural habitats is largely mediated by mobile genetic elements (MGEs), which together form the so-called 'horizontal gene pool' [4], also referred to as the 'flexible gene pool' [5]. These elements can be transferable plasmids, (conjugative) transposons, integrons, genomic islands or phage, which are all able to move within and/or between genomes, thus allowing 'evolution in quantum leaps' [5]. The ability of microbial communities to adapt to anthropogenic chemicals by HGT first became evident from the observation that rapid dissemination of antibiotic resistance determinants by means of MGEs accompanied the widespread use of antibiotics [6]. More recently, the importance of HGT in bacterial evolution has been confirmed by comparison of available sequence information for prokaryotic genomes, showing that as much as 24% of the *Escherichia coli* genome has been acquired due to HGT [7]. During the past 15 years, several groups have provided strong indications that MGEs also play an important role in the horizontal spread of existing catabolic pathways, as well as in the natural construction of novel ones. This transfer allows bacteria to rapidly adapt to new xenobiotic compounds entering their habitats (see also the review of van der Meer and Sentchilo

COR-AAD1-00001149

in this issue [8]). Here, we summarize the evidence by giving an overview of retrospective analyses of catabolic MGEs and of studies that have shown adaptation of communities to xenobiotics by means of HGT.

Retrospective evidence for HGT as a natural mechanism of bacterial adaptation

One approach to collect evidence about how gene transfer has contributed to the adaptation of bacterial populations to a specific xenobiotic is to characterize and compare the genes involved in degradation of identical or similar xenobiotic compounds in different closely or more distantly related bacterial isolates from geographically distinct locations. Four key findings have been made from such studies. Firstly, evolutionarily related catabolic genes and gene clusters are found in bacteria originating from very distant locations [9]. Secondly, in some cases, as shown for some 2,4-dichlorophenoxyacetic acid (2,4-D) degradation genes, the phylogeny of the catabolic genes is not congruent with that of the 16S rRNA genes of the corresponding hosts [10,11]. Thirdly, genes for the degradation of organic pollutants, both of natural and synthetic origin, are often associated with MGEs such as plasmids and transposons [12]. Finally, evolutionarily related catabolic genes and entire gene modules are involved in the degradation of structurally similar but different xenobiotic compounds [13–15]. These observations strongly indicate that horizontal gene transfer by MGEs plays a major role in the worldwide spread of catabolic pathways and in the formation of novel ones. Formation of novel pathways seems to occur by recruitment of catabolic genes or gene segments from different organisms into one suitable host. This was recently confirmed by the identification of relics of precursor genes encoding the degradation of toluene, flanking gene clusters for chlorobenzene degradation in two independently isolated bacterial strains [16,17]. The fact that synthetic pollutants did not appear and certainly did not accumulate in the environment until about half a century ago, suggests that the necessary pathway assembly events must have occurred quite recently.

Furthermore, it has been found that the same genes or gene clusters can be part of different replicons. For example, almost exact copies of the *xyl* operons encoding the enzymes for toluene/xylene degradation have been identified on plasmids of different incompatibility groups [18], and very similar gene modules encoding enzymes for chlorocatechol degradation have been identified on both plasmids and conjugative integrative elements [19••,20]. These findings indicate independent movement and acquisition of such operons, which in some cases can be explained by the fact that the genes are enclosed within transposons. All in all, these observations suggest that catabolic genes and entire gene clusters are being shuttled between different replicons, thereby facilitating their spread within and among microbial communities. As a consequence, it can be argued that *in situ* assembly of a xenobiotic catabolic pathway can occur independently by recruiting the same pre-existing catabolic genes or gene modules [16]. For example, the almost simultaneous detection of very similar atrazine degradative genes and gene combinations in different hosts on different continents [21] may be explained either by independent recruitment of the same genes or by long-distance transport of bacteria combined with HGT of the entire operon.

One important group of catabolic MGEs consists of the catabolic plasmids. Interestingly, among these, the IncP-2 and IncP-9 plasmids often encode degradation of naturally occurring compounds, whereas genes that encode degradation of mostly man-made compounds, such as several chloroaromatics, are often encoded by the well-known broad host range IncP-1 plasmids [22]. These plasmids are the most promiscuous self-transmissible plasmids characterized to date, with a host range that is much wider than that of IncP-2 and IncP-9 plasmids (see Table 1 for examples). The possible correlation between this high promiscuity of IncP-1 plasmids and their ability to acquire and transfer genes encoding degradation of recently introduced xenobiotics is very intriguing. Indeed, highly transferable and promiscuous plasmids such as those of the IncP-1 group that act as shuttle vectors for locally adapted genes may facilitate rapid local adaptation of both phylogenetically related and distinct populations present in the same bacterial community. More DNA sequence information for IncP-1 plasmids and a better understanding of their dynamics in response to recent pollution will be necessary to fully understand their role in bacterial adaptation to xenobiotics.

During the past decade, the types of MGEs that carry determinants for organic pollutant degradation have been extended to class I and class II transposons, which are often located on catabolic plasmids [23], and even to the increasing array of mosaic MGEs including genomic islands, conjugative transposons and integrative plasmids [8,19••,24–26]. An overview of some examples of the different types of catabolic MGEs, together with a brief description for each, is presented in Table 1.

Although the role of MGEs in the molecular adaptation to organic pollutants seems clear, some questions are still unanswered. For example, how are these genes recruited by MGEs? Only a few researchers have looked at the possibility of trapping genes with catabolic plasmids [27]. Interestingly, both on MGEs and in the chromosome, catabolic genes are often bordered by insertion sequences (IS). These IS elements may have played a role in recruitment of these genes by the replicon, but also increase the potential of further exchange of the genes between different hosts and replicons [28–30]. An intriguing case in this context is the identification of IS*1071*-like sequences, which flank many plasmid-borne catabolic genes in a diverse range of organisms including Gram-negative

COR-AAD1-00001150

Table 1

Examples of the different types of catabolic mobile elements.

Mobile element	Strain	Substrate(s)	Size of MGE (kb)	*Inc group	Relevant remarks and references
Plasmids					
pSS60	*Achromobacter* sp. LBS1C1	4-Chlorobenzoate	53	P1	Recruitment of a 4-chlorobiphenyl (4CBP) degradative gene cluster present on a chromosomally located mobile element by pSS60 has been observed, creating plasmid-encoded mineralization of 4CBP [24].
pBRC60	*Alcaligenes* sp. BR60	3-Chlorobenzoate	75	P1	Contains Tn*5271* carrying chlorobenzoate catabolic genotype *cba* (see below) [29].
pENH91	*Ralstonia eutropha* NH9	3-Chlorobenzoate	78	P1	Carries chlorocatechol catabolic composite transposon Tn*5707* (see below) with a gene module evolutionarily related to the chlorocatechol degradation gene modules present on the CB degradation plasmids pP51 and pPS12-1, on the *clc* element, and on the 2,4-D degradative plasmid pJP4 [20].
pJP4	*Ralstonia eutropha* JMP134	2,4-D	75	P1	Carries different gene modules the combined action of which leads to mineralization of 2,4-D. Chlorocatechol degradation genes are evolutionarily related to those on the *clc* element and pENH91, pP51 etc.; pJP4 contains a duplication of these genes and is prone to rearrangements. The plasmid contains IS elements flanking the chlorocatechol genes and IS*1071*-related elements. Other 2,4-D degraders carry similarly or differently organized genes [50].
pTSA	*Comamonas testosteroni* T-2	*p*-Toluenesulfonic acid	85	P1	Carries two copies of a *tsa* gene cluster encoding *p*-toluenesulfonic acid degradation. These are flanked by two IS*1071* elements, constituting a putative composite transposon; *tsa* genes were also found in other organisms from distant locations present on the same composite transposon, but only one strain contained a plasmid-encoded *tsa* gene [51].
pP51	*Pseudomonas* sp. P51	Chlorobenzene			Carries a chlorocatechol degradation gene module similar to that of pPS12-1, the *clc* element and pENH91. In addition, this plasmid encodes a composite transposon Tn*5280* (see below) carrying *tcbAB* genes for the degradation of CB into chlorocatechol. Thus, this pathway is a clear example of natural pathway assembly [17].
pPS12-1	*Burkholderia* sp. PS12	1,2,4,5- Tetrachlorobenzene	85	P1	Carries a chlorocatechol degradation gene module similar to that of pP51, the *clc* element and pENH91. In addition, it encodes a composite transposon carrying *tecAB* genes for the degradation of CB into chlorocatechol. These *tecbAB* genes are similar to the *tcbAB* genes of strain P51, but are flanked by different sized remnants of the toluene pathway [18].
Class I transposons					
Tn*5280*	*Pseudomonas* sp. P51	Chlorobenzene	8.5		Present on pP51 (see above), carries the *tcbAB* genes responsible for CB transformation into chorocatechol. The *tcbAB* genes are clearly derived from a toluene degradation pathway on the basis of the presence of gene relics belonging to the toluene degradation genes, which flank the *tcbAB* genes [17].
Tn*5271*	*Alcaligenes* sp. BR60 (pBRC60)	Chlorobenzoate	17		Present on pBRC60 (see above). Carries chlorobenzoate catabolic genotype *cba*. Homologous transposons have been found in other chlorobenzoate degraders from different geographical origins. In some cases these transposons carry gene relics of other pathways [29].
Tn*5707*	*Ralstonia eutropha* NH9 (pENH91)	3-Chlorobenzoate	15		Present on pENH91 (see above). Carries genes evolutionarily related to the chlorocatechol degradation genes on plasmids pP51 and pPS12-1, on the *clc* element and on pJP4 [20].
DEH	*Pseudomonas putida PP3*	Chlorinated aliphatic acids	9.74		This transposon contains a dehalogenase gene and is bordered by two copies of the IS element IS*Ppu12*, which has also been found flanking other catabolic genes [52].
Class II transposons					
Tn*4651*	*Pseudomonas putida* mt-2 (pWWO)	Toluene, xylene	56		Carries gene modules for toluene/xylene degradation via catechol, located between two IS elements. The transposon is located within another transposon Tn*4663*. The genes are similarly organized in many other TOL plasmids [18,53].
Tn*4653*	*Pseudomonas putida* mt-2 (pWWO)	Toluene, xylene	70		Carries Tn*4661* (see above) [53].

COR-AAD1-00001151

Table 1 Continued

Mobile element	Strain	Substrate(s)	Size of MGE (kb)	*Inc group	Relevant remarks and references
Tn*4655*	*Pseudomonas putida* G7 (NAH7)	Naphthalene	38		Carries three gene modules for naphthalene degradation via catechol. The module encoding catechol degradation is similar in organization to that of the pWWO plasmid [33].
Other elements					
Tn*4371*	*Ralstonia oxalatica* A5	Biphenyl/ 4-chlorobiphenyl	55		This element is an unusual combination of phage-like genes (including an integrase), catabolic genes and genes for plasmid transfer. Similar elements have been found in other polychorinated biphenyl degraders [34].
clc element	*Pseudomonas* sp. B13	Chlorocatechol	105		Carries a chlorocatechol degradation gene module similar to that of pP51, pPS12-1 and pENH91 [19••].

*Inc group, incompatibility group.

and Gram-positive bacteria [29,31]. Another study suggests integrase-dependent gene acquisition of catabolic genes [32]. New information on this topic will undoubtedly become available through analysis of whole plasmid genomes. Another interesting question is how the different genes constituting a complete pathway finally came together in one strain or even on one plasmid. Interestingly, for some xenobiotics, such as atrazine [33] and nitrotoluene [34], degradation is sometimes mediated by a combination of bacterial strains, which each carry different genes responsible for different parts of the pathway. The same genes are, however, in both cases also found combined in one bacterial strain, which is able to mineralize that same compound by itself. This might suggest that such gene combinations in one organism have evolved over time through independent gene exchange events between different strains, which were brought together in a synergistic community. HGT might have been facilitated by the formation of tight cell contacts (preferably on surfaces), which was required for efficient metabolic fluxes between the different cells of the consortia. It should be noted in this context that higher conjugative gene transfer frequencies have been reported in biofilms [35] and that natural conjugative plasmids can induce biofilm development [36]. Cell consortia and biofilms may therefore be very conducive to the assembly of new pathways by HGT and lead to final pathway assembly in one host if the community characteristics allow it. At first sight, it may seem more advantageous to any member of the consortium to contain the entire degradative pathway, instead of having to rely on neighboring cells. However, the fitness cost associated with maintaining this genetic information, often on a self-replicating self-transferable plasmid, may outweigh the benefit of being independent of the other members of the community for (part of the) resources. This delicate balance may explain why both bacterial consortia and single strains with very similar catabolic genes and phenotypes have been found.

These different pieces of information, although very informative, only provide retrospective evidence and questions still remain as to when, where, at what rates and between what range of 'species' these HGT events actually occur. We also need to address what the intensity of selection is for those catabolic genes under natural conditions. Our picture of the microbial populations involved in xenobiotic degradation may be very distorted by limitations such as poor culturability of important populations, selective enrichment for fast-growing bacteria under conditions of enrichment cultures in the laboratory, and possibly gene transfer during these enrichments. Therefore, these retrospective analyses, although very valuable, have to be complemented with studies that monitor the actual gene transfer processes and pathway assembly in natural habitats.

Direct evidence of microbial community adaptation by HGT

A direct way to examine if transfer of a degradative MGE among soil bacteria plays a role in bacterial adaptation to the corresponding organic contaminant is to monitor both the degradation of that xenobiotic compound and the dissemination of the degradative genes. Several studies have demonstrated that horizontal transfer of catabolic genes, mostly by means of plasmid-mediated conjugation, occurs in soil microcosms, bioreactors, lake mesocosms and so on after inoculation of a donor strain containing a natural catabolic MGE [12,37–39]. Some of these studies also showed that the number of transconjugants (i.e. indigenous bacteria that acquired the xenobiotic-degrading phenotype after uptake of the catabolic MGE) was higher when the corresponding xenobiotic was present in the habitat. Very few groups, however, have provided clear direct evidence of adaptation of the microbial community to this xenobiotic, that is, by showing that the overall degradative capacity improved due to this gene exchange. We will restrict our overview of these studies to the most recently published ones, and refer the reader to a recent minireview for a more complete and extensive description of case studies [12].

A series of studies on catabolic gene transfer in soils by Newby and coworkers [40–42] and Top and coworkers

COR-AAD1-00001152

[43••] have used the herbicide 2,4-D as a model compound, and the IncP-1β plasmid pJP4 or the non-IncP-1β plasmid pEMT1, both encoding 2,4-D degradation, as model plasmids. In some, but not all, cases a clear correlation was shown between transfer of the plasmids from inoculants to indigenous bacteria and adaptation of the soil community to degrade 2,4-D. Another study on gene transfer in soil showed adaptation of a rhizosphere bacterial community from pine seedlings mycorrhized with *Suillus bovinus* to a pollutant. The self-transferable TOL plasmid was acquired from an inoculated *Pseudomonas fluorescens*, thereby protecting the plant and fungus from meta-toluate [44]. *In situ* spread of the *tfdA* gene, encoding the first step in the degradation of 2,4-D and phenoxyacetic acid (PAA), from an inoculum into indigenous phenol degraders in soil was shown to result in vertical expansion of the phenol degradation pathway to now include PAA. Interestingly, a higher number of transconjugants was detected in the residue sphere (i.e. the interface between decaying plant material [barley leaves] and the soil matrix) than in the bulk soil, which was explained by an increased biomass of donors and recipients in this microhabitat [45••]. Similarly, Springael and coworkers [46••] also observed transfer of the *clc* element (see Table 1) in an inoculated model biofilm membrane reactor treating chlorobenzoate-containing influent. The original inoculum carrying the *clc* element was frequently found to be replaced by other chlorobenzoate-degrading bacteria. As control reactors without inoculum never developed a chlorobenzoate-degrading microbial community nor a detectable degradation activity, it was suggested that contaminant bacteria had acquired the *clc* element from the inoculum strain and became more competitive under the implied bioreactor conditions. This result is in agreement with earlier observations in a gnotobiotic biofilm, which showed that a new biochemical pathway offering a growth advantage can be introduced effectively in a biofilm community when carried on an MGE [47].

Several of these reports show that the spread of catabolic genes on MGEs enabled specific indigenous bacterial species to acquire new catabolic characteristics and thereby to become numerically dominant members of the community. In addition, habitats such as the rhizosphere residue sphere and biofilms might play a yet underestimated role in the adaptation of microbiota to pollutants. The disadvantages of these studies are the unnatural aspect of inoculation of high numbers of donor cells (often 10^6–10^7 colony-forming units [CFU]/g soil), and the relatively high concentrations of contaminants added. In addition, particularly in the case of soil, these manipulations almost always entail mixing and thus disturbing the samples. There is no direct evidence that HGT would be as clearly responsible for community adaptation without this disturbance and at lower, more environmentally relevant, xenobiotic concentrations. Another important aspect is the inclusion of all necessary control experiments to exclude the possibility that the detected 'transconjugants' are indigenous organisms that already contained the same or a similar catabolic gene or MGE as the one introduced via the donor strain. Often, transconjugants are only verified by limited screening methods to confirm the presence of the inoculated gene/MGE. It is not impossible that rare organisms with the same or similar catabolic MGEs become enriched upon manipulation of the ecosystem; for example, the inoculation of large cell numbers (often 10^6–10^8 CFU/g or ml) disturbs the system in an uncontrolled way and introduces nutrients once the cells die off. It may therefore be recommended to include a control microcosm inoculated with the donor strain without the catabolic MGE instead of the usual control without inoculum. Despite these shortcomings, overall these studies are very valuable, as they monitor a direct correlation between HGT and adaptation of soil bacterial populations to xenobiotics.

Field studies to assess microbial community adaptation by HGT

The most direct, non-disturbing and objective approach that could be used to evaluate whether MGEs have contributed to the adaptation of a microbial community to pollution, is to compare the degradation potential and the presence of degradative genes within bacterial communities in the field before and after a specific pollutant spill has occurred. A good indicator of gene exchange in response to a new pollutant entering the habitat would, for example, be the detection of a catabolic gene in only one population before the contamination, and in multiple phylogenetically distinct populations after contamination. Another valuable indication would be the assembly in one organism of two or more different catabolic genes or operons that were previously only found separately in distinct bacterial strains. To our knowledge there are no studies with such direct evidence, as data from the period before contamination are usually not available. There are, however, a few good examples of field studies that strongly suggest that the horizontal transfer of mobile catabolic genes has played an important role in community adaptation to the pollutant present in the site, in spite of the absence of data obtained before pollution. For example, six years after introduction of a plasmid-encoded *pheBA* operon (encoding enzymes for phenol degradation) for the purpose of bioremediation after an accidental fire in an Estonian oil shale mine, this gene cluster was detected in a few different *Pseudomonas* sp. isolates. These findings suggested that several bacterial populations had adapted to the polluted environment by acquiring and expressing the introduced genetic material that allowed them to use the pollutant as a substrate [48]. The authors were not able to completely exclude that this same *pheBA* gene cluster was already present in that site before the donor strain was released. They

COR-AAD1-00001153

argued, however, that this would have been very unlikely given the very high similarity of the detected *pheBA* operons with the unique features of the introduced one, which originated from a site in a different country and had been in a laboratory environment for a long time. In a coal tar contaminated field site in the US, strong evidence was provided that natural horizontal transfer of naphthalene dioxygenase genes had occurred between members of the soil bacterial community. A naphthalene catabolic plasmid may have played a role in adaptation of this community to the coal tar contamination [49] (see Update). Finally, a study by van der Meer *et al.* [15] found that chlorobenzene (CB)-degrading bacteria could be isolated from CB-contaminated groundwater, but not from the uncontaminated region outside. The study clearly showed that bacterial adaptation to the pollution, as demonstrated by enhanced CB removal, must have occurred by genetic recombination among bacteria in the aquifer, resulting in a novel CB degradation pathway. As pointed out above, important drawbacks in field studies that monitor the spread of a specific gene into various indigenous populations include the absence of perfect controls and problems with detection limits. The lack of detection of a specific gene or combination of genes in an organism before pollution or before inoculation of a specific donor strain, does not necessarily mean its absolute absence in that population. Therefore, detection of the gene afterwards is not per se due to HGT, but could also be a consequence of an increase in the population size of that previously undetectable catabolic gene(s)–host combination. Field studies that include the proper controls with sensitive detection methods will be indispensable to evaluate the relative importance of gene transfer as one of the mechanisms of local adaptation of bacterial populations to xenobiotics.

Conclusions

In conclusion, this review summarizes evidence that the catabolic operons known today have evolved in part thanks to MGEs, which brought diverse catabolic genes together in one host and thus allowed a wide range of organic compounds to be degraded through a few central metabolic pathways. Specific cases of clear correlations between the observed transfer of catabolic genes and microbial community adaptation to specific organic xenobiotics are still very limited. A few examples indicate that in cases where the limiting factor for biodegradation is the absence of the necessary genes, inoculated catabolic MGEs seem to spread and thereby allow rapid adaptation of the community, resulting in the accelerated removal of pollutants. It has become clear, however, that in spite of the many studies on gene transfer in soil and to some extent in bioreactors, our basic understanding of factors that determine the dissemination of genes in such systems needs to be improved. This includes understanding the conditions that stimulate conjugation and other gene transfer mechanisms in these ecosystems, and comprehending the role of IS elements such as IS*1071* in the assembly of catabolic pathways. In addition, we know very little about the natural host range of several catabolic MGEs and the expression range of their catabolic genes. Furthermore, we also need to learn more about the competitive fitness of transconjugants and the maintenance of MGEs, both in the absence and presence of selective pressure and at more environmentally relevant pollutant concentrations, which are often much lower than those applied so far in most studies.

Update

Recent work by Müller *et al.* [54] presents additional evidence on the involvement of HGT in the *de novo* construction of catabolic pathways in natural environments. A chlorobenzene catabolic pathway was shown to have been assembled in one bacterial strain. This was achieved by IS-mediated mobilization of a gene module that encodes conversion of chlorinated benzenes to chlorocatechol, onto an integrative element that encodes chlorocatechol degradation. In addition, very recent evidence for natural horizontal transfer of a naphthalene dioxygenase gene (*phnAc*) in a polycyclic aromatic hydrocarbon-contaminated site underlines the importance of HGT in bacterial adaptation to pollutants [55].

Acknowledgements

EMT is in part supported by the National Institutes of Health (NIH) and National Center for Research Resources (NCRR) Idaho Center of Biomedical Research Excellence (COBRE) (grant P20 RR16448-01). This review is dedicated to the memory of RC Wyndham, who contributed greatly to our current knowledge on catabolic MGEs.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Myneni SCB: **Formation of stable chlorinated hydrocarbons in weathering plant material**. *Science* 2002, **295**:1039-1041.
2. Leisinger T: **Microorganisms and xenobiotic compounds**. *Experientia* 1983, **39**:1183-1220.
3. Liu ST, Sulfita JM: **Ecology and evolution of microbial populations for bioremediation**. *Trends Biotechnol* 1993, **11**:344-352.
4. Thomas CM: *The Horizontal Gene Pool. Bacterial Plasmids and Gene Spread*. Amsterdam: Harwood Academic Publishers; 2000.
5. Hacker J, Carniel E: **Ecological fitness, genomic islands and bacterial pathogenicity**. *EMBO Rep* 2001, **2**:376-381.
6. Cohen ML: **Epidemiology of drug resistance: implications for a post-antimicrobial era**. *Science* 1992, **257**:1050-1055.
7. Lawrence JG, Ochman H: **Reconciling the many faces of lateral gene transfer**. *Trends Microbiol* 2002, **10**:1-4.
8. van der Meer JR, Sentchilo V: **Genomic islands and the evolution of catabolic pathways in bacteria.** *Curr Opin Biotechnol* 2003: in press.
9. Whyte LG, Smits THM, Labbe D, Witholt B, Greer CW, van Beilen JB: **Gene cloning and characterization of multiple alkane hydroxylase systems in *Rhodococcus* strains Q15 and NRRL B-16531**. *Appl Environ Microbiol* 2002, **68**:5933-5942.

COR-AAD1-00001154

10. McGowan C, Fulthorpe R, Wright A, Tiedje J: **Evidence for interspecies gene transfer in the evolution of 2,4-dichlorophenoxyacetic acid degraders**. *Appl Environ Microbiol* 1998, **64**:4089-4092.

11. Vallaeys T, Courde L, McGowan C, Wright AD, Fulthorpe RR: **Phylogenetic analyses indicate independent recruitment of diverse gene cassettes during assemblage of the 2,4-D catabolic pathway**. *FEMS Microbiol Ecol* 1999, **28**:373-382.

12. Top EM, Springael D, Boon N: **Catabolic mobile genetic elements and their potential use in bioaugmentation of polluted soils and waters**. *FEMS Microbiol Ecol* 2002, **42**:199-208.

13. Smejkal CW, Vallaeys T, Seymour FA, Burton SK, Lappin-Scott HM: **Characterization of (*R/S*)-mecoprop [2-(2-methyl-4-chlorophenoxy) propionic acid]-degrading *Alcaligenes* sp. CS1 and *Ralstonia* sp. CS2 isolated from agricultural soils**. *Environ Microbiol* 2001, **3**:288-293.

14. Johnson GR, Jain RK, Spain JC: **Origins of the 2,4-dinitrotoluene pathway**. *J Bacteriol* 2002, **184**:4219-4232.

15. van der Meer JR, Werlen C, Nishino SF, Spain JC: **Evolution of a pathway for chlorobenzene metabolism leads to natural attenuation in contaminated groundwater**. *Appl Environ Microbiol* 1998, **64**:4185-4193.

16. Beil S, Timmis K, Pieper D: **Genetic and biochemical analysis of the *tec* operon suggests a route for evolution of chlorobenzene degradation genes**. *J Bacteriol* 1999, **181**:341-346.

17. Werlen C, Kohler H-PE, van der Meer J-R: **The broad substrate chlorobenzene dioxygenase and *cis*-chlorobenzene dihydrodiol dehydrogenase of *Pseudomonas* sp. P51 are linked evolutionarily to the enzymes for benzene and toluene degradation**. *J Biol Chem* 1996, **271**:4009-4016.

18. Sentchillo VSPA, Zehnder AJB, van der Meer JR: **Molecular diversity of plasmids bearing genes that encode toluene and xylene metabolism in *Pseudomonas* strains isolated from different contaminated sites in Belarus**. *Appl Environ Microbiol* 2000, **66**:2842-2852.

19. •• van der Meer JR, Ravatn R, Sentchilo V: **The *clc* element of *Peudomonas* sp. strain B13 and other mobile degradative elements employing phage-like integrases**. *Arch Microbiol* 2001, **175**:79-85.

The *clc* element was, together with Tn*4371*, the first conjugative integrative element or gene island to be shown to encode degradation of xenobiotics and employs phage-like integrases. These elements extend the types of xenobiotic degradative MGEs beyond plasmids and transposons and might indicate a 'yet to come' diversity of 'undiscovered' catabolic MGEs. This review gives an excellent overview of such MGEs, emphasizing their relationship with similar elements that encode different functions.

20. Ogawa N, Miyashita K: **The chlorocatechol-catabolic transposon Tn*5707* of *Alcaligenes eutrophus* NH9, carrying a gene cluster highly homologous to that in the 1,2,4-trichlorobenzene-degrading bacterium *Pseudomonas* sp. strain P51, confers the ability to grow on 3-chlorobenzoate**. *Appl Environ Microbiol* 1999, **65**:724-731.

21. De Souza ML, Wacket LP, Sadowsky MJM: **The *atzABC* genes encoding atrazine catabolism are located on a self-transmissible plasmid in *Pseudomonas* sp. strain ADP**. *Appl Environ Microbiol* 1998, **64**:2323-2326.

22. Top EM, Moënne-Loccoz Y, Pembroke T, Thomas CM: **Phenotypic traits conferred by plasmids.** In *The Horizontal Gene Pool: Bacterial Plasmids and Gene Spread*. Edited by Thomas CM. Harwood Academic Publishers; 2000:249-285.

23. Tsuda M, Tan HM, Nishi A, Furukawa K: **Mobile catabolic genes in bacteria**. *J Biosci Bioeng* 1999, **87**:401-410.

24. Springael D, Ryngaert A, Merlin C, Toussaint A, Mergeay M: **Occurrence of Tn*4371* related sequences in (chloro-)biphenyl degrading bacteria**. *Appl Environ Microbiol* 2001, **67**:42-50.

25. Osborn M, Böltner D: **When phage, plasmids, and transposons collide: genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum**. *Plasmid* 2002, **48**:202-212.

26. Toussaint A, Merlin C: **Mobile elements as a combination of functional modules**. *Plasmid* 2002, **47**:26-35.

27. Ronchel MC, Ramos-Diaz MA, Ramos JL: **Retrotransfer of DNA in the rhizosphere**. *Environ Microbiol* 2000, **2**:319-323.

28. Springael D, van Thor J, Goorissen H, Ryngaert A, De Baere R, Van Hauwe P, Commandeur LCM, Parsons JR, De Wachter R, Mergeay M: **RP4::Mu3A mediated *in vivo* cloning and transfer of a chlorobiphenyl catabolic pathway**. *Microbiology* 1996, **142**:3283-3293.

29. Di Gioia D, Peel M, Fava F, Wyndham RC: **Structures of homologous composite transposons carrying *cbaABC* genes from Europe and North America**. *Appl Environ Microbiol* 1998, **64**:1940-1946.

30. Schmid-Appert M, Zoller K, Traber H, Vuilleumier S, Leisinger T: **Association of newly discovered IS elements with dichloromethane utilization genes of methylotrophic bacteria**. *Microbiology* 1997, **143**:2557-2567.

31. Rousseaux S, Soulas G, Hartmann A: **Plasmid localisation of atrazine-degrading genes in newly described *Chelatobacter* and *Arthrobacter* strains**. *FEMS Microbiol Ecol* 2002, **41**:69-75.

32. Poelarends G, Kulakov L, Larkin M, van Hylckama Vlieg J, Janssen D: **Roles of horizontal gene transfer and gene integration in evolution of 1,3-dichloropropene- and 1,2-dibromoethane-degradative pathways**. *J Bacteriol* 2000, **182**:2191-2199.

33. De Souza M, Newcombe D, Alvey S, Crowley D, Hay H, Sadowsky M, Wackett L: **Molecular basis of a bacterial consortium: interspecies catabolism of atrazine**. *Appl Environ Microbiol* 1998, **64**:178-184.

34. Snellinx Z, Taghavi S, Van Gronsveld J, van der Lelie D: **Isolation and characterization of a synergistic microbial community degrading 2,4-dinitrotoluene.** *Biodegradation* 2003: in press.

35. Hausner M, Wuertz S: **High rates of conjugation in bacterial biofilms as determined by quantitative *in situ* analysis**. *Appl Environ Microbiol* 1999, **65**:3710-3713.

36. Ghigo J: **Natural conjugative plasmids induce bacterial biofilm development**. *Nature* 2001, **412**:442-445.

37. Fulthorpe RR, Wyndham RC: **Transfer and expression of the catabolic plasmid pBRC60 in wild bacterial recipients in a freshwater ecosystem**. *Appl Environ Microbiol* 1991, **57**:1546-1553.

38. Fulthorpe RR, Wyndham RC: **Involvement of a chlorobenzoate-catabolic transposon Tn*5271*, in community adaptation to chlorobiphenyl, chloroaniline, and 2,4-dichlorophenoxyacetic acid in a freshwater ecosystem**. *Appl Environ Microbiol* 1992, **58**:314-325.

39. Hill KE, Top EM: **Gene transfer in soil systems using microcosms**. *FEMS Microbiol Ecol* 1998, **25**:319-329.

40. Newby DT, Josephson KL, Pepper IL: **Detection and characterization of plasmid pJP4 transfer to indigenous soil bacteria**. *Appl Environ Microbiol* 2000, **66**:290-296.

41. Newby DT, Gentry TJ, Pepper IL: **Comparison of 2,4-dichlorophenoxyacetic acid degradation and plasmid transfer in soil resulting from bioaugmentation with two different pJP4 donors**. *Appl Environ Microbiol* 2000, **66**:3399-3407.

42. Newby DT, Pepper IL: **Dispersal of plasmid pJP4 in unsaturated and saturated 2,4-dichlorophenoxyacetic acid contaminated soil**. *FEMS Microbiol Ecol* 2002, **39**:157-164.

43. •• Dejonghe W, Goris J, El Fantroussi S, Hofte M, De Vos P, Verstraete W, Top EM: **Effect of dissemination of 2,4-dichlorophenoxyacetic acid (2,4-D) degradation plasmids on 2,4-D degradation and on bacterial community structure in two different soil horizons**. *Appl Environ Microbiol* 2000, **66**:3297-3304.

For the first, time a very pronounced effect of inoculation and subsequent transfer of a catabolic plasmid on degradation of the herbicide 2,4-D was observed in the B-horizon of an agricultural soil, which was unable to degrade the added 2,4-D by itself. In addition 16S rRNA gene fingerprints confirmed the numerical dominance of the 2,4-D-degrading transconjugants within the soil community.

COR-AAD1-00001155

44. Sarand I, Haario H, Jørgensen KS, Romantschuk M: **Effect of inoculation of a TOL plasmid containing mycorrhizosphere bacterium on development of Scots pine seedlings, their mycorrhizosphere and the microbial flora in m-toluate-amended soil**. *FEMS Microbiol Ecol* 2000, **31**:127-141.

45. •• de Lipthay JR, Barkay T, Sørensen SJ: **Enhanced degradation of phenoxyacetic acid in soil by horizontal transfer of the *tfdA* gene encoding a 2,4-dichlorophenoxyacetic acid dioxygenase**. *FEMS Microbiol Ecol* 2001, **35**:75-84.
Not an entire plasmid-encoded pathway, but just one gene on an MGE can be sufficient to genetically complement other catabolic operons and thus expand the degradative capacities of indigenous soil bacterial populations by natural conjugative gene transfer.

46. •• Springael D, Peys K, Ryngaert A, Van Roy S, Hooyberghs L, Ravatn R, Heyndrickx M, van der Meer J-R, Vandecasteele C, Mergeay M *et al.*: **Community shifts in a seeded 3-chlorobenzoate degrading membrane biofilm reactor: indications for involvement of *in situ* horizontal transfer of the *clc* element from inoculum to contaminant bacteria**. *Environ Microbiol* 2002, **4**:70-80.
This paper shows how HGT of a catabolic conjugative integrative element, which is known to transfer in plate matings at very low frequencies, can nevertheless provoke important community changes in a reactor system. Horizontal transfer occurred to various bacteria, which were more competitive than the inoculant, although the latter was provided with a selective carbon source.

47. Christensen BB, Sternberg C, Andersen JB, Eberl L, Moller S, Givskov M, Molin S: **Establishment of new genetic traits in a microbial biofilm community**. *Appl Environ Microbiol* 1998, **64**:2247-2255.

48. Peters M, Heinaru E, Talpsep E, Wand H, Stottmeister U, Heinaru A, Nurk A: **Acquisition of a deliberately introduced phenol degradation operon, pheBA, by different indigenous *Pseudomonas* species**. *Appl Environ Microbiol* 1997, **63**:4899-4906.

49. Stuart-Keil KG, Hohnstock AM, Drees KP, Herrick JB, Madsen EL: **Plasmids responsible for horizontal transfer of naphthalene catabolism genes between bacteria at a coal tar-contaminated site are homologous to pDTG1 from *Pseudomonas putida* NCIB 9816-4**. *Appl Environ Microbiol* 1998, **64**:3633-3640.

50. Don RH, Pemberton JM: **Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus***. *J Bacteriol* 1981, **145**:681-686.

51. Tralau T, Cook AM, Ruff J: **Map of the IncPβ plasmid pTSA encoding widespread genes (*tsa*) for *p*-toluenesulfonate degradation in *Comamonas testosteroni* T-2**. *Appl Environ Microbiol* 2001, **67**:1508-1516.

52. Weightman AJ, Topping AW, Hill KE, Lee LL, Sakai K, Slater JH, Thomas AW: **Transposition of DEH, a broad-host-range transposon flanked by IS*Ppu12*, in *Pseudomonas putida* is associated with genomic rearrangements and dehalogenase gene silencing**. *J Bacteriol* 2001, **184**:6581-6591.

53. Tan HM: **Bacterial catabolic transposons**. *Appl Microbiol Biotechnol* 1999, **51**:1-12.

54. Müller TM, Werlen C, Spain J, van der Meer JR: **Evolution of a chlorobenzene degradative pathway among bacteria in a contaminated groundwater mediated by a genomic island in *Ralstonia***. *Environ Microbiol* 2003, **5**:163-173.

55. Wilson MS, Herrick JB, Jeon CO, Hinman DE, Madsen EL: **Horizontal transfer of *phnAc* dioxygenase genes within one of two phenotypically and genotypically distinctive naphthalene-degrading guilds from adjacent soil environments**. *Appl Environ Microbiol* 2003, **69**:2172-2181.

COR-AAD1-00001156

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, July 1996, p. 2470–2476 Vol. 62, No. 7
0099-2240/96/$04.00+0
Copyright © 1996, American Society for Microbiology

Capture of a Catabolic Plasmid That Encodes Only 2,4-Dichlorophenoxyacetic Acid:α-Ketoglutaric Acid Dioxygenase (TfdA) by Genetic Complementation

E. M. TOP,[1,2]* O. V. MALTSEVA,[1] AND L. J. FORNEY[1]

National Science Foundation Center for Microbial Ecology, Michigan State University, East Lansing, Michigan 47824,[1] and Laboratory of Microbial Ecology, Faculty of Agricultural and Applied Biological Sciences, University of Ghent, B-9000 Ghent, Belgium[2]

Received 4 December 1995/Accepted 5 April 1996

The modular pathway for the metabolism of 2,4-dichlorophenoxyacetic acid (2,4-D) encoded on plasmid pJP4 of *Alcaligenes eutrophus* JMP134 appears to be an example in which two genes, *tfdA* and *tfdB*, have been recruited during the evolution of a catabolic pathway. The products of these genes act to convert 2,4-D to a chloro-substituted catechol that can be further metabolized by enzymes of a modified *ortho*-cleavage pathway encoded by *tfdCDEF*. Given that modified *ortho*-cleavage pathways are comparatively common and widely distributed among bacteria, we sought to determine if microbial populations in soil carry *tfdA* on plasmid vectors that lack *tfdCDEF* or *tfdB*. To capture such plasmids from soil populations, we used a recipient strain of *A. eutrophus* that was rifampin resistant and carried a derivative of plasmid pJP4 (called pBH501aE) in which the *tfdA* had been deleted. Upon mating with mixed bacterial populations from soil treated with 2,4-D, transconjugants that were resistant to rifampin yet able to grow on 2,4-D were obtained. Among the transconjugants obtained were clones that contained a ca. 75-kb plasmid, pEMT8. Bacterial hosts that carried this plasmid in addition to pBH501aE metabolized 2,4-D, whereas strains with only pEMT8 did not. Southern hybridization showed that pEMT8 encoded a gene with a low level of similarity to the *tfdA* gene from plasmid pJP4. Using oligonucleotide primers based on known *tfdA* sequences, we amplified a 330-bp fragment of the gene and determined that it was 77% similar to the *tfdA* gene of plasmid pJP4 and 94% similar to *tfdA* from *Burkholderia* sp. strain RASC. Plasmid pEMT8 lacked genes that exhibited significant levels of homology to *tfdB* and *tfdCDEF*. Moreover, cell extracts from *A. eutrophus*(pEMT8) cultures did not exhibit TfdB, TfdC, TfdD, and TfdE activities, whereas cell extracts from *A. eutrophus*(pEMT8)(pBH501aE) cultures did. These data suggest that pEMT8 encodes only *tfdA* and that this gene can effectively complement the *tfdA* deletion mutation of pBH501aE.

Previous studies have shown that the ability to metabolize 2,4-dichlorophenoxyacetic acid (2,4-D) is widely distributed among phylogenetically diverse bacteria and that there is genetic and biochemical diversity in the pathways used for 2,4-D degradation (1, 9, 17, 19, 45–47). This suggests that distinct pathways for metabolism of 2,4-D may have evolved independently of one another. One of these pathways, the pathway found on plasmid pJP4 from *Alcaligenes eutrophus*, has been characterized most extensively (10, 24–26, 32, 35, 44). On this plasmid, the genes required for the metabolism of 2,4-D are organized in three transcriptional units. The *tfdA* and *tfdB* genes are transcribed separately and encode enzymes that convert 2,4-D into 3,5-dichlorocatechol, which is then further metabolized by a modified *ortho*-cleavage pathway. The *tfdCDEF* genes comprise the third transcriptional unit and encode the enzymes of the modified *ortho*-cleavage pathway. These genes exhibit moderate levels of DNA sequence similarity to *clcABC* and *tcbCDEF*, which are involved in the metabolism of 3-chlorobenzoic acid and 1,2,4-trichlorobenzene, respectively (40, 53). Likewise, the *tfdB* gene encodes a monooxygenase that converts 2,4-dichlorophenol (2,4-DCP) to 3,5-dichlorocatechol and exhibits sequence similarity with other bacterial genes that encode phenol hydroxylases (34). Thus, it appears that homologs of *tfdCDEF* and *tfdB* are comparatively common. In contrast, the *tfdA* gene does not exhibit a significant level of similarity with any other known gene and may be a unique gene that is required for initiation of 2,4-D metabolism (16). The *tfd* genes required for the conversion of 2,4-D to 3-oxoadipate are often (although not always [31, 45]) encoded on self-transmissible broad-host-range plasmids (9, 14, 50) and so are transferred together and can confer the ability to use 2,4-D as a sole carbon source. However, it is also conceivable that hosts with homologs of *tfdCDEF* and *tfdB* could metabolize 2,4-D if they were to acquire *tfdA*. The purpose of this study was to determine if certain plasmids in soil bacterial populations encode only *tfdA* and lack homologs of the other *tfd* genes found on plasmid pJP4. We found this to be the case for one of the three plasmids captured by genetic complementation.

* Corresponding author. Mailing address: Laboratory of Microbial Ecology, Department of Biochemical and Microbial Technology, University of Ghent, Coupure Links 653, B-9000 Ghent, Belgium. Phone: 32/9/264 5976. Fax: 32/9/264 6248. Electronic mail address: Eva.Top @rug.ac.be.

MATERIALS AND METHODS

Bacterial strains and plasmids. *A. eutrophus* JMP228(pBH501aE) was the recipient strain used as the genetic sink in the plasmid capture experiments (50); plasmid pBH501aE is a derivative of plasmid pJP4 (9) that was obtained through site-specific deletion of the 566-bp *Nru*I fragment of the *tfdA* gene into which an *nptII* (kanamycin resistance) cassette was inserted. This strain was used in experiments designed to capture plasmids from soil. It was found that these plasmids had to carry only a *tfdA*-like gene since the rest of the 2,4-D catabolic pathway is located on pBH501aE and the chromosome of *A. eutrophus* JMP228 (50). *A. eutrophus* BH501aE is JMP134 (9) in which plasmid pJP4 is replaced by pBH501aE. *A. eutrophus* JMP228n is a nalidixic acid-resistant mutant of JMP134

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001157

cured of plasmid pJP4 (38, 50). *Escherichia coli* CM120(RP4) is a *trpA229-tonB* mutant of W3110 (2) and harbors the IncP plasmid RP4 (Tra^+ Tc^r Km^r Ap^r) (6, 48). *E. coli* HB101(pRK2013) (13) was used as a helper strain to mobilize pSP329. Plasmid pSP329 (Tc^r) is a broad-host-range IncP cloning vector that was constructed and kindly provided by T. Tsoi. *E. coli* S17-1λpir(pUTmini-Tn*5Km1*) (8) and DH5α (12) were used in mating experiments performed to tag pEMT8 with a kanamycin resistance gene.

Media and culture conditions. The media and culture conditions used in the mating experiments and for maintenance of the strains were the same as the media described by us previously (50). All cultures to be used for plasmid DNA extraction and mating experiments were grown in Luria-Bertani (LB) broth containing the appropriate antibiotics when required. Cultures to be used for measurements of growth and 2,4-D disappearance and for the enzyme assays were grown in MMO mineral medium (43) supplemented with pyruvate (500 mg/liter) and 2,4-D (250 mg/liter).

Plasmid capture from soil via conjugation. The soil sample used in this study was obtained from an untreated control plot at the National Science Foundation long-term ecological research site at the Kellogg Biological Station in Hickory Corners, Mich.; this plot was adjacent to plots that had been regularly treated with 2,4-D since 1988 (20). The sample was collected in the fall of 1993 and was also used in our previous plasmid capture experiments (50). The sample was stored at 4°C until it was used and was preincubated at room temperature (23 ± 2°C) for 2 to 4 weeks prior to the start of the mating experiments. A 100-g soil sample in a 200-ml beaker was treated with 2,4-D in phosphate buffer at a final concentration of 100 mg/kg. One 100-g control sample received only phosphate buffer. The soil samples were incubated at room temperature (23 ± 2°C), and 18 days after 2,4-D was added, the plasmid capture experiment was performed as described previously (50) by using a 5-g subsample.

The total number of heterotrophic bacteria and the most probable number of 2,4-D-degrading bacteria in soil samples were determined as described previously (50). To show that the putative transconjugants were derived from the recipient, the Rep-PCR fingerprints (7, 57) were compared.

Plasmid DNA isolation and Southern hybridization. Plasmid DNA was isolated by using a modified method of Kado and Liu (21) with modifications that have been described previously (49, 50). Plasmid DNA digested with restriction enzymes was analyzed by Southern hybridization (37) by using DNA probes from internal regions of the six structural genes for 2,4-D metabolism (*tfdA*, *tfdB*, *tfdC*, *tfdD*, *tfdE*, and *tfdF*) encoded on plasmid pJP4 (19, 50). Hybridization experiments under high-, medium-, and low-stringency conditions were performed as described previously (17, 50).

Tagging of plasmid pEMT8 with a kanamycin resistance gene. To facilitate characterization of pEMT8, a kanamycin resistance gene was introduced into the plasmid. To do this, *A. eutrophus* JMP228(pEMT8)(pBH501aE) was first cured of plasmid pBH501aE through mobilization of pSP329 (Tc^r) into the strain by means of helper plasmid pRK2013. This yielded *A. eutrophus* JMP228 (pEMT8)(pSP329) transconjugants, which were then mated with *E. coli* S17-1λpir(pUTmini-Tn*5Km1*) (8). The mating mixture was resuspended in LB broth containing 100 μg of rifampin per ml and 50 μg of kanamycin per ml. After three transfers in the same broth, the cells were pelleted, resuspended in LB broth, and mated with *E. coli* DH5α. DH5α(pEMT8::mini-Tn*5Km1*) transconjugants were selected on LB agar containing 50 μg of kanamycin per ml that was incubated at 42°C, a temperature that is nonpermissive for the growth of *A. eutrophus* JMP228. The presence of pEMT8::mini-Tn*5Km1* was confirmed by the results of an analysis of plasmid DNA restricted with *Eco*RI.

Plasmid mating experiments. Plasmids were transferred among various bacterial strains by conjugation on LB agar plates incubated at 28°C overnight as described previously (28).

Cloning and sequence analysis of the *tfdA*-like gene of pEMT8. A 10.5-kb *Sac*I fragment of pEMT8 was ligated (37) into the *Sac*I site of pGEM3Zf(+) (Promega Corp., Madison, Wis.) and transformed by electroporation into *E. coli* XL1Blue (Stratagene) by using the instructions of the manufacturer (Bio-Rad Laboratories, Richmond, Calif.).

Oligonucleotide primers internal to *tfdA* (33, 51) were used to amplify a region of the *tfdA*-like gene of pEMT8 by PCR. The amplified DNA was electrophoretically separated on a 1.5% agarose gel, purified by using a GeneClean kit (Bio 101, Inc., La Jolla, Calif.), and sequenced by using an Applied Biosystems model 373A automated sequencer.

Dynamics of 2,4-D degradation. To measure rates of 2,4-D metabolism, *A. eutrophus* strains were grown in MMO mineral medium supplemented with 500 mg of pyruvate per liter and 250 mg of 2,4-D per liter. At 3-h intervals, the optical densities at 550 nm of the cultures were determined, and loss of 2,4-D and potential formation of 2,4-DCP were determined by high-performance liquid chromatography (HPLC) as described previously (19). After 24 h of incubation the cells were harvested, washed with 50 mM Tris-HCl buffer (pH 8.0), and used for enzyme assays.

Enzyme assays. 2,4-D dioxygenase (TfdA) activity was assayed by using a modification of a procedure described by Pieper et al. (36). Freshly harvested cells were resuspended to an optical density at 550 nm of about 4 and incubated with 1 mM 2,4-D at 30°C. The 2,4-D concentration was measured by HPLC (19).

2,4-Dichlorophenol hydroxylase (TfdB) activity was determined by the modified procedure of Beadle and Smith (3, 22). Freshly harvested bacterial cells were resuspended in 100 mM KH_2PO_4-Na_2HPO_4 (pH 7.6) containing 0.03 mM flavin adenine dinucleotide, 0.1 mM EDTA, and 1 mM dithiothreitol (buffer A), broken with a French pressure cell, and centrifuged at 50,000 × *g* for 40 min. The cell extract was added to reaction mixtures containing 0.15 mM NADPH in buffer A. After nonspecific NADPH oxidation had been recorded at 340 nm, 0.025 mM 2,4-dichlorophenol was added, and oxidation of NADPH was further monitored.

The cell extracts used to measure chlorocatechol 1,2-dioxygenase (TfdC), chloromuconate cycloisomerase (TfdD), and dienelactone hydrolase (TfdE) activities were prepared by resuspending cells in 50 mM Tris-HCl buffer (pH 8.0) containing 4 mM $MnSO_4$ and 0.1 mM dithiothreitol, breaking the cells with a French pressure cell, and centrifuging the preparations at 50,000 × *g* for 40 min. Chlorocatechol 1,2-dioxygenase (TfdC) activity was measured spectrophotometrically by the procedure of Dorn and Knackmuss (11). Chloromuconate cycloisomerase (TfdD) activity was determined as described previously (41) by using 2,4-dichloro-*cis*,*cis*-muconate (0.1 mM) which had been prepared by incubating 3,5-dichlorocatechol with partially purified chlorocatechol 1,2-dioxygenase (27) as the substrate. Dienelactone hydrolase (TfdE) activity was measured in the presence of 50 mM Tris-HCl buffer (pH 8.0) and *cis*-4-carboxymethylenebut-2-en-4-olide (*cis*-dienelactone) by using a modification of the procedure of Schmidt and Knackmuss (42).

Specific activities are expressed below as nanomoles of substrate converted or product formed per minute per milligram of protein; the values given are means of the values obtained in duplicate assays. Values similar to those presented in this paper were obtained in independent experiments. Protein contents were determined by using Bio-Rad Protein Assay Dye Reagent Concentrate and bovine serum albumin as the standard.

RESULTS

Isolation of 2,4-D-degradative plasmids from agricultural soil amended with 2,4-D. Plasmid-encoded *tfdA* genes were captured from microbial populations in an agricultural soil by using *A. eutrophus* JMP228 (Rif^r) that carried pBH501aE, a derivative of plasmid pJP4 in which the *tfdA* gene had been insertionally inactivated (50). Acquisition of a *tfdA* gene through horizontal plasmid transfer produced transconjugants that were able to use 2,4-D as a sole carbon source. Mating experiments done with the 2,4-D-amended soil gave 3.3×10^5 CFU of *A. eutrophus* JMP228(pBH501aE) transconjugants per ml of mating mixture. As in previous studies (50), the number of transconjugants was the same on medium/liter containing 250 mg of 2,4-D per liter and media containing 1,000 mg of 2,4-D per liter. When unamended soil was used, no 2,4-D-degrading transconjugants were found. The most probable number of 2,4-D-degrading organisms in the unamended soil was 1.1×10^2 organisms per g of soil, a value which was very low compared with the 3.2×10^7 organisms per g of soil found in soil that had been amended with 2,4-D.

A total of 32 *A. eutrophus* JMP228(pBH501aE) transconjugants from plates containing 1,000 mg of 2,4-D per liter were examined. Of these transconjugants, 28 contained pBH501aE and plasmid pEMT1, which has been described previously (50). One transconjugant lacked pBH501aE but had acquired a different plasmid, pEMT9. The remaining three transconjugants contained a previously undescribed plasmid, pEMT8, in addition to pBH501aE. Plasmid pEMT8 was ca. 75 kb long and had an *Eco*RI restriction pattern that clearly differed from the restriction patterns of pEMT1 and pEMT9 (Fig. 1).

Characterization of pEMT8. Two observations suggest that plasmids pEMT8 and pBH501aE were compatible with one another and were required for the metabolism of 2,4-D. First, cells that contained pEMT8 always contained pBH501aE, suggesting that these plasmids could be stably maintained within the same host and that pEMT8 was not an IncP plasmid. Second, efforts to transfer pEMT8 from *A. eutrophus* JMP228 (pBH501aE)(pEMT8) to *A. eutrophus* JMP228n always resulted in transconjugants which carried both pBH501aE and pEMT8 (data not shown). This suggests that both plasmids are required for 2,4-D degradation and that pEMT8 alone may not encode the entire complement of *tfd* genes. This was con-

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001158

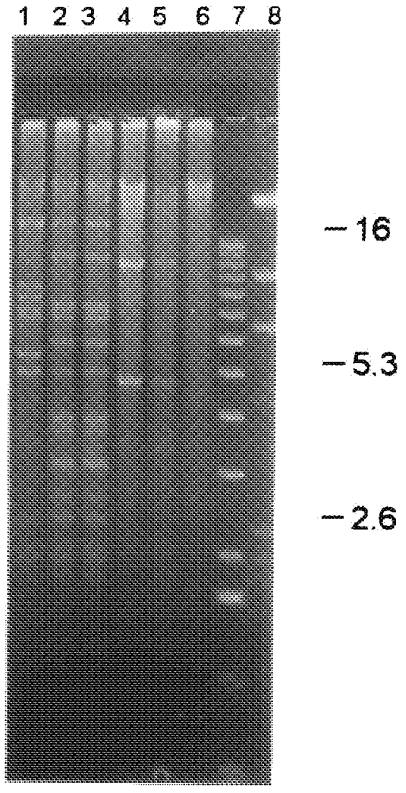


FIG. 1. Agarose gel electrophoresis of plasmid DNA after digestion with *Eco*RI. Lane 1, pEMT1 and RP4; lanes 2 and 3, pEMT8 and RP4; lanes 4 and 5, pEMT9 and RP4; lane 6, RP4; lane 7, 1-kb ladder; lane 8, λ *Hin*dIII. Plasmid band sizes (in kilobases) are indicated on the right. *Eco*RI digestion of RP4 yields one 60-kb fragment.

firmed by curing *A. eutrophus* JMP228(pBH501aE)(pEMT8) of pBH501aE by introducing IncPα plasmid RP4 by conjugation with *E. coli* CM120(RP4). Since RP4 and pBH501aE belong to the same incompatibility group (IncP), selection for RP4 (Tc^r) resulted in a loss of pBH501aE (Fig. 1). The resulting clones were unable to grow in mineral salts medium with 2,4-D as a sole carbon source, indicating that pEMT8 alone is not able to confer the ability to metabolize 2,4-D. In a control experiment, the same procedure was followed with *A. eutrophus* JMP228 (pBH501aE)(pEMT1). Previous studies had shown that pEMT1 alone confers the ability to metabolize 2,4-D (50). As expected, clones cured of plasmid pBH501aE retained the ability to metabolize 2,4-D.

The ability of the *tfdA* homolog on pEMT8 to complement a *tfdA* mutant was demonstrated in several ways. First, *A. eutrophus* JMP228(pEMT8)(RP4) (Rif^r) was mated with *A. eutrophus* BH501aE (carrying pBH501aE) (Rif^s), and Rif^r transconjugants able to grow with 2,4-D as a sole carbon source were obtained. These clones were shown to harbor plasmids pBH501aE and pEMT8. Second, JMP228(pEMT8)(RP4) was used as a donor in a mating experiment performed with *Burkholderia* sp. strain TFD6-1b, a mutant of the 2,4-D-degrading strain TFD6, which has a Tn*5* insertion in a chromosomally encoded *tfdA* gene (31). Transconjugants of *Burkholderia* sp. strain TFD6-1b able to metabolize 2,4-D were found to harbor pEMT8 (data not shown). Third, transfer of the tagged plasmid, pEMT8::mini-Tn*5Km1*, from *E. coli* DH5α to *A. eutrophus* JMP228(pBH501aE) yielded *A. eutrophus* transconjugants that metabolized 2,4-D. These results indicate that pEMT8 is a self-transmissible broad-host-range plasmid that encodes at least a TfdA-like function that is able to complement the *tfdA* mutations found on plasmid pBH501aE and in the chromosome of *Burkholderia* sp. strain TFD6-1b.

Characterization of the *tfdA* homolog of pEMT8. Southern hybridization analyses were done to determine if plasmid pEMT8 encoded genes that were homologs of the *tfd* genes of plasmid pJP4. The *tfdA* gene of plasmid pJP4 did not hybridize to restricted pEMT8 DNA under high-stringency conditions. However, under medium-stringency conditions, the *tfdA* gene

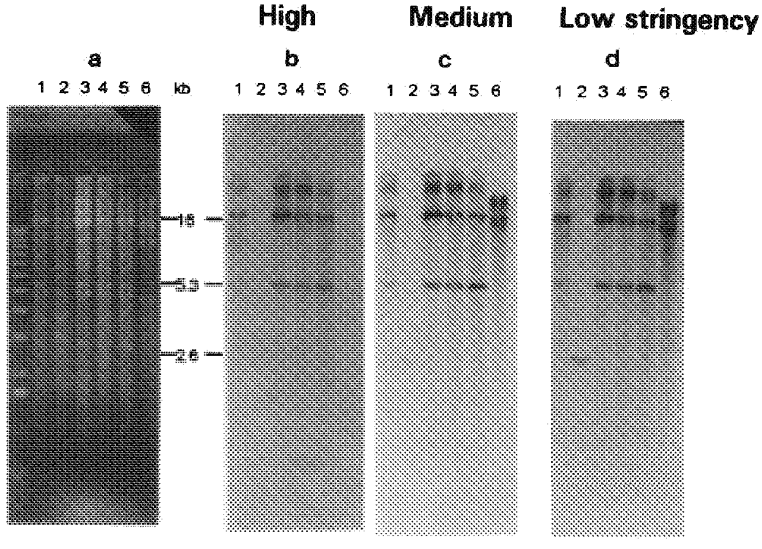


FIG. 2. (a) Agarose gel electrophoresis of plasmid DNA after digestion with *Eco*RI. (b through d) Southern blot hybridized with the *tfdA* gene probe of pJP4 under high-, medium-, and low-stringency conditions. Lanes 1, 3, 4, and 5, pEMT1 and pBH501aE; lanes 2, pEMT8 and pBH501aE; lanes 6, pEMT9. Plasmid band sizes (in kilobases) are indicated between panels a and b.

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001159

TABLE 1. Hybridization of *A. eutrophus* strains carrying different plasmids with probes specific for the *tfd* genes of plasmid pJP4 from *A. eutrophus* JMP134

Probe	Hybridization[a]				
	JMP134	JMP228 (pEMT8::mini-Tn*5Km1*)	JMP228 (pBH501aE) (pEMT8)	JMP228 (pBH501aE)	JMP228
tfdA	+++	+	+	NH	NH
tfdB	+++	NH	+++	+++	NH
tfdC	+++	NH	+++	+++	NH
tfdD	+++	NH	+++	+++	NH
tfdE	+++	NH	+++	+++	NH
tfdF	+++	NH	+++	+++	NH

[a] +++, hybridization occurred under high-stringency conditions, as described in Materials and Methods; +, hybridization occurred under low-stringency conditions, as described in Materials and Methods; NH, no hybridization occurred under low-stringency conditions.

of plasmid pJP4 hybridized weakly to a 2.3-kb *Eco*RI fragment, while under low-stringency conditions the gene hybridized to two *Eco*RI fragments (2.3 and 5.9 kb) (Fig. 2 and data not shown). This suggests there was an *Eco*RI site in the *tfdA* gene of pEMT8, as there is in the *tfdA* gene of pJP4. In contrast to pEMT8, restricted pEMT1 and pEMT9 hybridized to *tfdA* under high- and medium-stringency conditions (Fig. 2). Unlike the *tfdA* gene, the *tfdB*, *tfdC*, *tfdD*, *tfdE*, and *tfdF* genes of plasmid pJP4 did not hybridize to pEMT8, even under low-stringency conditions (Table 1).

A 10.5-kb *Sac*I fragment that hybridized with the *tfdA* gene probe (data not shown) was cloned into pGEM3Zf(+) and used as a template for PCR amplification of an internal region of the *tfdA* gene. The amplification product was ca. 330 bp long, which was similar to the length of the product obtained when the *tfdA* gene of plasmid pJP4 was used as the template. The nucleotide sequence of this fragment was determined and was found to be 77% similar to the partial *tfdA* sequence of plasmid pJP4, 94% similar to the partial *tfdA* sequences of *Burkholderia* sp. strains RASC and TFD6 (31, 45), and identical to the partial *tfdA* sequence of strain I-18, a member of the *Halomonadaceae* (30) (data not shown).

Tfd enzyme activities. In order to confirm that plasmid pEMT8 encodes only 2,4-D dioxygenase (TfdA) and no other enzymes of the 2,4-D degradation pathway, the activity of TfdA in whole cells and the activities of TfdB, TfdC, TfdD, and TfdE in cell extracts of *A. eutrophus* JMP228(pEMT8::mini-Tn*5Km1*) were determined and compared with the enzyme activities of several other strains (Table 2). Cells of *A. eutrophus* JMP228(pEMT8::mini-Tn*5Km1*) were shown to have TfdA activity. In contrast, no chlorophenol hydroxylase (TfdB) activity, chlorocatechol-1,2-dioxygenase (TfdC) activity, or chloromuconate cycloisomerase (TfdD) activity was detected (Table 2). The strain examined did exhibit a very low level of dienelactone hydrolase (TfdE) activity, but comparable levels of activity were seen in *A. eutrophus* JMP228 and *A. eutrophus* JMP228(pBH501aE), which were used as negative controls (Table 2). The levels of TfdC, TfdD, and TfdE activities were very low in *A. eutrophus* JMP228(pBH501aE) cell extracts, possibly because 2,4-dichloromuconate, the putative inducer of the modified *ortho*-cleavage pathway, was not formed (5, 29). The activities of all enzymes of the 2,4-D degradation pathway in *A. eutrophus* JMP228(pBH501aE)(pEMT8) were of the same order of magnitude as the enzyme activities in *A. eutrophus* JMP134 containing pJP4, which was used as a positive control. These data confirm that plasmid pEMT8 encodes TfdA, but does not encode a phenol hydroxylase (TfdB) or the enzymes of the modified *ortho*-cleavage pathway of chlorocatechol degradation (TfdC, TfdD, and TfdE).

Dynamics of growth and 2,4-D degradation. The growth of *A. eutrophus* JMP228(pEMT8::mini-Tn*5Km1*) in mineral salts medium supplemented with 2,4-D and pyruvate was determined and compared with the growth of the same strain carrying either pBH501aE and pEMT8, pJP4, pBH501aE, or no plasmid (Fig. 3). Complementation of the *tfdA* mutation by pEMT8 in *A. eutrophus* JMP228(pEMT8)(pBH501aE) resulted in complete mineralization of 2,4-D at a rate that was comparable to the rate observed with *A. eutrophus* JMP134 (which contains pJP4). In contrast, *A. eutrophus* JMP228(pEMT8::mini-Tn*5Km1*) converted only a small percentage of 2,4-D to stoichiometric amounts of 2,4-DCP, which was not further transformed. This also suggests that pEMT8 does not contain a *tfdB* homolog that allows the strain to convert 2,4-DCP to 3,5-dichlorocatechol.

DISCUSSION

Previous studies have shown that *tfdA*-like genes are widely distributed in phylogenetically diverse bacterial hosts from different parts of the world. These genes have been found as part of genetic mosaics with other *tfd* catabolic genes that differ in their degrees of similarity to each other and to the genes found on the canonical pathway encoded on plasmid pJP4 (17). These homologs of *tfd* genes probably underwent recombination with transcriptional units during the evolution of the catabolic genes to produce the genetic mosaics that are observed. This suggests that horizontal gene transfer, perhaps mediated in part by broad-host-range plasmids, serves to shuttle either

TABLE 2. Enzyme activities in cell extracts of *A. eutrophus* strains carrying different plasmids

Enzyme	Substrate	Enzyme activity (nmol/min/mg of protein)				
		JMP134	JMP228 (pEMT8:: mini-Tn*5Km1*)	JMP228 (pEMT8) (pBH501aE)	JMP228 (pBH501aE)	JMP228
TfdA[a]	2,4-D	38	13	36	<1	<1
TfdB	2,4-DCP	300	<5	ND[b]	<5	<5
TfdC	3,5-Dichlorocatechol	146	<2	158	<2	<2
	4-Chlorocatechol	118	<2	107	<2	<2
	Catechol	127	14	125	7	11
TfdD	2,4-Dichloromuconate	174	<2	182	<2	<2
TfdE	*cis*-Dienelactone	493	4	485	3	9

[a] TfdA activity was determined with whole cells.
[b] ND, not determined.

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001160

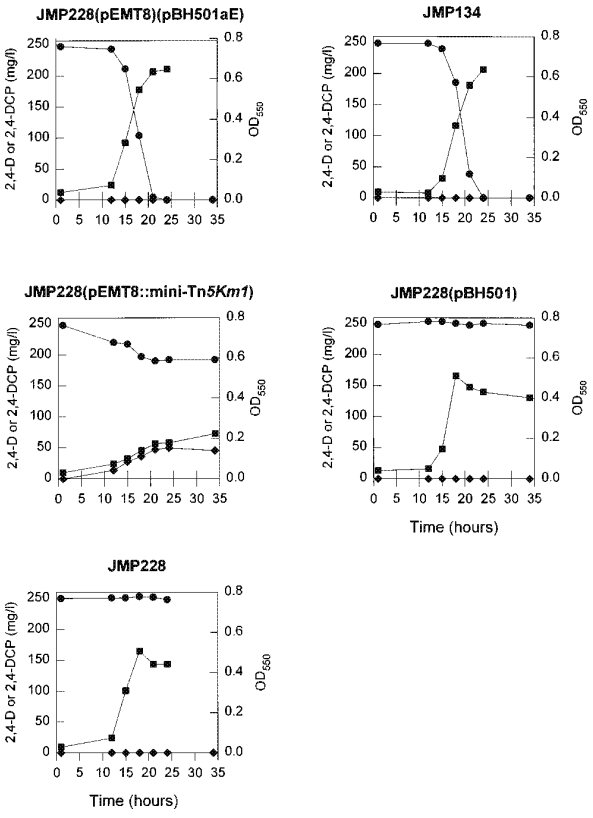


FIG. 3. Growth of different strains, disappearance of 2,4-D, and formation of 2,4-DCP in MMO mineral medium containing pyruvate (500 μg/ml) and 2,4-D (250 μg/ml). Symbols: ●, 2,4-D concentration; ■, optical density at 550 nm (OD_{550}); ◆, 2,4-DCP concentration.

single genes or more extensive genetic modules among members of the microbial gene pool. The results of our study provide direct evidence that there is a self-transferable broad-host-range vector, pEMT8, that encodes only *tfdA* and not the other *tfd* genes.

There are at least three phylogenetically distinct families of *tfdA* genes (33) that are widely distributed among strains of eubacteria (17, 47). Representatives of one gene family have high levels of DNA sequence homology to the *tfdA* gene of *A. eutrophus* JMP134 plasmid pJP4 and are often encoded on broad-host-range plasmids (50). Representatives of a second gene family have been found exclusively in strains of the genus *Burkholderia*. This gene has recently been cloned and sequenced from two distinctly different *Burkholderia* strains, strains RASC and TFD6 (31, 33, 45), and has been shown to be 77% similar to *tfdA* from pJP4 and to be chromosomally encoded. The results of analyses of partial DNA sequences of representatives of the third gene family suggest these genes are similar but not identical to the *tfdA* gene of *Burkholderia* sp. These genes have been found to be carried by strain I-18, a member of the family of *Halomonadaceae* (30), *Rhodoferax fermentans* TFD31, *R. fermentans* 6-9 (17), and *Alcaligenes paradoxus* TV1 (52). These strains represent the β and γ subgroups of the *Proteobacteria* and were isolated from geographically distinct regions, including Michigan, Oregon, Saskatchewan and Ontario in Canada, and France. The partial DNA sequence of the *tfdA* gene encoded by plasmid pEMT8, which is described in this paper, was very similar to the partial DNA sequences of members of the third gene family. These data suggest that the families of *tfdA* genes diverged from a common ancestor and have become disseminated among phylogenetically distinct species of bacteria. Often genes required for the metabolism of 2,4-D are found on conjugative broad-host-range plasmids and are horizontally transferred as a single entity (9, 14, 50). However, our data have shown that plasmid pEMT8 encodes only *tfdA* and does not contain other *tfd* genes; therefore, this plasmid represents an exception to this paradigm.

There are at least three potential explanations for why plasmid pEMT8 encodes only *tfdA*. First, it is possible that the remaining *tfd* genes were encoded on the chromosome or a second extrachromosomal element of the original host(s). Since pEMT8 was isolated only from soil treated with 2,4-D and not from untreated soil, it is likely that the original host(s) was enriched when 2,4-D was added and therefore was able to metabolize this resource. Thus, this strain(s) may have contained the whole complement of *tfd* genes encoded on more than one genetic element. Chromosomal *tfd* genes in addition to *tfdA* have been found in *Burkholderia* sp. strain RASC (23). A second possible explanation is that 2,4-D may be metabolized by microbial consortia in which the host(s) of pEMT8 converts 2,4-D to glyoxylate and 2,4-dichlorophenol, one or both of which are then metabolized by other microbial species. The results of surveys of the geographic distribution in pristine soils of bacteria able to metabolize 2,4-D suggest that such consortia may be commonly found in the environment. Fulthorpe et al. (18) used 672 soil samples collected from five continents and Hawaii as inocula for enrichment cultures and found that 2,4-D was mineralized in 63% of the cultures but that axenic cultures of 2,4-D-degrading strains were obtained from ca. 1% of these enrichments. Interestingly, an internal region of the *tfdA* gene with a nucleotide sequence nearly identical to that of plasmid pEMT8 was once amplified from a consortium by using the same PCR primers described in this study; this finding implicated homologs of the third *tfdA* gene family in the metabolism of 2,4-D by these enrichment cultures. A third potential explanation is that strains that carry plasmid pEMT8 may subsist solely through mineralization of glyoxylate that is formed upon cleavage of the 2,4-D ether bond (15).

Data from previous studies of the genetics of catabolic pathways for chlorinated aromatic compounds are consistent with the general model that peripheral enzymes are recruited for the purpose of converting the parent compounds into chloro-substituted catechols that are subsequently cleaved to produce compounds that enter bacterial intermediary metabolism (40, 53). The enzymes of various modified *ortho*-cleavage pathways are similar to one another in terms of DNA sequence, genetic organization, and the fact that they are commonly encoded on plasmids (4, 39, 40, 53). However, the peripheral enzymes of the pathways differ from one another. The evolutionary origins of these recruited genes are usually unknown, but there are a number of examples which show that they reside on transposable elements that facilitate their dissemination among bacterial populations. For example, *Pseudomonas* sp. strain P51, which metabolizes chlorobenzenes, harbors catabolic plasmid pP51, in which chlorobenzene dioxygenase (TcbA) and chlorobenzene glycol dehydrogenase (TcbB) serve as peripheral enzymes that convert chlorobenzene to a chloro-substituted catechol that is further metabolized by *ortho* cleavage catalyzed by TcbC, TcbD, TcbE, and TcbF (54, 55). The *tcbA* and *tcbB* genes are now encoded on transposon Tn*5280* (56), which suggests that the pathway evolved through the recruitment of

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001161

these genes. The evolution of the pathway used for the metabolism of 2,4-D may have evolved in an analogous manner; homologs of *tfdA* may have been recruited to the catabolic plasmids that encoded the ancestral genes of *tfdCDEF* and *tfdB*, and this may have eventually led to catabolic plasmids that encode the entire pathway for 2,4-D metabolism. This hypothesis is supported by our finding in this study that soil communities contain vectors such as pEMT8 that harbor only a *tfdA* gene and no other *tfd* genes.

The original host of plasmid pEMT8 in soil is unknown. The method which we used to examine the diversity of catabolic plasmids is dependent on the ability of the acquired plasmid to complement a genetic mutation in the *tfdA* gene of the recipient strain, thus conferring the ability to metabolize 2,4-D and providing a way to select those transconjugants that have the desired genotype. This obviates the need to cultivate the original host and may provide a more complete understanding of the genetic diversity present in the soil. Indeed, it is conceivable that pEMT8 resides in more than one bacterial host in the soil which was tested. The data presented in this paper and previously (50) show that our plasmid capture method is a useful approach for studying the genetic diversity of homologous catabolic pathways and could potentially be used to capture novel genes that could be used to broaden the substrate range or catalytic efficiency of an existing pathway or to construct a novel catabolic pathway.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation Center for Microbial Ecology (grant BIR-91-20006), by the Japan Research and Development Corporation, and by the European Community BIOTECH program (grant BIO2-CT92-0491). E.M.T. is indebted to the Belgian National Fund for Scientific Research for a postdoctoral researcher fellowship and for Research grant "Krediet aan Navorsers, 1995."

We are grateful to M. Schlömann and M. Vollmer for supplying partially purified chlorocatechol 1,2-dioxygenase. We thank all of the workers at the Research on Microbial Evolution Laboratory of the National Science Foundation Center for Microbial Ecology for many helpful suggestions and discussions.

REFERENCES

1. **Amy, P. S., J. W. Schulke, L. M. Frazierand, and R. J. Seidler.** 1985. Characterization of aquatic bacteria and cloning of genes specifying partial degradation of 2,4-dichlorophenoxyacetic acid. Appl. Environ. Microbiol. **49:**1237–1245.
2. **Bachmann, B. J.** 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol. Rev. **36:**525–557.
3. **Beadle, C. A., and A. R. W. Smith.** 1982. The purification and properties of 2,4-dichlorophenol hydroxylase from a strain of *Acinetobacter* species. Eur. J. Biochem. **123:**323–332.
4. **Burlage, R. S., L. A. Bemis, A. C. Layton, G. S. Sayler, and F. Larimer.** 1990. Comparative genetic organization of incompatibility group P degradative plasmids. J. Bacteriol. **172:**6818–6825.
5. **Coco, W. M., R. K. Rothmel, S. Henikoff, and A. M. Chakrabarty.** 1993. Nucleotide sequence and initial functional characterization of the *clcR* gene encoding a LysR family activator of the *clcABD* chlorocatechol operon in *Pseudomonas putida*. J. Bacteriol. **175:**417–427.
6. **Datta, N., R. W. Hedges, E. J. Shaw, R. B. Sykes, and M. H. Richmond.** 1971. Properties of an R factor from *Pseudomonas aeruginosa*. J. Bacteriol. **108:** 1244–1249.
7. **De Bruijn, F.** 1992. Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. Appl. Environ. Microbiol. **58:**2180–2187.
8. **de Lorenzo, V., M. Herrero, U. Jakubzik, and K. N. Timmis.** 1990. Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. J. Bacteriol. **172:**6568–6572.
9. **Don, R. H., and J. M. Pemberton.** 1981. Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J. Bacteriol. **145:**681–686.
10. **Don, R. H., A. J. Weightman, H.-J. Knackmuss, and K. N. Timmis.** 1985. Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4). J. Bacteriol. **161:**85–90.
11. **Dorn, E., and H.-J. Knackmuss.** 1978. Chemical structure and biodegradability of halogenated aromatic compounds. Two catechol 1,2-dioxygenases from a 3-chlorobenzoate grown pseudomonad. Biochem. J. **174:**73–84.
12. **Dower, W. J., J. F. Miller, and L. W. Ragsdale.** 1988. High efficiency of transformation of *Escherichia coli* by high voltage electroporation. Nucleic Acids Res. **16:**6127–6145.
13. **Figurski, D. H., and D. R. Helinski.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc. Natl. Acad. Sci. USA **76:**1648–1652.
14. **Friedrich, B., M. Meyer, and H. G. Schlegel.** 1983. Transfer and expression of the herbicide-degrading plasmid pJP4 in aerobic autotrophic bacteria. Arch. Microbiol. **134:**92–97.
15. **Fukumori, F., and R. P. Hausinger.** 1993. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetate monooxygenase" is an α-ketoglutarate-dependent dioxygenase. J. Bacteriol. **175:**2083–2086.
16. **Fukumori, F., and R. P. Hausinger.** 1993. Purification and characterization of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. J. Biol. Chem. **268:**24311–24317.
17. **Fulthorpe, R. R., C. McGowan, O. V. Maltseva, W. E. Holben, and J. M. Tiedje.** 1995. 2,4-Dichlorophenoxyacetic acid-degrading bacteria are mosaics of catabolic genes. Appl. Environ. Microbiol. **61:**3274–3281.
18. **Fulthorpe, R. R., A. N. Rhodes, and J. M. Tiedje.** 1996. Pristine soils mineralize 3-chlorobenzoate and 2,4-dichlorophenoxyacetate via different microbial populations. Appl. Environ. Microbiol. **62:**1159–1166.
19. **Holben, W. E., B. M. Schroeter, V. G. Calabrese, R. H. Olsen, J. K. Kukor, V. O. Biederbeck, A. E. Smith, and J. M. Tiedje.** 1992. Gene probe analysis of soil microbial populations selected by amendment with 2,4-dichlorophenoxyacetic acid. Appl. Environ. Microbiol. **58:**3941–3948.
20. **Ka, J. O., P. Burauel, J. A. Bronson, W. E. Holben, and J. M. Tiedje.** 1995. DNA probe analysis of the microbial community selected in the field by long-term 2,4-D application. Soil Sci. Soc. Am. J. **59:**1581–1587.
21. **Kado, C. I., and S. T. Liu.** 1981. Rapid procedure for detection and isolation of large and small plasmids. J. Bacteriol. **145:**1365–1373.
22. **Kamagata, Y.** Personal communication.
23. **Kamagata, Y., A. D. Wright, L. J. Forney, and J. M. Tiedje.** Unpublished data.
24. **Kaphammer, B., J. J. Kukor, and R. H. Olsen.** 1990. Regulation of *tfdCDEF* by *tfdR* of 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. J. Bacteriol. **172:**2280–2286.
25. **Kaphammer, B., and R. H. Olsen.** 1990. Cloning and characterization of *tfdS*, the repressor-activator gene of *tfdB*, from the 2,4-dichlorophenoxyacetic acid catabolic plasmid pJP4. J. Bacteriol. **172:**5856–5862.
26. **Kasberg, T., D. L. Daubaras, A. M. Chakrabarty, D. Kinzelt, and W. Reineke.** 1995. Evidence that operons *tcb*, *tfd*, and *clc* encode maleylacetate reductase, the fourth enzyme of the modified *ortho* pathway. J. Bacteriol. **177:**3885–3889.
27. **Kuhm, A. E., M. Schlömann, H.-J. Knackmuss, and D. H. Pieper.** 1990. Purification and characterization of dichloromuconate cycloisomerase from *Alcaligenes eutrophus* JMP134. Biochem. J. **266:**877–883.
28. **Lejeune, P., M. Mergeay, F. Van Gijsegem, M. Faelen, J. Gerits, and A. Toussaint.** 1983. Chromosome transfer and R-prime plasmid formation mediated by plasmid pULB113 (RP4::mini-Mu) in *Alcaligenes eutrophus* CH 34 and *Pseudomonas fluorescens* 6.2. J. Bacteriol. **155:**1015–1026.
29. **Leveau, J. H. J., W. M. de Vos, and J. R. van der Meer.** 1994. Analysis of the binding site of the LysR-type transcriptional activator TcbR on the *tcbR* and *tcbC* divergent promoter sequences. J. Bacteriol. **176:**1850–1856.
30. **Maltseva, O., C. McGowan, R. Fulthorpe, and P. Oriel.** Degradation of 2,4-dichlorophenoxyacetic acid by haloalkaliphilic bacteria. Microbiology, in press.
31. **Matheson, V. G., L. J. Forney, Y. Suwa, C. H. Nakatsu, A. J. Sextone, and W. E. Holben.** 1996. Evidence for acquisition in nature of chromosomal 2,4-dichlorophenoxyacetic acid etherase gene by different *Burkholderia* spp. Appl. Environ. Microbiol. **62:**2457–2463.
32. **Matrubutham, U., and A. R. Harker.** 1994. Analysis of duplicated gene sequences associated with *tfdR* and *tfdS* in *Alcaligenes eutrophus* JMP134. J. Bacteriol. **176:**2348–2353.
33. **McGowan, C.** 1995. Interspecies gene transfer in the evolution of 2,4-D degrading bacteria. Ph.D. thesis. Michigan State University, East Lansing.
34. **Nurk, A., L. Kasak, and M. Kivisaar.** 1991. Sequence of the gene (*pheA*) encoding phenol monooxygenase from *Pseudomonas* sp. EST1001: expression in *Escherichia coli* and *Pseudomonas putida*. Gene **102:**13–18.
35. **Perkins, E. J., M. P. Gordon, O. Caceres, and P. F. Lurquin.** 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. **172:**2351–2359.
36. **Pieper, D. H., K. Stadler-Fritzsche, K.-H. Engesser, and H.-J. Knackmuss.** 1993. Metabolism of 2-chloro-4-methylphenoxyacetate by *Alcaligenes eutrophus* JMP134. Arch. Microbiol. **160:**169–178.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001162

laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
38. **Sassanella, T., and M. Bagdasarian.** Unpublished data.
39. **Sayler, G. S., G. S. Hooper, A. C. Layton, and J. M. Henry King.** 1990. Catabolic plasmids of environmental and ecological significance. Microb. Ecol. **19:**1–20.
40. **Schlömann, M.** 1994. Evolution of chlorocatechol catabolic pathways: conclusions to be drawn from comparisons of lactone hydrolases. Biodegradation **5:**301–321.
41. **Schlömann, M., E. Schmidt, and H.-J. Knackmuss.** 1990. Different types of dienelactone hydrolases in 4-fluorobenzoate-utilizing bacteria. J. Bacteriol. **172:**5112–5118.
42. **Schmidt, E., and H.-J. Knackmuss.** 1980. Chemical structure and biodegradability of halogenated aromatic compounds. Halogenated muconic acids as intermediates. Biochem. J. **192:**331–337.
43. **Stanier, R. Y., N. J. Palleroni, and M. Doudoroff.** 1966. The aerobic pseudomonads: a taxonomic study. J. Gen. Microbiol. **43:**159–271.
44. **Streber, W. R., K. N. Timmis, and M. H. Zenk.** 1987. Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. **169:**2950–2955.
45. **Suwa, Y., A. D. Wright, F. Fukimori, K. A. Nummy, R. P. Hausinger, W. E. Holben, and L. J. Forney.** 1996. Characterization of a chromosomally encoded 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase from *Burkholderia* sp. strain RASC. Appl. Environ. Microbiol. **62:**2464–2469.
46. **Tiedje, J. M., J. M. Duxbury, M. Alexander, and J. E. Dawson.** 1969. 2,4-D metabolism: pathway of degradation of chlorocatechols by *Arthrobacter* species. J. Agric. Food Chem. **17:**1021–1026.
47. **Tonso, N. L., V. G. Matheson, and W. E. Holben.** 1995. Polyphasic characterization of a suite of bacterial isolates capable of degrading 2,4-D. Microb. Ecol. **30:**3–24.
48. **Top, E., I. De Smet, W. Verstraete, R. Dijkmans, and M. Mergeay.** 1994. Exogenous isolation of mobilizing plasmids from polluted soils and sludges. Appl. Environ. Microbiol. **60:**931–938.
49. **Top, E., M. Mergeay, D. Springael, and W. Verstraete.** 1990. Gene escape model: transfer of heavy metal resistance genes from *Escherichia coli* to *Alcaligenes eutrophus* on agar plates and in soil samples. Appl. Environ. Microbiol. **56:**2471–2479.
50. **Top, E. M., W. E. Holben, and L. J. Forney.** 1995. Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. Appl. Environ. Microbiol. **61:**1691–1698.
51. **Vallaeys, T., R. R. Fulthorpe, A. M. Wright, and G. Soulas.** The metabolic pathway of 2,4-dichlorophenoxyacetic acid degradation involves different families of *tfdA* and *tfdB* genes according to PCR-RFLP analysis. FEMS Microbiol. Ecol., in press.
52. **Vallaeys, T., and G. Soulas.** 1992. Gene *tfd* from the 2,4-D degrading bacteria *Alcaligenes paradoxus* TV1 could be used as probe for the research of soil microorganisms able to degrade this herbicide, p. 177–183. *In* Proceedings of the International Symposium on Environmental Aspects of Pesticide Microbiology. Department of Microbiology, Swedish University of Agricultural Sciences, Uppsala.
53. **van der Meer, J. R., W. M. de Vos, S. Harayama, and A. J. B. Zehnder.** 1992. Molecular mechanisms of genetic adaptation to xenobiotic compounds. Microbiol. Rev. **56:**677–694.
54. **van der Meer, J. R., R. I. L. Eggen, A. J. B. Zehnder, and W. M. de Vos.** 1991. Sequence analysis of the *Pseudomonas* sp. strain P51 *tcb* gene cluster, which encodes metabolism of chlorinated catechols: evidence for specialization of catechol 1,2-dioxygenase for chlorinated substrates. J. Bacteriol. **173:**3700–3708.
55. **van der Meer, J. R., A. R. W. van Neerven, E. J. de Vries, W. M. de Vos, and A. J. B. Zehnder.** 1991. Cloning and characterization of plasmid-encoded genes for the degradation of 1,2-dichloro-, 1,4-dichloro-, and 1,2,4-trichlorobenzene of *Pseudomonas* sp. strain P51. J. Bacteriol. **173:**6–15.
56. **van der Meer, J. R., A. J. B. Zehnder, and W. M. de Vos.** 1991. Identification of a novel composite transposable element, Tn*5280*, carrying chlorobenzene dioxygenase genes of *Pseudomonas* sp. strain P51. J. Bacteriol. **173:**7077–7083.
57. **Versalovic, J., T. Koeuth, and J. R. Lupski.** 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res. **19:**6823–6831.

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001163

ELSEVIER

FEMS Microbiology Ecology 28 (1999) 203–213

FEMS MICROBIOLOGY ECOLOGY

Downloaded from https://academic.oup.com/femsec/article/28/3/203/547273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

Methane oxidation as a method to evaluate the removal of 2,4-dichlorophenoxyactic acid (2,4-D) from soil by plasmid-mediated bioaugmentation

Eva M. Top [a,*], Mphekgo P. Maila [1,a], Marleen Clerinx [a], Johan Goris [b], Paul De Vos [b], Willy Verstraete [a]

[a] *Laboratory of Microbial Ecology, Department of Biochemical and Microbial Technology, Faculty of Agricultural and Applied Biological Sciences, University of Ghent, Coupure Links 653, B-9000 Ghent, Belgium*

[b] *Laboratory of Microbiology, Department of Biochemistry, Microbiology and Physiology, Faculty of Sciences, University of Ghent, Ledeganckstraat 25, B-9000 Ghent, Belgium*

Received 29 June 1998; received in revised form 27 October 1998; accepted 2 November 1998

Abstract

The herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) is known to inhibit methanotrophic bacteria. Methane oxidation was therefore used as a parameter to evaluate the residual 2,4-D after bioaugmentation of an agricultural soil. Several strains harbouring catabolic plasmids which code for the degradation of this pesticide, were compared for their potential to alleviate the negative impact of 2,4-D on methane oxidation by soil microorganisms. Three indigenous soil bacteria which contain the 2,4-D degradative plasmid pEMT1k, obtained from a donor by in situ plasmid transfer in previous experiments, were compared with *Ralstonia eutropha* JMP134, which harbours the well studied 2,4-D degradative plasmid pJP4. In addition a *Pseudomonas putida* UWC3(pEMT1k), which does not degrade 2,4-D, was used as donor to investigate the potential bioaugmentation through in situ transfer of the catabolic genes towards the indigenous soil bacteria. Both the strains that can degrade 2,4-D as well as the *P. putida* donor strain could enhance the recovery of methane oxidation by increasing the rate of degradation of 2,4-D and thus removing its toxic effect on the methane oxidising microbial populations. In all cases the time needed to oxidise methane was consistently shorter (4–10 days) in a 2,4-D treated soil inoculated with the strains, than in the non-inoculated 2,4-D treated soil, but still longer (5–10 days) than in the soil without 2,4-D. These data indicate that pesticide residues as well as their toxic effect on important soil microbial processes could be successfully removed from the soil by addition of well adapted specialised strains with the genetic information required to degrade the applied pesticides. © 1999 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.

Keywords: Methane oxidation; Bioaugmentation; 2,4-Dichlorophenoxyacetic acid; Catabolic plasmid; Soil

1. Introduction

The negative effects of pesticides on the environment caused by residues remaining in the soil and subsequently leaching to the groundwater, could be

* Corresponding author. Tel.: +32 (9) 264 5912; Fax: +32 (9) 264 6248; E-mail: eva.top@rug.ac.be

[1] Present address: CSIR (Environmentek), Building 21(B166), P.O. Box 395, Pretoria 0001, South Africa.

0168-6496/99/$20.00 © 1999 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.
PII: S0168-6496(98)00106-8

COR-AAD1-00001164

diminished by applying a short decontamination step after harvest of the crops. Microorganisms with specific genetic information for the degradation of the applied pesticides, inoculated in bulk soil or in the rhizosphere of resistant plants with dense roots, can be used to increase the degradation rate of the pesticide [1].

Methane oxidation in soil is an important ecological process. Methane, which plays a significant role as a greenhouse gas, is removed from the environment via reactions with the hydroxyl radicals in the troposphere [2,3], but also biologically through oxidation by methanotrophic and to a lesser extent ammonia oxidising bacteria in surface soils. Both aerobic and anaerobic oxidation of methane occur in nature [4]. Aerobic methanotrophs can oxidise methane to CO_2 and use it as sole source of carbon and energy. Of the several factors that have been shown to inhibit methane removal (ammonium, moisture content, temperature, etc.), pesticide use in agricultural soils has received little attention [5]. The herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) is an extensively used pesticide that has a clear negative effect on methane oxidation. In air-tight soil microcosms 2,4-D concentrations of 5 mg (kg soil)$^{-1}$ retarded methane removal, and at 50 mg kg^{-1} and higher concentrations methane oxidation was completely inhibited [5]. In another study, agricultural applications of 2,4-D at 4 kg ha^{-1} active ingredient (ca. 1.3 mg (kg soil)$^{-1}$) did not have an observable effect on soil respiration, suggesting that the microbial community was unaffected by this herbicide [6]. However, methane oxidisers represent only a small portion of the total microbial biomass and therefore their inhibition would not be observed by soil respiration tests. Because of the ecological importance of methane oxidation in soil, this microbial process could be an important parameter when evaluating the toxic effect of agrochemicals.

Bioremediation of soils containing recalcitrant chemicals by inoculation with specialised strains has been shown to fail quite often due to several reasons. A major reason is often the poor survival of the inoculated strain due to competition with the indigenous well established populations [7]. To overcome this problem, indigenous bacteria isolated from the same soil could be used, either as such or after being equipped in the laboratory with the necessary degradative genes. Another alternative is to stimulate in situ dissemination of the degradative genes, present in the applied strain, to other soil bacteria, which could result in several genera with the capability of degrading the novel pesticide.

The degradation of 2,4-D by bacteria has been studied intensively both at the biochemical, genetic, and ecological level [8]. The genes coding for 2,4-D degradation are often located on plasmids which can be transferred between a wide range of bacteria [8–10]. Recently bioaugmentation of a soil by dissemination of a 2,4-D degradative plasmid was shown to be feasible in some soils but not in all [11]. The 2,4-D removal was enhanced in inoculated soils compared to non-inoculated soils, and transconjugants that had received plasmid pEMT1k were found. The goal of the present study was to use the methane oxidation test to evaluate if bio-

Table 1
Strains and plasmids used in this study

Strain designation in this paper	Host	Plasmid[a]	Reference
S2(pEMT1k)	*Ralstonia eutropha*	pEMT1k	[8,11], this study
S4(pEMT1k)	*Ralstonia eutropha*	pEMT1k	[8,11], this study
S14(pEMT1k)	*Burkholderia* sp.	pEMT1k	[8,11], this study
JMP134	*Ralstonia eutropha* JMP134	pJP4	[12]
UWC3	*Pseudomonas putida* UWC3	–	[13]
UWC3(pEMT1k)	*Pseudomonas putida* UWC3	pEMT1k	[8,11,13]
JMP228(pEMT1k)	*Ralstonia eutropha* JMP228	pEMT1k	[8,11]
JMP228(pEMT3k)	*Ralstonia eutropha* JMP228	pEMT3k	[8,11]
JMP228(pEMT8k)	*Ralstonia eutropha* JMP228	pEMT8k	[9]
JMP228	*Ralstonia eutropha* JMP228	–	[8]
CM330	*Escherichia coli*	–	[14]

[a]– indicates strains that do not contain plasmids.

Downloaded from https://academic.oup.com/femsec/article/28/3/203/547273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

COR-AAD1-00001165

augmentation with strains harbouring 2,4-D degradative genes can accelerate the removal of 2,4-D, and, more importantly, as a consequence its toxic effect on methane oxidation in soil.

2. Materials and methods

2.1. Bacterial strains and plasmids

The strains and plasmids used in this study are listed in Table 1. Strains S2(pEMT1k), S4(pEMT1k) and S14(pEMT1k) were isolated as transconjugants from the same sandy soil Ardoyen, after inoculation of the donor strain *Escherichia coli* XL1-Blue-(pEMT1k) [11]. After identification by fatty acid methyl esters (FAME) and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (see below) the first two strains were identified as *Ralstonia eutropha* strains with different SDS-PAGE profiles, while S14(pEMT1k) is a *Burkholderia* sp. strain. They are all able to mineralise 2,4-D. The strain *Pseudomonas putida* UWC3(pEMT1k) does not mineralise 2,4-D even though it contains the plasmid pEMT1k, and the strain is also auxotrophic for isoleucine, leucine and valine. Therefore it is a good donor to examine the effect of plasmid transfer on the degradation of 2,4-D. Plasmids pEMT1k, pEMT3k and pJP4 code for the degradation of 2,4-D, while plasmid pEMT8k codes only for the first step in the degradation of 2,4-D, i.e., the conversion of 2,4-D to 2,4-dichlorophenol (DCP). Therefore the strain JMP228(pEMT8k) cannot mineralise 2,4-D [9]. Plasmids pEMT1k, pEMT3k and pEMT8k all contain the minitransposon mini-Tn*5*Km1 [15], which confers resistance to kanamycin. *R. eutropha* JMP228 is a plasmid free strain without 2,4-D degradative genes, and cannot mineralise 2,4-D [8].

2.2. Media and culture conditions

The mineral medium MMO [16], supplemented with 2,4-D (500 μg ml^{-1}), was used for growth of 2,4-D degrading strains in liquid medium and for plate counts. Kanamycin (Km; 50 μg ml^{-1}) was added when the plasmid contained a Km resistance gene, and cycloheximide (200 μg ml^{-1}) was added to solidified MMO medium (15 g l^{-1} Noble Agar, Difco) for all platings from soil, to inhibit fungal growth. Stock solutions (20 g l^{-1}) of analytical grade 2,4-D (Sigma) were prepared in 0.1 M phosphate buffer (pH 7) as described previously [8]. The strains that do not mineralise 2,4-D were grown in MMO with gluconic acid (1 g l^{-1}) instead of 2,4-D. All liquid cultures and plates were incubated at 28°C (except for *E. coli* counts, incubated at 43°C) and liquid cultures were shaken on a rotary shaker at ca. 140 rpm. For counts of *R. eutropha* JMP228, LB agar medium [17] with rifampicin (100 μg ml^{-1}) was used, *P. putida* UWC3(pEMT1k) was counted on MKB [17] supplemented with rifampicin (100 μg ml^{-1}) and kanamycin (50 μg ml^{-1}).

2.3. Soils

The Ardoyen soil is a sandy soil (pH 6.8; organic matter 0.9%; water holding capacity (WHC) 21%), collected at Zwijnaarde (Belgium) [11]. The soil was sieved (4 mm) to remove stones, roots and small branches, and the moisture content was adjusted to ca. 75% of its WHC. The soil was stored at room temperature for 1 month prior to use.

2.4. Growth curves

Strains S2(pEMT1k), S4(pEMT1k), S14(pEMT1k) and JMP134 were each grown in an MMO+2,4-D liquid medium at 28°C on a rotary shaker (140 rpm) for four days. The optical density of these cultures at 600 nm (OD_{600}) was measured and one ml of each culture was transferred to duplicate 250-ml Erlenmeyer flasks, containing 100 ml freshly prepared medium, which were incubated on a rotary shaker. The following days the OD_{600} was measured at several time points and samples for HPLC analysis (2,4-D concentration, see below) were taken at the same time. *P. putida* UWC3(pEMT1k) was tested for growth on 2,4-D in the same medium supplemented with filter-sterile solutions of isoleucine, leucine and valine (final concentration: 40 μg ml^{-1} each), and also in MMO with 2,4-D and gluconic acid (1 g l^{-1}).

Downloaded from https://academic.oup.com/femsec/article/28/3/203/547273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

COR-AAD1-00001166

2.5. Effect of inoculation of 2,4-D degraders on methane oxidation and 2,4-D degradation in soil

All strains were grown in the appropriate liquid media, as described above, and incubated overnight at 28°C on a rotary shaker. The OD_{600} and plate counts on the respective media described above were determined before inoculation in soil. In order to inoculate the soil with similar cell numbers for each strain, the amount of liquid culture to be used was calculated based on the OD_{600} of the culture. These volumes were centrifuged ($8000 \times g$, for 10 min at 20°C), washed once with saline solution (0.85% NaCl), and resuspended in 0.5 ml of this solution. Per treatment, this cell suspension, as well as 2,4-D (0.15 ml of the 20 mg ml^{-1} stock solution in phosphate buffer), and sterile demineralised water (up to 75% of WHC, i.e. 13% on wet soil) were added per 30 g of Ardoyen soil. Two types of control soils were prepared. To the inoculum-control (IC) 0.5 ml saline (without 2,4-D degraders) was added per 30 g soil instead of the cell suspension. To the herbicide-inoculum control (HC) 0.15 ml of 0.1 M phosphate buffer was added instead of the 2,4-D stock solution and saline instead of the cell suspension. The prepared soils were thoroughly mixed and in some experiments one gram of each soil preparation was set aside for analysis of the initial 2,4-D (see below). For the experiments described in Figs. 1, 2 and 4, 30 g of soil were subsequently divided into three portions that were each transferred in a 125-ml penicillin bottle (10 g of the prepared soil mixture per bottle, yielding three replicates per treatment). For the experiment represented in Fig. 5, 60 g of each soil treatment was prepared and divided into 6 bottles. This allowed to sacrifice one bottle each time 2,4-D analyses and/or plate counts were performed, still leaving three replicates for each methane measurement. As a result the 2,4-D measurements and plate counts represent single measurements. The experiment described in Fig. 3 was only set up in duplicate, and one bottle was already sacrificed at day 4 for 2,4-D analysis, leaving only one bottle per treatment. In this figure the methane and 2,4-D concentrations thus represent single measurements. In all other figures the datapoints represent the average of three replicates and the error bars represent the standard deviation. When error bars are invisible, they are hidden behind the symbols. The initial concentration of 2,4-D in the soils (except the HC control) was ca. 100 mg kg^{-1}, and the inoculation level of the strains in the inoculated soils was ca. 10^6 CFU (g soil)$^{-1}$. All penicillin bottles were sealed with rubber stoppers and aluminium caps. Subsequently 6.3 ml of methane (about 5%) was injected into each bottle through the rubber stopper to obtain a methane application of 0.025 μmol (g soil)$^{-1}$. Previous experiments have shown that there was no leakage of methane through the caps, even after 13 injections over 36 days.

In all but the first experiment, the soil was preincubated with methane. Pre-incubation was carried out by adding ca. 80 ml of methane in a 1-l bottle (ca. 8%). Consumption of methane was followed till complete depletion of methane, after which the soil was used for the experiment.

2.6. Effect of inoculation of P. putida donor strain UWC3(pEMT1k) on methane oxidation and 2,4-D degradation in soil

The same approach was used as described above for the 2,4-D degrading strains, except that *P. putida* UWC3 was grown overnight in liquid LB, and UWC3(pEMT1k) in LB supplemented with kanamycin. The cells were rinsed in saline and 0.5 ml of the cell suspension was inoculated per 30 g soil as described above, without any adjustment of the OD_{600}. The soil was always preincubated with methane (5% instead of 8%).

2.7. 2,4-D analysis in liquid cultures

One ml of liquid sample was centrifuged at $14\,000 \times g$ for 5 min, filtered through a 0.45-μm pore size filter (Acrodisc, Gelman) and transferred to special HPLC vials. Concentration of 2,4-D was determined by HPLC as described previously [11].

2.8. Conjugation and confirmation of plasmid transfer

To confirm the presence of plasmid pEMT1k in potential transconjugants formed in soil inoculated with UWC3(pEMT1k), the re-transfer of this plasmid was examined using *E. coli* CM330 as recipient in plate matings, performed as described previously

Downloaded from https://academic.oup.com/femsec/article/28/3/203/547273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

COR-AAD1-00001167

[18]. *E. coli* transconjugants were selected on LB agar containing kanamycin (50 mg l^{-1}), incubated at 43°C. The presence of plasmid pEMT1k in the *E. coli* transconjugants was confirmed by a modified plasmid extraction according to Kado and Liu [8,19], followed by gel electrophoresis.

2.9. Methane analysis

The methane concentration in each bottle was determined by injecting one ml gas of the headspace into an Intersmat IGC 120 MB gas chromatograph connected to a Hewlett-Packard 3395 integrator. The chromatograph was equipped with a dual-column arrangement, consisting of an in-series connected Poropak (50–80 mesh) column. Methane and carbon dioxide were led directly to a catharometer detector after separation in the Poropak column, by means of a mechanical 6-way valve. The column temperature was isothermic at 30°C, and the carrier gas was He at a flow rate of 20 ml min^{-1}.

2.10. Identification of transconjugants

GC analysis of fatty acid methyl esters (FAME) was performed as described previously [20]. The FAME profiles were identified using the Microbial Identification System software (MIS version 4.0). For sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), cells were harvested from tryptic soy agar (BBL) plates after 48 h of incubation at 37°C. Preparation of the cell-free extracts, gel electrophoresis and numerical comparison were performed as previously described [21] and by using the GelCompar 4.1 software package (Applied Maths, Kortrijk, Belgium).

3. Results

3.1. Growth and 2,4-D degradation in liquid medium

In a first stage the cell growth and degradation of 2,4-D in liquid medium by the *Ralstonia eutropha* strains S2(pEMT1k) and S4(pEMT1k), *Burkholderia* sp. S14(pEMT1k), and the well studied 2,4-D degrading *Ralstonia eutropha* strain JMP134 (carrying plasmid pJP4), were compared. Strain JMP134 was the fastest grower: the maximum OD_{600} was reached within 20 h and 500 mg 2,4-D l^{-1} was totally degraded within a 23-h period. Strain S2(pEMT1k) and S4(pEMT1k) were slightly slower but behaved in a very similar way; the same amount of 2,4-D was degraded within roughly the same period. Strain S14(pEMT1k), however, was clearly the slowest grower, and accordingly the degradation of 2,4-D by this culture was very slow (minimum 40 h needed for complete disappearance). It was repeatedly shown that S14 grew most slowly of all strains in the mineral medium with 2,4-D as sole carbon source (data not shown).

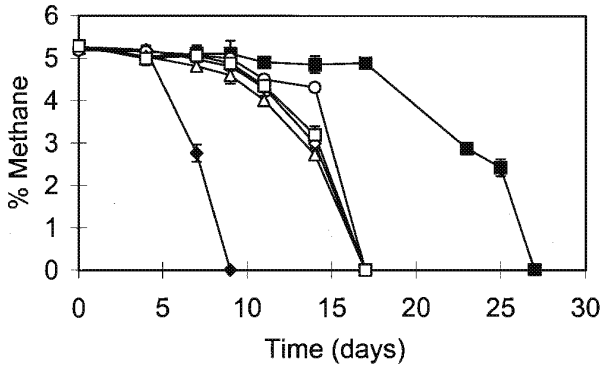


Fig. 1. Effect of strains harbouring 2,4-D degradative plasmids on methane oxidation in soil treated with 2,4-D; the soil was not preincubated with methane. JMP134 (◇); S2(pEMT1k) (□); S4(pEMT1k) (△); S14(pEMT1k) (○); HC (◆); IC (■). IC = inoculum control: uninoculated soil; HC = herbicide control: soil without 2,4-D and inoculum.

To confirm previous observations (on agar plates, unpublished results) that *P. putida* UWC3(pEMT1k) could not mineralise 2,4-D, the strain was grown in both MMO with 2,4-D and gluconic acid and in MMO with only 2,4-D. Results showed that there was only significant growth in the gluconic acid containing medium, and that there was no 2,4-D degradation in both cultures (data not shown). Hence this strain can be used as donor to investigate the effect of plasmid transfer on 2,4-D degradation in soil.

3.2. Effect of bioaugmentation with 2,4-D degrading strains on methane oxidation

The effect of addition of 2,4-D degrading strains on the methane oxidation capacity of a sandy soil was examined. First a soil not preincubated with

Downloaded from https://academic.oup.com/femsec/article/28/3/203/547273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

COR-AAD1-00001168

methane was used. Cells were inoculated at a concentration of ca. 2×10^6 CFU per g soil, as determined by plate counts of the liquid cultures. As shown in Fig. 1, the time needed for complete oxidation of methane was only ca. 9 days in the non-treated non-inoculated soil (HC soil, herbicide-inoculum control soil) while it was 27 days in the 2,4-D treated soil without inoculum (IC soil, inoculum-control soil). When the 2,4-D degrading strains were inoculated in the 2,4-D treated soil, the time period needed was 17 days for all strains, which was larger than in the HC soil but much shorter than in the IC soil. The experiment was repeated and similar results were obtained (data not shown).

In order to shorten the lag phase for methane oxidation, the experiment was repeated using a soil preincubated with methane for 10 days. The inoculum level of the strains varied between 2×10^6 and 7×10^7 CFU (g soil)$^{-1}$. Fig. 2 shows that the time required for methane removal was reduced in this preincubated soil compared to the non-preincubated soil. Again, the addition of all strains had a clear positive effect on methane oxidation in the 2,4-D treated soil. The experiment was repeated (Fig. 3A), and this time the 2,4-D concentration in the soil was monitored. From Fig. 3B it can be seen that the level of methane started to decrease one week after 2,4-D was no longer detectable in the soils. The presence of 2,4-D degraders enhanced the degradation of 2,4-D and therefore the removal

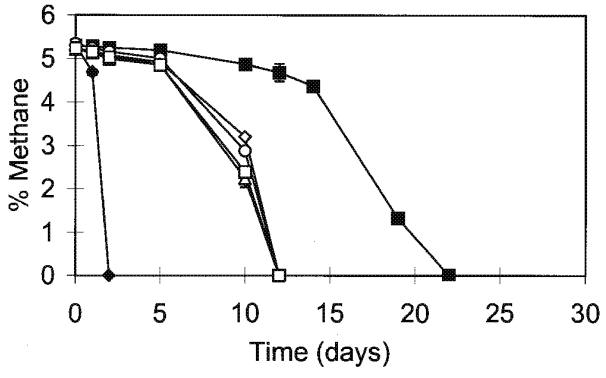


Fig. 2. Effect of strains harbouring 2,4-D degradative plasmids on methane oxidation in soil treated with 2,4-D; the soil was preincubated with methane. JMP134 (◇); S2(pEMT1k) (□); S4(pEMT1k) (△); S14(pEMT1k) (○); HC (◆); IC (■). IC = inoculum control: uninoculated soil; HC = herbicide control: soil without 2,4-D and inoculum.

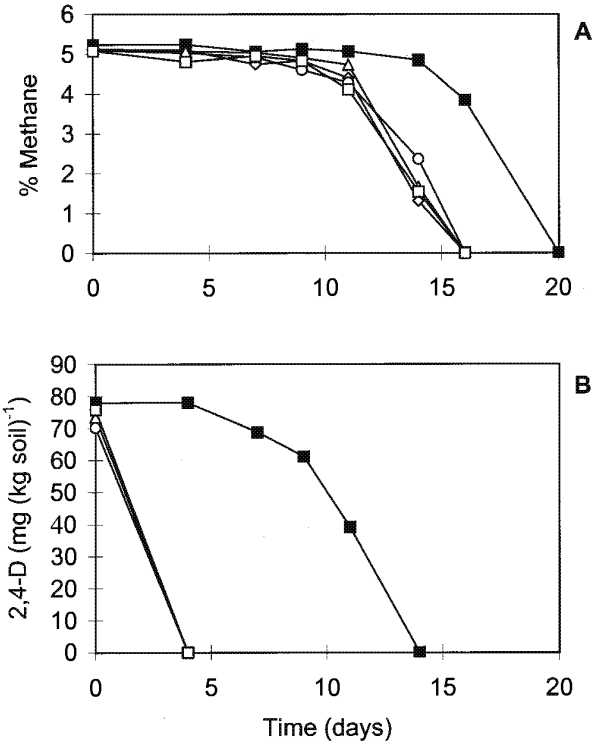


Fig. 3. Effect of strains harbouring 2,4-D degradative plasmids on methane oxidation (A) and 2,4-D degradation (B) in soil treated with 2,4-D; the soil was preincubated with methane. JMP134 (◇); S2(pEMT1k) (□); S4(pEMT1k) (△); S14(pEMT1k) (○); IC (■). IC = inoculum control: uninoculated soil.

of methane. The degradation of 2,4-D and the removal of methane both occurred faster in all soils inoculated with 2,4-D degraders compared to the IC soil. In contrast to the slower 2,4-D degradation by strain S14(pEMT1k) in liquid medium, in soil this strain seemed to degrade 2,4-D, and thus recover methane oxidation, as fast as the other strains.

3.3. *Effect of bioaugmentation through plasmid transfer on 2,4-D degradation and methane oxidation in soil*

In order to investigate whether transfer of the same 2,4-D degradative plasmid, pEMT1k, to the indigenous soil bacteria would also result in an increased removal of 2,4-D and hence a faster methane oxidation, a host, *P. putida* UWC3, was chosen which does not degrade 2,4-D in mineral medium, even when it harbours plasmid pEMT1k. Fig. 4

Downloaded from https://academic.oup.com/femsec/article/28/3/203/541273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

COR-AAD1-00001169

shows that methane oxidation was accelerated in the 2,4-D treated soils after addition of this donor strain, while the plasmid free host itself had no positive effect at all. The experiment was repeated with only a 2,4-D treated soil with and without donor, and this time the survival of the donor and the formation of transconjugants in soil, as well as the 2,4-D concentration, were monitored (Fig. 5). A high number of transconjugants was found (Fig. 5C) and no background growth was observed on the plates when the IC soil was plated on the same selective medium. It is clear from Fig. 5 that the degradation of 2,4-D from ca. 110 mg (kg soil)$^{-1}$ to ca. 50 mg (kg soil)$^{-1}$ during the first 5 days (Fig. 5B) coincided with the formation of 10^5 CFU of transconjugants per gram soil (Fig. 5C), and was soon followed by a steep decline in methane concentration, first noticed at day 8 (Fig. 5A). The results suggest that the transfer of the plasmid to the indigenous soil bacteria must be responsible for the faster recovery of methane oxidation, compared to the IC soil or soil with a plasmid free 2,4-D$^-$ donor strain.

From each set of transconjugant plates, 30 randomly picked transconjugants were purified and replicated on LB medium to compare their colony morphology. All picked transconjugants formed very slimy colonies on LB plates and showed almost identical FAME-profiles. The presence of the hydroxylated fatty acid 14:0 3OH indicated that they belong to the former rRNA group II of pseudomonads, now presented by the genera *Burkholderia* and *Ralstonia* [22]. The MIS-software identified all transconjugants as *Ralstonia solanacearum*, but with low probabilities. SDS-PAGE was performed on all 30 transconjugants and profiles were visually and numerically compared with *Burkholderia* and *Ralstonia* reference strains. All transconjugant patterns were very similar and clustered together with *R. eutropha* reference strains in the numerical comparison. We therefore conclude that they indeed belong to the species *R. eutropha*. The presence of the plasmid

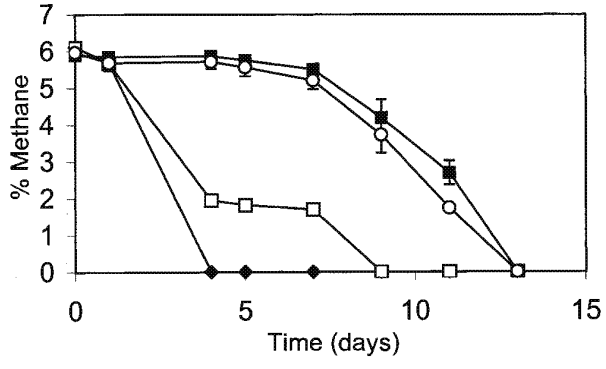


Fig. 4. Effect of transfer of plasmid pEMT1k from donor strain UWC3(pEMT1k) to soil bacteria on methane oxidation in soil treated with 2,4-D; the soil was preincubated with methane. UWC3(pEMT1k) (□); UWC3 (○); HC (◆); IC (■). IC = inoculum control: uninoculated soil; HC = herbicide control: soil without 2,4-D and inoculum.

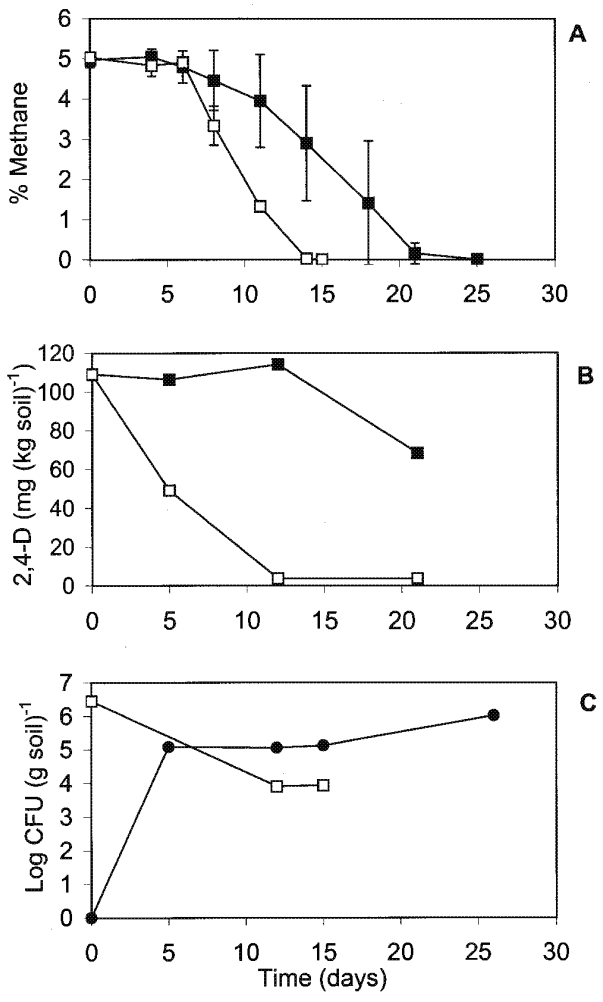


Fig. 5. Effect of transfer of plasmid pEMT1k from donor strain UWC3(pEMT1k) to soil bacteria on methane oxidation (A) and 2,4-D degradation (B) in soil treated with 2,4-D; the soil was preincubated with methane. Survival of the donor and appearance of transconjugants (C). UWC3(pEMT1k) (□); transconjugants (●); IC (■). IC = inoculum control: uninoculated soil.

Downloaded from https://academic.oup.com/femsec/article/28/3/203/547273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

COR-AAD1-00001170

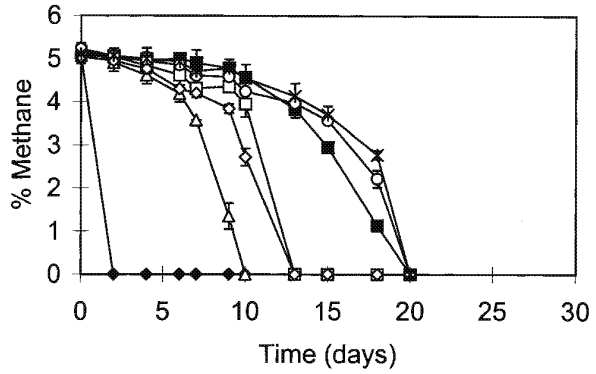


Fig. 6. Effect of strains JMP228 harbouring different or no 2,4-D degradative plasmids on methane oxidation in soil treated with 2,4-D; the soil was preincubated with methane. JMP134 (◇); JMP228(pEMT1k) (△); JMP228(pEMT3k) (□); JMP228-(pEMT8k) (○); JMP228 (×); HC (◆); IC (■). IC = inoculum control: uninoculated soil; HC = herbicide control: soil without 2,4-D and inoculum.

pEMT1k was confirmed by mating a randomly picked set of transconjugants with *E. coli* CM330 and selecting for the transfer of Km resistance, located on the plasmid. All soil transconjugants transferred the resistance marker and the presence of pEMT1k in *E. coli* CM330 was confirmed by plasmid extraction.

3.4. *Effect of inoculation of strains harbouring different plasmids on methane oxidation in soil containing 2,4-D*

In a last experiment, four *R. eutropha* strains JMP228, which each contained a different plasmid, as well as the plasmid free strain JMP228, were compared in their ability to restore the methane oxidation potential in a 2,4-D treated soil. Plasmids pEMT1k, pEMT3k, and pJP4 confer complete degradation of 2,4-D, while pEMT8k only contains the *tfdA* gene and does not code for complete 2,4-D mineralisation. Therefore strains JMP228(pEMT1k), JMP228(pEMT3k) and JMP134 were grown in MMO+2,4-D, while JMP228(pEMT8k) and JMP228 were grown in MMO+gluconic acid. The strains were inoculated at 2×10^6–6×10^6 CFU (g soil)$^{-1}$. The three strains able to mineralise 2,4-D enhanced the oxidation of methane compared to the IC soil, while the two strains incapable of mineralising 2,4-D did not (Fig. 6). The methane removal curves of the three degrading strains were very similar to those presented in the figures presented above for the S2, S4, and S14 strains with pEMT1k, and for JMP134. Of the three 2,4-D degradative plasmids, pEMT1k seemed to confer a slightly faster degradation, as reflected by a slightly faster removal of methane.

4. Discussion

This study aimed at evaluating a method for easy and straightforward monitoring of the effect of bioaugmentation of a pesticide treated soil with strains that contain degradative plasmids. The herbicide 2,4-D has been shown to inhibit methane oxidation, an important ecological process in soil [5]. As far as we know, the use of this methane oxidation capacity as a toxicity test to evaluate the recovery of a contaminated soil after bioaugmentation, has not yet been described in literature. This study successfully used the methane oxidation potential as an indication for removal of the herbicide 2,4-D from soil upon bioaugmentation with strains containing 2,4-D degradative plasmids.

All three strains used in the first experiments (*R. eutropha* S2(pEMT1k) and S4(pEMT1k) and *Burkholderia* sp. S14(pEMT1k)) showed the ability to use 2,4-D as a sole source of carbon and energy in liquid medium, indicating that they express the 2,4-D degradative genes located on plasmid pEMT1k. These strains are soil bacteria that have obtained plasmid pEMT1k from an *E. coli* donor strain via conjugative transfer in soil from the same location (Ardoyen) as the soil used in this study [11]. When their growth was compared with *R. eutropha* JMP134, which contains the well studied 2,4-D degradative plasmid pJP4, JMP134 was found to grow faster on 2,4-D compared to the other three strains, while *Burkholderia* sp. S14(pEMT1k) degraded 2,4-D at the slowest rate.

In all methane oxidation experiments methane removal in herbicide-inoculum control (HC) soil was much faster than in the inoculum control (IC). This confirms earlier findings [5] that 2,4-D has a negative effect on methane oxidation. In each case addition of different strains, harbouring the same 2,4-D degradative plasmid pEMT1k, reduced the time necessary

Downloaded from https://academic.oup.com/femsec/article/28/3/203/547273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

COR-AAD1-00001171

for removal of methane. The 2,4-D analyses showed that this rapid removal of the toxicity of 2,4-D towards methanotrophic bacteria was linked with the enhanced degradation of 2,4-D by all strains. Indeed, the level of methane started to decrease after the 2,4-D was partially or totally degraded in the soil (Figs. 3 and 6). This clearly shows that 2,4-D has a negative effect on methane oxidising bacteria and therefore its concentration has to be low enough (under 50 mg (kg soil)$^{-1}$, see Fig. 5B) before methanotrophs can start consuming methane. This threshold concentration of 50 mg kg^{-1} is in agreement with the results of Syamsul et al. [5], mentioned before, who also observed a complete inhibition of methane oxidation in soil at concentrations above 50 mg (kg soil)$^{-1}$, and a slight inhibition at 5 mg kg^{-1}. There was no significant difference between the treatments with the different strains S2(pEMT1k), S4(pEMT1k), S14(pEMT1k) and JMP134. Even inoculation of the soil with *Burkholderia* sp. strain S14(pEMT1k), which grew very slowly in liquid MMO medium resulted in removal of 2,4-D, and thus its toxicity towards methanotrophs, at a rate which was similar to those of the other strains. This could mean that degradation results obtained in synthetic liquid media are not always directly relevant for the rate of degradation in soil. On the other hand, it cannot be excluded that transfer of pEMT1k to indigenous soil bacteria and subsequent expression of the *tfd* genes was partially or totally responsible for this fast 2,4-D degradation. The fact that 2,4-D degradation and methane oxidation was also observed in the IC soil, although after a longer time, confirms that 2,4-D is readily biodegradable in soil [23].

Interestingly, the addition to the soil of a donor of the 2,4-D catabolic plasmid pEMT1k also resulted in an enhanced removal of 2,4-D and its toxicity. Since the donor strain was shown not to degrade 2,4-D in liquid medium, even in the presence of gluconic acid as alternative carbon source (unpublished data), its bioaugmentation effect is probably due to transfer of pEMT1k to indigenous soil bacteria, followed by expression of the genes and thus degradation of 2,4-D in soil. Indeed, many transconjugants were found and they grew very well on 2,4-D containing medium. All colonies examined had the same morphology and showed very similar FAME- and SDS-PAGE-profiles. They were identified as *Ralstonia eutropha*. The transfer of the same 2,4-D degradative plasmid as a means of bioaugmentation, has been shown before with *E. coli* Xl1Blue(pEMT1k) as donor strain in the same sandy soil Ardoyen [11]. In this previous study only the 2,4-D concentration in soil, as determined by HPLC after extraction with water, was used as a parameter for successful removal of 2,4-D. Indeed, faster 2,4-D disappearance was observed in the soil inoculated with the donor strain compared to the non-inoculated control soil. In the current study, first of all a more realistic approach was chosen by using a different donor strain, i.e., *P. putida,* often encountered in soil and rhizosphere, instead of *E. coli*, which is a bacterium whose normal habitat is not soil but the gastrointestinal tract. Moreover, the success of bioaugmentation was not only evaluated by measuring the 2,4-D in the soil, but also by monitoring the toxicity of 2,4-D towards methanotrophs, an ecologically important group of soil microorganisms.

The dissemination and subsequent expression of the catabolic genes to soil bacteria could mean that the genes will survive longer in the soil microbial community than if they would be contained in the inoculated strain, which most often does not survive in soil over extended periods. This can be positive if one wants to maintain the biodegradative potential in the soil, but must also be considered as a possible disadvantage in the case of pesticides. If one wants to reapply the pesticide, the persistence of such degradative genes in soil could lead to 'resistance' of this soil to this pesticide, which could lead to a lack of pesticide effect. However, it may be envisioned that during a certain period without selective pressure (i.e., without pesticide), the genes would eventually disappear as well.

When the ability to degrade 2,4-D and therefore enhance methane oxidation was compared between different 2,4-D degradative plasmids, pEMT1k, pEMT3k and pJP4, in the same host, *R. eutropha* JMP228, no real difference could be observed. It should be noted that strain JMP228(pEMT3k), containing plasmid pEMT3k, has been shown to grow much more slowly in liquid as well as on solid MMO medium with 2,4-D as sole C-source, compared to the same strain containing pEMT1k or pJP4 [8,11]. In soil, however, JMP228(pEMT3k) apparently contributed to the removal of the 2,4-D toxicity almost

Downloaded from https://academic.oup.com/femsec/article/28/3/203/547273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

COR-AAD1-00001172

as well as the strain JMP134 with plasmid pJP4 (all methane was removed within 13 days). Again, as shown for the *Burkholderia* sp. S14(pEMT1k), this shows that biodegradation and growth in synthetic media is not necessarily a good indication for the degradation behaviour of that strain in a soil environment. The strain JMP228(pEMT8k), containing plasmid pEMT8k, was included in order to examine if the first and only step in the degradation of 2,4-D, encoded by this plasmid, i.e., transformation into 2,4-dichlorophenol, could be complemented by pathways present in the soil microbial populations. Although complementation with the rest of the pathway, present on a different plasmid in the same strain, has been shown to result in complete 2,4-D degradation [9], this complementation obviously did not occur in soil as this strain was not able to take away the toxic effect of 2,4-D. Since 2,4-dichlorophenol is toxic to the cell [9], the inoculated strain may have died off rather soon, not allowing possible exchange of plasmids.

All these data show that bioaugmentation of a 2,4-D treated soil with strains with the necessary genetic catabolic information shortens the time needed for complete removal of 2,4-D and its toxicity towards methanotrophs. The transfer of plasmid pEMT1k, resulting in increased 2,4-D removal, confirms our previous results which indicate that dissemination of catabolic genes can be an alternative to the classical bioaugmentation approach, which relies on the activity of the inoculated strain. Most importantly, this study has shown that measuring methane oxidation in soil can be a very good bio-assay to evaluate the success of bioaugmentation if the pollutant or a possible metabolite inhibits this microbial process.

Acknowledgments

E.M. Top is indebted to the Fund for Scientific Research of Flanders (FWO–Vlaanderen) for a position as Research Associate (Onderzoeksleider), for an equipment grant 'Krediet aan Navorsers, 1995', and for an FWO project grant (1998–2001). This work was also partially supported by the grant GOA 1997–2002 of the 'Ministerie van de Vlaamse Gemeenschap, Bestuur Wetenschappelijk Onderzoek' (Belgium). We are grateful to S. Maertens for her excellent technical assistance and help with the figures, T. Tanghe and P. De Boever for their advice and help with the HPLC analyses, and to W. De Jonghe and S. El Fantroussi for their useful suggestions on the manuscript.

References

[1] Crowley D.E., Brennerova, M.V., Irwin, C., Brenner, V. and Focht, D.D. (1996) Rhizosphere effects on biodegradation of 2,5-dichlorobenzoate by a bioluminescent strain of root-colonizing *Pseudomonas fluorescens*. FEMS Microbiol. Ecol. 20, 79–89.

[2] Watson R.T., Rodhe, H., Oeschger, H. and Siegenthaler, U. (1990) Greenhouse gases and aerosols. In: Climate Change: The IPCC Scientific Assessment (Houghton, J.T., Jenkins, G.J. and Ephraums, J.J., Eds.), pp. 1–41. Cambridge University Press, Cambridge.

[3] Prinn, R.G. (1994) Global atmospheric-biospheric chemistry. In: Global Atmospheric-Biospheric Chemistry (Prinn, R.G., Ed.), pp. 1–18. Plenum Press, New York, NY.

[4] Conrad, R. (1996) Soil microorganisms as controllers of atmospheric trace gases (H_2, CO, CH_4, OCS, N_2O and NO). Microbiol. Rev. 60, 609–640.

[5] Syamsul Arif, M.A., Houwen, F. and Verstraete, W. (1996) Agricultural factors affecting methane in arable soil. Biol. Fertil. Soils 21, 95–105.

[6] Wardle, D.A. and Parkinson, D. (1992) Influence of the herbicides 2,4-D and glyphosate on soil microbial biomass and activity: a field experiment. Soil Biol. Biochem. 24, 185–186.

[7] Alexander, M. (1994) Biodegradation and bioremediation, 233 pp. Academic Press, California.

[8] Top, E.M., Holben, W.E. and Forney, L.J. (1995) Characterization of diverse 2,4-dichlorophenoxyacetic acid degradative plasmids isolated from soil by complementation. Appl. Environ. Microbiol. 61, 1691–1697.

[9] Top, E.M., Maltseva, O.V. and Forney, L.J. (1996) Capture of a catabolic plasmid that encodes only 2,4-dichlorophenoxyacetic acid:α-ketoglutaric acid dioxygenase (TfdA) by genetic complementation. Appl. Environ. Microbiol. 62, 2470–2476.

[10] Don, R.H. and Pemberton, J.M. (1985) Genetic and physical map of the 2,4-dichlorophenoxyacetic acid degradative plasmid pJP4. J. Bacteriol. 161, 466–468.

[11] Top, E.M., Van Daele, P., De Saeyer, N. and Forney, L.J. (1998) Enhancement of 2,4-dichlorophenoxyacetic acid (2,4-D) degradation in soil by dissemination of catabolic plasmids. Antonie van Leeuwenhoek 73, 87–94.

[12] Don, R.H. and Pemberton, J.M. (1981) Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J. Bacteriol. 145, 681–686.

[13] Hill, K.E., Weightman, A.J. and Fry, J.C. (1992) Isolation and screening of plasmids from the epilithon which mobilize recombinant plasmid pD10. Appl. Environ. Microbiol. 58, 1292–1300.

Downloaded from https://academic.oup.com/femsec/article/28/3/203/547273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

COR-AAD1-00001173

[14] Springael, D., Kreps, S. and Mergeay, M. (1993) Identification of a catabolic transposon, Tn*4371*, carrying biphenyl and 4-chlorobiphenyl degradation genes in *Alcaligenes eutrophus* A5. J. Bacteriol. 175, 1674–1681.

[15] de Lorenzo, V., Herrero, M., Jakubzik, U. and Timmis, K.N. (1990) Mini-Tn5 transposon derivatives for insertion mutagenesis, promotor probing, and chromosomal insertion of cloned DNA in Gram-negative eubacteria. J. Bacteriol. 172, 6568–6572.

[16] Stanier, R.Y., Palleroni, N.J. and Doudoroff, M. (1966) The aerobic pseudomonads: a taxonomic study. J. Gen. Microbiol. 43, 159–271.

[17] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual, 2nd Edn., 1626 pp. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[18] Lejeune, P., Mergeay, M., Van Gijsegem, F., Faelen, M., Gerits, J. and Toussaint, A. (1983) Chromosome transfer and R-prime plasmid formation mediated by plasmid pULB113 (RP4::mini-Mu) in *Alcaligenes eutrophus* CH 34 and *Pseudomonas fluorescens* 6.2. J. Bacteriol. 155, 1015–1026.

[19] Kado, C.I. and Liu, S.T. (1981) Rapid procedure for detection and isolation of large and small plasmids. J. Bacteriol. 145, 1365–1373.

[20] Vauterin, L., Yang, P., Hoste, B., Vancanneyt, M., Civerolo, E.L, Swings, J. and Kersters, K. (1991) Differentiation of *Xanthomonas campestris* pv. *citri* strains by sodium dodecyl sulfate-polyacrylamide gel electrophoresis of proteins, fatty acid analysis, and DNA-DNA hybridization. Int. J. Syst. Bacteriol. 41, 535–542.

[21] Pot, B., Vandamme, P. and Kersters, K. (1994) Analysis of electrophoretic whole organism protein fingerprints. In: Modern Microbial Methods. Chemical Methods in Prokaryotic Systematics (Goodfellow, M. and O'Donnell, A.G., Eds.), pp. 493–521. Wiley, Chichester.

[22] Vancanneyt, M., Witt, S., Abraham, W.-R., Kersters, K. and Fredrickson, H.L. (1996) Fatty acid content in whole-cell hydrolysates and phospholipid fractions of pseudomonads: a taxonomical evaluation. Syst. Appl. Microbiol. 19, 528–540.

[23] Holben, W.E., Schroeter, B.M., Calabrese, V.G., Olsen, R.H., Kukor, J.K., Biederbeck, V.O., Smith, A.E. and Tiedje, J.M. (1992) Gene probe analysis of soil microbial populations selected by amendment with 2,4-dichlorophenoxyacetic acid. Appl. Environ. Microbiol. 58, 3941–3948.

Downloaded from https://academic.oup.com/femsec/article/28/3/203/471273 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

COR-AAD1-00001174

Characterization of an intradiol dioxygenase involved in the biodegradation of the chlorophenoxy herbicides 2,4-D and 2,4,5-T

V.M. Travkin[a], A.P. Jadan[a], F. Briganti[b], A. Scozzafava[b], L.A. Golovleva[a,*]

[a]*Institute of Biochemistry and Physiology of Microorganisms, Prospect Nauka 5, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia*

[b]*University of Florence, Department of Chemistry, Via Capponi 7, 50121 Florence, Italy*

Received 30 January 1997; revised version received 25 February 1997

Abstract Hydroxyquinol 1,2-dioxygenase, an intradiol dioxygenase, which catalyzes the cleaving of the aromatic ring of hydroxyquinol, a key intermediate of 2,4-D and 2,4,5-T degradation, was purified from *Nocardioides simplex* 3E cells grown on 2,4-D as the sole carbon source. This enzyme exhibits a highly restricted substrate specificity and is able to cleave hydroxyquinol (K_m for hydroxyquinol as a substrate was 1.2 μM, V_{max} 55 U/mg, K_{cat} 57 s^{-1} and K_{cat}/K_m 47.5 μM s^{-1}), 6-chloro- and 5-chlorohydroxyquinol. Different substituted catechols and hydroquinones are not substrates for this enzyme. This enzyme appears to be a dimer with two identical 37-kDa subunits. Protein and iron analyses indicate an iron stoichiometry of 1 iron/65 kDa homodimer, α_2 Fe. Both the electronic absorption spectrum which shows a broad absorption band with a maximum at 450 nm and the electron paramagnetic resonance spectra are consistent with a high-spin iron(III) ion in a rhombic environment typical of the active site of intradiol cleaving enzymes.

© 1997 Federation of European Biochemical Societies.

Key words: Hydroxyquinol 1,2-dioxygenase; Intradiol dioxygenase; Chlorophenoxyalkanoic herbicide; 2,4-D; 2,4,5-T; *Nocardioides simplex*

1. Introduction

2,4-Dichlorophenoxyacetate (2,4-D) is at present the most commonly used herbicide since another phenoxy-herbicide, 2,4,5-T, has been banned in many countries due to its high toxicity. Furthermore, their biodegradation intermediates 2,4-dichloro- and 2,4,5-trichloro-phenols (2,4-DCP and 2,4,5-TCP) are still used as biocides and constitute a major group of environmental pollutants [1].

Nocardioides simplex 3E, isolated by enrichment culture techniques under selective pressure of 2,4,5-T, completely mineralizes both 2,4-D and 2,4,5-T and the corresponding chlorophenols [2].

Based on identification of 2,4-D and 2,4,5-T degradation pathways and activity of key enzymes it was suggested that degradation of both herbicides by *N. simplex* 3E proceeds through the formation of trihydroxylated benzene [3].

We report here the identification, purification and preliminary characterization of an enzyme from this strain which catalyzes the splitting of the aromatic ring of hydroxyquinol with high specificity.

2. Material and methods

2.1. Enzyme production and purification

Nocardioides simplex 3E isolated by enrichment culture techniques from soils treated with chloroaromatic pesticides was grown in a 10-l fermentor with 7 liters of a mineral medium containing 2,4-D as the sole source of carbon. Conditions of cultivation and preparation of cell free extracts were as previously described [4].

Cell free extract was brought to 37% saturation with powdered $(NH_4)_2SO_4$ and stirred for 1 h at 4°C before centrifugation. The pellet was discharged and the supernatant solution was brought to 67% $(NH_4)_2SO_4$ saturation. The precipitate was collected by centrifugation, dissolved in 40 ml of 50 mM Tris-HCl buffer pH 7.8 (buffer A) and dialyzed against 5 liters of the same buffer.

After removal of $(NH_4)_2SO_4$ the enzyme solution was loaded onto a DEAE-Toyopearl column (25/30) previously equilibrated with buffer A, washed with 200 ml of the same buffer, and a linear gradient from buffer A to buffer A plus 0.5 M NaCl (2 l) with a flow rate of 4 ml/min was applied. Active fractions were pooled and an equal volume of buffer A was added. The enzyme solution was then loaded onto a Q-Sepharose FF column (28/16) previously equilibrated with buffer A plus 0.2 M NaCl, washed with 50 ml of the same buffer, and a linear gradient from 0.2–0.5 M NaCl in the same buffer (800 ml) with a flow rate of 1.5 ml/min was applied. The enzyme eluted with 0.31–0.35 M NaCl. Active fractions were pooled and 3.5 M solution of ammonium sulfate was added to the final concentration of 0.7 M. The enzyme solution was loaded onto a column of phenyl-Superose CL-4B (20/4) previously equilibrated with 0.7 M of ammonium sulfate in buffer A. The column was washed with 10 ml of the same buffer with ammonium sulfate and a linear gradient of 0.7–0 M ammonium sulfate (100 ml) with a flow rate of 1 ml/min was applied. The enzymatic activity eluted at 0.6–0.1 M of ammonium sulfate. Active fractions were pooled and concentrated to 0.5 ml by ultrafiltration with Amicon membrane PM 10 and Centricon 10. The concentrated solution was loaded onto a column of Superdex 200 (16/70) equilibrated with buffer A plus 0.2 M of NaCl and eluted with the same buffer and a flow rate of 0.5 ml/min. Active fractions were combined and loaded onto a Mono-Q 5/5 column in a Pharmacia FPLC system. The column was previously equilibrated with buffer A plus 0.2 M NaCl. The elution was performed by broken gradient of 0.2–0.5 M NaCl in buffer A (35 ml). The enzyme was eluted with 0.32–0.35 M NaCl. Hydroxyquinol was used as a substrate for enzyme assays during the purification.

2.2. Enzyme activity assay

The activity of HQ 1,2-DO was assayed either polarographically with a YSI 5300 device (USA) as described by Sze and Dagley [5] in 50 mM MES buffer (pH 6.5) or spectrophotometrically with a Shimadzu UV 160 spectrophotometer (Japan) according to the modified method of Tiedje et al. [6]. The reaction mixture contained 0.1 M MES buffer (pH 6.5), 0.1 mM hydroquinol and the enzyme. The reaction was started by the addition of substrate. Activity was measured by determining the increase in absorption at 243 nm which correlates with the formation of maleylacetate, molar extinction 4440 M^{-1} cm^{-1} [6].

*Corresponding author. Fax: (7) (95) 9233602.
E-mail: GOLOVLEVA@ibpm.serpukhov.su

Abbreviations: 2,4-D, 2,4-dichlorophenoxyacetate; 2,4,5-T, 2,4,5-trichlorophenoxyacetate; HQ, hydroxyquinol; CHQ, chlorohydroxyquinol; HQ 1,2-DO, hydroxyquinol 1,2-dioxygenase; CHQ 1,2-DO, chlorohydroxyquinol 1,2-dioxygenase; 2,4,5-TCP, 2,4,5-trichlorophenol; EPR, electron paramagnetic resonance

0014-5793/97/$17.00 © 1997 Federation of European Biochemical Societies. All rights reserved.
PII S0014-5793(97)00297-4

COR-AAD1-00001175

One unit of enzyme activity was defined as the amount (mg) of enzyme that catalyzed the formation of 1 μmol of product per min.

Protein concentration, during purification, was determined using the A_{280}/A_{260} method of Warburg and Christian [7], and the concentration of the purified enzyme was determined using the A_{280}/A_{205} method of Scopes [8].

2.3. Analytical methods

The absorption spectrum of the enzyme was recorded using a Shimadzu UV 160 or a Varian Cay 3 spectrophotometer.

The content of metal ions in the purified enzyme was determined using atomic absorption spectroscopy with flame atomization, spectrophotometer model AAS 5100/Zeeman, Perkin Elmer.

X band EPR spectra were recorded at 4.2 K on a Bruker ER200 spectrometer equipped with an Oxford continuous flow ESR 900 cryostat.

Electrophoresis in 12.5% polyacrylamide gel in the presence of SDS was performed according to a modification of the Laemmli method [9]. The protein content in samples was 2–5 μg. The low molecular weight kit MW-SDS-70L from Sigma (USA) was used as standard. The proteins were stained with Coomassie R-250 [10].

For the estimation of molecular mass under non-denaturing conditions protein was applied on a Superdex-200 (16/70) and eluted with buffer A plus 0.2 M NaCl. The molecular weight was calculated from a standard linear regression curve for the reference proteins (M_r 17 800–150 000).

The substrate specificity of the enzyme was determined in 50 mM MES buffer pH 6.5 by polarography with a Clark DW 1 oxygen electrode using a YSI 5300 device (USA) and spectrophotometrically with freshly prepared solutions of di- and trihydroxylated compounds. Trihydroxylated substrated solutions were prepared under anoxic (nitrogen) conditions.

2.4. Analysis of kinetic data

K_m and V_{max} values were calculated from 1/S versus 1/V plots. K_{cat} values were determined using a molecular mass of 65 kDa for the enzyme. Inhibition constants were estimated graphically using the Dixon method [11].

Chemicals, biochemicals and reagents were purchased from Sigma (USA) and Serva (Germany); HQ was from Aldrich (USA).

6-Chloro-, 2-methoxy-6-chloro- and 5-chlorohydroxyquinols were synthesized by Yu. Shuruchin according to a modification of the method described in [12]. Q-Sepharose FF and phenyl-Sepharose-CL-6B were from Pharmacia (Sweden), DEAE-Toyopearl-650M was from Toyo Soda (Japan), and reagents for electrophoresis from Bio-Rad (USA).

3. Results and discussion

The purification procedure is summarized in Table 1. Six purification steps led to an electrophoretically homogeneous preparation of HQ 1,2-DO. The pure enzyme with specific activity of 55 U/mg was obtained after a 785-fold enrichment, with a yield of about 60%.

The purified enzyme retained 100% activity after 2 months storage at 4°C in the pH range 7.0–9.5. HQ 1,2-DO was active from pH 6.5 to 10.5 with the optimal activity at pH 7.5. For pH values higher than 8 it was impossible to determine the pH dependence of the enzyme reaction because of the rapid autoxidation of hydroxyquinol and chlorohydroxyquinols under basic conditions.

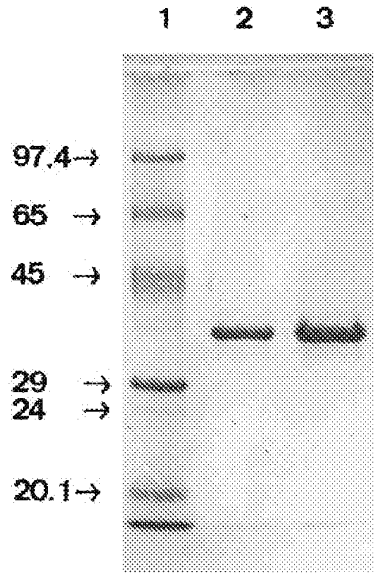


Fig. 1. SDS-PAGE (12% polyacrylamide) of HQ 1,2-DO. Lane 1 contains the low molecular mass protein standards of 20 100, 24 000, 29 000, 45 000, 65 000, and 97 400 Da. Lanes 2 and 3 correspond to HQ 1,2-DO, 3.0 and 6.0 μg.

Kinetic parameters for HQ 1,2-DO from *N. simplex* 3E were estimated in relation to HQ: $K_m = 1.2$ μM, $V_{max} = 55$ U/mg, $K_{cat} = 57$ s^{-1} and $K_{cat}/K_m = 47.5$ μM^{-1} s^{-1}. (K_{cat} was calculated for a molecular weight of 65 kDa).

The enzyme is active at 10–60°C with optimal activity at 50–55°C under test conditions. Heating of the enzyme at 30, 40, and 45°C for 10 min led to 20, 74 and 94% decrease of activity respectively.

The enzyme is remarkably specific. Only HQ, 5CHQ, and 6CHQ are the substrates. The oxidation rates of 6CHQ and 5CHQ by HQ 1,2-DO are about 5 and 2.4% of the oxidation rates of unsubstituted HQ, 2.75 and 1.3 U/mg respectively. The range of halogen-, methyl- and methoxy-substituted catechols, 4-nitrocatechol, chloro- and 2,6-dichloroquinol, 3,4-dihydroxybenzaldehyde tested are not substrates, but were found to be inhibitors of HQ turnover (Table 2). 4,5-Dibromocatechol and chloroquinol were the strongest inhibitors of HQ 1,2-DO. Hydroquinol, 4-fluorocatechol, and 2,6-dichlorohydroquinol were weak inhibitors.

The molecular mass of the native enzyme was estimated to be 65 kDa by gel filtration as reported in Section 2. The

Table 1
Purification of *N. simplex* 3E hydroxyquinol 1,2-dioxygenase

Purification step	Volume (ml)	Total protein (mg)	Total activity (units)	Specific activity (units/mg)	Purification (fold)	Yield (%)
Crude extract	150	4500	16.7	0.004		
Ammonium sulfate fractionation	140	2100	315.0	0.15	2.14	100
DEAE-Toyopearl TSK-Gel	140	610	262.3	0.43	6.13	83.3
Q-Sepharose FF	100	98	234.2	2.39	34.07	74.3
Phenyl-Sepharose	36	41	219.4	5.35	76.26	69.7
Superdex-200	4.5	7.0	204.1	29.16	416	64.8
Mono-Q 5/5	0.82	3.46	190.3	55	784.6	60.4

COR-AAD1-00001176

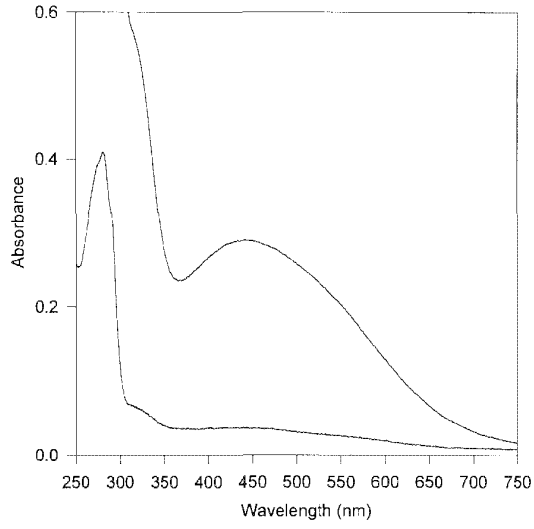


Fig. 2. Electronic absorption spectrum of the purified HQ 1,2-DO from *Nocardioides simplex* 3E 0.08 mM in 50 mM Tris-HCl buffer pH 7.8. The lower spectrum is taken from the same sample diluted 8 times.

molecular mass of the denatured protein as determined by SDS-PAGE was found to be 37 kDa (Fig. 1). On the basis of these results HQ 1,2-DO appears to be a homodimer.

The AAS of the HQ 1,2-DO showed the presence of 1 mol iron/mol enzyme (420 mg of protein contained 0.35 mg of Fe).

The enzyme solution is reddish-brown and in the visible spectrum the enzyme shows a broad, low intensity band centered at 450 nm ($\varepsilon_{450} = 3646\ M^{-1}\ cm^{-1}$, $\varepsilon_{450} = 0.0588$ mg ml^{-1} cm^{-1}) (Fig. 2).

The extinction coefficients at 280 nm and 205 nm were determined to be 0.65 and 31.05 mg ml^{-1} cm^{-1}, respectively, using the method of Scopes [8].

The EPR spectrum of the purified enzyme in frozen buffer at 4.2 K is shown in Fig. 3. The signal at $g = 4.3$ is characteristic of high spin iron(III) ions in rhombic environments similar to those observed for other intradiol cleaving dioxygenases.

Several HQ 1,2-DOs have been purified to date from different organisms like *Trichosporon cutaneum* [13], *Phanerochaete chrysosporium* [14], *Azotobacter* sp. [15], *Streptomyces rochei* 303 [16]. Recently HQ 1,2-DO was purified from *Burkholderia cepacia* AC 1100, degrading the herbicide 2,4,5-T, but this enzyme was specific for HQ and was not able to use CHQ [17].

Table 2
Inhibition constants (K_i) of different aromatic compounds for hydroxyhydroquinol 1,2-dioxygenase from *N. simplex* 3E

Inhibitor	Inhibition constant (mM)
Hydroquinone	150.0
Chlorohydroquinone	2.35
2,6-Dichlorohydroquinone	20.0
Methoxyhydroquinone[a]	
Catechol	10.0
3-Chlorocatechol	7.0
4-Chlorocatechol	10.0
3,5-Dichlorocatechol	2.0
4,5-Dichlorocatechol	7.8
4-Nitrocatechol	10.0
4,5-Dibromocatechol	4.8
4-Fluorocatechol	100.0
3,4-Dihydroxybenzaldehyde	9.9

[a]Does not inhibit.

Enzyme from *N. simplex* 3E, degrading both phenoxyalkanoic herbicides 2,4,-D and 2,4,5-T, catalyzed the *ortho* cleavage of both HQ and CHQ.

So, at the first time HQ 1,2-DO, splitting both substituted and non-substituted HQ, was isolated from the culture degrading phenoxyalkanoic herbicides. Like the corresponding enzymes HQ 1,2-DO from *Azotobacter* sp. [15] and 6-CHQ 1,2-DO from *Streptomyces rochei* 303 [16], the enzyme from *N. simplex* 3E was found to be a homodimer with 37 kDa subunits, the pH and T° optima for these enzymes were also very similar.

Preliminary spectroscopic studies (visible spectra and EPR) of the purified enzyme indicate the involvement of high spin Fe(III) ions (1 mol of iron per homodimer) in the catalytic process. The broad band centered at 450 nm in the electronic absorption spectrum is indicative of a ligand to Fe(III) charge transfer transition, typical of tyrosinate coordination to the metal, whereas the EPR spectrum suggests the presence of high spin iron(III) in a rhombic environment. Therefore both spectroscopic techniques indicate an active metal site very similar to those observed for other intradiol dioxygenases like catechol 1,2-dioxygenase and protocatechuate 3,4-dioxygenase [18–21].

Acknowledgements: This work was supported by a NATO Linkage Grant HT 951032, the Russian Program 'New Methods of Bioengineering (Environmental Biotechnology)'. We appreciate the contribution of the CNR Gruppo Nazionale per la Difesa dia Rischi.

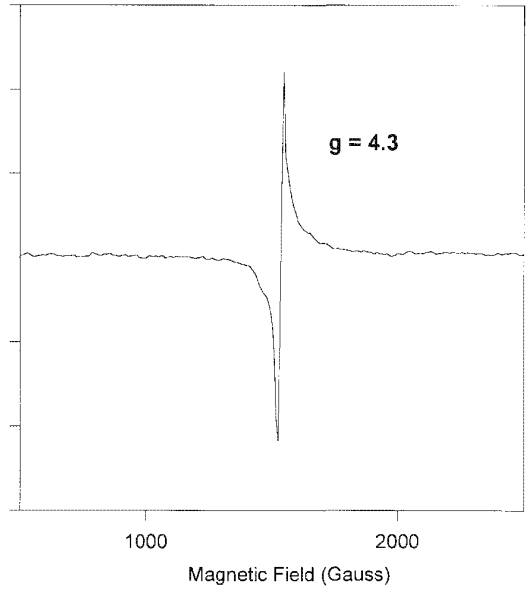


Fig. 3. EPR spectrum of the purified HQ 1,2-DO from *Nocardioides simplex* 3E 0.8 mM in 50 mM Tris-HCl buffer pH 7.8. Instrumental conditions: scan rate 1000 G/min, modulation amplitude 2 Gauss, modulation frequency 100 kHz, microwave frequency 9.24 GHz, microwave power 2 mW, receiver gain 800, temperature 4.2 K.

COR-AAD1-00001177

References

[1] Haggblom M, Valo R. In: Microbial Transformation and Degradation of Toxic Chemicals. New York: Wiley-Liss, 1995:389–439.
[2] L.A. Golovleva, R.N. Pertsova, L.I. Evtushenko, B.P. Baskunov, Biodegradation I (1990) 263–271.
[3] L.P. Kozyreva, L.A. Golovleva, Microbiology (translated from Microbiologiya) 62 (1993) 136–138.
[4] I.P. Solyanikova, Ya. Protopopova, V.M. Travkin, L.A. Golovleva, Biochemistry (translated from Biokhimiya) 61 (1996) 635–642.
[5] I.S. Sze, S. Dagley, J Bacteriol 159 (1984) 353–359.
[6] J.M. Tiedje, J.M. Duxlury, M. Alexander, J.E. Dawson, J Agric Food Chem 17 (1969) 1021–1026.
[7] O. Warburg, W. Christian, Biochem J 310 (1941) 384–421.
[8] R.K. Scopes, Anal Biochem 59 (1974) 277–282.
[9] U.K. Laemmli, Nature 227 (1970) 680–685.
[10] I.B. Holbrook, A.G. Leaver, Anal Biochem 75 (1976) 634–636.
[11] M. Dixon, Biochem J 55 (1953) 170–171.
[12] H.D. Dakin, Am Chem J 42 (1909) 477–498.
[13] I.S.-Y. Sze, S. Dagley, J Bacteriol 159 (1984) 353–359.
[14] S. Rieble, D.K. Joshi, M.H. Gold, J Bacteriol 176 (1994) 4838–4844.
[15] M. Latus, H.-J. Seitz, Ju. Eberspacher, F. Lingens, Appl Environ Microbiol 61 (1995) 2453–2460.
[16] O.E. Zaborina, M. Latus, J.U. Ebershpacher, L.A. Golovleva, F. Lingens, J Bacteriol 177 (1995) 229–234.
[17] D.L. Daubaras, K. Saido, A.M. Chakrabarty, Appl Environ Microbiol 62 (1996) 4276–4279.
[18] I. Bertini, F. Briganti, S. Mangani, H. Nolting, A. Scozzafava, Coordin Chem Rev 144 (1995) 321–345.
[19] I. Bertini, F. Briganti, C. Luchinat, A. Scozzafava, New J Chem 20 (1996) 187–193.
[20] Lipscomb JD, Orville AM. In: Sigel H, Sigel A, editors. Metal Ions in Biological Systems, Vol 28, Chapter 8. New York: Marcel Dekker, 1992:243-298.
[21] Que L Jr. in: Loehr T, editor. Iron Carriers and Iron Proteins. New York: VCH, 1989:467-524.

COR-AAD1-00001178

ELSEVIER

FEMS Microbiology Ecology 24 (1997) 269–278

FEMS MICROBIOLOGY ECOLOGY

PCR-RFLP analysis of 16S rRNA, *tfdA* and *tfdB* genes reveals a diversity of 2,4-D degraders in soil aggregates

Tatiana Vallaeys *, Fabienne Persello-Cartieaux [1], Nadine Rouard, Christine Lors, Gisèle Laguerre, Guy Soulas

CMSE-INRA Laboratoire de Microbiologie des Sols, B.V. 1540, 21034 Dijon Cedex, France

Received 8 April 1997; revised 24 July 1997; accepted 28 July 1997

Abstract

The 2,4-dichloropbenoxyacetic acid (2,4-D) degrading potential of an agricultural gleyic luvisol from the Dijon area (France) was assessed. 8% of the soil aggregates tested were shown to degrade chain labelled [^{14}C]2,4-D. Mineralization was observed either immediately or after a 4 week lag period. The diversity of the 2,4-D degrading microflora of three aggregates was further investigated. 16S rRNA genes from 68 isolates were therefore amplified, which led to the discrimination of 17 distinguishable RFLP patterns. Instability of the side chain 2,4-D degrading capabilities was observed amongst 25% of the isolates. None of the unstable 2,4-D degraders amplified when the *tfdA* and *tfdB* primer sets were used. These strains presented different 16S patterns by the *Rsa*I restriction enzyme and therefore belonged to various microbial species. Amongst the stable 2,4-D degrading isolates, two strains which were isolated from different aggregates did amplify using both *tfdA* and *tfdB* primer sets. These strains presented identical *tfdA* and *tfdB* PCR-RFLP patterns which were of the *A. eutrophus* JMP134 (pJP4) and *Burkholderia* sp RASC type, respectively, but showed different 16S rDNA PCR-RFLP patterns demonstrating that the degrading genes were distributed amongst various bacterial groups prior to any enrichment. Six strains amplified using the *tfdB* primer set but not using the *tfdA* primer set. These strains presented either a *Burkholderia* sp. RASC *tfdB* pattern or an undescribed new *tfdB* PCR-RFLP pattern but all shared a common 16S PCR-RFLP pattern. A third group contained stable degraders in which the presence of *tfdA* and *tfdB*-like genes could not be detected. These strains presented 12 distinct 16S PCR-RFLP patterns which differed from those observed in the first two groups of stable degraders. Therefore our results demonstrate the involvement of diverse bacterial species and genes in 2,4-D degradation within a soil.

Keywords: Alpha-ketoglutarate dependent dioxygenase; 2,4-Dichlorophenol hydroxylase; *tfdA*; *tfdB*; 2,4-Dichlorophenoxyacetic acid; Diversity; Soil aggregate

1. Introduction

The herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) has been used for the treatment against dicotyledonous plants for around 50 years. It had become a model for studying the degradation of chlorinated xenobiotics and assessing mechanisms by which de-

* Corresponding author. Tel.: +33 3 80 63 32 98; Fax: +33 3 80 63 32 24; E-mail: vallaeys@dijon.inra.fr

[1] Present address: Laboratoire d'Ecophysiologie Végétale et de Microbiologie, C.E.A. Centre de Cadarache, 13108 Saint Paul Lez Durance, France.

0168-6496/97/$17.00 © 1997 Federation of European Microbiological Societies. Published by Elsevier Science B.V.
PII S0168-6496(97)00068-8

COR-AAD1-00001179

Downloaded from https://academic.oup.com/femsec/article/24/3/269/548657 by Pierre Mendès France University user on 30 September 2020

grading potentialities evolve amongst soil microorganisms. A diversity of bacterial species involved in the degradation of this herbicide have been reported. 2,4-D degrading capabilities have been described in bacteria belonging to a number of genera including *Achromobacter*, *Alcaligenes*, *Arthrobacter*, *Corynebacterium*, *Flavobacterium*, *Pseudomonas*, and *Sphingomonas* [1–11]. In most cases, the genes involved in 2,4-D degradation were observed on conjugative plasmids [2,9,12]. Amongst these, the pJP4 plasmid has been extensively characterized with its degradative genes *tfdABCDEF* precisely localized and sequenced [13–17]. On this plasmid, the *tfdA* gene is involved in the degradation of the side chain of the herbicide: *tfdA* encodes an α-ketoglutarate dependent dioxygenase involved in the transformation of 2,4-D into 2,4-dichlorophenol [14] while *tfdB* encodes a 2,4-dichlorophenol hydroxylase involved in the transformation of 2,4-dichlorophenol into 3,5-dichlorocatechol [15]. The *tfdCDEF* operon encodes enzymes involved in the *ortho*-cleavage of the aromatic ring and subsequent reactions [15].

Under laboratory conditions, the plasmid pJP4 has been successfully transferred to a broad range of recipient species including *Rhizobium* and *Pseudomonas* [18–21]. In addition, early studies described 2,4-D degrading plasmids identical or with important similarities to pJP4 in the degradative genes area, and presenting similar size and restriction patterns [2,22]. As a consequence, soil adaptation to pesticide degradation was initially described as resulting from interspecies transfer of degradative plasmids [23]. However, recent studies reported an increasing diversity of degradative genes [10,11,24,25] and also of the gene locations in the bacterial genome. *Tfd*-like genes have been described on different plasmid backbones [8,26] but also in some cases with chromosomal locations [27], suggesting that different mechanisms are involved in the dissemination of degrading capabilities amongst soil microorganisms. However, most studies have been carried out on strains of different origins, isolated from various areas and using different cultivation conditions and isolation procedures [10,28]. Few results are available at the population level [11,24].

PCR-RFLP of 16S rRNA genes appeared a useful technique to distinguish phylogenetically distant species or genera [29,30] and has been successfully applied to estimate the diversity of microorganisms from environmental samples [31,32]. In addition, we recently developed *tfdA* and *tfdB* primer sets [25] which were shown to be more sensitive than hybridization probes in detecting *tfdA* and *tfdB*-like genes. These primers allowed the characterization of a diversity of *tfd*-like genes amongst collection strains. Therefore, the aim of this study was to characterize simultaneously the specific diversity and the distribution of the *tfdA* and *tfdB*-like gene types amongst 2,4-D degrading bacteria isolated from an agricultural soil.

2. Materials and methods

2.1. Soil samples preparation

The gleyic luvisol from the Citeaux area previously described by Nelson et al. [33] is characterized by a low biomass, and also a low natural 2,4-D degrading potential [34]. Although this is an agricultural soil, no historic account of 2,4-D treatment on the experimental field has been recorded. This soil was sampled using the coring system presented by Nelson et al. [33]. However, only the top layer (0–20 cm) was used in this study. Soil samples were sieved (< 2 mm) and individual aggregates of approximately 20 mg of soil were placed separately in 96 well microtitration plates.

2.2. Enrichment and isolation of 2,4-D degrading microorganisms

The cleavage of the ether bond and formation of 2,4-dichlorophenol is characteristic of the 2,4-D catabolic pathway. This step involves an α-ketoglutarate dependent dioxygenase encoded by the *tfdA* gene which presents no significant similarities to any other known sequences [35]. Therefore, chain labelled [^{14}C]2,4-D was used in this study to determine the 2,4-D degrading ability of the soil aggregates and to enrich the positive samples in 2,4-D degrading microorganisms. Strains were screened on their ability to degrade chain labelled [^{14}C]2,4-D. In addition, both chain labelled [^{14}C]2,4-D and uniformly ring labeled [U-^{14}C]2,4-D were used to determine the ability of the isolates to degrade the

COR-AAD1-00001180

Downloaded from https://academic.oup.com/femsec/article/24/3/269/[illegible] by Pierre Renard Library user on 03 September 2020

herbicide after storage on minimal 2,4-D agar slopes: the method developed by Tabor et al. [36] has been adapted by Fulthorpe et al. [37] to the isolation of 2,4-D degrading strains. This method was used with a few modifications. The enrichment medium was the 2,4-D minimal medium described by Vallaeys et al. [25]. In addition, three (instead of two) enrichment steps were performed as follows: 20 mg soil aggregates were used. Their ability to degrade 2,4-D was determined using the autoradiography method of Tabor et al. [36]. The content of the wells corresponding to degrading aggregates was diluted with 5 ml of minimal 2,4-D medium and aliquots of 10 μl were transferred to 384 well microtitration plates. Following 1, 2, 3 and 4 weeks incubation at 20°C, 2,4-D degradation was determined using the autoradiography method developed by Tabor et al. [36]. Positive wells were diluted 20-fold and transferred onto 96 well microtitration plates which were incubated using identical conditions. The content of positive wells was finally transferred onto 2,4-D minimal agar plates made with minimal 2,4-D medium [25] containing 18 g l^{-1} of minimal bacteriological agar type E (Biokar Diagnostics, France). Colonies were purified by re-streaking on 2,4-D minimal agar and side chain 2,4-D mineralization capabilities of the selected isolates were determined using the autoradiography method of Tabor et al. [36]. Considering the large number of side chain degrading isolates obtained, strains were stored on 2,4-D minimal agar slopes up to the DNA extraction stage; they were then re-streaked onto 2,4-D minimal agar plates. Fresh colonies were used to inoculate flasks containing liquid 2,4-D minimal medium or microtitration plates which were used respectively for DNA purification and for determination of 2,4-D (side chain and ring) degrading capabilities, using the Tabor et al. [36] autoradiography method.

2.3. Molecular methods

Bacterial cultures were grown for 5–20 days at 20°C on 2,4-D minimal liquid medium [25]. 5 ml of Luria Broth liquid medium was added to the 2,4-D minimal medium in order to increase biomass production 24 h prior to DNA extraction. Total genomic DNA extraction was carried out as described in Sambrook et al. [38]. DNA was resuspended in sterile distilled water at a concentration of 0.05 μg ml^{-1} as determined by UV absorbance at 260 nm.

Universal prokaryotic primers presented by Lane et al. [39] were used for the amplification of the 16S rRNA genes according to the conditions described by Laguerre et al. [29]. Aliquots of 5–10 μl of the PCR solutions were analyzed on horizontal, 1% agarose gels (Sigma) in TAE buffer (40 mM Tris-HCl, 4 mM sodium acetate, 1 mM EDTA, pH 7.9) using the Horizon 58 apparatus (Gibco-BRL, France). Electrophoresis was carried out for 1 h at 60 V. Ethidium bromide stained gels were then photographed under UV light using TYPE 665 positive negative films (Polaroid, France). The *Rsa*I restriction enzyme was shown to be amongst the most efficient at differentiating bacterial rRNA genes [40]. The obtained PCR fragments were therefore digested with *Rsa*I (Boehringer-Mannheim) as specified by the manufacturer. Restriction fragment analyses were performed on a horizontal 4% Nu-Sieve agarose gel (FMC Bioproducts). Boehringer VIII molecular mass marker (Boehringer-Mannheim, France) was used as a ladder. Electrophoresis was carried out at 80 V for 3 h.

Specific primers for the 2,4-D etherase gene (*tfdA*) and the dichlorophenol hydroxylase gene (*tfdB*) from *Alcaligenes eutrophus* JMP134 (pJP4) previously described by Vallaeys et al. [25] were used for selective amplification of the corresponding genes from purified bacterial DNA obtained from degrading isolates. The PCR conditions used were those described by Vallaeys et al. [25]. *A. eutrophus* JMP134 and its cured derivative *A. eutrophus* JMP222 were used as positive and negative control, respectively. The PCR products obtained were digested with the restriction enzyme *Hae*III (Boehringer-Mannheim) according to the conditions described by the manufacturer. Restriction fragments were visualized under UV light after migration on ethidium bromide stained 1.25% Nu-Sieve agarose gel (FMC Bioproducts) prepared in TAE buffer.

3. Results

The 2,4-D mineralization activity of a soil profile at the experimental station of Citeaux (France) has

COR-AAD1-00001181

Downloaded from https://academic.oup.com/femsec/article/24/3/269/554057 by Pierre Marie Curie University user on 30 September 2020

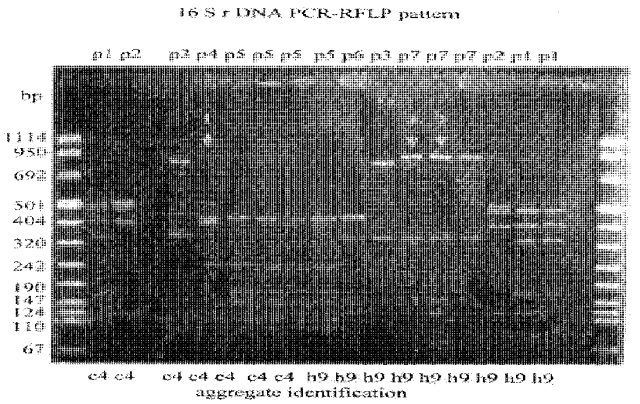


Fig. 1. Diversity of 16S PCR-RFLP patterns by *Rsa*I (p1–p7) obtained from 2,4-D degrading strains isolated from two different soil aggregates (c4 and h9). Arrows indicate the patterns of strains which yielded a PCR product using the *tfdA* and/or the *tfdB* primer sets. First and last lanes contain Boehringer VIII molecular mass markers. Sizes of the ladder fragments are indicated in bp on the left-hand side of the gel.

been previously studied by Dictor [34] and was shown to be maximal in the surface (0–20 cm) layer. The 2,4-D degrading capability of individual soil aggregates was therefore assessed in the top layer. Approximately 8% of the tested aggregates were capable of degrading chain labelled [^{1} C]2,4-D prior to any enrichment. 2,4-D degradation was observed either immediately or after a lag period of up to 4 weeks.

The successive dilution/enrichment cycles from eight individual soil aggregates, which presented different intensities of 2,4-D degrading activity, led to the isolation of 186 individual 2,4-D degrading strains. No correlation could be observed between the intensity of 2,4-D degradation in the aggregate and the number of isolated strains presenting 2,4-D degrading activity.

Three aggregates (c4, 14 strains; h9, 64 strains; e6, 45 strains) which differed in the range of 2,4-D degrading isolates obtained and intensity of degrading activity were chosen for a further characterization of their 2,4-D degrading microflora (i.e. a total of 123 isolates). In particular, we wanted to know firstly, if the enrichment method used led to the isolation of multiple representatives from a particular clone in each aggregate or if the isolated 2,4-D degrading strains reflected to some extent the initial diversity of the degrading microflora and secondly, if this diversity could explain the degrading behavior of the corresponding aggregates.

When using the 16S rDNA primers of Lane et al. [39], amplification products were successfully obtained from 68 isolates amongst the 123 isolates tested. Fifty-five isolates presented poor (26 strains) or no growth (15 strains) on the 2,4-D minimal liquid medium used or could not be amplified (14 strains). RFLP analysis of the PCR products using the restriction enzyme *Rsa*I led to the characterization of 17 distinguishable RFLP patterns suggesting a wide diversity of the microbial groups involved in 2,4-D degradation in the Citeaux soil (Fig. 1). A diversity of 16S rDNA patterns could even be observed amongst the strains isolated from the same aggregate: the three aggregates c4, h9, and e6 presented 6, 12 and 10 distinguishable 16S PCR-RFLP

Table 1
Distribution of the 16S PCR-RFLP patterns by the *Rsa*I restriction enzyme amongst the isolates from three aggregates of an agricultural soil from the Citeaux (France) area

Pattern number	Number of isolates		
	Aggregate c4	Aggregate h9	Aggregate e6
1	1	3	4
2	1	6	7
3	1	1	–
4	2 (*,1)	–	–
5	3	6	1
6	–	1	–
7	1	6 (*,1) (+,3)	6 (+,3)
8	–	–	1
9	–	–	1
10	–	–	1
11	–	1	1
12	–	4	–
13	–	1	–
14	–	4	1
15	–	1	–
16	–	1	–
17	–	–	1
Total number of patterns observed	6	12	10

(*,*n*) indicates the number *n* of isolates presenting both detectable *tfdA* and *tfdB*-like genes by PCR. (+,*n*) indicates the number of isolates presenting a *tfdB*-like gene detectable by PCR.

COR-AAD1-00001182

Downloaded from https://academic.oup.com/femsec/article/24/3/269/548057 by Pharmacy Hill Library user on 03 September 2020

*Rsa*I patterns, respectively (Table 1). A majority of patterns were observed only in one aggregate (patterns 4, 6, 8, 9, 10, 12, 13, 15, 16 and 17) and had a low number of representatives while others were observed amongst the isolates from the three aggregates (patterns 1, 2, 5 and 7).

The individual 2,4-D degrading isolates were checked for the degradation of side chain and ring ^{1} C-labelled 2,4-D after storage on minimal 2,4-D agar slopes. Although 2,4-D was used as carbon source during all cultivation and purification steps, the isolates presented various intensities of side chain degradative activity. An instability of the 2,4-D degradation capabilities was observed as approximately 25% of the subcultures tested presented a complete lack of degradation of side chain labelled 2,4-D (Fig. 2A). A high proportion of the strains (an additional

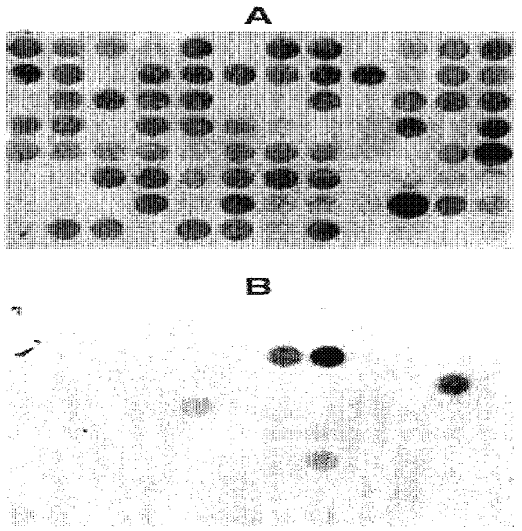


Fig. 2. Analysis of the stability of the 2,4-D side chain degradation capabilities of the isolates (A) and of their ability to degrade ring 2,4-D labelled molecules (B). Black spots indicate release of $^{14}CO_2$ from ^{14}C labelled molecules by the individual isolates inoculated in separate wells. Gray spots indicate slow degradation of 2,4-D. White spots indicate that no release of $^{14}CO_2$ could be detected for the corresponding culture.

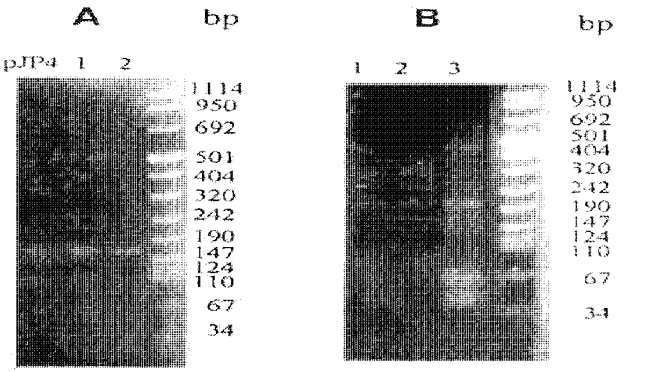


Fig. 3. PCR-RFLP analysis by *Hae*III of the *tfdA* (A) and *tfdB*-like (B) genes from three strains isolated from the Citeaux soil. Isolates 1 and 2 gave amplification signals with both *tfdA* and *tfdB* primer sets. Isolate 3 yielded a PCR product using the *tfdB* primer set only. The lane labelled pJP4 contains the control sample (RFLP analysis of the *A. eutrophus* JMP134 (pJP4) PCR fragment). Right-hand lanes in A and B contain Boehringer VIII molecular mass markers.

25%) showed a more or less weak positive signal. Lack of side chain degrading activity and poor degradation could be correlated with lack of growth or poor growth on the liquid 2,4-D minimal medium used in this study (results not shown), suggesting that these strains were unable to use 2,4-D as sole source of carbon and energy and were rather using agar carbon traces as substrate and co-metabolizing 2,4-D. In addition, approximately 6% of the isolates presented significant ring ^{14}C-labelled 2,4-D mineralizing activity (Fig. 2B).

PCR-RFLP analysis of the *tfdA* and *tfdB*-like genes was performed. A PCR product could be amplified from two strains using both *tfdA* and *tfdB* primer sets and six strains yielded a PCR product using the *tfdB* primer set but not the *tfdA* primer set. One *tfdA* and two different *tfdB* PCR-RFLP patterns were obtained, respectively (Fig. 3A,B). The strains which yielded a PCR product using both *tfdA* and *tfdB* primer sets showed identical PCR-RFLP patterns for both genes. These patterns corresponded to those described previously for *tfdA* and *tfdB* genes in *A. eutrophus* JMP134 and *Burkhol-*

COR-AAD1-00001183

Downloaded from https://academic.oup.com/femsec/article/24/3/269/4636157 by Pierre Mendes France University user on 03 September 2020

deria sp. RASC, respectively [25]. However, the isolates presented two different 16S PCR-RFLP patterns (numbers 4 and 7 which also differed from the *A. eutrophus* JMP134 and the *Burkholderia* sp. RASC ones) suggesting that they belonged to different bacterial species (Table 1, Fig. 1). As these strains were isolated from different soil aggregates, this result could not be the consequence of a gene transfer event during the enrichment step and reflects the natural distribution of degrading capabilities amongst soil microorganisms prior to any 2,4-D application. Conversely, strains isolated from the same soil aggregate (h9) and presenting identical 16S PCR-RFLP patterns (number 7) but different *tfdB* PCR-RFLP patterns were also observed (Fig. 1, Fig. 3B) which excludes a clonal origin of the isolates.

In addition, we found that strains which amplified using the *tfdB* primer set only presented two types of *tfdB* PCR-RFLP patterns. Their *tfdB* pattern type was in most cases identical to that described for *Burkholderia* sp. RASC [25]. However, one isolate showed a *tfdB* PCR-RFLP pattern which appeared distinct from all PCR-RFLP patterns of *tfdB*-like genes previously described [25]. In this case, the amplified PCR fragment was of a correct size and shared common bands with *tfdB* PCR-RFLP patterns previously described by us, thus suggesting the existence of an evolutionary relationship between this fragment and *tfdB*-like genes.

Four groups of strains could therefore be discriminated according to their degrading behavior and the results of the PCR-RFLP analysis of their *tfd*-like genes. The first group corresponded to the unstable 2,4-D degrading isolates in which the presence of *tfd*-like genes using the *tfdA* and *tfdB* primer sets could not be detected. These strains showed weak side chain degradation signal immediately after purification and lack of both side chain and ring degradation after storage and subculturing steps. They belonged to various phylogenetic groups as they presented three different 16S PCR-RFLP patterns (patterns 2, 12 and 16). Amongst the stable degraders (strains which maintained a side chain degrading activity after storage and subculturing steps), a first group of two isolates yielded a PCR product using both *tfdA* and *tfdB* primer sets and seemed to carry identical degradative genes. However, they belonged to different bacterial groups according to the results of the PCR-RFLP analysis of their 16S rRNA genes (patterns 4 and 7). A second group presented a stable side chain degrading activity but non-amplifiable *tfdA* genes. The *tfdB* genes from these strains could, however, be detected by PCR. The 16S PCR-RFLP analysis showed that they all belonged to the same phylogenetic group (pattern 7). A last group of stable degraders showed both non-amplifiable *tfdA* and *tfdB* genes. It included strains from 12 distinct bacterial groups as resulting from the restriction analysis of their 16S rRNA genes with the *Rsa*I enzyme (patterns 1, 3, 5, 6, 8, 9, 10, 11, 13, 14, 15, and 17).

4. Discussion

While there have been many studies reporting the isolation and characterization of biodegrading strains, only few authors have focused their interest on understanding biodegradation processes at the population level [11,41,42]. In a recent work, Fulthorpe et al. [37], while trying a similar approach on pristine soils, underlined the difficulty of isolating 2,4-D microorganisms from soil which had no previous contact with pesticides. These authors succeeded in isolating only five 2,4-D degraders whereas 50% of the soil samples showed 2,4-D degradative capabilities. Conversely, the Citeaux soil, which also has no history of 2,4-D application, allowed the isolation of a large number of strains able to degrade either partly or completely the 2,4-D molecule. In addition, while a lag period of 2 weeks was systematically observed before 2,4-D degradation could be measured in pristine soil aggregates [37], degradation at Citeaux could be observed either immediately or after a lag time of up to 4 weeks. This suggests that 2,4-D degrading capabilities were not evenly distributed amongst the aggregates but that some aggregates presented naturally a full degrading potential against the herbicide. Therefore, our work focused on understanding these differences by the characterization of the 2,4-D degrading microflora of the soil aggregates. This allowed us to demonstrate that an agricultural soil with no history of 2,4-D treatment may harbor various degrading bacterial strains in terms of both degradation character-

COR-AAD1-00001184

Downloaded from https://academic.oup.com/femsec/article/24/3/269/5103457 by Pierre Marie Curie University user on 03 September 2020

istics and phylogenetic groups involved prior to the application of the herbicide. This soil was characterized firstly, by a low proportion of isolates presenting both side chain and ring degrading capabilities (i.e. able to completely mineralize 2,4-D), secondly, by a large number of unstable 2,4-D degrading isolates (which presented side chain degrading activity immediately after streaking on minimal 2,4-D agar but lost this ability after storage and subculture), and finally, by a low percentage of strains which yielded a PCR product when using the *tfdA* and *tfdB* primer sets.

The small number of isolates presenting ring degradation activity observed in our study is certainly linked to the isolation procedure we used. We chose to screen strains on the basis of their chain degrading capabilities. However, the fact that only a small proportion of the isolates presented both ring and side chain mineralizing activities confirms the results previously reported by Fulthorpe et al. [37] for pristine soils and therefore suggests that in soils with no history of 2,4-D application, consortia may play a major part in the degradation of the herbicide. The existence of degrading communities has already been reported by Senior et al. [43]. In a recent study, Top et al. [24] reported the isolation of a plasmid carrying only a part of the 2,4-D degradation pathway, demonstrating thereby that catabolic genes may be spread out among various soil microorganisms. The results of the PCR-RFLP analysis of the *tfd*-like genes seem to support this hypothesis as some strains yielded a PCR product when using the *tfdB* primer set only. However, a key issue is whether or not the primer sets used in this study for amplifying and studying the diversity of *tfd*-like genes are indeed specific enough for, or in contrast restricted to, certain subfamilies of 2,4-D degrading genes. Primer specificity has been discussed elsewhere [25] and has been shown to be satisfactory when applied to the study of collection strains. In addition, sequence data analysis of PCR amplified DNA fragments obtained while using these *tfdA* and *tfdB* primer sets showed significant sequence similarities to the canonical *tfdA* and *tfdB* genes from *A. eutrophus* JMP134, respectively [44,45]. In addition, *tfdA* primer sets were also successfully used in the amplification of *tfdA*-like genes directly from certain pristine soils (unpublished data). However, while *tfdA* and *tfdB* primer sets have been shown to amplify most *tfdA* and *tfdB*-like genes when applied to collection strains [25], various 2,4-D degrading isolates belonging mainly to the *Sphingomonas* group did not yield a PCR product with these primers. In the Citeaux soil, a large number of 2,4-D degrading isolates were not amplified when these primer sets were used. Amongst these there appeared to be both stable and unstable side chain 2,4-D degraders.

In the case of strains presenting stable side chain degrading activity, the existence of an active enzyme was confirmed while the presence of the corresponding *tfdA*-like gene could not be detected by PCR. For these strains, the 16S PCR-RFLP analysis showed a diversity of patterns suggesting that *tfd* genes presumably non-homologous to the *A. eutrophus* degradative genes are widespread amongst various bacterial species and not restricted to the *Sphingomonas* genus [25,46].

The absence of any amplification signal amongst unstable 2,4-D degrading isolates or strains presenting weak side chain activity while using *tfdA* and *tfdB* primer sets may have a different origin. In this case, the possibility that strains had lost the corresponding genes during the cultivation steps is unlikely as 2,4-D was always included in all culture media in order to maintain selective pressure. For these isolates, the side chain mineralization initially observed could be either the fact of a 2,4-D dioxygenase phylogenetically unrelated to *tfdA* and for which 2,4-D is not the inducer (which could explain the loss of activity after prolonged storage on 2,4-D medium) or the consequence of co-metabolism or non-specific degradation of the pesticide by enzymes of low substrate specificity and which are induced by natural soil compounds but not 2,4-D. In a recent study, Filer and Harker [47] demonstrated that the inducing agent of the 2,4-D pathway in *A. eutrophus* JMP134 was dichloromuconate, an intermediate metabolic product and that induced levels of the *tfdA* encoded 2,4-D dioxygenase and the *tfdB* encoded 2,4-dichlorophenol hydroxylases were also detected when strains were grown on 3-chlorobenzoate, which is an alternative substrate for the catechol 1,2-dioxygenase involved in the 2,4-D catabolic pathway. Other studies [48,49] also demonstrated for various aromatic compounds that intermediate or end products acted as inducers of their catabolic path-

COR-AAD1-00001185

Downloaded from https://academic.oup.com/femsec/article/24/3/269/5107157 by Pierre and Marie Curie University user on 30 September 2020

ways. The inducing scheme in strains presenting only part of the pathway (which constitute the major part of the Citeaux isolates) is completely unknown. However, in the hypothesis of induction by a by-product or a metabolic intermediate, it is likely that strains presenting only side chain mineralization activity would behave, in their natural habitat, as members of active 2,4-D degrading consortia, and once removed from the close influence of the other community members, would progressively lose their activity. Cases of instability amongst pesticide degrading bacteria have already been reported. For instance, Parekh et al. [50] observed instability of degrading behavior in a number of isolates while studying carbofuran degrading microorganisms. Similar observations have been reported for atrazine degrading strains (Houot, personal communication). However, no systematic investigation has been conducted to explain the origin of this phenomenon.

In conclusion, a diversity of degrading behavior as well as a diversity of 16S rDNA PCR-RFLP patterns was observed amongst isolates from the Citeaux soil which confirms previous results by Ka et al. [11] and Fulthorpe et al. [37]. A number of 16S PCR-RFLP patterns obtained amongst the strains isolated from the Citeaux soil appeared similar to those previously observed in a collection of 2,4-D degrading strains from various origins (unpublished results) suggesting that the corresponding strains could be related to the *Alcaligenes xylosoxidans* and *Pseudomonas* species. However, we also obtained new types of 16S PCR-RFLP patterns which means that the range of species or genera involved in the degradation of the herbicide 2,4-D has not been extensively defined.

Similar conclusions could be drawn with respect to the degradative genes, as two different *tfdB* PCR-RFLP patterns were shown to co-exist amongst the isolates, one of which has never been described. In addition, two stable 2,4-D degrading isolates from the Citeaux soil presenting both identical 16S and *tfdB* PCR-RFLP patterns showed different behavior when the *tfdA* primer set was used. One led to an amplification signal while the other could not be amplified which suggests the independent recruitment of two different *tfdA*-like gene cassettes.

The PCR-RFLP analysis of 16S, *tfdA* and *tfdB* genes of isolates from different soil aggregates allowed us to demonstrate the existence of a diversity of strains and genes involved in the degradation of a herbicide at the aggregate level. In addition, it allowed us to assess the distribution of degradative capabilities and degrading genes in different groups of bacteria prior to the enrichment step. Therefore, the screening method developed here, which is based on the isolation of degrading strains from different soil aggregates, is suitable for assessing the structure of degrading populations in natural environments. However, the results obtained in this study appeared insufficient to explain the difference in 2,4-D degrading behavior (intensity of side chain activity) of the soil aggregates studied.

Therefore, although we previously demonstrated that PCR was more sensitive than hybridization analysis in the detection of *tfdA* and *tfdB* genes related phylogenetically to the canonical *tfdA* and *tfdB* genes from *A. eutrophus* JMP134 [25], the primer sets described failed to detect the presence of *tfd*-like genes in most of the isolates from the Citeaux soil. However, considering the high level of instability of 2,4-D degrading capabilities observed amongst the isolates, it is not clear whether our observations result from unspecific degradation of a related substrate by enzymes presenting low substrate specificity, the loss of *tfd*-like genes during subculturing steps, or the presence of genes presenting no region of homology to the previously described *tfd* genes [14,15,27]. It is therefore important to point out that most of the molecular tools were designed to detect and characterize the 2,4-D degrading *tfd* gene of the pJP4 type. Holben et al. [28], Farhana et al. [51], and Vallaeys et al. [25,52] developed detection techniques of *tfd*-like genes or enzymes based on hybridization, PCR or ELISA. However, the use of these techniques may in some cases lead to a considerable underestimation of the size but also the diversity of the 2,4-D degrading microflora and care should be taken while interpreting results.

Our results therefore emphasize the need to develop tools to characterize the non-homologous *tfd*-like gene group.

Acknowledgments

We thank Mohamed El Hamdaoui for technical assistance, Rémi Chaussod for comments and discus-

COR-AAD1-00001186

Downloaded from https://academic.oup.com/femsec/article/24/3/269/570157 by Pierre Mendes France University user on 30 September 2020

sions. We acknowledge Roberta Fulthorpe's ongoing support. We also thank Philippe Butler for his linguistic advice. This work was supported by EEC Grant EV5V-CT92-0216.

References

[1] Bell, G.R. (1957) Some morphological and biochemical characteristics of a soil bacterium which decomposes 2,4-dichlorophenoxyacetic acid. Can. J. Microbiol. 3, 821–840.

[2] Don, R.H. and Pemberton, J.M. (1981) Properties of six pesticide degradation plasmids isolated from *Alcaligenes eutrophus* and *Alcaligenes paradoxus*. J. Bacteriol. 145, 681–686.

[3] Evan, W.C., Smith, B.S.W., Fernley, H.N. and Davies, J.I. (1971) Bacterial metabolism of 2,4-dichlorphenoxyacetate. Biochem. J., 122, 543–551.

[4] Jensen, H.L. and Petersen, H.I. (1952) Detoxication of hormone herbicides by soil bacteria. Nature 170, 39–40.

[5] Loos, M.A., Roberts, R.N. and Alexander, M. (1967) Phenols as intermediates in the decomposition of phenoxyacetates by *Arthrobacter* species. Can. J. Microbiol. 13, 679–690.

[6] Rogoff, M.H. and Reid, J.J. (1956) Bacterial decomposition of 2,4-dichlorophenoxyacetic acid. J. Bacteriol. 71, 303–307.

[7] Walker, R.L. and Newman, A.S. (1956) Microbial decomposition of 2,4-dichlorophenoxyacetic acid. Appl. Environ. Microbiol. 4, 201–206.

[8] Chaudrhry, G.R. and Huang, G.H. (1988) Isolation and characterization of a new plasmid from a *Flavobacterium* sp. which carries the genes for degradation of 2,4-dichlorophenoxyacetic acid. J. Bacteriol. 170, 3897–3902.

[9] Bhat, M.A., Tsuda, M., Horiike, K., Nozaki, M., Vaidyanathan, C.S. and Nakazawa, T. (1994) Identification and characterization of a new plasmid carrying genes for degradation of 2,4-dichlorophenoxyacetate from *Pseudomonas cepacia* CSV90. Appl. Environ. Microbiol. 60, 307–312.

[10] Fulthorpe, R.R., McGowan, C., Maltseva, O.V., Holben, W.E. and Tiedje, J.M. (1995) 2,4-Dichlorophenoxyacetic acid degrading bacteria contain mosaics of catabolic genes. Appl. Environ. Microbiol. 61, 3274–3281.

[11] Ka, J.O., Holben, W.E. and Tiedje, J.M. (1994) Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D) degrading bacteria isolated from 2,4-D treated field soils. Appl. Environ. Microbiol. 60, 1106–1115.

[12] Mae, A.A., Martis, R.O., Ausmees, N.R., Kõiv, V.M. and Heinaru, A.L. (1993) Characterization of a new 2,4-dichlorophenoxyacetic acid degrading plasmid pEST4011: physical map and localization of catabolic genes. J. Gen. Microbiol. 139, 3165–3170.

[13] Don, R.H., Weightman, A.J., Knackmuss, H.H. and Timmis, K.N. (1985) Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4) J. Bacteriol. 161, 85–90.

[14] Streber, W.R., Timmis, K.N. and Zenk, M.H. (1987) Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetic mono-oxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. 169, 2950–2955.

[15] Perkins, E.J., Gordon, M.P., Caceres, O. and Lurquin, P.F. (1990) Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol., 172, 2351–2359.

[16] Kaphammer, B. and Olsen, R.H. (1990) Cloning and characterization of *tfdS*, the repressor-activator gene of *tfdB*, from the 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. J. Bacteriol. 172, 5856–5862.

[17] Kaphammer, B., Kukor, J.J. and Olsen, R.H. (1990) Regulation of *tfdCDEF* by *tfdR* of the 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. J. Bacteriol. 172, 2280–2286.

[18] Friedrich, B., Meyer, M. and Schlegel, H.G. (1983) Transfer and expression of the herbicide-degrading plasmid pJP4 in aerobic autotrophic bacteria. Arch. Microbiol. 134, 92–97.

[19] Kinkle, B.K., Sadowsky, M.J., Schmidt, E.L. and Koskinen (1993) Plasmids pJP4 and r68.45 can be transferred between populations of Bradyrhizobia in nonsterile soil. Appl. Environ. Microbiol. 59, 1762–1766.

[20] Neilson, J.W., Josephson, K.L., Pepper, I.L., Arnold, R.B., DiGiovanni, G.D. and Sinclair, N.A. (1994) Frequency of horizontal gene transfer of a large catabolic plasmid (pJP4) in soil. Appl. Environ. Microbiol. 60, 4053–4058.

[21] DiGiovanni, G.D., Neilson, J.W., Pepper, I.L. and Sinclair, N.A. (1996) Gene transfer of *Alcaligenes eutrophus* JMP134 plasmid pJP4 to indigenous soil recipients. Appl. Environ. Microbiol. 62, 2521–2526.

[22] Amy, P.S., Schulke, J.W., Frazier, L.M. and Seidler, R.J. (1985) Characterization of aquatic bacteria and cloning of genes specifying partial degradation of 2,4-dichlorophenoxyacetic acid. Appl. Environ. Microbiol. 49, 1237–1245.

[23] Fournier, J.C., Codaccioni, P. and Soulas, G. (1981) Soil adaptation to 2,4-D degradation in relation to the application rates and the metabolic behaviour of the degrading microflora. Chemosphere 10, 977–984.

[24] Top, E.M., Holben, W.E. and Forney, L.J. (1995) Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. Appl. Environ. Microbiol. 61, 1691–1698.

[25] Vallaeys, T., Fulthorpe, R.R., Wright, A.M. and Soulas, G. (1996) The metabolic pathway of 2,4-dichlorophenoxyacetic acid degradation involves different families of *tfdA* and *tfdB* genes according to PCR-RFLP analysis. FEMS Microbiol. Ecol. 20, 163–172.

[26] Vallaeys, T. (1992) Isolement d'une communauté microbienne dégradant l'acide 2,4-dichlorophénoxyacétique à partir d'un sol de Dijon. Caractérization cinétique et génétique des souches impliquées. PhD Thesis, Université de Lille I, Lille.

[27] Suwa, Y., Wright, A.M., Fukimori, F., Nummy, K.A., Hausinger, R.P., Holben, W.E. and Forney, L.J. (1996) Characterization of a chromosomally encoded 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase from *Burkholderia* sp. Strain RASC. Appl. Environ. Microbiol. 62, 2464–2469.

[28] Holben, W.E., Schroeter, B.M., Calabrese, V.G.M., Olsen, R.H., Kukor, J.K., Biederbeck, V.O., Smith, A.E. and Tiedje, J.M. (1992) Gene probe analysis of soil microbial populations

COR-AAD1-00001187

selected by amendment with 2,4-dichlorophenoxyacetic acid. Appl. Microbiol. Environ. 58, 3941–3948.
[29] Laguerre, G., Allard, M.R., Revoy, F. and Amarger, N. (1994) Rapid identification of Rhizobia by restriction fragment length polymorphism analysis of PCR-amplified 16S rDNA genes. Appl. Environ. Microbiol. 60, 56–63.
[30] Laguerre, G., Rigottier-Gois, L. and Lemanceau, P. (1994) Fluorescent *Pseudomonas* species categorized by using polymerase chain reaction (PCR)/restriction fragment analysis of 16S rDNA. Mol. Ecol. 3, 479–487.
[31] Moyer, C.L., Dobbs, F.C. and Karl, D.M. (1994) Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. Appl. Environ. Microbiol. 60, 871–879.
[32] Martinez-Murcia, A.J., Acinas, S.G. and Rodriguez-Valera, F. (1995) Evaluation of prokaryotic diversity by restrictase digestion of 16S rDNA directly amplified from hypersaline environments. FEMS Microbiol. Ecol. 17, 247–256.
[33] Nelson, P.N., Dictor, M.C. and Soulas, G. (1994) Availability of organic carbon in soluble and particle-size fractions from a soil profile. Soil Biol. Biochem. 26,1549–1555.
[34] Dictor, M.C. (1994) Caractérisation de la distribution et du comportement métabolique de la microflore indigene dans un profil de sol. PhD Thesis, INPL-ENSAIA, France.
[35] Fukumori, F. and Hausinger, R.P. (1993) *Alcaligenes eutrophus* JMP134 2,4-dichlorophenoxyacetate monooxygenase is an alpha-ketoglutarate-dependant dioxygenase. J. Bacteriol. 175, 2083–2086.
[36] Tabor, H., White-Tabor, C. and Hafner, E.W. (1976) Convenient method for detecting $^{14}CO_2$ in multiple samples: Application to rapid screening for mutants. J. Bacteriol. 128, 485–486.
[37] Fulthorpe, R.R., Rhodes, A.N. and Tiedje, J.M. (1996) Pristine soils mineralize 3-chlorobenzoate and 2,4-dichlorophenoxyacetate via different microbial populations. Appl. Environ. Microbiol. 62, 1159–1166.
[38] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning, A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[39] Lane, D.J. (1991) 16S/23S rRNA sequencing. In: Nucleic Acid Techniques in Bacterial Systematics (Stackebrandt, E. and Goodfellow, M., Eds.), Wiley and Sons, Chichester.
[40] Moyer, C.L., Tiedje, J.M., Dobbs, F.C. and Karl, D.M. (1996) A computer-simulated fragment length polymorphism analysis of bacterial small subunit rRNA genes: Efficacy of selected tetrameric restriction enzymes for studies of microbial diversity in nature. Appl. Environ. Microbiol. 62, 2501–2507.
[41] Soulas, G. (1993) Evidence for the existence of different physiological groups in the microbial community responsible for 2,4-D mineralization in soil. Soil Biol. Biochem. 25, 443–449.
[42] Tonso, N.L., Matheson, V.G. and Holben, W.E. (1995) Polyphasic characterization of a suite of bacterial isolates capable of degrading 2,4-D. Microb. Ecol. 30, 3–24.
[43] Senior, E., Bull, A.T. and Slater, J.H. (1976) Enzyme evolution in a microbial community growing on the herbicide Dalapon. Nature 263, 476–479.
[44] Wright, A.M., McGowan, C., Fulthorpe, R., Vallaeys, T., Tiedje, J. and Snyder, L. (1995) Comparison of *tfdA* sequences from 2,4-D degrading isolates and communities. 7th International Symposium on Microbial Ecology, Santos, Sao Paulo.
[45] Vallaeys, T., Fulthorpe, R.R. and Wright, A.D. (1997) Comparison of *tfdB* sequences from bacterial isolates degrading 2,4-D and 2,4-dichlorophenol. 97th ASM General Meeting Q403.
[46] Kamataga, Y., Fulthorpe, R.R., Tamura, K., Takami, H., Forney, L.J. and Tiedje, L.J. (1997) Pristine environments harbour a new group of oligotrophic 2,4-dichlorophenoxyacetic degrading bacteria. Appl. Environ. Microbiol. 63, 2266–2272.
[47] Filer, K. and Harker, A.R. (1997) Identification of the inducing agent of the 2,4-dichlorophenoxyacetic acid pathway encoded by plasmid pJP4. Appl. Environ. Microbiol. 63, 317–320.
[48] Parsek, M.R., Shinabarger, D.L., Rothmel, R.K. and Chakrabarty, A.M. (1992) Roles of CatR and *cis,cis*-muconate in activation of the *catBC* operon which is involved in benzoate degradation in Pseudomonas putida. J. Bacteriol. 174, 7798–7806.
[49] Van der Tweel, W.J., ter Burg, J.N., Kok, J.B. and de Bont, A.M. (1986) Biotransformations of 4-hybroxybenzoate from 4-chlorobenzoate by *Alcaligenes eutrophus* NTB-1. Appl. Microbiol. Biotechnol. 25, 289–294.
[50] Parekh, N.R., Hartmann, A. and Fournier, J.C. (1996) PCR detection of the *mcd* gene and evidence of sequence homology between the degradative genes and plasmids from diverse carbofuran-degrading bacteria. Soil Biol. Biochem. 28, 1797–1804.
[51] Farhana, L. (1996) Use of Serological Techniques to Detect and Purify 2,4-Dichlorophenol Hydroxylase from *Alcaligenes eutrophus*. PhD Thesis, University of Sydney, Sydney.
[52] Vallaeys, T. and Soulas, G. (1992) Gene *tfd* from the degrading bacterium *Alcaligenes paradoxus* TV1 could be used as probe for the research of soil microorganisms able to degrade this herbicide. Proceedings of the International Symposium on Environmental Aspects of Pesticide Microbiology, Sigtuna, pp. 177–183.

COR-AAD1-00001188

Downloaded from https://academic.oup.com/femsec/article/24/3/269/537857 by Pharmacia Hall Library user on 30 September 2020

ELSEVIER

FEMS Microbiology Ecology 20 (1996) 163–172

FEMS MICROBIOLOGY ECOLOGY

The metabolic pathway of 2,4-dichlorophenoxyacetic acid degradation involves different families of *tfdA* and *tfdB* genes according to PCR-RFLP analysis

Tatiana Vallaeys [a,*], Roberta R. Fulthorpe [b,1], Alice M. Wright [b], Guy Soulas [a]

[a] *Centre de Microbiologie du Sol et de l'Environnement, Laboratoire de Microbiologie des Sols, B.V. 1540, 21034 Dijon, France*
[b] *Center for Microbial Ecology, Michigan State University, East Lansing MI, USA*

Received 30 October 1995; revised 18 March 1996; accepted 21 March 1996

Abstract

Twenty-five 2,4-dichlorophenoxyacetic acid (2,4-D) degrading bacteria from geographically diverse locations and presenting various degrees of similarity or no similarity to the *tfdA* and *tfdB* genes from *Alcaligenes eutrophus* JMP134 were analysed by PCR-RFLP (restriction length fragment polymorphism). Primers for the 2,4-D etherase gene were derived by sequence alignment of the *tfdA* genes from *A. eutrophus* JMP134 and *Burkholderia* sp. RASC. Primers for the 2,4-dichlorophenolhydroxylase gene were based on the *tfdB* gene sequence from *A. eutrophus* JMP134 by taking codon degeneration and variations in amino acid residue sequences into consideration. PCR amplification using the *tfdA* primer set produced fragments of 0.3 kb from 17 strains which showed varying degrees of similarity to the *tfdA* gene probe from *A. eutrophus* JMP134. Significant variations in the gene sequences were confirmed by PCR-RFLP analysis. DNA amplification using the *tfdB* primer set produced a 1.1 kb fragment from 19 strains. Amongst them, two did not show any similarity to the *tfdB* gene probe. The size and restriction pattern of the products obtained from *A. eutrophus* JMP134 were in accordance with the expected size calculated from the *A. eutrophus tfdA* and *tfdB* gene sequence and their theoretical PCR-RFLP patterns. Some strains which did not amplify using the *tfdA* primer set did however amplify with the *tfdB* primer set. These results suggest the independent evolution of these two genes in the construction of the 2,4-D metabolic pathway. Our *tfdA* and *tfdB* primer sets could be used for the detection of similar sequences in bacteria and soils. Moreover, PCR-RFLP patterns could also be used to select subsets of strains for sequencing to study the phylogeny of the *tfdA* and *tfdB* genes.

Keywords: *Alcaligenes eutrophus* JMP134; *Burkholderia* sp. RASC; PCR-primer; *tfdA* gene; *tfdB* gene; 2,4-Dichlorophenolhydroxylase; Alpha-ketoglutarate dependent dioxygenase

1. Introduction

Soil microorganisms are involved in the biodegradation of a number of xenobiotic compounds in natural habitats. The literature reports an increasing need to develop new methods to detect, isolate and

* Corresponding author. Tel: +33 80 63 32 98; Fax: +33 80 63 32 24; E-mail: vallaeys@dijon.inra.fr
[1] Present address: Scarborough Campus, University of Toronto, Scarborough, Ont. M1C 1A4, Canada.

0168-6496/96/$15.00 © 1996 Federation of European Microbiological Societies. Published by Elsevier Science B.V.
PII S0168-6496(96)00027-X

COR-AAD1-00001189

Downloaded from https://academic.oup.com/femsec/article/20/2/163/478565 by Purdue University Libraries user on 03 September 2020

characterize the strains playing a part in these degradation processes. Selective media containing a xenobiotic compound as a sole carbon source can be used to isolate heterotrophic strains capable of the complete mineralization of the xenobiotic; but, cometabolizing strains or microbial communities involved in the xenobiotic's biodegradation are neglected by such methods. Recent studies confirm that strains may carry parts of a specified metabolic pathway without being able to degrade a xenobiotic molecule entirely [1,2].

The chlorinated herbicide, 2,4-D (2,4-dichlorophenoxyacetic acid), has been used and studied for over 40 years. This compound is degraded by various species of bacteria and fungi, a number of which have been isolated and characterized [3–10].

Alcaligenes eutrophus JMP134 [4] is a 2,4-D degrading bacterium isolated from an Australian soil which carries the genes involved in the 2,4-D degradation pathway on a well studied plasmid, pJP4. The 2,4-D pathway and catabolic genes from *A. eutrophus* JMP134 (pJP4) have been extensively characterized [11–16]. On the plasmid pJP4, the six genes involved in 2,4-D metabolism are grouped into three operons: *tfdA*, *tfdB* and *tfdCDEF*.

The *tfdB* gene encodes 2,4-dichlorophenolhydroxylase which converts 2,4-dichlorophenol to 3,5-dichlorocatechol. Nurk et al. [17] found that the *tfdB* sequence exhibits similarity to the phenol monooxygenase gene, *pheA*. However, the 2,4-dichlorophenolhydroxylase protein is expressed only in 2,4-D induced *A. eutrophus* JMP134 cells [13] and seems to be specifically involved in the orientation of 2,4-D catabolic intermediates to the ortho-cleavage route.

The *tfdA* gene encodes the protein involved in the first step of the 2,4-D catabolic pathway, i.e. the rupture of the ether bond and the formation of 2,4-dichlorophenol. This enzyme, previously known as a 2,4-dichlorophenoxyacetate monooxygenase has recently been shown to be an alpha-ketoglutarate dependent dioxygenase [18]. This enzyme also accepts phenoxyacetate and MCPA (2-methyl-4-chlorophenoxyacetate) as other substrates [12,19]. However, Pieper et al. [19] showed that it was not induced by unsubstituted phenoxyacetates. Moreover, contrary to the *tfdBCDEF* genes, the *tfdA* sequence does not show significant similarities to any other known sequences [18] and therefore could be specific to the 2,4-D metabolic pathway.

New approaches have been developed for the detection of 2,4-D degrading strains, based on molecular methods [20–23]: molecular probes or primers have been developed to identify, enumerate and isolate microbes capable of degrading xenobiotic compounds. Specific genes involved in degradative pathways of such compounds are used as probes to detect the strains of interest in the environment. These techniques eliminate problems associated with culture techniques on agar media, such as non-specific growth on trace carbon sources or the inability to grow degrading communities on such media. However, such methods are based on the assumption that high similarities exist between genes and pathways involved in the biodegradation of these compounds. In recent work, Fulthorpe et al. [24] showed that similarities amongst different *tfdA* and *tfdB* gene sequences could be as low as 60%. The detection of sequences presenting low similarity to a probe requires reduction in stringency during hybridization and leads to non-specific binding. As primers could be chosen amongst short conserved regions of a gene sequence, the PCR approach appears more suitable, reducing the screening of false positive degraders while increasing the ability to detect divergent genes. *tfdB* primers have been described previously [22,23], however, none of the primer sets previously described take into consideration codon redundancy, and may therefore lead to an underestimation of the genes present in the environment.

The advantage of the PCR approach is to provide a rapid, sensitive and specific tool for direct detection of microorganisms or genes in the environment. Therefore, we developed specific primer sets to study the diversity of 2,4-D degradative genes involved in the first two steps of 2,4-D metabolism.

Comparisons of gene families involved in the biodegradation of the herbicide 2,4-D may allow for a better understanding of the origin and evolution of catabolic pathways for this herbicide.

2. Materials and methods

Bacterial species used in this study are presented in Table 1. *A. eutrophus* JMP134 (pJP4) has been

Downloaded from https://academic.oup.com/femsec/article/20/2/163/583869 by Pasteur Institut International, Inc. user on 08 September 2020

COR-AAD1-00001190

COR-AAD1-00001191

Table 1
Bacterial strains used in this study and correlations between hybridization and PCR/RFLP analysis of the *tfdA* and *tfdB* regions

Strain		Origin	Hybridisation with *tdfA*	Hybridisation with *tfdB*	PCR results				Reference
					PCR: *tfdA* amplification	*tfdA* PCR/RFLP group	PCR: *tfdB* amplification	*tfdB* PCR/RFLP pattern	
Alcaligenes eutrophus	JMP134	Australia	+	+	+	1	+	I	[4]
Burkholderia sp.	RASC	Oregon (USA)	+	+	+	2	+	II	[24]
Burkholderia mallei	TFD6	Michigan(USA)	+	+	+	2	+	II	[24]
Rhodoferax fermentans	TFD23	Michigan (USA)	+	NH	+	3	±	nd	[24]
Rhodoferax fermentens	TFD31	Canada	+	NH	+	4	+	III	[24]
Alcaligenes eutrophus	TFD41	Michigan(USA)	+	+	+	1	+	I	[24]
Rhodoferax fermentens	B6.9	Ontario (Canada)	+	NH	+	4	+	III	[24]
Variovorax paradoxus	TV1	Dijon (France)	+	+	+	4	+	II	[25]
undetermined	JD6	Michigan(USA)	+	+	+	1	+	I	Dunbar (unpubl.)
"	JD12	Michigan (USA)	+	+	+	1	+	I	Dunbar (unpubl.)
"	JD16	Michigan (USA)	+	+	+	1	+	IV	Dunbar (unpubl.)
"	JD17	Michigan (USA)	+	+	+	1	+	IV	Dunbar (unpubl.)
"	ASD6	Canada	+	+	+	1	+	V	Smith (unpubl.)
Alcaligenes xylosoxidans	AXD1	Dijon (France)	+	+	+	1	+	V	[35]
Alcaligenes xylosoxidans	POD3	Dijon (France)	+	+	+	1	+	V	[36]
Alcaligenes xylosoxidans	D2D9	Versailles(France)	+	+	+	1	+	V	[36]
Alcaligenes xylosoxidans	CH_3D11	Dijon (France)	+	nd	+	1	−	−	[36]
undetermined	AW5	Great Britain	NH	nd	−	−	+	I	Weightman (unpubl.)
Variovorax paradoxus	JMP116	Australia	+	nd	±	nd	−	−	[4]
Alcaligenes eutrophus	JMP132	Australia	+	nd	±	nd	−	−	[4]
Sphingomonas sp.	B6.5	Ontario (Canada)	NH	NH	−	−	−	−	[24]
Sphingomonas sp.	B6.10	Ontario (Canada)	NH	NH	−	−	−	−	[24]
Sphingomonas sp.	TFD44	Montana (USA)	NH	+	−	−	+	VI	[24]
Sphingomonas sp.	EML146	Oregon (USA)	NH	+	−	−	+	VI	[24]
Burkholderia cepacia	PCD4	Indonesia	NH	nd	−	−	−	−	[36]
Variovorax paradoxus	JMP118(*)	Australia	NH	NH	−	−	−	−	[4]
Alcaligenes eutrophus	JMP138(*)	Australia	NH	nd	−	−	−	−	[4]
Alcaligenes eutrophus	JMP139(*)	Australia	NH	nd	−	−	−	−	[4]
Alcaligenes eutrophus	JMP222(*)	Australia	NH	NH	−	−	−	−	[4]

*, non-degrading derivative; +, positive amplification or hybridisation; ±, weak amplification signal; −, negative amplification; nd, not determined; NH, no homology detected.

Downloaded from https://academic.oup.com/femsec/article/20/3/163/4785553 by Ronse Hubert International, Inc. user on 03 September 2020

used as the positive control in our study. Its cured derivative, strain JMP222, lacking the genes involved in 2,4-D metabolism, as well as JMP118, JMP138, JMP138 [4] have been used as negative controls in all our PCR and hybridization experiments. Twenty five 2,4-D degrading strains isolated from diverse locations by various laboratories, including our own, were screened for the presence of genes similar to *tfdA* and *tfdB* by PCR. Culture media described by Vallaeys [25] have been used: 2,4-D degrading strains were cultivated on 2,4-D minimal medium (per litre distilled water: 1.5 g K_2HPO_4; 0.5 g KH_2PO_4; 1 g $(NH_4)_2SO_4$; 0.2 g $MgSO_4$, $7(H_2O)$; 5 mg yeast extract; 500 μl of a salt solution containing 0.4 g l^{-1} $ZnSO_4$, $7(H_2O)$; 20 mg l^{-1} $MgCl_2$, $4(H_2O)$; 10 mg l^{-1} H_3BO_3; 50 mg l^{-1} $CoCl_2$, $6(H_2O)$; 200 mg l^{-1} $CuSO_4$; 10 mg l^{-1} $NiCl_2$, $6(H_2O)$; 500 mg l^{-1} $FeSO_4$, $7(H_2O)$; 250 mg l^{-1} EDTA). Control strains were grown on Luria Broth [26].

Total genomic DNA extraction was carried out as described in Sambrook et al. [26] from cultures grown on 2,4-D minimal medium supplemented with 5 mg l^{-1} yeast extract. DNA was resuspended in sterile distilled water at a concentration of 0.05 μg ml^{-1} as determined by UV absorbance at OD_{260nm}.

Specific primers for the 2,4-D etherase gene (*tfdA*) were derived by sequence alignment of the *tfdA* genes from *A. eutrophus* JMP134 [12] and *Burkholderia* sp. RASC [10] taking codon degeneration into consideration. Primers for the 2,4-dichlorophenolhydroxylase gene (*tfdB*) were based on the *tfdB* gene sequence from *A. eutrophus* JMP134 [13]. They were designed by considering codon degeneration and variations in amino acid residue sequences. PCR amplification was carried out on purified total genomic DNA using the Promega Amplification PCR kit and the Hybaid (Omnigene, France) PCR thermocycler. Each 100 μl reaction mixture contained 1 × amplification buffer, 1 U of *Taq* polymerase, 200 μM (each) deoxynucleotide triphosphate, 50 μM of each primer, 50 mM of $MgCl_2$ and 10 μM of DNA template. A hot start (3 min, 94°C) was done before the addition of the *Taq* DNA polymerase. Thirty-five PCR cycles (denaturation: 45 s, 94°C; annealing: 30 s, 59°C (*tfdA*) or 57°C (*tfdB*); extension: 2 min, 72°C) were followed by a 6 min extension at 72°C. PCR products were revealed by electrophoresis in a 1% agarose gel (Sigma, France) and DNA fragments were transferred to a Nytran membrane (Schleicher and Schuell, France).

Hybridizations were performed as described by Fulthorpe et al. [24] using *tfdA* and *tfdB* probes from *A. eutrophus* JMP134 under medium stringency conditions (20% formamide in hybridization solution). Internal fragments from the *tfdA* and *tfdB* genes used as molecular probes have been described by Holben et al. [22]. Probes were labelled using the non-radioactive labelling kit (Boehringer-Mannheim) according to the instructions provided by the manufacturer.

The resulting amplified PCR fragments were digested by appropriate restriction enzymes (Boehringer-Mannheim) according to the supplier's instructions. Restriction enzymes were chosen according to the frequency of their restriction sites on the sequence of the *tfdA* and *tfdB* genes of *A. eutrophus* JMP134 [12,13]. Restriction fragments were separated by electrophoresis in a horizontal 4% Nusieve agarose gel (FMC Bioproducts) stained with ethidium bromide. Boehringer-Mannheim weight marker VIII was used as a size ladder.

3. Results

3.1. Description of the primers

Since cysteine is involved in disulfur bond formation, this amino acid frequently plays an important part in the primary structure of proteins. The *tfdA* primer set has been chosen at the position of the two unique cysteine residues in the *tfdA* amino acid sequences. In addition, both positions correspond to a conserved part of the nucleotide sequences between the two sequenced *tfdA* genes [10,12]. *Taq* polymerase is extremely sensitive to mismatches at the 3′ end [27]. Therefore, codon redundancy has been taken into consideration in order to respect a strict complementary nucleotide sequence at the 3′ end of the primers. To avoid extreme degeneration of the primers, 3′ ends were chosen amongst amino acid residues encoded by one or two different codons. Resulting sequences for the *tfdA* primer set were consequently designed as 5′-AAC-GCA-GCG-(G/A)TT-(G/A)TC-CCA-3′ and 5′-ACG-GAG-

COR-AAD1-00001192

Downloaded from https://academic.oup.com/femsec/article/20/2/163/630800 by guest on 08 September 2020

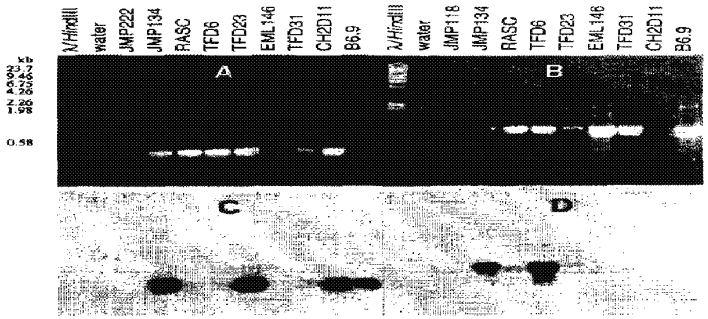


Fig. 1. Hybridization analysis of the 300 bp *tfdA*-like PCR product (**A**) and the 1.1 kb *tfdB*-like PCR product (**B**) obtained from 2,4-D degrading strains from geographically diverse origins using the *tfdA* and the *tfdB* primer set respectively. Similarity can be detected at low stringency with the *tfdA* gene probe (**C**) and the *tfdB* gene probe (**D**) from *A. eutrophus* JMP134.

TTC-TG(C/T)-GA(C/T)-ATG-3′ for the lower and the upper *tfdA* primers, respectively.

Similar considerations resulted in the choice of the following sequences for the *tfdB* lower and upper primers: 5′-ATA-GCG-CTG-(A/G)TT-CAT-(C/T)TC-3′ and 5′-CG-CA(C/T)-ATC-ACC-AA(C/T)-CA(G/A)-C-3′.

Table 2
Restriction fragment length polymorphism of the *tfdA* PCR products

Pattern number [+]	Restriction fragment length (bp) [*]		
	I	II	III
*Msp*I	155	208	
	52 (53, 54)	154	
	31		
	(14, 6)		
*Hha*I	145	274	274
	99	88	55
	63		(20,14)
	37		
	20		
*Hae*III	167	164	167
	107	124	107
	74	74	52
	(11, 7)		(22, 17, 7)

[*], values in brackets indicate triplet bands and small fragments which were undistinguishable or not visualised on agarose gel.
[+], all restriction fragments of patterns I and II agreed with theorical values calculated from *tfdA* gene sequence of *A. eutrophus* JMP134 and *Burkholderia sp.* RASC. Pattern III showed fragment size differing from predicted theorical values.

Table 3
PCR/RFLP groups of the *tfdA* PCR products from the 2,4-D degrading strains

PCR/RFLP group	*Msp*I	*Hha*I	*Hae*III
1	I	I	I
2	II	II	II
3	I	I	III
4	II	III	II

3.2. Amplification and restriction patterns

A collection of twenty five 2,4-D degrading strains from geographically diverse origins and presenting various degrees of similarity to the *tfdA* and *tfdB* probes from *A. eutrophus* JMP134 were used to check the specificity of two primer sets for the *tfdA* and *tfdB* genes. Plasmid-cured derivatives *A. eutrophus* JMP222, JMP138, JMP139 and *Variovorax paradoxus* JMP 118 used as negative controls were not amplified in the presence of either the *tfdA* or *tfdB* primer set.

PCR amplification using the *tfdA* primer set gave rise to fragments of 300 bp from 17 strains which

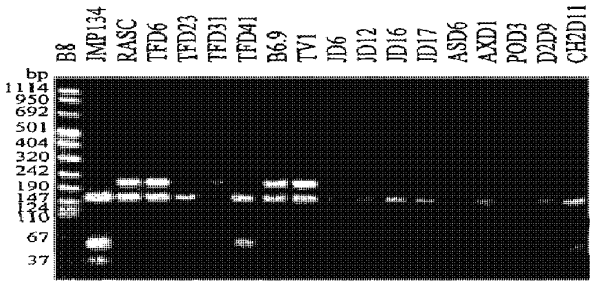


Fig. 2. PCR-RFLP analysis by the *Msp*I restriction enzyme of the 300 bp PCR fragment obtained from the strain collection using the *tfdA* primer set.

COR-AAD1-00001193

Downloaded from https://academic.oup.com/femsec/article/20/2/163/633193 by guest on 03 September 2025

showed similarity to the *tfdA* gene probe (Table 1). The size of the obtained fragments was in accordance with the size predicted from the *tfdA* gene sequences (Table 2). However, when *tfdA* PCR products were hybridized with the *tfdA* gene probe, there was no correlation between the quantity of the PCR product and the hybridization signal obtained (Fig. 1A,C). Differences in the hybridization signals suggest significant variations in the amplified gene sequences which were confirmed by PCR-RFLP analysis. The digestion of PCR fragments with 3 four-base restriction enzymes *Hae*III, *Msp*I, and *Hha*I generated 3, 2, and 3 distinguishable patterns, respectively. An illustration of this PCR-RFLP analysis by the restriction enzyme *Msp*I is given in Fig. 2. Strains presenting different combinations of patterns by the various restriction enzymes were classified in different PCR-RFLP groups. This PCR-RFLP analysis using 3 restriction enzymes resulted in the observation of 4 different groups of patterns (Table 3). The first group corresponded to the *A. eutrophus* JMP134 *tfdA* theoretical pattern as obtained from the sequence [12]. The second group of patterns were in agreement with the theoretical values calculated from the *Burkholderia* sp. RASC *tfdA* gene [10]. The third and fourth groups showed patterns differing from the predicted theoretical values for at least one

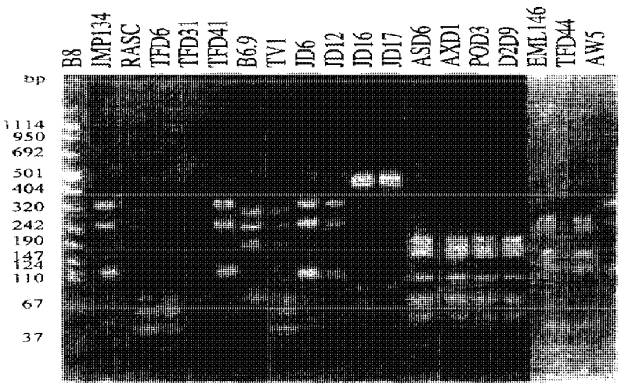


Fig. 3. PCR-RFLP analysis by the *Hae*III restriction enzyme of the 1.1 kb PCR fragment obtained from geographically diverse 2,4-D degrading isolates using the *tfdB* primer set.

Table 4
Restriction fragment length polymorphism of *tfdB* PCR products

Pattern number	Restriction fragment length (bp)					
	I	II	III	IV	V	VI
*Hae*III *	338	315	315	425 (420)	200	270
	260	250	250		180	250
	132 (126,	200	200	105	165	165
	118)	85 (80)	85 (80)		110	135
	60	65 (60)		70	75 (67)	80
	53		55	60	55	65 (60)
	47	43 (40)	45 (40)	45	50	43 (40)
	37				40 (35)	

*, values in brackets indicate triplet bands and small fragments which were undistinguishable or not visualised on agarose gel.

restriction enzyme. The use of additional enzymes did not improve the differentiation between the PCR products. There was no correlation between PCR-RFLP groups and geographic origins of the 2,4-D degrading strains.

The *tfdB* primer set gave rise to a 1.1 kb PCR product from 19 strains presenting varying degrees of similarity to the canonical *tfdB* gene of *A. eutrophus* JMP134 (Table 1, Fig. 1B,D). Amongst them, two strains (TFD31 and B6.9) did not show any similarity to the *tfdB* gene probe in previous hybridization studies [24]. The *Hae*III enzyme was chosen for the PCR-RFLP analysis of the *tfdB*-like fragment because of the high frequency of its restriction sites on the *A. eutrophus* JMP134 *tfdB* sequence. The size and restriction pattern by *Hae*III of the product obtained from *A. eutrophus* JMP 134 was in accordance with the expected size calculated from the *A. eutrophus tfdB* gene sequence [13] and the theoretical PCR-RFLP pattern of the *A. eutrophus tfdB* gene by this enzyme (Fig. 3). Six different restriction patterns could be distinguished amongst the digested PCR products obtained using the *tfdB* primer set (Table 4). The two strains presenting no similarity to the *tfdB* gene probe showed an identical restriction pattern which was not observed amongst the strains presenting similarity to this probe (Fig. 3).

Some strains which did not amplify using the *tfdA* primer set did however amplify with the *tfdB* primer set. Moreover, there was no correlation between the *tfdA* and *tfdB* PCR-RFLP groups (Table 1). These results suggest the independent evolution of these

COR-AAD1-00001194

Downloaded from https://academic.oup.com/femsec/article/20/2/163/547856 by Harvard Medical International, Inc. user on 03 September 2022

two genes in the construction of the 2,4-D metabolic pathway.

4. Discussion

This study clearly demonstrates the utility of using PCR amplification and restriction digests as a means of studying the diversity of genes within 2,4-D degraders. This method allows for the rapid screening of a number of strains, allowing for the reduction of the redundancy represented in strain collections prior to more time consuming and expensive analysis.

A key issue is whether or not the amplification products are indeed fragments of genes isofunctional to *tfdA* and *tfdB*. We believe there are several pieces of evidence to show this is a parsimonious assumption. First, the amplifications were carried out at highly stringent annealing temperatures (59 or 57°C). Secondly, the primers consistently amplify DNA of the correct size from 2,4-D degraders which are known to carry *tfdA* and *tfdB* like genes (by hybridization analysis). Conversely, with only two exceptions, there is no amplification from 2,4-D degrading strains which do not show hybridization to the *tfdA* and *tfdB* of JMP134, and never any amplification from negative control strains. All of the amplification products hybridized, to some extent, with the *tfdA* and *tfdB* gene probes under low stringency conditions. Recently, numerous fragments, amplified using the *tfdA* primers, have been sequenced. To date, three distinctly different *tfdA*-like genes have been partially sequenced; in each case, the fragments encode polypeptides of 99% amino acid similarity [28–30]. Perhaps most importantly, the amplification products, after restriction digestion, fall into distinct groups, and each of the restriction patterns shares at least one band with the other restriction patterns. This would be highly unlikely to occur if DNA had been randomly amplified from unrelated genes. Such results suggest the existence of phylogenetic relations between the amplified fragments obtained from the 2,4-D degraders using our primer sets.

The PCR-RFLP analysis places the various genes into groups that correspond with the phylogeny or the plasmid content of their hosts. Type I *tfdA*'s identical to the canonical *tfdA* are found in association with *Alcaligenes* or *Variovorax* species known or suspected of carrying incP plasmids (JMP134, TFD41) [31]. Type 2 *tfdA*'s are found in two *Burkholderia* strains for which complete *tfdA* sequences confirm 100% DNA similarity, and chromosomal gene locations have been established [32,33]. Type 4 *tfdA*'s are found in two species, *Rhodoferax fermentens* (B6-9, TFD31) and *Variovorax paradoxus* TV1. What is common to these strains is the observation of one or two large plasmids where the hybridization signal is detected [24,25]. Organisms carrying type 4 sequences have been found in various strains. Maltseva et al. [29], recently report a type 4 *tfdA* in a salt-tolerant alkalophile that degrades 2,4-D. Vallaeys et al. (unpublished) detected this gene in a strain isolated from a chemostat. Top et al. (personal communication) isolated this gene from 2,4-D enriched soils using a plasmid capture technique—their study illustrates clearly that the *tfd* genes are capable of independent movement. Our type 3 *tfdA* appears to be unique amongst this group, in spite of the fact that it was amplified from a *Rhodoferax* strain. In summary, the version of *tfdA* that is carried by strains seems more dependant on the presence of mobile agents such as plasmids than on the phylogenetic origin of the strain.

When we look at *tfdB* patterns, there are three things to note. First, a higher degree of diversity has been uncovered than for the *tfdA* gene but this may be related to the higher size of the amplified fragment. Secondly, there is some correspondence between *tfdB* type and phylogeny of the bacterial strains carrying them. The distinct pattern VI is found only in two *Sphingomonas* strains. Patterns I, II, and III appear in strains in conjunction with *tfdA* patterns 1, 2 and 4 respectively suggesting that these three pairs of genes are linked. However, patterns IV and V also can be found with type 1 *tfdA*'s, and in *Variovorax paradoxus* TV1, *tfdA* pattern 4 is observed in conjunction with *tfdB* pattern II. This indicates that, perhaps, there is a mechanism for *tfdA* and *tfdB* genes of different types to become reassorted into different combinations.

There were three cases where amplification of DNA occurred using the *tfdB* primers in spite of the absence of hybridization between the target and the canonical *tfdB* genes. In two cases the amplification was strong. We expected to see at least some exam-

COR-AAD1-00001195

Downloaded from https://academic.oup.com/femsec/article/20/2/163/478650 by Pioneer Hi-Bred International, Inc. user on 03 September 2020

ples of this. The lowest possible DNA similarity that can be detected using the low stringency method described in Fulthorpe et al. [24] is in the range of 60% DNA similarity. This does not preclude the successful amplification from less similar genes, in which we would expect to find conserved active sites, but divergent intervening sequences. This was the kind of result we hoped to obtain with these primers. The amplified DNA is clearly a *tfdB*-like gene because although it has a unique PCR-RFLP pattern, it shares six digestion fragments with the type II *tfdB*. In this case, the utility of PCR-RFLP in providing a means of analysing undetected genes is clear. It is likely that by further decreasing the annealing temperature, even more distantly related genes may be amplified, but products would have to be sequences if their RFLP patterns diverged significantly from known isofunctional genes.

The high diversity of both *tfdA* and *tfdB* like genes, their broad host range, and geographic homogeneity all suggest that much evolutionary time has elapsed since they evolved. Since we know little about the rates of gene divergence in soil bacteria, and even less about the rates of gene transfer and bacterial global transport, it is difficult to speculate about when the *tfd* genes might have first evolved and why. Either the 30 or 40 years that have elapsed since 2,4-D came into wide use is sufficient for this level of diversity and mixing to have occurred, or a natural substrate is responsible for the selection of this of *tfd* elements. Fulthorpe et al. [34] have recently shown that 2,4-D degradation is common in the soils of uncontaminated ecosystems, providing some support for the latter hypothesis.

When using PCR-RFLP, we were able to discriminate different families of 2,4-D related (i.e. similar) amongst 2,4-D degrading isolates. The PCR-RFLP technique could therefore be useful to study the presence and diversity of 2,4-D related genes directly on DNA extracted from treated and untreated soils thus avoiding problems linked with the use of selective media. Moreover, our results suggest that our method for PCR detection of the *tfdA* and *tfdB* genes is a more sensitive method than hybridization for the detection of similar sequences in bacteria and soils. Sequence divergence observed by Fulthorpe et al. [24] amongst *tfd* genes and confirmed by this study show the limitations of the hybridization method when applied to the detection of similar sequences amongst soil microorganisms.

Moreover, this method allowed for a finer level of analysis than previous hybridization studies: the PCR-RFLP patterns could be used to select subsets of strains for sequencing and therefore establish phylogenetic relations between the different *tfdA* and *tfdB* genes amongst our 2,4-D degrading isolates. By taking codon degeneration and variations in amino acid residues sequences into consideration, we succeeded in designing PCR primers able to detect a significant proportion of *tfdA* and *tfdB*-like genes amongst 2,4-D degrading isolates. Consequently this method constitutes a powerful tool for studying 2,4-D degraders or members of 2,4-D degrading communities in the environment.

Acknowledgements

We thank Mohamed El Hamdaoui and Bernard Lagacherie for technical assistance and Gisèle Laguerre for interesting discussions, suggestions and comments. We also thank Dawn Hand for her linguistic advise.

References

[1] Top, E.M., Holben, W.E. and Forney, L.J. (1995) Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. Appl. Environ. Microbiol. 61, 1691–1698.
[2] Van der Meer, J.R., Werlen, C. and Leveau, J.H.J. (1995) Natural evolution of bacteria degrading chlorinated benzenes. Abstracts, 7th International Symposium on Microbial Ecology. S4-3.3. Sao Paulo, Brazil.
[3] Fisher, P.R., Appleton, J. and Pemberton, J.M. (1978) Isolation and characterization of the pesticide-degrading plasmid pJP1 from *Alcaligenes paradoxus*. J. Bacteriol. 135, 798–804.
[4] Don, R.H. and Pemberton, J.M. (1981) Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J. Bacteriol. 145, 681–686.
[5] Amy, P.S., Schulke, J.W., Frazier, L.M. and Seidler, R.J. (1985) Characterization of aquatic bacteria and cloning of genes specifying partial degradation of 2,4-dichlorophenoxyacetic acid. Appl. Environ. Microbiol. 49, 1237–1245.
[6] Chaudhry, G.R. and Huang, G.H. (1988) Isolation of a new plasmid from a *Flavobacterium* sp. which carries the genes

COR-AAD1-00001196

Downloaded from https://academic.oup.com/femsec/article/20/2/163/618810 by guest on 08 September 2022

for degradation of 2,4-dichlorophenoxyacetate. J. Bacteriol. 170, 3897–3902.
[7] Yadav, J.S. and Reddy, C.A. (1993) Mineralization of 2,4-dichlorophenoxyacetic acid (2,4-D) and 2,4,5-trichlorophenoxyacetic acid by *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. 59, 2904–2908.
[8] Bhat, M.A., Tsuda, M., Horiike, K., Nozaki, M., Vaidyanathan, C.S. and Nakazawa, T. (1994) Identification and characterization of a new plasmid carrying genes for degradation of 2,4-dichlorophenoxyacatate from *Pseudomonas cepacia* CSV90. Appl. Environ. Microbiol. 60, 307–312.
[9] Ka, J.O., Holben, W.E. and Tiedje, J.M. (1994) Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D) degrading bacteria isolated from 2,4-D treated field soils. Appl. Environ. Microbiol. 60, 1106–1115.
[10] Suwa, Y., Holben, W.E. and Forney, L.J. (1994) Cloning of a novel 2,4-D catabolic gene isofunctional to *tfdA* from *Pseudomonas* sp. TFD3. Abstr. 94th Gen. Meet. Am. Soc. Microbiol. Q-403, p. 293. American Society for Microbiology, Washington, DC, USA.
[11] Don, R.H., Weightman, A.J., Knackmuss, H.H. and Timmis, K.N. (1985) Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4). J. Bacteriol. 161, 85–90.
[12] Streber, W.R., Timmis, K.N. and Zenk, M.H. (1987) Analysis, cloning and high-level expression of 2,4-dichlorophenoxyacetic acid monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. 169, 2950–2955.
[13] Perkins, E.J., Gordon, M.P., Caceres, O. and Lurquin, P.F. (1990) Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. 172, 2351–2359.
[14] Kukor, J.J., Olsen, R.H. and Siak, J.-S. (1989) Recruitment of a chromosomally encoded maleylacetate reductase for degradation of 2,4-Dichlorophenoxyacetic acid by plamid pJP4. J. Bacteriol. 171, 3385–3390.
[15] Kaphammer, B. and Olsen, R.H. (1990) Cloning and characterization of *tfdS*, the repressor-activator gene of *tfdB*, from the 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. J. Bacteriol. 172, 5856–5862.
[16] Kaphammer, B., Kukor, J.J. and Olsen, R.H. (1990) Regulation of *tfdCDEF* by *tfdR* of the 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. J. Bacteriol. 172, 2280–2286.
[17] Nurk, A.L., Kasak, L. and Kivisaar, L. (1991) Sequence of the gene *pheA* encoding phenol monooxygenase from *Pseudomonas* sp. EST 1001: expression in *Escherichia coli* and *Pseudomonas putida*. Gene 102, 13–18.
[18] Fukumori, F. and Hausinger, R.P. (1993) *Alcaligenes eutrophus* JMP134 2,4-dichlorophenoxyacetate monooxygenase is an α-ketoglutarate-dependant dioxygenase. J. Bacteriol. 175, 2083–2086.
[19] Pieper, D.H., Reineke, W., Engesser, K.H. and Knackmuss, H.J. (1989) Regulation of catabolic pathways of phenoxyacetic acids and phenols in *Alcaligenes eutrophus* JMP134. Arch. Microbiol. 151, 365–371.
[20] Holben, W.E., Jansson, J.K., Chelm, B.K. and Tiedje, J.M. (1988) DNA probe method for the detection of specific microorganisms in the soil bacterial community. Appl. Environ. Microbiol. 58, 3941–3948.
[21] Amy, P.S., Staudaher, M.V. and Seidler, R.J. (1990) Comparison of a gene probe with classical methods for detecting 2,4-dichlorophenoxyacetic acid (2,4-D)-biodegrading bacteria in natural waters. Curr. Microbiol. 21, 95–101.
[22] Holben, W.E., Schroeter, B.M., Calabrese, V.G.M., Olsen, R.H., Kukor, J.K., Biederbeck, V.O., Smith, A.E. and Tiedje, J.M. (1992) Gene probe analysis of soil microbial populations selected by amendment with 2,4-dichlorophenoxyacetic acid. Appl. Environ. Microbiol. 58, 3941–3948.
[23] Neilson, J.W., Josephson, K.L., Pillai, S.D. and Pepper, I.L. (1992) Polymerase chain reaction and gene probe detection of the 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. Appl. Environ. Microbiol. 56, 1271–1275.
[24] Fulthorpe, R., McGowan, C., Maltseva, O., Holben, W. and Tiedje, J.M. (1995) 2,4-D degrading bacteria contain mosaics of catabolic genes. Appl. Environ. Microbiol. 61, 3274–3281.
[25] Vallaeys, T. (1992) Isolement d'une communauté microbienne dégradant l'acide 2,4-dichlorophénoxyacétique à partir d'un sol de Dijon. Caractérisation cinétique et génétique des souches impliquées. Thèse de doctorat. Université Lille I. France.
[26] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning, a Laboratory Manual, Second Edition. Cold Spring Harbor Laboratory Press, New York, USA.
[27] Kwok, S., Kellogg, D.E., McKinney, N., Spasic, D., Goda, L., Levenson, C. and Sninsky, J.J. (1990) Effects of primer-template mismatches on the polymerase chain reaction: Human immunodefiency virus type 1 model studies. Nucleic Acids Res. 18, 999–1005.
[28] Wright, A.D., McGowan, C., Fulthorpe, R., Vallaeys, T., Tiedje, J. and Snyder, L. (1995) Comparison of *tfdA* sequences from 2,4-D degrading isolates and communities. 7th International Symposium on Microbial Ecology. Santos, Sao Paulo, Brazil.
[29] Maltseva, O.V., McGowan, C., Fulthorpe, R. and Oriel, P. (1996) Degradation of 2,4-dichlorophenoxyacetic acid by a haloalkalophilic bacteria. Microbiology, in press.
[30] McGowan, C. (1994) Interspecies gene transfert in the evolution of 2,4-dichlorophenoxyacetic acid degrading bacteria. PhD. Thesis, Michigan State University.
[31] Tonso, N.L., Matheson, V.G. and Holben, W.E. (1995) Polyphasic characterization of a suite of bacterial isolates capable of degrading 2,4-D. Microb. Ecol. 30, 3–24.
[32] Calabrese, G., Forney, L., Suwa, Y. and Holben, W. (1994) Localization and characterization of a novel chromosomal 2,4-D catabolic gene isofunctional to *tfdA*. Abstr. 94th Gen. Meet. Am. Soc. Microbiol. Q-405 p. 459. American Society for Microbiology, Washington, DC, USA.
[33] Matheson, V.G., Forney, L.J., Suwa, Y., Nakatsu, C.H., Sextone, A.J. and Holben, W.E. (1996) Cloning and sequence analysis of a chromosomal 2,4-D catabolic gene isofunctional to *tfdA*. Appl. Environ. Microbiol., in press.
[34] Fulthorpe, R.R., Rhodes, A.N. and Tiedje, J.M. (1996) Pris-

COR-AAD1-00001197

Downloaded from https://academic.oup.com/femsec/article/20/2/163/553978 by guest on 16 September 2021

tine soils mineralize 3-chlorobenzoate and 2,4-dichlorophenoxyacetate via different microbial populations. Appl. Environ. Microbiol. 62, 1159–1166.
[35] Fournier, J.C. (1989) Aspects du comportement de la microflore dégradant les produits phytosanitaires dans le sol. Thèse de Doctorat ès Sciences Naturelles. Université de Perpignan, France.
[36] Gunalan (1991) Etude du comportement dans le sol d'un inoculum bactérien dégradant le 2,4-D. Thèse de Doctorat. Université Claude Bernard, Lyon I, France.

COR-AAD1-00001198

Downloaded from https://academic.oup.com/femsec/article/20/3/163/547858 by Pennsylvania State University, Inc. user on 03 September 2020

ELSEVIER

Gene 255 (2000) 281–288

GENE
AN INTERNATIONAL JOURNAL ON GENES, GENOMES AND EVOLUTION

www.elsevier.com/locate/gene

Analysis of the 2,4-dichlorophenoxyacetic acid-degradative plasmid pEST4011 of *Achromobacter xylosoxidans* subsp. *denitrificans* strain EST4002

Eve Vedler *, Viia Kõiv, Ain Heinaru

Institute of Molecular and Cell Biology, Tartu University, 23 Riia Street, 51010 Tartu, Estonia

Received 5 January 2000; received in revised form 7 May 2000; accepted 17 July 2000
Received by D.L. Court

Abstract

The 2,4-dichlorophenoxyacetic acid (2,4-D)-degradative bacterium *Achromobacter xylosoxidans* subsp. *denitrificans* strain EST4002, isolated in Estonia more than 10 years ago, was found to contain the 70 kb plasmid pEST4011 that is responsible for the bacterium having had obtained a stable 2,4-D$^+$ phenotype. The *tfd*-like genes for 2,4-D degradation of the strain EST4002 were located on a 10.5 kb region of pEST4011, but without functional genes coding for chloromuconate cycloisomerase and chlorodienelactone hydrolase. The latter two genes are probably encoded by homologous, *tcb*-like genes, located elsewhere on pEST4011. We also present evidence of two copies of insertion element IS*1071*-like sequences on pEST4011. IS*1071* is a class II (Tn*3* family) insertion element, associated with different catabolic genes and operons and globally distributed in the recent past. We speculate that this insertion element might have had a role in the formation of plasmid pEST4011. The 28 kb plasmid pEST4012 is generated by deletion from pEST4011 when cells of *A. xylosoxidans* EST4002 are grown in the absence of 2,4-D in growth medium. We propose that this is the result of homologous recombination between the two putative copies of IS*1071*-like sequences on pEST4011. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Chloroaromatic catabolism; Genes coding for enzymes of 2,4-dichlorophenoxyacetic acid degradation; Sequence analysis

1. Introduction

Many microorganisms are able to degrade the widely used herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) (Loos et al., 1979; Fournier, 1980; Don and Pemberton, 1981; Ou, 1984; Ka and Tiedje, 1994; Fulthorpe et al., 1995; Maltseva et al., 1996). The 2,4-D degraders that belong to the β and γ subdivisions of *Proteobacteria* contain *tfd*-like (pJP4-like) genes for 2,4-D degradation, often encoded on transmissible plasmids and horizontally spread among the representatives of these two subdivisions (Ka et al., 1994a,b,c; Top et al., 1995). Plasmid pJP4 from the well-characterized bacterium *Ralstonia eutropha* (formerly *Alcaligenes eutrophus*) JMP134 (Don and Pemberton, 1985; Streber et al., 1987; Kaphammer et al., 1990; Perkins et al., 1990; Matrubutham and Harker, 1994; Leveau and van der Meer, 1996; Leveau et al., 1998), originally isolated in Australia, carries a 22 kb DNA region encoding *tfd* genes responsible for converting 2,4-D to a common intermediate metabolite 3-oxoadipate (Figs. 1 and 2). This DNA region has a mosaic-like structure and has apparently been composed of DNA fragments of different bacterial origins (Fulthorpe et al., 1995). The catabolic genes of pJP4 — *tfdA*, *tfdB* and *tfdCDEF*, as well as the module 2 genes $tfdD_{II}C_{II}E_{II}F_{II}B_{II}$ and *tfdK* — are all equally expressed in the presence of 2,4-D (Leveau et al., 1999). There is strong homology between the *tfdCDEF* 2,4-dichlorocatechol degradative operon and other chlorocatechol degradative operons — *tcbCDEF* from 1,2,4-trichlorobenzene degradative plasmid pP51 of *Pseudomonas* sp. strain P51 (van der Meer

Abbreviations: A, adenosine; bp, base pair(s); C, cytosine; 2,4-D, 2,4-dichlorophenoxyacetic acid; G, guanosine; IR, inverted repeat sequence; IS, insertion sequence; kb, kilobase(s) or 1000 bp; LB, Luria–Bertani (medium); MAR, maleylacetate reductase; ORF, open reading frame; PCR, polymerase chain reaction; RBS, ribosome-binding site(s); rRNA, ribosomal RNA; S, sedimentation constant; SDS, sodium dodecyl sulfate; SSC, 0.15 M NaCl/0.015 M Na_3·citrate pH 7.6; T, thymidine.

* Corresponding author. Tel.: +3727-375014; fax: +3727-420286.
E-mail address: evedler@ebc.ee (E. Vedler)

0378-1119/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0378-1119(00)00329-2

COR-AAD1-00001199

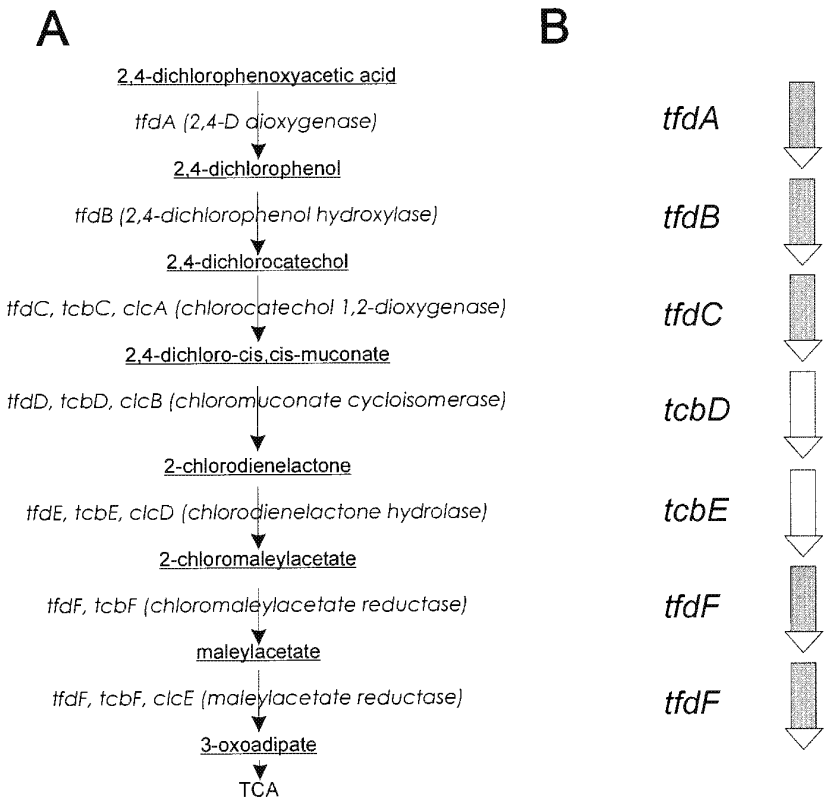


Fig. 1. (**A**) Degradation pathway of 2,4-D along with homologous genes coding for enzymes of different chlorocatechol degradation pathways. (**B**) Putative 2,4-D degradation pathway coded by the genes of plasmid pEST4011.

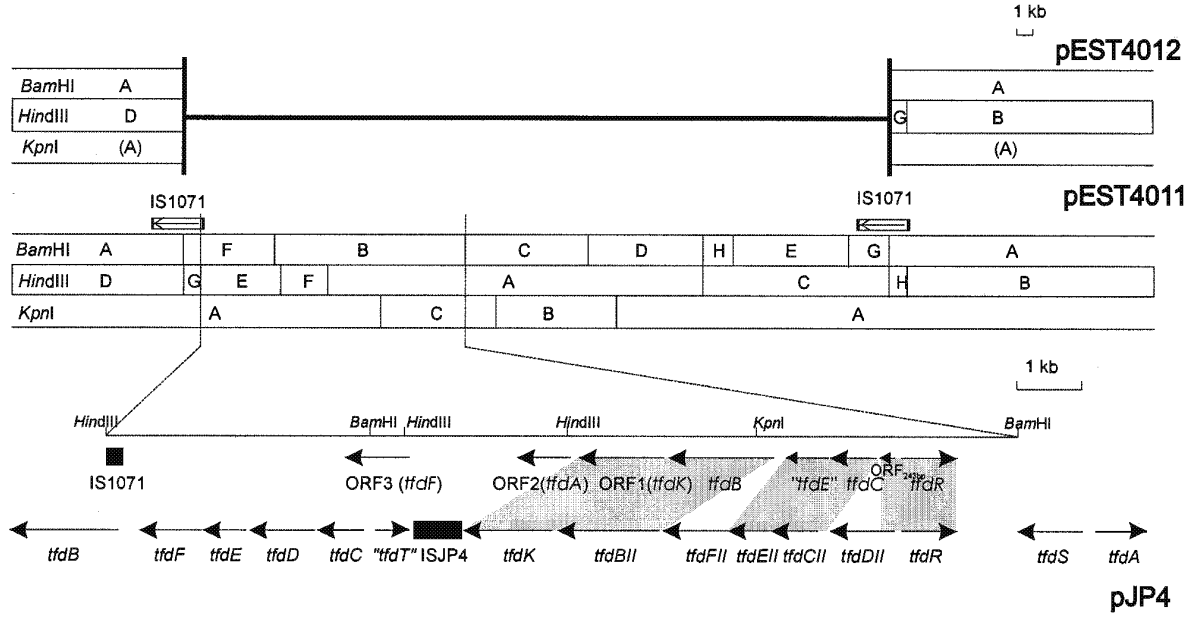


Fig. 2. Restriction map of plasmids pEST4011 and pEST4012 and putative positions of IS elements on pEST4011. Organization of *Bam*HI-B and *Hin*dIII-E fragments of pEST4011 and the region of pJP4 encoding *tfd* genes (Leveau and van der Meer, 1996; GenBank accession number U16782; see Sections 1 and 3.2). Directions of transcription are shown for each gene and ORF. The truncated *tfdD* gene of pEST4011 is designated ORF_{243bp}. The gray regions represent significant homology between plasmids pEST4011 and pJP4.

COR-AAD1-00001200

et al., 1991) and *clcABDE* from 3-chlorobenzoate degradative plasmid pAC27 of *Pseudomonas putida* (Ghosal and You, 1989) (Fig. 1).

The stable 2,4-D degradative plasmid pEST4011 (70 kb) is a natural deletant plasmid generated from unstable pEST4002 (78 kb) of the bacterium (originally named *Pseudomonas* sp. EST4002) isolated in Estonia (Ausmees and Heinaru, 1990). In this study we compared the 2,4-D degradative strain EST4002, which now was found to contain the stable plasmid pEST4011, with databank strains by using Biolog identification and 16S rRNA gene sequencing.

The pEST4011 genes *tfdR*, *tfdC* and *tfdB* have been sequenced and it has been shown that TfdR is a LysR-type transcriptional activator, its effector molecule is 2,4-dichloro-*cis*,*cis*-muconate and it activates the expression of the divergently transcribed *tfdCB* operon (Kõiv et al., 1996; Vedler et al., 2000) (Fig. 2). In this study we finished determination of the nucleotide sequence of the pEST4011 region, which had been shown to hybridize to pJP4 *tfd* genes (Mäe et al., 1993). It was shown that a 10.5 kb region of pEST4011 contains almost the whole pathway of *tfd*-like genes for 2,4-D degradation (Figs. 1 and 2), but there are no functional *tfd*-like genes coding for chloromuconate cycloisomerase (*tfdD*) and chlorodienelactone hydrolase (*tfdE*) either on plasmid pEST4011 or chromosome of the strain EST4002. Instead, these two enzymes are probably encoded by homologous, *tcb*-like (pP51-like) genes (*tcbD* and *tcbE*), located on pEST4011. Also found were sequences on pEST4011 similar to the insertion element IS*1071*, associated with catabolic genes or operons from different geographically distinct isolates (Di Gioia et al., 1998).

2. Material and methods

2.1. Bacterial strains and plasmids used, media and culture conditions

The bacterial strains used in this study were *Escherichia coli* DH5α (BLR), *Achromobacter xylosoxidans* subsp. *denitrificans* EST4002 [pEST4011] with 2,4-D$^+$ phenotype and *A. xylosoxidans* subsp. *denitrificans* EST4003 [pEST4012] with 2,4-D$^-$ phenotype. Vector plasmid pBluescript(SK) (Stratagene) was used to clone pEST4011 *Bam*HI-B and *Hin*dIII-E fragments for sequence analysis. *E. coli* and *A. xylosoxidans* were grown in LB at 37°C and 30°C respectively.

2.2. DNA manipulations and sequence analysis

Plasmid and chromosomal DNA preparations and other DNA manipulations were carried out by established procedures (Sanger et al., 1977; Sambrook et al., 1989). 16S rRNA gene of EST4002 was PCR-amplified using the forward primer PCRI (5′ AGAGTTTGATCATGGCTCAG 3′), complementary to *E. coli* 16S rRNA gene positions 6 to 26, and the reverse primer PCRII (5′ TACGGTTACCTTGTTACGACTT 3′), complementary to *E. coli* 16S rRNA gene positions 1513 to 1492. The PCR product (about 1.5 kb) was sequenced with T7 Sequenase v2.0 PCR product sequencing kit (Amersham Life Science, Inc. 1997), using primers PCRI, PCRII and the internal reverse primers SEQ1 (5′ GTATTACCGCGGCTGCTGG 3′), SEQ2 (5′ TTGCGCTCGTTGCGGGACT 3′), SEQ3 (5′ ACGGGCGGTGTGTACAAG 3′) and S4 (5′ CCAGGGTATCTAATCC 3′), complementary to conserved regions of different eubacterial 16S rRNA genes. Fragments of putative chloromuconate cycloisomerase and chlorodienelactone hydrolase were respectively PCR-amplified using primer pairs TCBD1 (5′ GGCACAATCGGTTCAAGGT 3′)–TCBD2 (5′ TACAACGTCTCGAACTCGC 3′), complementary to regions in *Variovorax paradoxus* plasmid pTV1 putative *tfdD* gene conserved between different cycloisomerase genes, and TCBE1 (5′ GTCGCGGTGATCGGCTATTG 3′)–TCBE2 (5′ CGAACGAGTGCCCGGCGTC 3′), complementary to regions in *Pseudomonas* sp. strain P51 plasmid pP51 *tcbE* gene conserved between different hydrolase genes, and the plasmid pEST4011 DNA as a template. The fragment of putative chloromuconate cycloisomerase gene was sequenced with Taq DNA Sequencing Kit for Cycle Sequencing (Boehringer Mannheim GmbH), using the same radioactively labeled primers that were used in PCR amplification, and TCBD PCR product as a template. T4 polynucleotide kinase (Fermentas) and [γ-^{32}P]dATP were used to end-label primers TCBD1 and TCBD2 for sequence analysis and primers HINDE (5′ CAATGAAATCAACGGTCTA 3′), TCBD2 and TCBE2 for Southern hybridization analyses. Computer analysis of DNA and amino acid sequences was performed with the program PC/GENE (Genofit, Geneva, Switzerland). Database searches were performed with the BLASTN and BLASTX programs at the National Center for Biotechnology Information (NCBI).

2.3. Southern hybridization analyses

Plasmids pEST4011, pEST4012 and the chromosome of strain EST4002 were digested with restrictional endonuclease *Bam*HI (Fermentas) overnight at 37°C, fractionated by 0.7% agarose gel electrophoresis and transferred to a Hybond-N+ nylon filter (Amersham). When the 191 bp *Nde*I–*Stu*I fragment of the truncated *tfdD* sequence and the 305 bp *Pfl*MI–*Pvu*I fragment of dysfunctional *tfdE* sequence of pEST4011 *Bam*HI-B fragment were used as the probes, filters were hybridized overnight at +65°C in 6× SSC, 0.5% SDS, 5× Denhardt's and 200 μg of denatured Salmon sperm

COR-AAD1-00001201

DNA per milliliter. Probes were labeled with [α-^{32}P]dCTP by random priming (Amersham). After hybridization, membranes were washed in standard washes (6× SSC–0.1% SDS, 2× SSC–0.1% SDS, 1× SSC–0.1% SDS at +65°C for 30 min). When the end-labeled primers HINDE, TCBD2 and TCBE2 were used as the probes, filters were hybridized for 2 h at +50°C in 6× SSC, 0.5% SDS, 5× Denhardt's and 200 μg of denatured Salmon sperm DNA per milliliter. After hybridization, membranes were washed twice in 4× SSC–0.1% SDS at +37°C for 10 min. Images were made with a PhosphorImager and Image Quant software (Molecular Dynamics, Sunnywale, CA).

3. Results and discussion

3.1. Isolation of pEST4011 from strain EST4002 and identification of this 2,4-D degradative strain

Ausmees and Heinaru (1990) reported on the isolation of a 2,4-D degradative bacterium, containing a 78 kb plasmid pEST4002, in Estonia. This strain had a very unstable 2,4-D$^+$ phenotype: 80% of the cells had lost their ability to grow on 2,4-D after growth in a non-selective LB medium for 4 h. To overcome this instability and difficulties with DNA extraction from the wild-type strain, pEST4002 was transferred by conjugation into the plasmid-free recipient strain *P. putida* PaW340 (Fig. 3). The achieved 2,4-D$^+$ transconjugants of PaW340 contained a 70 kb plasmid pEST4011. Subsequently, *P. putida* strain PaW85 was transformed with pEST4011 and a stable 2,4-D$^+$ transformant (pEST4011) was obtained and named EST4021 (Mäe et al., 1993). The original 2,4-D degradative strain EST4002 was maintained on minimal medium agar plates containing 0.1% 2,4-D. In 1999, we found that this strain contained a plasmid that had the same size and restrictional pattern as pEST4011 did, and this plasmid was also named pEST4011. We found that pEST4011 is very stable in EST4002: growth of EST4002 in LB medium for 24 h resulted only in 1–2% cells with irreversible 2,4-D$^-$ phenotype, which all contained the same 28 kb deletant plasmid. The latter strain was named EST4003. The corresponding deletant plasmid pEST4012 was also achieved by Mäe et al. (1993) from 2,4-D$^-$ *P. putida* strain EST4022 when they cultivated *P. putida* EST4021 [pEST4011] cells in LB for 30 generations. The correlated restrictional maps of pEST4011 and pEST4012 are given in Fig. 2.

The 2,4-D degradative strain EST4002 was originally named *Pseudomonas* sp. (Mäe et al., 1993). In this study this strain was analyzed at a physiological level by microtiter plate-bound substrate utilization assay (Biolog; Biolog Inc., Hayward, CA) and at the phylogenetic level by sequencing and comparison of the nearly complete PCR-amplified 16S rRNA gene. The closest relative by Biolog identification (readings were taken after 24 h) appeared to be *A. xylosoxidans* subsp. *denitrificans* (formerly *Alcaligenes xylosoxidans* subsp. *denitrificans*; belongs to the β subdivision of *Proteobacteria*) with similarity coefficient 0.951 (1.0 indicates complete homology to a databank species). The comparison of the 16S rRNA gene of strain EST4002 with databank strains supported the result of Biolog identification, as the most relative strain appeared to be *A. xylosoxidans* subsp. *denitrificans* (GenBank accession number M22509) (only 12 mismatches between these two genes). In this alignment, 21 unspecified nucleotides of the GenBank sequence M22509 were not used. According

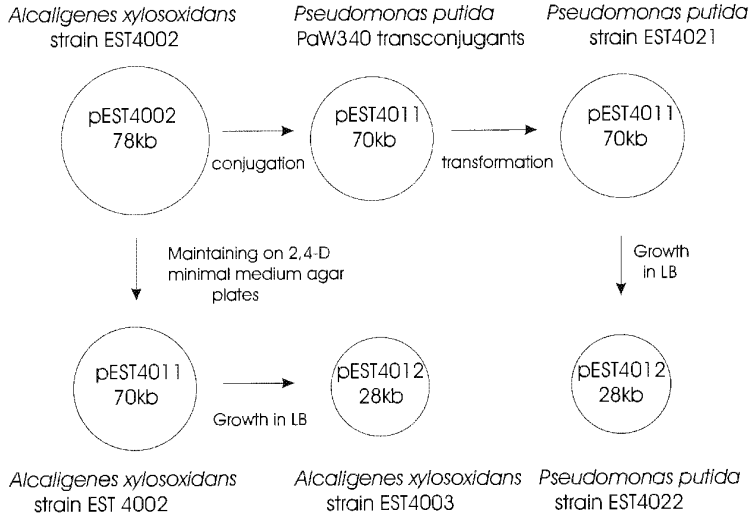


Fig. 3. Schematic presentation of different strains and plasmids used and discussed in this study.

COR-AAD1-00001202

to the results given above, we named our 2,4-D degradative strain *A. xylosoxidans* subsp. *denitrificans* EST4002.

3.2. Sequencing and databank comparison of 2,4-D degradation genes of pEST4011

We have previously shown (Mäe et al., 1993) that different fragments of pJP4 containing at least parts of the genes *tfdA*, *tfdS*, *tfdR*, $tfdD_{II}$, *tfdT*, *tfdCDEF*, *tfdB* and *ISJP4* hybridized only with the 10.8 kb *Bam*HI-B restriction fragment of pEST4011 (Fig. 2). Sequence analysis of the 5 kb *Xho*I–*Stu*I fragment of this pEST4011 region and expression studies (Kõiv et al., 1996; Vedler et al., 2000) revealed three functional genes necessary for 2,4-D degradation: *tfdR*, *tfdC* and *tfdB*. The latter two genes form one operon transcribed from *tfdC* to *tfdB* and they code for chlorocatechol 1,2-dioxygenase and 2,4-dichlorophenol hydroxylase respectively. The *tfdR* gene is located upstream and transcribed divergently from the *tfdCB* operon, and its product, a LysR-type transcriptional activator, activates the expression of the *tfdCB* operon. *tfdR*, *tfdC* and *tfdB* of pEST4011 were shown to be the most similar to the pJP4 module 2 genes *tfdR*, $tfdC_{II}$ and $tfdB_{II}$. Although the pEST4011 nucleotide sequence between *tfdC* and *tfdB* is very similar to the pJP4 module 2 $tfdE_{II}$ gene, deletion of GC dinucleotide (corresponding to the positions 73 and 74 in the pJP4 $tfdE_{II}$ gene) in the coding region leads to frameshifting and no functional protein could be translated from this region (GenBank accession number U32188). We also found a 243 bp nucleotide sequence in the *tfdR–tfdC* intergenic region similar to the beginning of pJP4 $tfdD_{II}$ gene, which encodes chloromuconate cycloisomerase. This region was designated ORF_{243bp}.We have shown that ORF_{243bp} is the most upstream part of the *tfdCB* operon, but it is not a translated ORF since there is no RBS in front of it because the first nucleotide in its ATG translation initiation codon is the transcription start site of the *tfdCB* operon (Vedler et al., 2000).

In order to find all genes necessary for 2,4-D degradation we finished sequencing the 10.8 kb *Bam*HI-B and the 5 kb *Hin*dIII-E restriction fragments of pEST4011 (Fig. 2). We found three additional ORFs downstream of the *tfdB* gene, all preceded by potential RBS (AGGAG in all cases). ORF1 was found 99 bp downstream of *tfdB*, it was 1392 bp long and its amino acid sequence was most similar to pJP4 TfdK (73% identity over the 456 amino-terminal amino acids). It has been shown that the pJP4 *tfdK* gene product facilitates uptake of 2,4-D by *R. eutropha* JMP134 (Leveau et al., 1998). The predicted ORF1 polypeptide bears all the putative transmembrane domains and motifs, which is characteristic for the members of the major facilitator superfamily of transporter proteins (including the pJP4 TfdK) (Jessen-Marshall et al., 1995). Thus we speculate that the pEST4011 TfdK may be a functional transporter of 2,4-D. ORF2 was found 149 bp downstream of ORF1; it was 864 bp long and its amino acid sequence was similar to pJP4 TfdA, the α-ketoglutarate-dependent dioxygenase (Fukumori and Hausinger, 1993) (81% identity). As we have shown that 2,4-D dioxygenase (TfdA) activity could be detected in *P. putida* PaW85 cells carrying the 10.8 kb *Bam*HI fragment of pEST4011 cloned into the vector plasmid pKT240 (Mäe et al., 1993), we assume that ORF2 codes for an active 2,4-D dioxygenase. ORF3 was found 1811 bp downstream of ORF2; it was 1074 bp long and its amino acid sequence was most similar to pJP4 TfdF, the maleylacetate reductase (MAR) (Kasberg et al, 1995) (48% identity over the 352 amino-terminal amino acids). Alignment of the deduced ORF3 polypeptide with different MARs showed that ORF3 contains all patterns typical for this enzyme (Seibert et al., 1998). These results propose that ORF3 may be a functional MAR.

Alignment results, obtained previously (Kõiv et al., 1996; Vedler et al., 2000) and in this study, show that the *tfd* genes, present on the 10.5 kb region of pEST4011, obviously comprise genes of different origins. The genes *tfdR* to *tfdK* of pEST4011 (Fig. 2) are very similar to, and organized in the same way as, the pJP4 module 2 genes *tfdR* to *tfdK* (GenBank accession number U16782). Thus it is very likely that these regions of pEST4011 and pJP4 have evolved from the same origin. In pJP4, it has been suggested that the module 2 *tfd* genes have been incorporated into an early version of this plasmid as one unit by two copies of the insertion element IS*JP4* (Leveau and van der Meer, 1997). In pEST4011, the regions corresponding to the pJP4 $tfdF_{II}$ gene and the downstream part of the $tfdD_{II}$ gene, as well as the GC dinucleotide in the middle of the $tfdE_{II}$ gene, have been lost during evolution.

McGowan et al. (1998) have sequenced internal conserved regions of *tfdA* genes of phylogenetically different 2,4-D degraders. When we aligned the corresponding region of pEST4011 *tfdA* (bases 453–765) to these sequences, we found that this fragment was identical to class III *tfdA* sequences that have been found in phylogenetically distant 2,4-D degraders of *Comamonas–Rhodoferax* group (β-*Proteobacteria*) and *Halomonas* group (γ-*Proteobacteria*). The pJP4 *tfdA* gene belongs to class I sequences that have been found in *Burkholderia*, *Comamonas–Rhodoferax* and *Alcaligenes–Ralstonia* groups (all β-*Proteobacteria*). Considering these results, we presume that the pEST4011 *tfdA* gene has a different origin than the pJP4 *tfdA* gene does, and this gene has been recruited into a precursor plasmid of pEST4011 independently from other *tfd* genes.

While the (putative) products of the genes *tfdA*, *tfdB*, *tfdC*, '*tfdE*' (frameshift eliminated for alignment) and *tfdK* of pEST4011 are about 75–90% similar to corre-

COR-AAD1-00001203

sponding proteins coded by pJP4 module 2 genes, the deduced amino acid sequence of pEST4011 *tfdF* gene is only about 48% identical to pJP4 module 1 *tfdF* gene product and about 44% identical to module 2 $tfdF_{II}$ gene product. With each other, the known MARs share about 45–60% identity, thus being more heterogeneous than other enzymes for degradation of chlorocatechols. The pEST4011 *tfdF* gene has obviously evolved from an origin not shared with other known genes coding for MARs and has apparently been independently incorporated into a precursor plasmid of pEST4011.

3.3. Detection of two copies of the insertion element IS1071-like sequences on plasmid pEST4011

About 3.7 kb downstream of ORF3 we found a 280 bp sequence (we sequenced this region only to the *Hin*dIII restriction site) (Fig. 2) that was 100% identical at the nucleotide level to IS*1071* region containing the 110 bp inverted repeat sequence (IR) and the beginning of the transposase gene. IS*1071* is a unique class II (Tn*3* family) insertion sequence: it is 3.2 kb long and contains one ORF of 2.91 kb coding for a transposase protein (TnpA) (Nakatsu et al., 1991). It has been shown that this insertion element has an important role in the mobilization of different catabolic genes and operons into composite transposon structures on plasmids and it has been distributed globally in a number of host genera in the recent past (Di Gioia et al., 1998).

In order to find out whether there may be more copies of IS*1071* on pEST4011, we performed Southern hybridization using the radioactively labeled oligonucleotide HINDE, complementary to a part of the 110 bp IR of IS*1071*, as the probe. The analysis showed that this oligonucleotide hybridized to three different *Bam*HI restriction fragments of pEST4011, namely *Bam*HI-A, *Bam*HI-F and *Bam*HI-G, and to the *Bam*HI-A restriction fragment of pEST4012 (Fig. 2). The nucleotide sequence of IS*1071* has one *Bam*HI site and two *Hin*dIII sites (at a distance of 1.1 kb) (GenBank accession number M65135). The Southern hybridization result proposes that there may be two copies of IS*1071* on pEST4011, the IRs of one element being on *Bam*HI-A and *Bam*HI-F, and the IRs of the other element being on *Bam*HI-A and *Bam*HI-G restriction fragments. This result is supported by the presence of two 1.1 kb *Hin*dIII restriction fragments on pEST4011 at the same positions where the two elements might be according to the Southern analysis. Although these results do not show whether these two putative IS elements of pEST4011 are complete and functional transposable elements, forming a composite transposon, we can speculate that they have had a role in the formation of the 2,4-D degradative plasmid pEST4011. As the Southern analysis showed that there (is) are (one) two IRs of IS*1071* on the *Bam*HI-A fragment of the deletant plasmid pEST4012 and that pEST4012 actually consists only, or mainly, of this *Bam*HI-A fragment (Fig. 2), there is a possibility that the formation of pEST4012 in the absence of 2,4-D in the growth medium may be the result of homologous recombination between the two putative direct copies of this IS element on pEST4011. If this is the case, the 2,4-D degradation genes of pEST4011 must be coded on the 39 kb region between the IS*1071*-like elements (Fig. 2).

3.4. Determination of genes coding for chloromuconate cycloisomerase and chlorodienelactone hydrolase

As described in Section 3.2, all pJP4 *tfd* genes hybridized only with pEST4011 *Bam*HI-B fragment (Fig. 2). Sequence analysis of this region revealed that there are no functional genes coding for chloromuconate cycloisomerase (*tfdD*) and chlorodienelactone hydrolase (*tfdE*) either on the *Bam*HI-B or the *Hin*dIII-E fragment. In order to find out whether the other copies of *tfdD* and/or *tfdE* genes, similar to the ones on pEST4011 *Bam*HI-B fragment, might be on pEST4011 or chromosome of EST4002, we undertook Southern hybridization using the 191 bp *Nde*I–*Stu*I fragment of the truncated *tfdD* sequence, ORF_{243bp}, in front of the *tfdC* gene of pEST4011 and the 305 bp *Pfl*MI–*Pvu*I fragment of the dysfunctional *tfdE* sequence of pEST4011 as the probes. These studies showed that the two probes did not hybridize to any *Bam*HI fragment of pEST4011 or the chromosome of EST4002 other than to the *Bam*HI-B fragment of pEST4011. As isolation of EST4002 chromosome also produced small amounts of plasmid DNA, in the case of hybridization with EST4002 chromosomal DNA only the fragment corresponding to the plasmid hybridized. Thus it seems very unlikely that there are functional genes for chloromuconate cycloisomerase and chlorodienelactone hydrolase in the strain EST4002, either on plasmid or chromosome, the nucleotide sequences of which are most similar to the corresponding genes of the *tfd* pathway.

Database searches done with pEST4011 *tfd* genes revealed that the published nucleotide sequence of *V. paradoxus* plasmid pTV1 (GenBank accession number AB028643) is nearly 100% identical to the corresponding sequence of pEST4011. 2.6 kb downstream of pTV1 *tfdR* there is an ORF coding for putative chloromuconate cycloisomerase (*tfdD*), the nucleotide sequence of which is most similar to the *tcbD* gene of *Pseudomonas* sp. strain P51 plasmid pP51 (81% over 769 bp). This finding led us to an idea to design a pair of oligonucleotides for both putative cycloisomerase and hydrolase genes of pEST4011 — TCBD1 and TCBD2, complementary to regions in pTV1 putative *tfdD* gene conserved between different cycloisomerase genes, and TCBE1 and TCBE2, complementary to regions in pP51 *tcbE* gene conserved between different hydrolase genes. Southern

COR-AAD1-00001204

hybridization, done with radioactively labeled TCBD2 and TCBE2 primers, showed that both primers hybridized to *Bam*HI-C, *Bam*HI-D and/or *Bam*HI-E fragments of pEST4011 (*Bam*HI-D and *Bam*HI-E fragments are indistinguishable by gel electrophoresis as they are of the same size). TCBE2 primer hybridized very weakly to the pEST4011 *Bam*HI-B fragment as well. We performed PCR amplification with these two pairs of primers using the plasmid pEST4011 as a template. The results showed a PCR product of about 300 bp (as expected) only with TCBD primers. Sequence analysis of the TCBD PCR product (293 bp of it was sequenced) showed that it was 100% identical to the corresponding fragment of *V. paradoxus* pTV1 *tfdD* gene and 80% identical to the *tcbD* gene of *Pseudomonas* sp. strain P51 plasmid pP51 at the nucleotide level. Amino acid sequence (97 amino acids) of TCBD PCR product was 85% identical to TcbD of pP51 and CbnB of *R. eutropha* NH9 plasmid pENH91 (Ogawa and Miyashita, 1999), 75% identical to ClcB of *P. putida* plasmid pAC27, 59% identical to TfdD of pJP4 and only 43% identical (over 66 amino acids) to $TfdD_{II}$ of pJP4. The TCBD PCR product must have been amplified from putative *tcb*-like gene(s) somewhere on the unsequenced region of pEST4011, because the region that could bind the TCBD primers is not present in ORF_{243bp} of the pEST4011 *Bam*HI-B fragment. As the sequenced TCBD PCR product was amplified from the whole plasmid pEST4011 as a template, it shows that this is the only variant of this type of gene on pEST4011, and according to the Southern analysis there is probably at least one copy of this gene on pEST4011 *Bam*HI-C, *Bam*HI-D (and/or *Bam*HI-E) fragments. In the case of TCBE primers, we expected that the product would be amplified from putative *tcb*-like gene(s), and not from the dysfunctional *tfdE* sequence of the *Bam*HI-B fragment. Nevertheless, TCBE primers were obviously not complementary enough to anneal with (any) putative chlorodienelactone hydrolase gene(s) on pEST4011. As the Southern analysis showed that TCBE2 primer hybridized much more strongly with pEST4011 *Bam*HI-C and *Bam*HI-D/E fragments than with *Bam*HI-B fragment, we believe that there are at least two copies of a gene coding for chlorodienelactone hydrolase on pEST4011 which are more similar to *tcbE* than to *tfdE* gene.

The 39 kb region of pEST4011 between the two putative IS*1071*-like elements contains genes of different origins which all serve for 2,4-D degradation. All these genes are deleted from pEST4011 when cells, harboring this plasmid, are grown in the absence of 2,4-D in growth medium. We have sequenced only about 16 kb of this region on which a part of the 2,4-D degradation pathway is coded by the *tfd*-like genes. In this study we have found evidence that the remaining region of about 23 kb does not contain any *tfd*-like genes and that genes, more similar to *tcb*-like sequences, have been recruited into pEST4011 to code for at least two enzymes of the 2,4-D degradation pathway — chloromuconate cycloisomerase and chlorodienelactone hydrolase (Fig. 1). It will be interesting to finish analysis of the whole 39 kb region of pEST4011 in order to determine all genes present on that fragment.

The nucleotide sequence of EST4002 16S rRNA gene has been assigned GenBank accession number AF232712. The nucleotide sequences of the plasmid pEST4011, discussed and partly sequenced in this study, have been assigned GenBank accession numbers U32188 (the genes *tfdR*, *tfdC*, '*tfdE*', *tfdB*, *tfdK*, *tfdA* and *tfdF*), AF233095 (a fragment of putative element IS*1071*) and AF233096 (a fragment of putative gene for chloromuconate cycloisomerase).

4. Conclusions

(1) The 2,4-D degradative bacterium, isolated in Estonia more than 10 years ago, is *A. xylosoxidans* subsp. *denitrificans* strain EST4002. Maintaining this strain on minimal medium agar plates containing 2,4-D for several years has resulted in a much more stable $2,4\text{-}D^{+}$ phenotype due to the deletion of about 8 kb fragment of pEST4002 which gave rise to the plasmid pEST4011.

(2) A 10.5 kb region of pEST4011 contains almost the whole pathway of *tfd*-like genes for 2,4-D degradation without functional genes for chloromuconate cycloisomerase and chlorodienelactone hydrolase. The latter two genes are obviously coded by *tcb*-like genes located elsewhere on pEST4011.

(3) The two putative copies of insertion element IS*1071*-like sequences could have had a role in the formation of a precursor plasmid of the 2,4-D degradative plasmid pEST4011. There is also a possibility that the generation of the deletant plasmid pEST4012 in the absence of 2,4-D in the growth medium is the result of homologous recombination between the two IS*1071*-like sequences.

Acknowledgement

This work was supported by the Institute of Molecular and Cell Biology, by grant 2334 from the Estonian Science Foundation.

References

Ausmees, N.R., Heinaru, A.L., 1990. New plasmids of herbicide 2,4-D degradation. Genetika 26, 770–772. (in Russian).

Di Gioia, D., Peel, M., Fava, F., Campbell, R.C., 1998. Structures of homologous composite transposons carrying *cbaABC* genes from

COR-AAD1-00001205

Europe and North America. Appl. Environ. Microbiol. 64, 1940–1946.

Don, R.H., Pemberton, J.M., 1981. Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J. Bacteriol. 145, 681–686.

Don, R.H., Pemberton, J.M., 1985. Genetics and physical map of the 2,4-dichlorophenoxyacetic acid degratative plasmid pJP4. J. Bacteriol. 161, 466–468.

Fournier, J.C., 1980. Enumeration of the soil microorganisms able to degrade 2,4-D by metabolism or cometabolism. Chemosphere 9, 169–174.

Fukumori, F., Hausinger, R.P., 1993. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetate monooxygenase" is an α-ketoglutarate-dependent dioxygenase. J. Bacteriol. 175, 2083–2086.

Fulthorpe, R.R., McGowan, C., Maltseva, O.V., Holben, W.E., Tiedje, J.M., 1995. 2,4-Dichlorophenoxyacetic acid degrading bacteria contain mosaics of catabolic genes. Appl. Environ. Microbiol. 61, 3274–3281.

Ghosal, D., You, I.-S., 1989. Operon structure and nucleotide homology of the chlorocatechol oxidation genes of plasmids pJP4 and pAC27. Gene 83, 225–232.

Jessen-Marshall, A.E., Paul, N.J., Brooker, R.J., 1995. The conserved motif, GXXX(D/E)(R/K)XG[X](R/K)(R/K), in hydrophilic loop 2/3 of the lactose permease. J. Biol. Chem. 270, 16 251–16 257.

Ka, J.O., Tiedje, J.M., 1994. Integration and excision of a 2,4-dichlorophenoxyacetic acid-degradative plasmid in *Alcaligenes paradoxus* and evidence of its natural intergenic transfer. J. Bacteriol. 176, 5284–5289.

Ka, J.O., Holben, W.E., Tiedje, J.M., 1994a. Analysis of competition in soil among 2,4-dichlorophenoxyacetic acid-degrading bacteria. Microbiology 60, 1121–1128.

Ka, J.O., Holben, W.E., Tiedje, J.M., 1994b. Use of gene probes to aid in recovery and identification of functionally dominant 2,4-dichlorophenoxyacetic acid-degrading populations in soil. Appl. Environ. Microbiol. 60, 1116–1120.

Ka, J.O., Holben, W.E., Tiedje, J.M., 1994c. Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D-treated soils. Appl. Environ. Microbiol. 60, 1106–1115.

Kaphammer, B., Kukor, J.J., Olsen, R.H., 1990. Regulation of *tfdCDEF* by *tfdR* of the 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. J. Bacteriol. 172, 2280–2286.

Kasberg, T., Daubaras, D.L., Chakrabarty, A.M., Kinzelt, D., Reineke, W., 1995. Evidence that operons *tcb*, *tfd*, and *clc* encode maleylacetate reductase, the fourth enzyme of the modified *ortho* pathway. J. Bacteriol. 177, 3885–3889.

Kõiv, V., Marits, R., Heinaru, A., 1996. Sequence analysis of the 2,4-dichlorophenol hydroxylase gene *tfdB* and 3,5-dichlorocatechol 1,2-dioxygenase gene *tfdC* of 2,4-dichlorophenoxyacetic acid degrading plasmid pEST4011. Gene 174, 293–297.

Leveau, J.H.J., van der Meer, J.R., 1996. The *tfdR* gene can successfully take over the role of the insertion element-inactivated TfdT protein as a transcriptional activator of the *tfdCDEF* gene cluster, which encodes chlorocatechol degradation in *Ralstonia eutropha* JMP134(pJP4). J. Bacteriol. 178, 6824–6832.

Leveau, J.H., van der Meer, J.R., 1997. Characterization of insertion sequence ISJP4 on plasmid pJP4 from Ralstonia eutropha JMP134. Gene 202, 103–114.

Leveau, J.H.J., Zehnder, A.J.B., van der Meer, J.R., 1998. The *tfdK* gene product facilitates uptake of 2,4-dichlorophenoxyacetate by *Ralstonia eutropha* JMP134(pJP4). J. Bacteriol. 180, 2237–2243.

Leveau, J.H.J., König, F., Föchslin, H., Werlen, C., van der Meer, J.R., 1999. Dynamics of multigene expression during catabolic adaptation of *Ralstonia eutropha* JMP134 (pJP4) to the herbicide 2,4-dichlorophenoxyacetate. Mol. Microbiol. 33, 396–406.

Loos, M.A., Schlosser, I.F., Mapham, W.R., 1979. Phenoxy herbicide degradation in soils: quantitive studies of 2,4-D- and MCBA-degrading microbial populations. Soil Biol. Biochem. 11, 377–385.

Mäe, A., Marits, R., Ausmees, N., Kõiv, V., Heinaru, A., 1993. Characterization of a new 2,4-dichlorophenoxyacetic acid degrading plasmid pEST4011: physical map and localization of catabolic genes. J. Gen. Microbiol. 139, 3165–3170.

Maltseva, O., McGowan, C., Fulthorpe, R., Oriel, P., 1996. Degradation of 2,4-dichlorophenoxyacetic acid by haloalcalophilic bacteria. Microbiology 142, 1115–1122.

Matrubutham, U., Harker, A.R., 1994. Analysis of duplicated gene sequences associated with *tfdR* and *tfdS* in *Alcaligenes eutrophus* JMP134. J. Bacteriol. 176, 2348–2353.

McGowan, C., Fulthorpe, R., Wright, A., Tiedje, J.M., 1998. Evidence for interspecies gene transfer in the evolution of 2,4-dichlorophenoxyacetic acid degraders. Appl. Environ. Microbiol 64, 4089–4092.

Nakatsu, C., Ng, J., Singh, R., Straus, N., Wyndham, R.C., 1991. Chlorobenzoate catabolic transposon Tn*5271* is a composite class I element with flanking class II insertion sequences. Proc. Natl. Acad. Sci. U. S. A. 88, 8312–8316.

Ogawa, N., Miyashita, K., 1999. The chlorocatechol-catabolic transposon Tn*5707* of *Alcaligenes eutrophus* NH9, carrying a gene cluster highly homologous to that in the 1,2,4-trichlorobenzene-degrading bacterium *Pseudomonas* sp. strain P51, confers the ability to grow on 3-chlorobenzoate. Appl. Environ. Microbiol. 65, 724–731.

Ou, L.T., 1984. 2,4-D degradation and 2,4-D degradation microorganisms in soils. Soil Sci. 137, 100–107.

Perkins, E.J., Gordon, M.P., Caceres, O., Lurquin, P.F., 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. 172, 2351–2359.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: a Laboratory Manual. second edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467.

Seibert, V., Kourbatova, E.M., Golovleva, L.A., Schlömann, M., 1998. Characterization of the maleylacetate reductase MacA of *Rhodococcus opacus* 1CP and evidence for the presence of an isofunctional enzyme. J. Bacteriol. 180, 3503–3508.

Streber, W.R., Timmis, K.N., Zenk, M.H., 1987. Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. 169, 2950–2955.

Top, E.M., Holben, W.E., Forney, L.J., 1995. Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. Appl. Environ. Microbiol. 61, 1691–1698.

Van der Meer, J.R., Eggen, R.I.L., Zehnder, A.J.B., Devos, W.M., 1991. Sequence analysis of the *Pseudomonas* sp. strain P51 *tcb* gene cluster, which encodes metabolism of chlorinated catechols: evidence for specialization of catechol 1,2-dioxygenases for chlorinated catechols substrates. J. Bacteriol. 173, 2425–2434.

Vedler, E., Kõiv, V., Heinaru, A., 2000. TfdR, the LysR-type transcriptional activator, is responsible for the activation of the *tfdCB* operon of *Pseudomonas putida* 2,4-dichlorophenoxyacetic acid degradative plasmid pEST4011. Gene 245, 161–168.

COR-AAD1-00001206

ELSEVIER

Gene 245 (2000) 161–168

GENE
AN INTERNATIONAL JOURNAL ON GENES, GENOMES AND EVOLUTION

www.elsevier.com/locate/gene

TfdR, the LysR-type transcriptional activator, is responsible for the activation of the *tfdCB* operon of *Pseudomonas putida* 2,4-dichlorophenoxyacetic acid degradative plasmid pEST4011

Eve Vedler *, Viia Kõiv, Ain Heinaru

Institute of Molecular and Cell Biology, Tartu University, 23 Riia Street, 51010 Tartu, Estonia

Received 8 September 1999; received in revised form 27 December 1999; accepted 10 January 2000
Received by A. Nakazawa

Abstract

In *Pseudomonas putida* EST4021, the *tfdCB* operon of plasmid pEST4011 encodes enzymes involved in 2,4-dichlorophenoxyacetic acid degradation. We have identified a gene, *tfdR*, important for the regulation of the *tfdCB* operon. Sequence analysis of the *tfdR* gene revealed an open reading frame with amino acid sequence similar to the LysR family of transcriptional activators. The *tfdR* gene is located upstream and transcribed divergently from the *tfdCB* operon. Utilizing primer extension analysis, the transcription initiation sites of the gene *tfdR* and the *tfdCB* operon were localized 85 (84) bp and 292 bp upstream from the coding sequences of these genes, respectively. Multiple sequence analysis revealed that the genes *tfdR*, *tfdC* and *tfdB* of plasmid pEST4011 are most similar to the regulatory gene *tfdR* and the module 2 genes $tfdC_{II}$ and $tfdB_{II}$ of pJP4, respectively. The promoter–operator sequences of *tfdR* and its target *tfdCB* operon of pEST4011 have regions with highly conserved nucleotides characteristic for the catechol-subgroup LysR-type transcriptional activators. We showed that the pEST4011 *tfdR* gene product activates the expression of the *tfdCB* operon and the effector molecule for TfdR is 2,4-dichloro-*cis*,*cis*-muconate. Our data indicate that the structure and the mode of regulation of *tfd* genes are similar, despite the bacteria being isolated from different geographical regions. © 2000 Published by Elsevier Science B.V. All rights reserved.

Keywords: Chloroaromatic catabolism; Effector molecule; Regulation of gene expression; Sequence analysis; Transcription initiation

1. Introduction

Plasmid pJP4 from *Ralstonia eutropha* JMP134, originally isolated from Australia, has been well characterized and is a model to study the degradation of 2,4-dichlorophenoxyacetic acid (2,4-D) (Don et al., 1985; Leveau and van der Meer, 1996). This plasmid (Fig. 1) contains six catabolic genes *tfdA*, *tfdB* to *tfdF* of module 1, which code for enzymes that convert 2,4-D to a common intermediate metabolite 3-oxoadipate (Perkins et al., 1990; Streber et al., 1987), three regulatory genes *tfdS*, *tfdR* (Matrubutham and Harker, 1994; You and Ghosal, 1995) and *tfdT* (Leveau and van der Meer, 1996), *tfd* genes of module 2-$tfdB_{II}$ to $tfdD_{II}$, and *tfdK* gene whose product was shown to be an active transporter of 2,4-D (Leveau et al., 1998). It has been demonstrated that in the presence of 2,4-D, expression of *tfdA*, *tfdB*, *tfdCDEF* and *tfdK*, as well as the module 2 genes $tfdD_{II}C_{II}E_{II}F_{II}B_{II}$, occurs in concert from promoters upstream from the genes *tfdA*, *tfdC* and $tfdD_{II}$ (Leveau et al., 1999). The nucleotide sequences of *tfdS* and *tfdR* are identical and both code for TfdR. This regulatory protein is involved in the activation of pJP4 *tfd* genes (Kaphammer et al., 1990; Leveau et al., 1999; Leveau and van der Meer, 1996; You and Ghosal, 1995). However, the role of putative enzymes encoded by pJP4 module 2 *tfd* genes in the degradation of 2,4-D by *R. eutropha* JMP134 is unknown.

The physical map of the 2,4-D degradative plasmid pEST4011, a derivative of pEST4002 from *Pseudomonas*

Abbreviations: 2,4-D, 2,4-dichlorophenoxyacetic acid; 2,4-DCPH, 2,4-dichlorophenol hydroxylase; CC12O, chlorocatechol 1,2-dioxygenase; 2,4-DCP, 2,4-dichlorophenol; 3,5-DCC, 3,5-dichlorocatechol; bp, base pair(s); cDNA, DNA complementary to RNA; IPTG, isopropyl β-D-thiogalactopyranoside; kb, kilobase(s) or 1000 bp; kDa, kilodalton(s); LB, Luria–Bertani (medium); ORF, open reading frame; PAGE, PA-gel electrophoresis; RBS, ribosome-binding site(s); SDS, sodium dodecyl sulfate.

* Corresponding author. Tel.: +37-27-375014; fax: +37-27-420286.
E-mail address: evedler@ebc.ee (E. Vedler)

0378-1119/00/$ - see front matter © 2000 Published by Elsevier Science B.V. All rights reserved.
PII: S0378-1119(00)00017-2

COR-AAD1-00001207

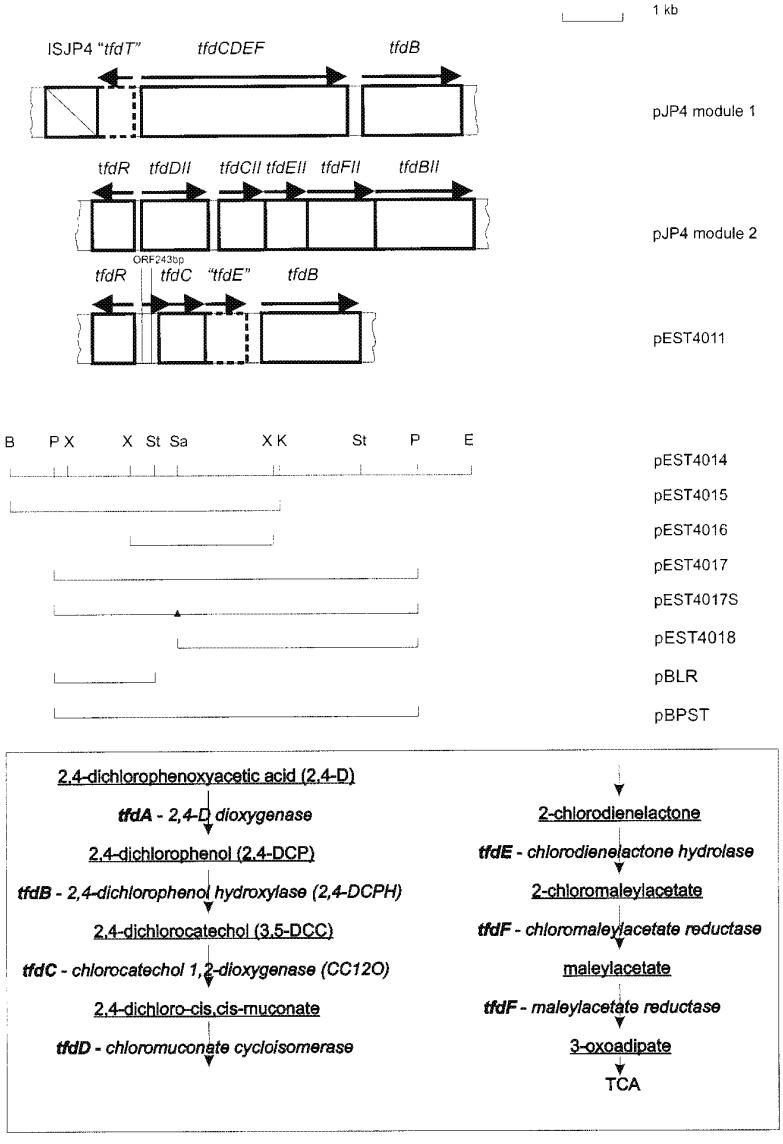


Fig. 1. Organization of the pJP4 module 1 and module 2 (Leveau and van der Meer, 1996; van der Meer, 1996 [U16782], see Introduction) and the 4–7 kb fragment of pEST4011 (Kõiv et al., 1996; this study, see Section 3.1). Directions of transcription are shown for each gene and $ORF_{243\ bp}$. The genes '*tfdT*' of pJP4 and '*tfdE*' of pEST4011 (marked by dotted lines) represent disfunctional genes. Restriction map of the plasmid pEST4014 (7–8 kb *Bam*HI–*Eco*RI fragment of the plasmid pEST4011 cloned into vector plasmid pKT240) and subclones used in this study. Relevant restriction sites used for cloning are shown (B, *Bam*HI; P, *Pst*I; X, *Xho*I; St, *Stu*I; Sa, *Sal*I; K, *Kpn*I; E, *Eco*RI). Construction of plasmids pEST4014, pEST4015, pEST4016 (Mäe et al., 1993) and pEST4017 (Kõiv et al., 1996) has been described previously. The plasmids pEST4017S, pEST4018, pBLR and pBPST were constructed as described in Sections 2.4, 2.7 and 3.2. The black triangle in plasmid pEST4017S shows the lost *Sal*I site. The degradation pathway of 2,4-dichlorophenoxyacetic acid is also presented.

sp. EST4002, isolated from Estonia, was presented (Mäe et al., 1993). The 6.1 kb *Pst*I fragment of this plasmid was subcloned into pKT240 to give pEST4017 (Fig. 1), which encodes information for the regulated expression of 2,4-dichlorophenol hydroxylase (2,4-DCPH) and chlorocatechol 1,2-dioxygenase (CC12O). Sequence analysis of 4.2 kb *Xho*I–*Stu*I fragment of pEST4017 was presented (Kõiv et al., 1996). This fragment was shown to contain the genes *tfdC* and *tfdB* which form one operon transcribed from *tfdC* to *tfdB*. In this study, the open reading frame (ORF) of the regulatory gene *tfdR* was identified and the product of this gene was shown on SDS–PAGE. The role of TfdR in the expression of the catabolic genes *tfdC* and *tfdB* was established. We localized the transcriptional starting points of the *tfdR* gene and the *tfdCB* operon and performed database

COR-AAD1-00001208

searches with the region of pEST4011 containing the genes *tfdR*, *tfdC* and *tfdB*.

2. Materials and methods

2.1. Bacterial strains and plasmids used

Bacterial strains used in this study were *Escherichia coli* DH5α (BLR) and *Pseudomonas putida* PaW85 (Bayley et al., 1977). Vector plasmids pKT240 (Bagdasarian et al., 1983) and pBluescript(SK) (Stratagene) were used to construct plasmids for different analyses (Fig. 1). Plasmids pEST4014, pEST4015, pEST4016 and pEST4017 were described earlier (Kõiv et al., 1996; Mäe et al., 1993).

2.2. Media and culture conditions

E. coli was grown in Luria–Bertani medium (LB) at 37°C, *P. putida* was grown in minimal medium containing 0.3% casamino acids with or without 2,4-dichlorophenol (2,4-DCP) or 3,5-dichlorocatechol (3,5-DCC) (both 20 μM final concentration) at 30°C.

2.3. DNA manipulations and sequence analysis

Plasmid DNA preparations, transformations, and other DNA manipulations were carried out by established procedures (Sambrook et al., 1989; Sanger et al., 1977). Computer analysis of DNA and amino acid sequences was performed with the program PC/GENE (Genofit, Geneva, Switzerland). Database searches were performed with the BLASTN and BLASTX programs at the National Center for Biotechnology Information (NCBI).

2.4. Expression of the gene tfdR and analysis of its product on SDS–PAGE

In order to express the *tfdR* gene, plasmid pBLR (Fig. 1) was constructed as follows: the 1.6 kb *Pst*I–*Stu*I fragment of pEST4017 containing the *tfdR* gene was cloned under the control of *lac* promoter into the plasmid pBluescript(SK). Cells of *E. coli* DH5α containing either pBluescript(SK) or pBLR were grown overnight in LB in the presence of ampicillin (100 μg ml^{-1} final concentration). The cultures were diluted 100 times with fresh LB + ampicillin and grown for 1 h. After that the cultures were grown for an additional 8 h with or without IPTG (2 mM final concentration) as an inducer. Samples of the cultures were spun down, the cells were suspended in sample buffer containing 2% SDS and analyzed on SDS–PAGE.

2.5. Preparation of cell extracts and enzyme assays

Cells of *P. putida* PaW85 containing different plasmids for expression analysis of the genes *tfdB* and *tfdC* were pregrown without 2,4-DCP or 3,5-DCC and then transferred to a fresh medium containing or not containing 2,4-DCP or 3,5-DCC. The crude extracts for enzyme assays were prepared and the specific activities of 2,4-DCPH and CC12O were determined as described earlier (Kõiv et al., 1996).

2.6. Isolation of total RNA

Total RNA for primer extension analysis was isolated from *P. putida* PaW85 cells harboring pEST4017 by using Rneasy Total RNA Purification Protocols (QIAGEN, 1994). The growth conditions were the same as in the case of crude extract preparation for enzyme assays. The cells were grown in the presence or absence of 2,4-DCP.

2.7. Mapping of the transcription initiation points of tfdR gene and tfdCB operon

30 and 20 μg of total RNA was used as template for the mapping of the 5′ ends of the *tfdR* and *tfdCB* transcripts, respectively. Primer extensions of the RNA samples were carried out with avian reverse transcriptase (Promega, Madison, WI) as described by Nurk et al. (1993). Oligonucleotide 5′-TTTCTCCACCTTCATTC-GCTC-3′, complementary to the nucleotide sequences at positions −2 to −22 relative to the start codon of *tfdR* gene, and oligonucleotide 5′-CGACAATGGCT-GTTTCCACGGC-3′, complementary to the nucleotide sequences at positions −202 to −223 relative to the start codon of *tfdCB* operon, were used in primer extension analysis as well as to generate the DNA sequence ladders. Plasmid pBPST, carrying the 6.1 kb *Pst*I fragment of pEST4011 cloned into pBluescript(SK) (Fig. 1), was used as a template for DNA sequencing.

3. Results and discussion

3.1. Sequence analysis of tfdR and detection of its product on SDS–PAGE

We mapped the regulatory gene *tfdR* on plasmid pEST4011 upstream from the gene *tfdC* using Southern hybridization with pJP4 regulatory gene *tfdR* as the probe (Mäe et al., 1993). Our present work completed analysis of the region upstream from *tfdC*; only a partial sequence of this area was previously reported (Kõiv et al., 1996). This analysis revealed a 915 bp ORF starting with an ATG start codon preceded by a potential ribosome-binding site (RBS). This ORF is transcribed

COR-AAD1-00001209

divergently from *tfdC* (Fig. 1) and it predicts a polypeptide of 32.99 kDa. In order to detect the product of the gene *tfdR* on SDS–PAGE, plasmid pBLR was constructed (Fig. 1) as described in Materials and methods, and expressed in *E. coli*. SDS–PAGE showed that the cells containing plasmid pBLR produced a protein of approximately 33 kDa (Fig. 2), corresponding to the size of TfdR, estimated from the nucleotide sequence.

The size and amino acid sequence of TfdR was found to be similar to the members of the catechol-subgroup of the LysR family of activator proteins (Henikoff et al., 1988; Schell, 1993), which all activate transcription from homologous downstream divergent promoters of homologous genes encoding for oxidation of chlorinated or non-chlorinated catechols: TfdR of *R. eutropha* pJP4 (90% identity over the 292 amino-terminal amino acids) (Leveau and van der Meer, 1996; Matrubutham and Harker, 1994), ClcR of *P. putida* pAC27 (58% identity over the 286 amino-terminal amino acids) (Coco et al., 1993, 1994), TcbR of *Pseudomonas* sp. pP51 (52% identity over the 290 amino-terminal amino acids) (Leveau et al., 1994; van der Meer et al., 1991) and CatR of *P. putida* (37% identity over 265 amino acids) (Rothmel et al., 1990, 1991). As there exists similarity in the 61 amino-terminal amino acid residues of TfdR (pEST4011), TfdR (pJP4), ClcR, TcbR and CatR [these residues include the proposed DNA binding domain of LysR family proteins (Henikoff et al., 1988)], TfdR of pEST4011 could bind to the promoter DNA and act as a positive regulator of the downstream *tfdCB* operon. The similarities in the amino acid sequences of TfdR (pEST4011) and other LysR proteins also indicate probable relatedness in their requirement for effector metabolites: for TfdR of pJP4, ClcR, TcbR and CatR it has been shown or proposed that they require *cis,cis*-muconate or its chlorinated derivative for their action (Coco et al., 1993; Filer and Harker, 1997; Parsek et al., 1992; van der Meer et al., 1991). Therefore we assumed that the effector molecule for TfdR could be the product of CC12O activity, 2,4-dichloro-*cis,cis*-muconate.

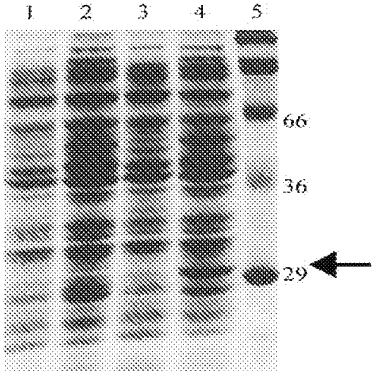


Fig. 2. The expressed gene *tfdR* in *E. coli* DH5α cell extracts visualized on a Coomassie blue stained 0–1% SDS–10% PAGE (see Sections 2.4 and 3.1). Lanes: 1, *E. coli* [pBluescriptSK] total protein; 2, *E. coli* [pBluescriptSK] total protein plus IPTG; 3, *E. coli* [pBLR] total protein; 4, *E. coli* [pBLR] total protein plus IPTG; 5, molecular mass marker. The arrow indicates the *tfdR* gene product of approximately 33 kDa.

3.2. Determination of the role of TfdR in the expression of the tfdCB operon and the effector molecule for TfdR

We determined the role of TfdR in the expression of the *tfdCB* operon and the effector molecule for TfdR by measuring the specific activities of 2,4-DCPH (TfdB) and CC12O (TfdC) in *P. putida* PaW85 carrying different plasmids (Fig. 1). We grew the cells 30 min with and without 2,4-DCP and 3,5-DCC as the substrates of 2,4-DCPH and CC12O, respectively. The results are shown in Fig. 3. *P. putida* PaW85 harboring plasmids pEST4014 and pEST4017 (Fig. 1) (Kõiv et al., 1996; Mäe et al., 1993) had basal 2,4-DCPH and CC12O activities without 2,4-DCP or 3,5-DCC, and induced 2,4-DCPH and CC12O activities in the presence of these compounds (in the case of 3,5-DCC induction occurred to a lesser extent). In order to determine whether TfdR is responsible for *tfdCB* activation, we used plasmids without the intact *tfdR*: pEST4016 (Mäe et al., 1993) and pEST4018 (4 kb *SalI–PstI* fragment of pEST4017 was cloned into pKT240) (Fig. 1). *P. putida* PaW85 [pEST4018] showed no detectable activities of either

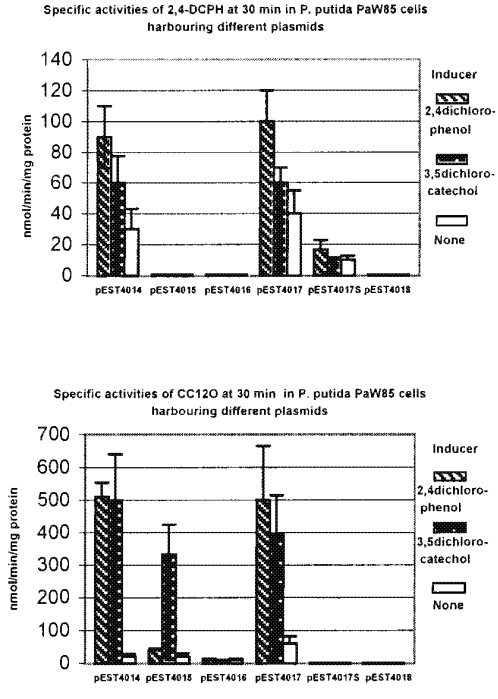


Fig. 3. The specific activities of 2,4-DCPH and CC12O at 30 min in *P. putida* PaW85 cells harboring different plasmids (Fig. 1, see Sections 2.5 and 3.2).

COR-AAD1-00001210

2,4-DCPH or CC12O with or without 2,4-DCP or 3,5-DCC (Fig. 3). The cells harboring plasmid pEST4016 had no detectable 2,4-DCPH and only low basal CC12O activities in all growth conditions. These results showed that without TfdR no induction of *tfdCB* operon was achieved.

Experiments with *P. putida* PaW85 [pEST4017] showed that the effector for TfdR is either 2,4-DCP, 3,5-DCC or 2,4-dichloro-*cis*,*cis*-muconate [the product of *tfdC* (CC12O) converts 3,5-DCC to 2,4-dichloro-*cis*,*cis*-muconate]. At the determination of the effector molecule for TfdR we eliminated the *tfdC* and *tfdB* genes in order to achieve the situation where the generation of the possible effector molecule would be impossible. For this purpose we used plasmids pEST4015 (Mäe et al., 1993) and pEST4017S (Fig. 1), in which ORF of *tfdC* was interrupted as a result of lost *Sal*I restriction site (plasmid pEST4017 was digested with *Sal*I followed by S1 nuclease treatment). *P. putida* PaW85 [pEST4015] showed no detectable 2,4-DCPH activities, with or without 2,4-DCP or 3,5-DCC (Fig. 3). However, we detected low constitutive levels of CC12O in the absence of 2,4-DCP or 3,5-DCC or in the presence of 2,4-DCP, and induced CC12O activity in the presence of 3,5-DCC. *P. putida* PaW85 [pEST4017S] had no detectable CC12O activities and only low constitutive levels of 2,4-DCPH in all growth conditions. These results (see also Fig. 1) strongly suggested that the activator molecule for TfdR was the product of CC12O activity, 2,4-dichloro-*cis*,*cis*-muconate.

3.3. Mapping of the transcription initiation points of tfdR gene and tfdCB operon

Primer extension analysis of mRNAs of *tfdR* and *tfdCB* demonstrated that identical *tfdR* extension products were detected when RNA was isolated from cultures grown with or without 2,4-DCP and the transcription starting point of *tfdR* mapped to the positions −85 and −84 relative to the translation starting point of TfdR (Fig. 4A). The nucleotide sequence CAGTCTN$_{16}$TGGTCT was found 6 (7) bp upstream from the transcription initiation site. This sequence only very slightly resembles the *E. coli* RNA polymerase sigma^{70}-recognized promoter TTGACAN$_{16-18}$TATAAT (two matches in both −35 and −10 boxes). In contrast, the transcript of *tfdCB* was observed with RNA isolated only from cultures grown on 2,4-DCP. The transcription starting point of *tfdCB* mapped to the position −292 relative to the translation starting point of TfdC (Fig. 4B). Additional minor bands seen on the gel are probably decay products. The nucleotide sequence TTAGACN$_{18}$TAGACT, found 6 bp upstream from the transcription initiation site of this operon, resembles the *E. coli* RNA polymerase sigma^{70}-recognized promoter more than the *tfdR* promoter does (two matches in −35

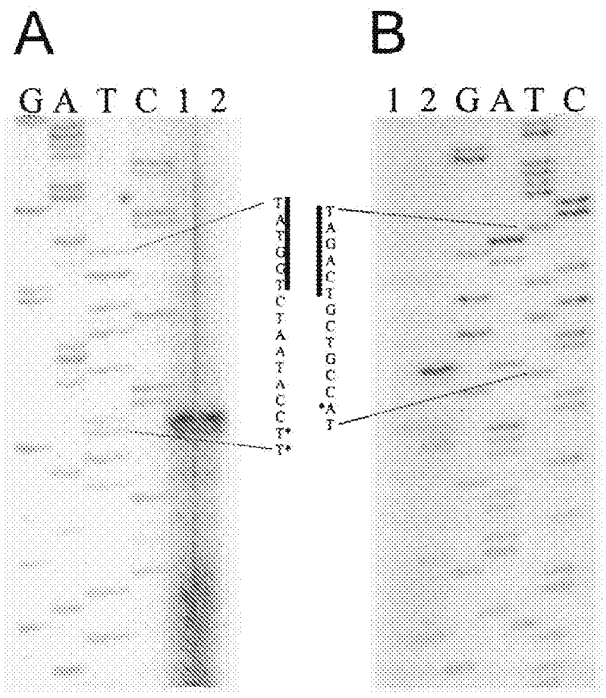


Fig. 4. Mapping of the transcription initiation points of the (**A**) *tfdR* gene and (**B**) *tfdCB* operon (see Section 2.7). The nucleotide sequences of the *tfdR* and *tfdCB* promoter regions (see Section 3.3), including the −10 sequences (marked by vertical lines) and the transcription initiation points (indicated by asterisks), are shown alongside. RNA isolated from *P. putida* PaW85 [pEST4017] grown in the absence of 2,4-DCP (see Section 2.6) is shown in lanes 1 and in the presence of 2,4-DCP in lanes 2. Double-stranded DNA sequence ladders of the *tfdR* and *tfdCB* promoter regions are shown in lanes G, A, T and C.

and four matches in −10 box). These results indicated that transcription of *tfdR* gene is constitutive, while in the case of *tfdCB* operon transcription is induced by growth on 2,4-DCP and the transcription level of this operon must be higher than that of *tfdR* gene. Primer extension analyses were also made to detect a separate transcription starting point in front of *tfdB*, but no cDNA bands were detected in these cases.

3.4. Database searches with the region of pEST4011 containing the genes tfdR, tfdC and tfdB

It has been shown that the intervening sequences of gene pairs *tfdS–tfdA*, *tfdT–tfdCDEF*, *tfdR–tfdD$_{II}$* of pJP4, *clcR–clcABD* of pAC27, *tcbR–tcbCDEF* of pP51 and *catR–catBC* of *P. putida* chromosome have regions with high conservation of nucleotides (Matrubutham and Harker, 1994). These regions are shown or proposed to contain the respective promoter sequences and transcription initiation sites. A database search done in this study revealed regions with high conservation of nucleotides in pEST4011 *tfdR–tfdC* and pJP4 *tfdR–tfdD$_{II}$*, *tfdS–tfdA*, *tfdT–tfdCDEF*, pP51 *tcbR–tcbCDEF* and

COR-AAD1-00001211

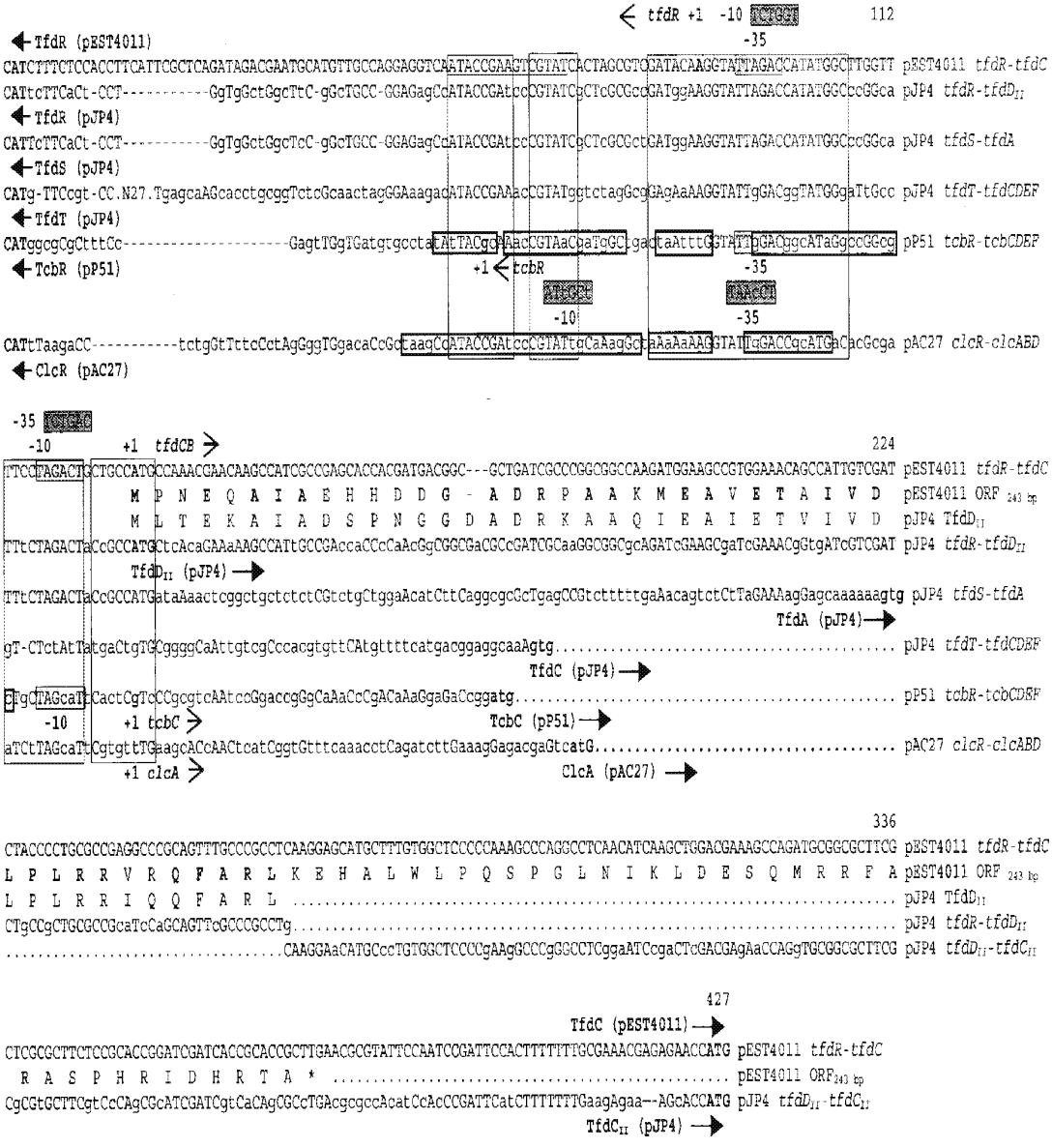


Fig. 5. Alignment of the nucleotide sequence of the pEST4011 *tfdR–tfdC* intergenic region with the nucleotide sequences between the genes *tfdR–tfdD*$_{II}$ of pJP4 (Matrubutham and Harker, 1994), *tfdS–tfdA* of pJP4 (Streber et al., 1987), *tfdT–tfdCDEF* of pJP4 (Leveau and van der Meer, 1996), *tcbR–tcbCDEF* of pP51 (van der Meer et al., 1991) and *clcR–clcABD* of pAC27 (Coco et al., 1993) (see Section 3.4). −10 and −35 promoter sequences of *tfdCB* and *tcbC* are in pale gray boxes. −10 and −35 promoter sequences of *tfdR* and *tcbR* are in dark gray boxes. Arrows indicate the direction of transcription and translation. Transcription initiation points and translation initiation codons are in bold letters. Common nucleotides for pEST4011 and other plasmids are in capital letters. The conserved regions of pEST4011 and other aligned plasmids are boxed. The amino acid sequence, where necessary, is in capital letters below or above the nucleotide sequence. Amino acids identical for pEST4011 ORF$_{243\ bp}$ and pJP4 TfdD$_{II}$ are in bold letters. Regions in the nucleotide sequences of pP51 and pAC27 which have been shown to be protected by TcbR (Leveau et al., 1994) and ClcR (Coco et al., 1994), respectively, are bold boxed. Gaps needed for optimal alignment are shown (—). Regions which have no sequence similarities with the aligned pEST4011 are dotted. N, any nucleotide. The T-N$_{11}$-A binding consensus found in pEST4011 is underlined.

COR-AAD1-00001212

pAC27 *clcR–clcABD* intergenic nucleotide sequences (Fig. 5). The transcription starting points and promoter sequences of pEST4011 *tfdR* and *tfdCB* were localized in these regions (Figs. 4 and 5). This conservation also contains binding sites for TcbR (Leveau et al., 1994) and ClcR (Coco et al., 1994) with the T–N_{11}–A interrupted dyad binding consensus for LysR type regulators (Goethals et al., 1992). This alignment probably indicates similar molecular mechanisms of regulation in the case of pEST4011 and the well studied cases of pP51 and pAC27.

Alignment of the nucleotide sequences of pEST4011 and pJP4 showed that the pEST4011 genes *tfdR*, *tfdC* and *tfdB* were the most similar to the pJP4 module 2 genes *tfdR*, $tfdC_{II}$ and $tfdB_{II}$, respectively (van der Meer, 1996 [U16782]) (Fig. 1). Although the pEST4011 nucleotide sequence between *tfdC* and *tfdB* was found to be very similar to the pJP4 module 2 $tfdE_{II}$ gene (van der Meer, 1996 [U16782]) (Fig. 1), deletion of GC dinucleotide in the coding region leads to frameshifting. SDS–PAGE analysis (Kõiv et al., 1996) revealed no protein band that could have been translated from this region. Amino acid sequences of pEST4011 TfdC, 'TfdE' and TfdB are 75–85% identical to pJP4 $TfdC_{II}$, $TfdE_{II}$ and $TfdB_{II}$, and only about 60% identical to pJP4 module 1 gene products TfdC, TfdE and TfdB, respectively (van der Meer, 1996 [U16782], [M35097]) (Fig. 1). We also found a 243 bp nucleotide sequence in the *tfdR–tfdC* intergenic region, the predicted 43 amino-terminal amino acids of which had 65% identity with the respective 44 amino acids of pJP4 putative chloromuconate cycloisomerase, $TfdD_{II}$ (van der Meer, 1996 [U16782]) (Figs. 1 and 5). This region was designated $ORF_{243\ bp}$, it starts 292 bp and ends 48 bp upstream from the coding sequence of pEST4011 *tfdC*. Finding this $ORF_{243\ bp}$ was actually the reason why we made primer extension of *tfdCB* using the oligo, complementary to the nucleotide sequence so far upstream from the coding sequence of *tfdC*. No cDNA bands were detected with oligos targeting regions up- or downstream of the one used in this study (data not shown). These results indicated that $ORF_{243\ bp}$ is the most upstream part of the *tfdCB* operon, but it is not a translated ORF since there is no RBS in front of it because the first nucleotide in its ATG translation initiation codon is the transcription start site of the *tfdCB* operon (Fig. 5). Thus, our data showed that the *tfdR*, *tfdC* and *tfdB* genes are very similar to the module 2 set of genes of the archetypical 2,4-D plasmid pJP4. Unfortunately, no biochemical data about the products of the latter genes as well as their role in the degradation of 2,4-D by the cells carrying plasmid pJP4 has been available. Despite all that one can speculate that the pJP4 module 2 and the discussed region of pEST4011 have apparently evolved from the same origin. It is interesting that geographically distinct bacteria have obtained similarly organized sets of genes that give rise to the same phenotype. This is the reason why we will continue to study the plasmid pEST4011 in order to determine the whole set of 2,4-D catabolic genes as well as regulatory and structural units present in pEST4011.

The nucleotide sequence of pEST4011, discussed and partly sequenced in this study, has been assigned GenBank accession number U32188.

4. Conclusions

1. The size and amino acid sequence of the pEST4011 *tfdR* gene product is similar to the members of the catechol-subgroup of the LysR-family of transcriptional activators.
2. The pEST4011 *tfdR* gene product activates the expression of the *tfdCB* operon and the effector molecule for TfdR regulatory protein is 2,4-dichloro-*cis*,*cis*-muconate.
3. Database search revealed high conservation of nucleotides in pEST4011 *tfdR–tfdC* promoter–operator region and the respective sequences of other well studied LysR family catechol-subgroup regulatory systems. This alignment strongly suggests similar molecular mechanisms of regulation in all these cases.
4. Database searches also revealed that the whole region of pEST4011 which contains the genes *tfdR*, *tfdC* and *tfdB* is the most similar to the pJP4 region which contains the regulatory gene *tfdR* and the module 2 set of 2,4-D degradative genes $tfdD_{II}$ to $tfdB_{II}$, thus those two regions might have the same origin.

Acknowledgement

This work was supported by the Institute of Molecular and Cell Biology, by Grant 2334 from the Estonian Science Foundation.

References

Bagdasarian, M.M., Amann, E., Lurz, R., Rueckert, B., Bagdasarian, M., 1983. Activity of the hybrid *trp.lac(tac)* promoter of *Escherichia coli* in *Pseudomonas putida*. Construction of broad-host-range controlled-expression vectors. Gene 26, 273–282.

Bayley, S.A., Duggleby, C.J., Worsey, M.J., Williams, P.A., Hardy, K.G., Broda, P., 1977. Two models of loss of the TOL function from *Pseudomonas putida* mt-2. Mol. Gen. Genet. 154, 203–204.

Coco, W.M., Rothmel, R.K., Henikoff, S., Chakrabarty, A.M., 1993. Nucleotide sequence and initial functional characterization of the *clcR* gene encoding a LysR family activator of the *clcABD* chlorocatechol operon in *Pseudomonas putida*. Proc. Natl. Acad. Sci. USA 90, 1033–1037.

Coco, W.M., Parsek, M.R., Chakrabarty, A.M., 1994. Purification of the LysR family regulator, ClcR, and its interaction with the *Pseu-*

COR-AAD1-00001213

domonas putida clcABD chlorocatechol operon promoter. J. Bacteriol. 176, 5530–5533.

Don, R.H., Weightman, A.J., Knackmuss, H.J., Timmis, K.N., 1985. Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4). J. Bacteriol. 161, 85–90.

Filer, K., Harker, A.R., 1997. Identification of the inducing agent of the 2,4-dichlorophenoxyacetic acid pathway encoded by plasmid pJP4. Appl. Environ. Microbiol. 63, 317–320.

Goethals, K., van Montagu, M., Holsters, M., 1992. Conserved motifs in a divergent *nod* box of *Azorhizobium caulinodans* ORS571 reveal a common structure in promoters regulated by LysR-type proteins. Proc. Natl. Acad. Sci. USA 89, 1646–1650.

Henikoff, S., Haughn, G.W., Calvo, J.M., Wallace, J.C., 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA 85, 6602–6606.

Kaphammer, B., Kukor, J.J., Olsen, R.H., 1990. Regulation of *tfdCDEF* by *tfdR* of the 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. J. Bacteriol. 172, 2280–2286.

Kõiv, V., Marits, R., Heinaru, A., 1996. Sequence analysis of the 2,4-dichlorophenol hydroxylase gene *tfdB* and 3,5-dichlorocatechol 1,2-dioxygenase gene *tfdC* of 2,4-dichlorophenoxyacetic acid degrading plasmid pEST4011. Gene 174, 293–297.

Leveau, J.H.J., van der Meer, J.R., 1996. The *tfdR* gene can successfully take over the role of the insertion element-inactivated TfdT protein as a trancriptional activator of the *tfdCDEF* gene cluster, which encodes chlorocatechol degradation in *Ralstonia eutropha* JMP134(pJP4). J. Bacteriol. 178, 6824–6832.

Leveau, J.H.J., de Vos, W.M., van der Meer, J.R., 1994. Analysis of the binding site of the LysR-type transcriptional activator TcbR on the *tcbR* and *tcbC* divergent promoter sequences. J. Bacteriol. 176, 1850–1856.

Leveau, J.H.J., Zehnder, A.J.B., van der Meer, J.R., 1998. The *tfdK* gene product facilitates uptake of 2,4-dichlorophenoxyacetate by *Ralstonia eutropha* JMP134(pJP4). J. Bacteriol. 180, 2237–2243.

Leveau, J.H.J., König, F., Füchslin, H., Werlen, C., van der Meer, J.R., 1999. Dynamics of multigene expression during catabolic adaptation of *Ralstonia eutropha* JMP134 (pJP4) to the herbicide 2,4-dichlorophenoxyacetate. Mol. Microbiol. 33, 396–406.

Mäe, A., Marits, R., Ausmees, N., Kõiv, V., Heinaru, A., 1993. Characterization of a new 2,4-dichlorophenoxyacetic acid degrading plasmid pEST4011: physical map and localization of catabolic genes. J. Gen. Microbiol. 139, 3165–3170.

Matrubutham, U., Harker, A.R., 1994. Analysis of duplicated gene sequences associated with *tfdR* and *tfdS* in *Alcaligenes eutrophus* JMP134. J. Bacteriol. 176, 2348–2353.

Nurk, A., Tamm, A., Hõrak, R., Kivisaar, M., 1993. In vivo-generated fusion promoters in *Pseudomonas putida*. Gene 127, 23–29.

Parsek, M.R., Shinabarger, D.L., Rothmel, R.K., Chakrabarty, A.M., 1992. Roles of CatR and *cis,cis*-muconate in activation of the *catBC* operon, which is involved in benzoate degradation in *Pseudomonas putida*. J. Bacteriol. 174, 7798–7806.

Perkins, E.J., Gordon, M.P., Caceres, O., Lurquin, P.F., 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. 172, 2351–2359.

Rothmel, R.K., Aldrich, T.L., Houghton, J.E., Coco, W.M., Ornston, L.N., Chakrabarty, A.M., 1990. Nucleotide sequencing and characterization of *Pseudomonas putida catR*: a positive reguator of the *catBC* operon is a member of the LysR family. J. Bacteriol. 172, 922–931.

Rothmel, R.K., Chinabarger, D.L., Parsek, M.R., Aldrich, T.L., Chakrabarty, A.M., 1991. Functional analysis of the *Ps. putida* regulatory protein CatR: transcriptional studies and determination of the CatR DNA-binding site by hydroxyl-radical footprinting. J. Bacteriol. 173, 4717–4724.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. second ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Schell, M.A., 1993. Molecular biology of the LysR family of transcriptional regulators. Annu. Rev. Microbiol. 47, 597–626.

Streber, W.R., Timmis, K.N., Zenk, M.H., 1987. Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. 169, 2950–2955.

van der Meer, J.R., 1996. Direct submission to the databases.

van der Meer, J.R., Frijters, A.C.J., Leveau, J.H.J., Eggen, R.I.L., Zehnder, A.J.B., de Vos, W.M., 1991. Characterization of the *Pseudomonas* sp. strain P51 gene *tcbR*, a LysR-type transcriptional activator of the *tcbCDEF* chlorocatechol oxidative operon and analysis of the regulatory region. J. Bacteriol. 173, 3700–3708.

You, I.S., Ghosal, D., 1995. Genetic and molecular analysis of a regulatory region of the herbicide 2,4-dichlorophenoxyacetate catabolic plasmid pJP4. Mol. Microbiol. 16, 321–331.

COR-AAD1-00001214

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Jan. 1999, p. 264–269
0099-2240/99/$04.00+0

Vol. 65, No. 1

Pleiotropic Effects of Adaptation to a Single Carbon Source for Growth on Alternative Substrates

GREGORY J. VELICER

Center for Microbial Ecology, Michigan State University, East Lansing, Michigan 48824

Received 17 August 1998/Accepted 8 October 1998

It is frequently assumed that populations of genetically modified microorganisms will perform their intended function and then disappear from the environment due to inherent fitness disadvantages resulting from their genetic alteration. However, modified organisms used in bioremediation can be expected to adapt evolutionarily to growth on the anthropogenic substrate that they are intended to degrade. If such adaptation results in improved competitiveness for alternative, naturally occurring substrates, then this will increase the likelihood that the modified organisms will persist in the environment. In this study, bacteria capable of degrading the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) were used to test the effects of evolutionary adaptation to one substrate on fitness during growth on an alternative substrate. Twenty lineages of bacteria were allowed to evolve under abundant resource conditions on either 2,4-D or succinate as their sole carbon source. The competitiveness of each evolved line was then measured relative to that of its ancestor for growth on both substrates. Only three derived lines showed a clear drop in fitness on the alternative substrate after demonstrable adaptation to their selective substrate, while five derived lines showed significant simultaneous increases in fitness on both their selective and alternative substrates. These data demonstrate that adaptation to an anthropogenic substrate can pleiotropically increase competitiveness for an alternative natural substrate and therefore increase the likelihood that a genetically modified organism will persist in the environment.

Biological traits that are not important components of fitness in a particular environment may nonetheless evolve by a variety of mechanisms. Such mechanisms include pleiotropic side effects of adaptive mutations on unselected traits (43), the random fixation of effectively neutral alleles by genetic drift (19), and "hitchhiking" of neutral mutations due to genetic linkage with an adaptive mutation (16, 31). Along with adaptation by natural selection, these more indirect mechanisms of evolution are relevant to such issues as the genetic divergence of populations within a species (8, 43) and the fate of genetically modified organisms in natural environments (25, 36, 37, 40).

Several studies have shown that isolated populations of the same species may evolve different adaptations even to the same selective conditions. Some of the clearest evidence for this phenomenon comes from laboratory experiments using fruitflies (9, 18) and bacteria (42, 43). Alternative pathways of adaptation to a particular selective environment may lead to heterogeneous changes in traits that are not important for fitness in that environment. For example, several populations of organisms that are under selective pressure to use a particular substrate more efficiently may each find a different physiological mechanism to do so. The effects of these different adaptations on the ability of organisms to use some alternative substrate may be positive in some cases, neutral in others, and negative in yet others. As a consequence of these different correlated responses, some populations may be fortuitously well adapted and others may be poorly adapted to environments that contain this alternative substrate.

Of particular relevance to these concerns is the effect that evolutionary adaptation by bioremediative organisms to their intended anthropogenic substrate has on their abilities to grow on, and compete for, naturally occurring substrates. If adaptation to an anthropogenic substrate always leads to a reduction in fitness on natural substrates, then this reduced fitness provides a measure of safety for the release of organisms modified for bioremediation. If, however, such adaptation indirectly enhances fitness on natural substrates, then this enhanced fitness increases the likelihood that the released organisms (and any adverse effects they may cause) will persist in the environment.

In this study, evolving populations of bacteria were used to investigate the effect of adaptation to one carbon source on competitive performance on an alternative substrate. The bacterial strains employed were two natural isolates of the β-proteobacteria (genus *Burkholderia*), each capable of degrading the anthropogenic herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) (32, 41). Experimental evolution occurred in two separate stages, one in which succinate was the sole carbon source and one in which 2,4-D was the limiting substrate (Fig. 1). This two-stage design was employed to allow a period during which any major adaptations to general selective conditions other than the carbon substrate could occur (stage I on succinate) prior to subsequent evolution on the selective substrate of primary interest, 2,4-D. Allowing any such general adaptation first should maximize the degree of substrate-specific adaptation on 2,4-D. Previous evolution experiments also using 2,4-D-degrading bacteria have been conducted to investigate the role of environmental structure in adaptation and the genetic divergence of multiple populations evolving independently in identical habitats (21, 22, 23).

In stage I of the experimental evolution, two ancestral strains were used to found replicate lines that were propagated in batch culture with succinate (a naturally occurring substrate) as the sole carbon source (Fig. 1). One of these strains (strain F) grows relatively fast under laboratory conditions, whereas the other (strain S) grows more slowly (42). In stage II of the evolution experiment, clonal isolates derived from the stage I batch culture lines were used to found replicate lines that were further propagated in batch culture with 2,4-D as the sole

* Mailing address: Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824. Phone: (517) 353-0809. Fax: (517) 353-9334. E-mail: velicerg@pilot.msu.edu.

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001215

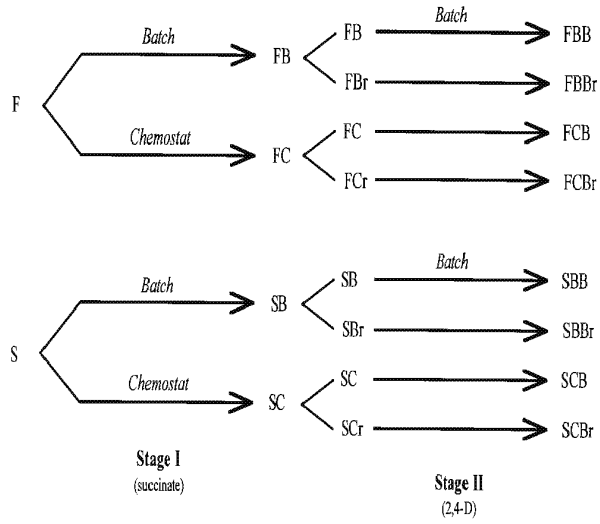


FIG. 1. Derivation of bacterial strains. Horizontal lines indicate periods of evolution, with the selective regimen and sole carbon source indicated above and below, respectively. Each horizontal line represents two independently evolved replicate populations. F and S represent the fast and slow ancestral strains, respectively. For stage I-derived strains, the first letter (F or S) indicates the ancestor and the second letter (B or C) indicates whether stage I evolution occurred in batch (B) or chemostat (C) cultures. The letter r indicates streptomycin resistance (see Materials and Methods). The first two letters for stage II evolution strains indicate their stage I-proximate ancestor, and the third letter (B) indicates stage II evolution in batch culture.

carbon source (Fig. 1). Additional stage II batch lines were founded with clones from stage I lines that had undergone evolution in a different selective regimen, chemostat culture (45). (These stage I lines were also involved in a sibling study of the effects of adaptation to either batch culture [an abundant-resource regimen] or chemostat culture [a scarce-resource regimen] on fitness in the alternative selective regimen) (45). The competitive performances of all of these derived lines were then measured relative to those of their immediate ancestors on both succinate and 2,4-D. The results of these experiments demonstrated that adaptation to either substrate can be associated with both improvements and losses in competitive performance on the alternative substrate.

MATERIALS AND METHODS

Strains. For simplicity, the designations of the two ancestral strains in this study were changed from those used in previous literature (strain F was TFD3, and strain S was TFD20), in which their 2,4-D catabolic pathways were genetically and phylogenetically characterized (32, 41).

Evolution experiments. (i) Stage I batch evolution. Two clones each from strains F and S were inoculated independently into 10 ml of a mineral salts medium (45) containing 500 μg of succinate per ml and grown to stationary phase while being shaken at 120 rpm. These lines were designated FB1, FB2, SB1, and SB2, with the first and second letters indicating ancestral strain (F or S) and selective regimen (batch), respectively, and with the numbers distinguishing replicate lines (Fig. 1). The four lines were then diluted daily into fresh medium for 75 days (see Table 1 for dilution factors and numbers of generations evolved for all selection experiments). Dilution factors were set to allow complete population replacement in 24 h with minimal time spent in stationary phase. All cultures were maintained at 25°C in all experiments.

(ii) Stage I chemostat evolution. Two initially clonal lines from each ancestor were also evolved in chemostats for 75 days during stage I, where fresh medium containing succinate as the sole carbon source continually flowed into culture vessels and culture volume was kept constant by continuous vacuum removal of excess culture (45). These lines were designated FC1, FC2, SC1, and SC2, where C indicates a chemostat selection history. These chemostat-evolved lines were competed against their ancestors in a different study (which tested for a trade-off between fitness at high and low substrate concentrations [45]) but not in this study. For logistical reasons, only strains that experienced the batch culture regimen during their most recent stage of evolution were competed against their ancestors in this study. Thus, the stage I chemostat-evolved strains served as the proximate ancestors for half of the stage II batch lines but not as evolved strains for analysis in their own right. Use of these strains as stage II ancestors (rather than simply initiating more replicate lines from the stage I batch clones) allowed a greater chance of detecting patterns of adaptation specific to distinct selective histories.

TABLE 1. Dilution factors for batch evolution and competition experiments and numbers of generations evolved (during the most recent stage of evolution)[a]

Line set	Daily dilution factor	No. of generations evolved
FB	0.01	500
SB	0.20	185
FBB	0.01	500
FCB (FCBr)	0.01 (0.08)	500 (275)
SBB	0.20	175
SCB	0.33	120

[a] Daily dilution factors and numbers of generations evolved are shown for each set of replicate lines. Dilution factors indicate the proportion of day-old cultures that were transferred daily into fresh medium (to a final volume of 10 ml). Except for the FCBr lines, values for streptomycin-resistant replicate lines within a set are identical to values for streptomycin-sensitive lines.

(iii) Stage I clones. Single clones were isolated from each of the eight stage I evolved lines. The four batch-evolved clones (FB1, FB2, SB1, and SB2) were compared to their ancestors in competition experiments described below. In addition, a spontaneous streptomycin-resistant sister clone was selected from each of the four clones FB1, FC1, SB1, and SC1 (denoted by the letter r). Each of these eight clones (FB1, FBr, FC1, FCr, SB1, SBr, SC1, and SCr) was then used to initiate two new lines (distinguished by the numbers 1 and 2) for stage II evolution.

(iv) Stage II batch evolution. Stage II batch lines were propagated as described above for stage I lines except that the growth medium contained 500 μg of 2,4-D per ml rather than succinate and dilution factors varied. After evolution, single clones from each line were isolated for use in competition experiments. Stage II clones were named by adding the letter B to the name of their stage I-proximate ancestor, indicating stage II evolution in batch culture (Fig. 1; Table 2). All strains were stored in medium containing 10% glycerol at −80°C. The identity of each clone as a true evolved descendant of the founding clone (rather than as a contaminant) was confirmed both by streptomycin marker type and by repetitive-extragenic-palindromic-PCR diagnostic fingerprints (46).

Competition experiments. The competitive performance of each clone that evolved in batch culture during its most recent stage of evolution was measured by directly competing a clone with its proximate ancestor of the opposite marker type. These experiments were performed on both selective and alternative substrates under the same conditions in which the clone evolved. Both strains in each pairwise competition were preconditioned for two growth cycles in the competition medium. The strains were then mixed and grown together for several generations (one to three daily dilution cycles). Initial and final relative frequencies of competing strains were estimated by dilution plating onto selective and nonselective agar plates. These frequencies were then used to calculate the selection rate constant (SRC) for each evolved clone (see below), which expresses the amount of change in competitive performance undergone by a clone relative to the competitive performance of its proximate ancestor (see reference 45 for a more detailed description of competition experiments).

Control competition experiments were also performed to factor out any effect of streptomycin resistance on competitive performance (45). For example, for comparison of a stage II strain with its stage I-proximate ancestor (e.g., FBB1 versus FB1), the stage I reciprocally marked ancestor (FBr) was competed both against its oppositely marked sister clone (FB1) and the stage II clone (FBB1). The SRC of the direct proximate ancestor relative to that of its sister clone (FB1 versus FBr) was then subtracted from the SRC of the stage II clone relative to that of its reciprocally marked ancestor (FBB1 versus FBr). This process gives the SRC for the stage II strain relative to that of its direct proximate ancestor (FBB1 versus FB1). Therefore, the SRC values shown for each evolved strain (Table 2) reflect the amount of performance change by each strain relative to the performance of its immediate proximate ancestor of the same marker type (e.g., SB1 relative to S, FCB1 relative to FC1, and FBBr2 relative to FBr2). A positive SRC indicates competitive superiority of a strain over its ancestor, whereas a negative value indicates inferiority. The SRC between two competitors, i and j, (s_{ij}) is the difference between their actualized Malthusian parameters: $s_{ij} = (1/t)\{\ln[N_i(t)/N_i(0)] - \ln[N_j(t)/N_j(0)]\}$, where t is the number of 24-h transfer cycles in the competition, $N_i(0)$ and $N_j(0)$ are the initial population densities, and $N_i(t)$ and $N_j(t)$ are the final population densities (measured at the same point in

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001216

TABLE 2. Results of competition experiments for evolved lines from stage I selection in succinate and stage II selection in 2,4-D

Line	Selective substrate	SRC (per day)[a] on:	
		Selective substrate	Alternative substrate
FB1	Succinate	0.0260	0.0236
FB2	Succinate	0.0910	**−4.6052***
SB1	Succinate	**0.2531**	0.0366
SB2	Succinate	**0.3759**	**0.1582**
FBB1	2,4-D	**0.4026**	−0.0929
FBB2	2,4-D	**0.9675**	0.0147
FBBr1	2,4-D	**0.3116**	**0.3932**
FBBr2	2,4-D	**0.3439**	**−0.8023**
FCB1	2,4-D	0.4171	**−2.7843**
FCB2	2,4-D	**0.6390**	**−1.1000**
FCBr1	2,4-D	**0.6788**	**1.9561**
FCBr2	2,4-D	−0.6600	−0.5012
SBB1	2,4-D	0.0379	−0.0287
SBB2	2,4-D	0.1073	**0.1892**
SBBr1	2,4-D	**0.2230**	−0.0880
SBBr2	2,4-D	0.0701	−0.0762
SCB1	2,4-D	**0.6924**	**0.2455**
SCB2	2,4-D	**0.7670**	0.0291
SCBr1	2,4-D	**0.4899**	**0.4913**
SCBr2	2,4-D	**0.3680**	**−1.3205**

[a] SRCs of evolved lines (relative to those of their appropriate ancestors) for both the substrate in which evolution occurred and the alternative substrate are presented. SRC values that are significantly different from those of the appropriate ancestor are presented in boldface type (two-tailed t test, $P < 0.05$). The asterisk marking strain FB2 for performance on 2,4-D indicates that the maximum SRC for FB2 on 2,4-D was inferred, rather than measured directly. The value is based on the ability of FB2's competitor (FBr) to grow at least 100-fold per day on 2,4-D, since FB2 is unable to grow on 2,4-D.

the daily growth cycle) (see references 28, 42, and 45 for more detailed descriptions and derivations of SRCs). Each SRC value reported in Table 2 is a mean value obtained from results of multiple independent replicate competitions (usually five, but a minimum of three) performed in parallel for each strain. Ninety-five percent significance levels were determined by t tests.

RESULTS

Adaptation to selective substrate. Of the stage I lines propagated on succinate in batch culture, the two S-derived lines showed significantly greater adaptation than did the two F-derived lines (Table 2; Fig. 2). Moreover, this difference is conservative, since the S-derived lines underwent fewer than half the number of doublings (185 generations) during stage I as did the F-derived lines (500 generations). This suggests that slower-growing strain S had more room for improvement in the batch regimen than did strain F.

Six of the eight stage II F-derived lines significantly improved their competitiveness on 2,4-D, which was the substrate supplied during stage II (Table 2). In one set of lines (FBB) all four lines demonstrably improved (Table 2; Fig. 3). Similarly, five of the eight S-derived stage II lines show significant fitness increases (Table 2; Fig. 4). All four SCB lines (which were subjected to changes in both culture regimen and growth substrate between stages I and II) improved significantly during stage II.

Competitive performance on alternative substrate. Among the 13 lines that showed significant improvement on their selective substrate, 5 also improved significantly on the alternative substrate (Table 3). Three additional lines had nonsignificant performance increases on the alternative substrate. In contrast, 3 of the 13 significantly adapted lines experienced significant losses of competitive ability on the alternative substrate and two others were suggestive of such losses (Table 3). Overall, 10 lines either showed or suggested correlated improvements on both substrates while 9 showed or suggested fitness losses on the alternative substrate (Table 3). Among the six sets of evolved lines, only the two SB replicate lines exhibited a homogeneous pattern of performance on both selective and alternative substrates. The other five sets showed (or suggested) both performance losses and performance gains among their lines on the alternative substrate.

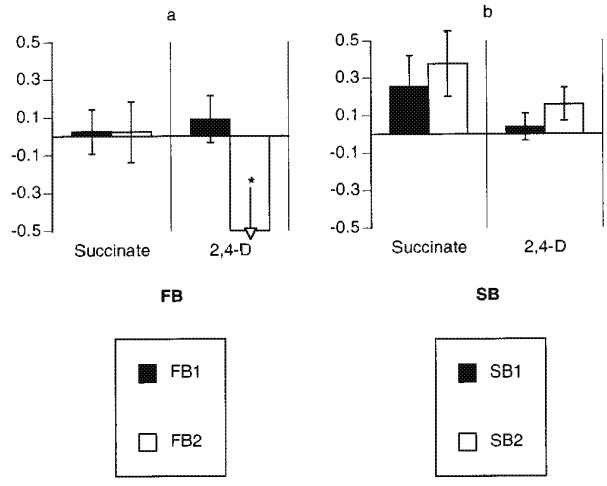


FIG. 2. Performance changes of stage I lines. (a) FB lines; (b) SB lines. Mean SRC values are shown, along with 95% confidence intervals, for both the selective substrate (succinate, left column) and the alternative substrate (2,4-D, right column). *, see Table 2.

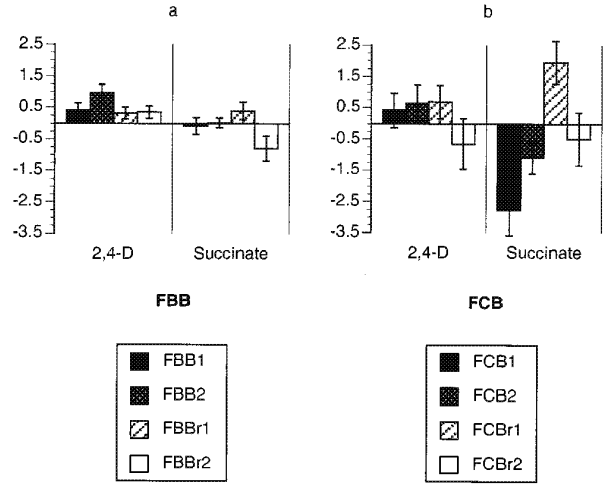


FIG. 3. Performance changes of stage II F-derived lines. (a) FBB lines; (b) FCB lines. Mean SRC values are shown, along with 95% confidence intervals, for both the selective substrate (2,4-D, left column) and the alternative substrate (succinate, right column).

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001217

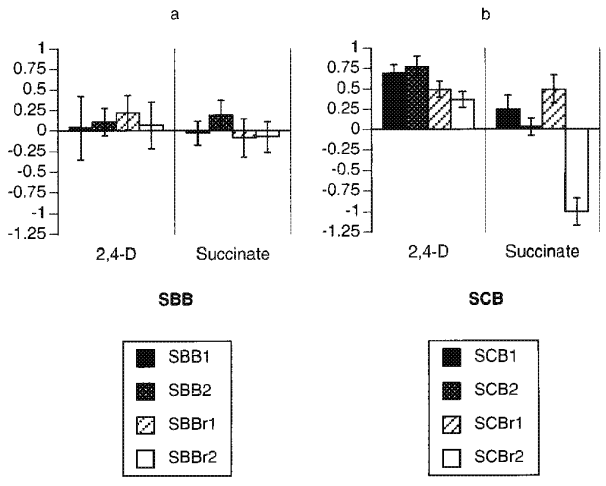


FIG. 4. Performance changes of stage II S-derived lines. (a) SBB lines; (b) SCB lines. Mean SRC values are shown, along with 95% confidence intervals, for both the selective substrate (2,4-D, left column) and the alternative substrate (succinate, right column).

Substrate-specific adaptation. The lines in this study may have increased their fitness by adaptations largely specific to their growth substrate during evolution, or alternatively, they may have improved by adapting to some other aspect of the selective regimen, such as the concentrations of various minerals in the medium. The latter possibility is especially relevant to the stage I lines. If most adaptation was not substrate specific, then evolved strains should be approximately equally superior to their ancestors on both selective and alternative substrates. In this study, however, many lines showed negative responses on their alternative substrate, indicating that their adaptation was largely substrate specific. Moreover, among those lines that improved on both substrates, in several cases the degrees of improvement were quite different for the two substrates, which also belies the general adaptation hypothesis.

DISCUSSION

Population dynamics of adaptation. The evolutionary changes in competitive performance on nonselective substrates (resulting from adaptation to a specific selective substrate) may be caused either by pleiotropic effects of adaptive mutations (43) or by the random drift of effectively neutral alleles (19). The latter possibility can be readily dismissed with respect to these experiments. Only a very small minority of mutations that affect fitness in a given environment are expected to be beneficial rather than harmful. It is much easier to disrupt than to improve the performance of a complex entity. That is, if one considers mutations that are selectively neutral in one environment (say, 2,4-D medium) but that affect fitness in an alternative environment (say, succinate medium), then the vast majority of the mutations should reduce, rather than improve, fitness in the alternative environment.

In this study, however, a majority of lines (five of eight) that underwent significant competitive changes (positive or negative) on both their selective and alternative substrates actually improved on both substrates. This outcome is extremely unlikely under the drift hypothesis. Even after performance of a sequential Bonferroni correction (a very conservative measure that corrects both for significance attributions due to chance and for any nonindependence of data points [38]), there remained four cases in which both performance responses were significant, and these were split evenly between trade-off and non-trade-off patterns. These data indicate that improvements in the alternative substrates for most lines were not due to substitution of random mutations by drift but instead were caused by the same mutations responsible for improvements in the selective regimens. In other words, the selected mutations were often beneficial in both regimens, indicating positive pleiotropy. This being the case, it seems likely that the cases of significant performance loss on the alternative substrate (associated with significant performance improvement on the selective substrate) were also due to pleiotropy, in this case negative pleiotropy.

It is also likely that each line's improvement in its selective environment was due to only one or a few adaptive mutations. This small number is because the duration of experimental evolution in this study was relatively short (maximum of 500 generations during either stage of evolution). Mathematical models indicate that many generations (often several hundred) are required for each beneficial mutation to sweep sequentially through a large, clonal population (28). The results of previous evolution experiments with *E. coli* support the predictions of these models (14, 29).

Multiple adaptive pathways. The results of this study clearly show that replicate evolving populations often achieved adaptation to their selective substrate by different mechanisms. This inference follows from the facts that in three of the four stage II sets (each containing four independent replicate lines), at least one line showed significant correlated improvements on both the selective and the alternative substrate and that at least one other line showed a significant loss of performance on its alternative substrate after adaptation to its selective substrate (Table 2; Fig. 3 and 4). In other words, the variation in competitive fitness among replicate populations on alternative substrates (standard deviation = 1.3394) was significantly greater ($F = 14.581$, 38 df, $P < 0.0001$ [F test statistic to compare variances]) than such variation on selective substrates (standard deviation = 0.3508) (43).

Even for replicate lines within a set that showed the same pattern of performance changes on both the selective and the alternative substrate, there was sometimes additional evidence that these changes were caused by different underlying physiological mechanisms. For example, lines SCB1 and SCBr1 both adapted significantly to their selective substrate (2,4-D) and they both showed significant correlated improvements on their alternative substrate (succinate). However, data presented by Velicer and Lenski (45) suggest that one of these lines, SCB1,

TABLE 3. Frequency of trade-off versus non-trade-off patterns of adaptation

Adaptive pattern	No. of lines showing[a]:	
	Trade-off	Non-trade-off
Significant responses in both selective and alternative regimens	3	5
Significant response in selective regimen only	2	3
Significant response in alternative regimen only	2	1
Significant response in neither regimen	2	1

[a] Each adaptive pattern, grouped according to whether the corresponding SRCs are significant (Tables 1 and 2).

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001218

had improved performance in the chemostat culture regimen during adaptation to the batch regimen but that the other strain, SCBr1, appeared to have reduced performance in chemostat culture. (The difference in chemostat performance between each strain and its proximate ancestor was not significant, but the resulting difference between the two evolved strains was highly significant [$t = 8.298$; 8 df; two-tailed t test, $P < 0.0001$] [Student's t test statistic]). Assuming that these changes in chemostat performance reflect the pleiotropic effects of mutations that were adaptive in the batch regimen, then these data imply that dual improvements of SCB1 and SCBr1 had occurred by different underlying mechanisms.

The stage I selection history of clones used to initiate stage II lines does not appear to have had a significant effect on the pattern of adaptive changes for stage II lines. Thus, among the eight stage II lines with a stage I batch history, three lines showed or suggested improvement on their alternative substrate (succinate) after stage II evolution whereas five showed or suggested performance decreases. Similarly, among stage II lines with a stage I chemostat history, four lines improved and four became worse on their stage II alternative substrate. (There were not enough line replicates during stage I evolution to test for an ancestor effect on adaptation patterns.)

The replicate lines that indicate different mechanisms of adaptation to 2,4-D in batch culture were founded from base populations that were initially isogenic. This finding corroborates the results of Travisano et al. (43), which showed that natural selection can cause significant divergence of populations within a species even when the populations are founded from the same progenitor and experience identical environments. Such divergence is presumably due to the fact that the replicate lines incur different sequences of random mutations during evolution, thus providing distinct patterns of genetic variation across populations upon which natural selection may act. Potential physiological mechanisms of adaptation have been discussed by Velicer (44).

Relevance to the adaptation of genetically modified organisms in natural habitats. In the discussion of relative benefits and risks of releasing genetically modified organisms into nature, it has often been argued that genetically altered microorganisms will not persist because they are inherently less fit than indigenous competitors (6, 10, 11). Lenski (25) reviewed various arguments for why genetically modified organisms might theoretically be less fit than their natural counterparts, and he summarized some experiments with bacteria that bear on those arguments.

Most of the studies reviewed by Lenski (25) support the view that modified microorganisms are competitively inferior to the strains from which they are derived. Mutations that alter basic metabolic functions (20), the expression of additional functions (1, 12, 24), the carriage of accessory genetic elements (27), and the domestication of bacteria as they adapt to laboratory conditions (11, 36) all tend to decrease the fitness of modified organisms relative to that of their progenitors under natural conditions, although important exceptions to this generality exist (2, 3, 13, 17). Beyond the immediate fitness effects of genetic manipulations that occur prior to release, there is the further possibility that genetically modified organisms adapt evolutionarily to the local ecosystem in which they are released. Such postrelease adaptation seems more likely than fortuitous preadaptation prior to release, and it may lead to the indefinite persistence of modified organisms in the environment (along with any adverse effects they might cause).

Several studies, including three with bacteria in the laboratory (4, 24, 34) and one with insects in the field (33), indicate that organisms may evolve so as to mitigate the fitness costs associated with other genetic changes. In one case, a population of *E. coli* carried a plasmid, pACYC184, that bore genes that encoded resistance to two antibiotics but that reduced the bacterium's fitness in the absence of antibiotic (4). These plasmid-bearing cells were then propagated in medium that contained antibiotic, in order to prevent spontaneous plasmid-free segregants from outcompeting the plasmid-bearing cells and taking over the population. After only 500 generations, the plasmid-bearing cells were, surprisingly, more fit than both their plasmid-free progenitor and their isogenic plasmid-free derivatives even in the absence of antibiotic (4, 26). In other words, evolution of a genetically modified organism in one environment (with antibiotic) led to a correlated improvement in fitness in an alternative environment (without antibiotic).

The above results point towards a second concern over the evolution of microorganisms that have been genetically modified for bioremediation in the environment. What is the effect of adaptation to an anthropogenic substrate, such as 2,4-D, on their competitive fitness for naturally occurring substrates, such as succinate? If such adaptation may have pleiotropic benefits for growth on natural substrates, then this increases the likelihood that these genetically modified organisms will persist in the environment. For example, adaptations that enhance bacterial growth on exogenous 2,4-D in the rhizosphere (5) may simultaneously improve bacterial fitness during growth on succinate, an endogenous exudate of nitrogen-fixing nodules (30). The results of this study clearly show that evolutionary adaptation of bacteria to growth on the common toxic herbicide 2,4-D often improves their competitive fitness on the natural substrate succinate. This outcome may have been facilitated by the fact that succinate and 2,4-D are catabolically linked (the 2,4-D catabolic pathway enters the tricarboxylic acid cycle via succinate). If so, a similar evolutionary scenario can be expected for bacterial adaptation to other xenobiotics that also are degraded via an intermediate compound that also occurs as a natural growth substrate.

The strains employed in this investigation were already capable of degrading 2,4-D, without genetic modification in the laboratory. However, there is no obvious reason that qualitatively similar results might not be obtained for strains that are genetically modified to degrade chlorinated aromatic compounds (or perform some other bioremediation function) more efficiently than natural strains (7, 35). In fact, such modified organisms may have even more opportunities for general improvements in their overall vigor, owing to the fitness costs associated with their modification (see above). Therefore, when considering the fate of genetically modified organisms released into the environment for bioremediation of anthropogenic substrates, it seems unrealistic to assume that these organisms will simply do their job and then disappear.

The use of environmentally important substrates increases the relevance of these results to concerns about the evolutionary adaptation of genetically modified organisms subsequent to their release in the environment. One concern is that certain environmental conditions may cause bacteria that are degrading a toxic substrate to produce an even more toxic metabolite in the process. For example, bacteria may produce highly toxic vinyl chloride during the biodegradation of tetrachloroethylene under methanogenic conditions (47). In a similar vein, genetically modified bacteria that are used to degrade some toxic compound may evolve novel degradative pathways that yield a more toxic metabolite even as they continue to degrade the intended substrate. Concerns over these and other potential adverse effects are magnified if there is a significant likelihood that the released organisms may persist indefinitely in the environment (25, 37, 40). And as the organisms adapt to

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001219

the natural environment in which they were released, the likelihood that they can persist will increase.

ACKNOWLEDGMENTS

I thank E. Smalley for assistance with the experiments. I also thank R. Lenski, M. Travisano, A. de Visser, and an anonymous reviewer for advice and comments.

This research was funded by the NSF Center for Microbial Ecology (DEB-9120006).

REFERENCES

1. **Andrews, K. J., and G. D. Hegeman.** 1976. Selective disadvantage of nonfunctional protein synthesis in *Escherichia coli*. J. Mol. Evol. **8:**317–328.
2. **Biel, S. W., and D. L. Hartl.** 1983. Evolution of transposons: natural selection for Tn*5* in *Escherichia coli* K12. Genetics **103:**581–592.
3. **Blot, M., B. Hauer, and G. Monnet.** 1994. The Tn*5* bleomycin resistance gene confers improved survival and growth advantage on *Escherichia coli*. Mol. Gen. Genet. **242:**595–601.
4. **Bouma, J. E., and R. E. Lenski.** 1988. Evolution of a bacteria/plasmid association. Nature **335:**351–352.
5. **Boyle, J. J., and J. R. Shann.** 1995. Biodegradation of pheno, 2,4-DCP, 2,4-D, and 2,4,5-T in field collected rhizosphere and nonrhizosphere soils. J. Environ. Qual. **24:**782–785.
6. **Brill, W. J.** 1985. Safety concerns and genetic engineering in agriculture. Science **227:**381–384.
7. **Chaudhry, G. R., and S. Chapalamadugu.** 1991. Biodegradation of halogenated organic compounds. Microbiol. Rev. **55:**59–79.
8. **Cohan, F. M.** 1984. Can uniform selection retard random genetic divergence between isolated conspecific populations? Evolution **38:**495–504.
9. **Cohan, F. M., and J.-D. Graf.** 1985. Latitudinal cline in *Drosophila melanogaster* for knockdown resistance to ethanol fumes and for rates of response to selection for further resistance. Evolution **39:**278–293.
10. **Davies, J.** 1988. Genetic engineering: processes and products. Trends Ecol. Evol. **3:**S7–S11.
11. **Davis, B. D.** 1987. Bacterial domestication: underlying assumptions. Science **235:**1329–1335.
12. **Dykhuizen, D. E.** 1978. Selection for tryptophan auxotrophs of *Escherichia coli* in glucose limited chemostats as a test of the energy conservation hypothesis of evolution. Evolution **32:**125–150.
13. **Edlin, G., R. C. Tait, and R. L. Rodriguez.** 1984. A bacteriophage Lambda cohesive ends (*cos*) DNA fragment enhances the fitness of plasmid-containing bacteria growing in energy-limited chemostats. Bio/Technology **2:**251–254.
14. **Elena, S. F., V. S. Cooper, and R. E. Lenski.** 1996. Punctuated evolution caused by selection of rare beneficial mutations. Science **272:**1802–1804.
15. **Forney, L.** 1997. Personal communication.
16. **Guttman, D. S., and D. E. Dykhuizen.** 1994. Detecting selective sweeps in naturally occurring *Escherichia coli*. Genetics **138:**993–1003.
17. **Hartl, D. L., D. E. Dykhuizen, R. D. Miller, L. Green, and J. deFramond.** 1983. Transposable element IS*50* improves growth rate of *E. coli* cells without transposition. Cell **35:**503–510.
18. **Hoffman, A. A., and F. M. Cohan.** 1987. Genetic divergence under uniform selection. III. Selection for knockdown resistance to ethanol in *Drosophila pseudoobscura* populations and their replicate lines. Heredity **58:**425–433.
19. **Kimura, M.** 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, United Kingdom.
20. **Koch, A. L.** 1983. The protein burden of *lac* operon products. J. Mol. Evol. **19:**455–462.
21. **Korona, R.** 1996. Adaptation to structurally different environments. Proc. R. Soc. Lond. B **263:**1665–1669.
22. **Korona, R.** 1996. Genetic divergence and fitness convergence under uniform selection in experimental populations of bacteria. Genetics **143:**637–644.
23. **Korona, R., C. H. Nakatsu, L. J. Forney, and R. E. Lenski.** 1994. Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat. Proc. Natl. Acad. Sci. USA **91:**9037–9041.
24. **Lenski, R. E.** 1988. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. I. Variation in competitive fitness among mutants resistant to virus T4. Evolution **42:**425–432.
25. **Lenski, R. E.** 1993. Evaluating the fate of genetically modified microorganisms in the environment: are they inherently less fit? Experientia **49:**201–209.
26. **Lenski, R. E.** 1997. Personal communication.
27. **Lenski, R. E., and J. E. Bouma.** 1987. Effects of segregation and selection on instability of plasmid pACYC184 in *Escherichia coli* B. J. Bacteriol. **169:**5314–5316.
28. **Lenski, R. E., M. R. Rose, S. C. Simpson, and S. C. Tadler.** 1991. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. Am. Nat. **138:**1315–1341.
29. **Lenski, R. E., and M. Travisano.** 1994. Dynamics of adaptation and diversification: a 10,000 generation experiment with bacterial populations. Proc. Natl. Acad. Sci. USA **91:**6808–6814.
30. **Lipton, D. S., R. W. Blanchar, and D. G. Blevins.** 1987. Citrate, malate, and succinate concentration in exudates from phosphorus-sufficient and phosphorus-stressed *Medicago sativa L.* seedlings. Plant Physiol. (Bethesda) **85:**315–317.
31. **Maynard-Smith, J.** 1991. The population genetics of bacteria. Proc. R. Soc. Lond. B **245:**37–42.
32. **McGowan, C.** 1995. Interspecies gene transfer in the evolution of 2,4-dichlorophenoxyacetate degrading bacteria. Ph.D. dissertation. Michigan State University, East Lansing.
33. **McKenzie, J. A., M. J. Whitten, and M. A. Adena.** 1982. The effect of genetic background on the fitness of diazinon resistance genotypes of the Australian sheep blowfly, *Lucilia cuprina*. Heredity **49:**1–9.
34. **Modi, R. I., and J. Adams.** 1991. Coevolution in bacteria-plasmid populations. Evolution **45:**656–667.
35. **Ramos, J. L., A. Wasserfallen, K. Rose, and K. N. Timmis.** 1987. Redesigning metabolic routes: manipulation of TOL plasmid pathway for catabolism of alkylbenzoates. Science **235:**593–596.
36. **Regal, P. J.** 1988. The adaptive potential of genetically engineered organisms. Trends Ecol. Evol. **3:**S36–S38.
37. **Regal, P. J.** 1993. The true meaning of 'exotic species' as a model for genetically engineered organisms. Experientia **49:**225–234.
38. **Rice, W. R.** 1989. Analyzing tables of statistical tests. Evolution **43:**223–225.
39. **Sassanella, T.** 1997. Personal communication.
40. **Tiedje, J. M., R. K. Colwell, Y. L. Grossman, R. Hodson, R. E. Lenski, R. N. Mack, and P. J. Regal.** 1989. The planned introduction of genetically engineered organisms: ecological considerations and recommendations. Ecology **70:**298–315.
41. **Tonso, N. L., V. G. Matheson, and W. E. Holben.** 1995. Polyphasic characterization of a suite of bacterial isolates capable of degrading 2,4-D. Microb. Ecol. **30:**3–24.
42. **Travisano, M., and R. E. Lenski.** 1996. Long-term experimental evolution in *Escherichia coli*. IV. Targets of selection and the specificity of adaptation. Genetics **143:**15–26.
43. **Travisano, M., F. Vasi, and R. E. Lenski.** 1995. Long-term experimental evolution in *Escherichia coli*. III. Variation among replicate populations in correlated responses to novel environments. Evolution **49:**189–200.
44. **Velicer, G. J.** 1997. Ph D. thesis. Michigan State University, East Lansing.
45. **Velicer, G. J., and R. E. Lenski.** Evolutionary tradeoffs under conditions of resource abundance and scarcity: experiments with bacteria. Ecology, in press.
46. **Versalovic, J., T. Koeuth, and J. R. Lupski.** 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res. **19:**6823–6831.
47. **Vogel, T. M., and P. L. McCarty.** 1985. Biotransformation of tetrachloroethylene to trichloroethylene, dichloroethylene, vinyl chloride, and carbon dioxide under methanogenic conditions. Appl. Environ. Microbiol. **49:**1080–1083.

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001220

Metal Coordination Environment of a Cu(II)-Substituted α-Keto Acid-Dependent Dioxygenase That Degrades the Herbicide 2,4-D

Adam K. Whiting and Lawrence Que, Jr.*

Department of Chemistry and Center for Metals in Biocatalysis, University of Minnesota Minneapolis, Minnesota 55455

Ruth E. Saari and Robert P. Hausinger*

Departments of Microbiology and Biochemistry Michigan State University East Lansing, Michigan 48824

Matthew A. Fredrick and John McCracken*

Department of Chemistry, Michigan State University East Lansing, Michigan 48824

Received December 30, 1996

Downloaded via CORTEVA AGRISCIENCE on September 3, 2020 at 21:26:09 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

α-Keto acid-dependent enzymes utilize a mononuclear nonheme iron center to activate dioxygen for a diverse range of metabolic processes involving the oxidation of an unactivated C−H bond.[1] One member of this class, the (2,4-dichlorophenoxy)acetate/α-ketoglutarate dioxygenase (TfdA) encoded on the pJP4 plasmid of *Alcaligenes eutrophus* JMP134 catalyzes the conversion of the herbicide (2,4-dichlorophenoxy)acetate (2,4-D) to 2,4-dichlorophenol and glyoxylate, coupled with the oxidative decarboxylation of α-ketoglutarate (α-KG) to succinate and CO_2, as shown in Scheme 1.[2] Previously, we have cloned and expressed the *tfdA* gene in *Escherichia coli* and purified the 64 kDa homodimeric protein to homogeneity.[3] Chemical modification studies indicate that multiple histidines, but no cysteines, are necessary for activity; otherwise there is scant information on the active site.[3] In this paper, we describe EPR and ESEEM studies of Cu(II)-substituted TfdA that provide the first insights into the metal coordination environment of the activated complex of an α-keto acid-dependent enzyme.

Steady-state kinetic analysis of TfdA demonstrates that Cu(II) is a very good competitive inhibitor versus Fe(II) (K_i = 0.1−0.3 μM), strongly implying that Cu(II) occupies the same binding site. Indeed, its K_i, which is 100-fold lower than the concentration of Fe(II) required for half-maximal activity (10−30 μM),[3] suggests tighter binding by Cu(II). Thus, spectroscopic analysis of the Cu-substituted enzyme is likely to reflect important features of the native Fe(II) enzyme.

Figure 1 shows X-band EPR spectra of Cu(II)-substituted TfdA (Cu-TfdA) alone and in the presence of the cofactor α-KG and the substrate 2,4-D. The EPR spectrum of Cu-TfdA has $g_{\|}$ (2.29) > $g_{\perp}$ (2.07) and $A_{\|}$ = 16.1 mK (Figure 1A), indicating that the Cu(II) center resides in a type 2 environment with a mixture of N and O ligands in the equatorial plane.[4] These EPR parameters are altered upon addition of the cofactor α-KG (Figure 1B; $g_{\perp}$ = 2.07, $g_{\|}$ = 2.36, $A_{\|}$ = 14.9 mK); however, the addition of the substrate 2,4-D does not affect the Cu-TfdA spectrum unless α-KG is also present (Figure 1C). The ternary Cu-TfdA/α-KG/2,4-D complex exhibits a more rhombic EPR signal with $g_{\|}$ = 2.37 and a significantly decreased $A_{\|}$ value of 12.0 mK. Very similar EPR spectra are observed when either of the α-KG analogs, pyruvate (K_m = 1.02 mM,[3] $g_{\|}$ = 2.36, $A_{\|}$ = 12.8 mK) or glutarate (K_i = 0.45 mM, $g_{\|}$ = 2.37, $A_{\|}$ = 12.5 mK) is used in forming the ternary complex. These comparisons show that only the 1-carboxylate of the α-KG cofactor is crucial for producing the distinctive EPR features of the ternary complex. For type 2 Cu(II) centers, these high $g_{\|}$ values (>2.35) are associated with an oxygen-rich environment,[4] while these small $A_{\|}$ values (<14.0 mK) indicate a significant distortion from planarity.[5,6] Comparable EPR parameters have been observed for Cu-substituted enolase complexed with 2-phosphoglycerate[7] and Cu-substituted ribulose-1,5-bisphosphate carboxylase/oxygenase complexed with 2-carboxyarabinitol bisphosphate.[8] The crystal structures of both of these complexes[9,10] show metal centers with distorted carboxylate-rich 6O ligand environments, which rationalize their unusually sharp signals (~1 mT fwhm), high $g_{\|}$ values (~2.4), and small $A_{\|}$ values (~12 mK).[7,8] The similarly small $A_{\|}$ value of the Cu-TfdA/α-KG/

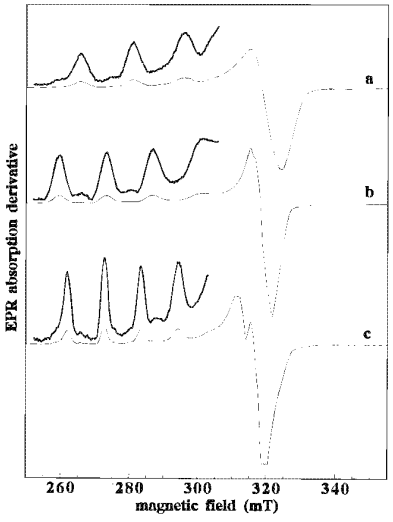
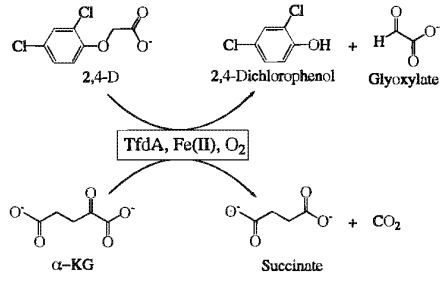


Figure 1. EPR spectra of (A) Cu-TfdA ($g_{\perp}$ = 2.07, $g_{\|}$ = 2.29, $A_{\|}$ = 16.1 mK, 4.2 mT fwhm at 273 mT), (B) Cu-TfdA + 5 mM α-KG ($g_{\perp}$ = 2.07, $g_{\|}$ = 2.36, $A_{\|}$ = 14.9 mK, 4.2 mT fwhm at 273 mT), and (C) Cu-TfdA + 5 mM α-KG + 5 mM 2,4-D ($g_{\perp}$ = 2.08, $g_{\|}$ = 2.37, $A_{\|}$ = 12.0 mK, 2.2 mT fwhm at 273 mT). Insets show $g_{\|}$ with intensity magnified 5 times. Spin quantitation indicates that all of the added Cu is bound in all three cases. Samples were prepared by slow addition of $CuCl_2$ (450 μM in H_2O) to apo-TfdA (500 μM subunit concentration in 25 mM MOPS buffer, pH 6.9) followed by addition of substrate solution (50 mM substrate in 25 mM MOPS buffer, pH 6.9). Samples were then inserted in 0.3 mm i.d. quartz tubes and frozen by slow immersion in liquid N_2. EPR spectra were obtained at 20 K as single 4 min scans from 220 to 380 mT using 0.20 mW power at 9.23 GHz, 1.0 mT modulation amplitude.

Scheme 1



(1) For reviews, see: (a) Que, L., Jr.; Ho, R. Y. N. *Chem. Rev.* **1996**, *96*, 2607−2624. (b) Prescott, A. G.; John, P. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1996**, *47*, 245−271. (c) Udenfriend, S.; Cardinale, G. In *Oxygenases and Oxygen Metabolism*; Nozaki, M., Yamamoto, S., Ishimura, Y., Coon, M. J., Ernster, L., Estabrook, R. W., Eds.; Academic Press: New York, 1982; pp 99−110.

(2) Fukumori, F.; Hausinger, R. P. *J. Bacteriol.* **1993**, *175*, 2083−2086.

(3) Fukumori, F.; Hausinger, R. P. *J. Biol. Chem.* **1993**, *268*, 24311−24317.

(4) Peisach, J.; Blumberg, W. E. *Arch. Biochem. Biophys.* **1974**, *165*, 691−708.

(5) Rosenberg, R. C.; Root, C. A.; Bernstein, P. K.; Gray, H. B. *J. Am. Chem. Soc.* **1975**, *97*, 2092−2096.

(6) Addison, A. W. In *Copper Coordination Chemistry: Biochemical and Inorganic Perspectives*; Karlin, K. D., Zubieta, J., Eds.; Adenine Press: New York, 1983; pp 109−128.

(7) Dickinson, L. C.; Rose, S. L.; Westhead, E. W. *J. Inorg. Biochem.* **1980**, *13*, 353−366.

S0002-7863(96)04449-6 CCC: $14.00 © 1997 American Chemical Society

COR-AAD1-00001221

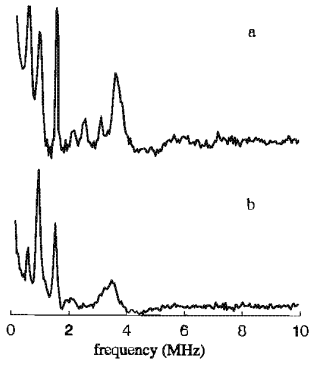


Figure 2. Three-pulse ESEEM data of (A) Cu-TfdA + 5 mM α-KG and (B) Cu-TfdA + 5 mM α-KG + 5 mM 2,4-D. Samples were prepared as described in Figure 1. Conditions for these measurements: (A) microwave frequency 8.98 GHz, magnetic field strength 280 mT, τ value 375 ns, number of scans averaged 1; (B) microwave frequency 8.92 GHz, magnetic field strength 275 mT, τ value 380 ns, number of scans averaged 4. Conditions common to both measurements were sample temperature (4.2 K), microwave pulse power (32 W), pulse width (20 μs, fwhm), and each time point (represents the average of 34 events).

2,4-D complex indicates a comparably distorted environment; however, its somewhat lower $g_{||}$ and broader signals (2.2 mT fwhm) suggest the presence of an endogenous nitrogen ligand.

ESEEM experiments have proven useful for establishing the presence of imidazole ligands strongly bound to a Cu center.[11] Three-pulse ESEEM spectra of Cu-TfdA complexes all exhibit sharp peaks at 0.6, 0.9, and 1.5 MHz together with a broader feature at 3.5 MHz, typical of the distal nitrogen of imidazole bound equatorially to Cu (Figure 2).[11] In addition, the spectra of Cu-TfdA (not shown) and Cu-TfdA/α-KG (Figure 2A) show weak "combination" peaks at 2.1, 2.5, and 3.1 MHz indicative of the presence of at least two imidazole ligands in this complex.[12] The ESEEM spectrum of the ternary complex (Figure 2B) shows a substantial decrease in modulation intensity and a lack of the "combination" peaks observed for the Cu-TfdA and Cu-TfdA/α-KG samples. This finding shows that the binding of 2,4-D has altered the hyperfine coupling to one of the histidyl imidazole ligands. While this might indicate displacement of one of the bound histidines, there is also a change in the relative intensities of the "fundamental" ESEEM peaks at 0.6, 0.9, and 1.5 MHz, which suggests a shift in the relative orientation of the *g*-tensor principal axis system. Hence, the formation of the ternary complex may redefine the tetragonal axis so that one of the histidyl imidazole ligands is now coordinated in an axial position and contributes minimally to the spectrum.

Figure 3 shows our working model for the metal coordination sphere in Cu-TfdA: a facial array of three endogenous protein ligands together with three coordination sites available for exogenous ligands. The endogenous His ligands in the equatorial plane are deduced from the ESEEM data, consistent with chemical modification studies of Fe-TfdA.[3] The axial Asp ligand, not yet corroborated by experiment, is proposed on the basis of its appearance in the TfdA sequence at a position (D115) between two His residues (H113 and H167) that fits well into the His-X-Asp-$X_{(53-57)}$-His motif that is common to α-keto acid-dependent dioxygenases and isopenicillin N synthase (IPNS).[13] Such a facial triad of 1 carboxylate and 2 His ligands binds the metal centers in IPNS[14] and 2,3-dihydroxybiphenyl 1,2-dioxygenase (an extradiol cleaving catechol dioxygenase),[15] and is emerging as the minimal coordinative scaffolding for mononuclear nonheme iron(II) centers that activate oxygen.[1a]

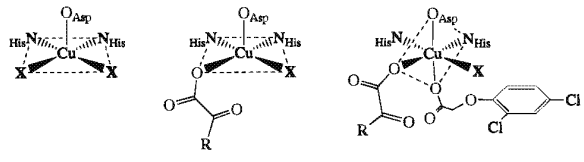


Figure 3. Working model for the metal coordination sphere in Cu-TfdA. The X marks coordination sites proposed to be occupied by solvent and/or endogenous ligands which are available for binding by exogenous ligands. The dashed lines define the equatorial plane of the Cu(II) center which gives rise to the observed EPR and ESEEM spectral features. This equatorial plane is proposed to reorient by 90° upon formation of the ternary complex.

As found for IPNS and extradiol-cleaving catechol dioxygenases,[1a,16,17] we propose that the remaining three coordination sites for Cu-TfdA are occupied by solvent (and/or other endogenous ligands) and are displaced upon binding α-KG, 2,4-D, and other exogenous ligands.[18] The highly distinctive spectral properties of the ternary complex suggest coordination of the carboxylates of both α-KG and 2,4-D affording a Cu equatorial ligand set consisting of a His residue and three carboxylates. This ligand set, which includes the putative Asp ligand, is proposed to result from a reorientation of the equatorial plane upon formation of the ternary complex (Figure 3) and poises the metal center to bind O_2 at the remaining sixth coordination site during catalysis. The fact that only the combination of cofactor and substrate elicits the characteristic ternary complex EPR signal is consistent with the tight coupling between the oxidative decarboxylation of α-KG and the cleavage of 2,4-D during catalysis.[2] It also suggests the existence of a synergistic interaction that very likely activates the native Fe(II) site for dioxygen binding as proposed in the enzyme mechanism.[19,20] These results thus provide the first direct glimpse of the metal coordination environment of an α-keto acid-dependent enzyme as it goes through its catalytic cycle; the insights we have obtained lay the foundation for future spectroscopic studies of the iron enzyme.

Acknowledgment. This research was supported by NIH grants GM 33162 (L.Q.) and GM 54065 (J.M.), NSF Science and Technology Center Grant BIR 9120006 (R.P.H.), and the Michigan State University Agricultural Experiment Station (R.P.H.). We also thank Professor Stenbjorn Styring for providing timely insights on the interpretation of the EPR spectra.

JA964449X

(8) Styring, S.; Branden, R. *Biochemistry* **1985**, *24*, 6011−6019.
(9) Andersson, I. *J. Mol. Biol.* **1996**, *259*, 160−174.
(10) Larsen, T. M.; Wedekind, J. E.; Rayment, I.; Reed, G. H. *Biochemistry* **1996**, *35*, 4349−4358.
(11) Mims, W. B.; Peisach, J. *J. Chem. Phys.* **1978**, *69*, 4921−4930.
(12) McCracken, J.; Pember, S.; Benkovic, S. J.; Villafranca, J. J.; Miller, R. J.; Peisach, J. *J. Am. Chem. Soc.* **1988**, *110*, 1069−1074.
(13) Borovok, I.; Landman, O.; Kreisberg-Zakarin, R.; Aharonowitz, Y.; Cohen, G. *Biochemistry* **1996**, *35*, 1981−1987.
(14) Roach, P. L.; Clifton, I. J.; Fulop, V.; Harlos, K.; Barton, G. J.; Hajdu, J.; Andersson, I.; Schofield, C. J.; Baldwin, J. E. *Nature* **1995**, *375*, 700−704.
(15) Sugiyama, K.; Senda, T.; Narita, H.; Yamamoto, T.; Kimbara, K.; Fukuda, M.; Yano, K.; Mitsui, Y. *Proc. Jpn. Acad.* **1995**, *71(B)*, 32−35. Han, S.; Eltis, L. D.; Timnuis, K. N.; Muchmore, S. W.; Bolin, J. T. *Science* **1995**, *270*, 976−980. Senda, T.; Sugiyama, K.; Narita, H.; Yamamoto, T.; Kimbara, K.; Fukuda, M.; Sato, M.; Yano, K.; Mitsui, Y. *J. Mol. Biol.* **1996**, *255*, 735−752.
(16) Lipscomb, J. D.; Orville, A. M. In *Metal Ions in Biological Systems*; Sigel, H., Sigel, A., Eds.; Marcel Dekker, Inc.: New York, 1992; Vol. 28, pp 243−298.
(17) Arciero, D. M.; Lipscomb, J. D. *J. Biol. Chem.* **1986**, *261*, 2170−2178.
(18) The EPR spectra of both Cu-TfdA and Cu-TfdA/α-KG are altered upon addition of 1.8 mM imidazole in a manner that suggests the displacement of solvent molecules by these exogenous N-ligands (Cu-TfdA/imidazole $g_{||} = 2.28$, $A_{||} = 16.9$ mK; Cu-TfdA/α-KG/imidazole $g_{||} = 2.30$, $A_{||} = 16.2$ mK). In contrast, the Cu-TfdA/α-KG/2,4-D EPR spectrum is completely unchanged in the presence of imidazole suggesting that access to the active site is limited by the formation of the ternary complex.
(19) Hanauske-Abel, H. M.; Günzler, V. *J. Theor. Biol.* **1982**, *94*, 421−455.
(20) Chiou, Y.-M.; Que, L., Jr. *J. Am. Chem. Soc.* **1995**, *117*, 3999−4013.

COR-AAD1-00001222

ELSEVIER

FEMS Microbiology Ecology 31 (2000) 53–60

FEMS MICROBIOLOGY ECOLOGY

www.fems-microbiology.org

Phylogenetic and phenotypic diversity of 4-chlorobenzoate-degrading bacteria isolated from soils

Haak-Rho Yi [a], Kyung-Hee Min [b], Chi-Kyung Kim [c], Jong-Ok Ka [a,*]

[a] *Department of Agricultural Biology, Seoul National University, Suwon 441-744, South Korea*
[b] *Department of Biology, Sookmyung Women's University, Seoul 140-742, South Korea*
[c] *Department of Microbiology, Chungbuk National University, Cheongju 360-763, South Korea*

Received 16 August 1999; received in revised form 25 September 1999; accepted 6 October 1999

Abstract

Twenty numerically dominant 4-chlorobenzoate (4-CBA)-degrading bacteria were isolated from agricultural soils. The isolates were able to utilize 4-CBA as a sole source of carbon and energy. A total of 65% of the isolates was identified to the species level by fatty acid methyl ester (FAME) analysis, and the isolates were strains of *Micrococcus*, *Pseudomonas*, *Oerskovia*, *Cellulomonas*, and *Arthrobacter* species. The chromosomal DNA patterns of the isolates obtained by polymerase chain reaction (PCR) amplification of repetitive extragenic palindromic (REP) sequences were distinct from each other. Most of the isolates grew rapidly in 4-CBA medium, but their substrate utilization capabilities were generally restricted. Plasmid DNAs were detected from 55% of the isolates, and one strain, HR7, was shown to have self-transmissible, 4-CBA degradative plasmids. 4-CBA degradative enzymes were inducible by the presence of 4-CBA and most of the isolates appeared to mineralize it through 4-hydroxybenzoate rather than 4-chlorocatechol. © 2000 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.

Keywords: 4-Chlorobenzoate; Biodegradation; Bacterial diversity

1. Introduction

Large amounts of man-made chlorinated organic chemicals have been introduced into the environment as agricultural pesticides and industrial agents. These compounds are classified among the most persistent and toxic group due to their chemical structures. Among them, chlorobenzoates are released into the environment as intermediates during biodegradation of polychlorinated biphenyls [1] and herbicides [2]. Most bacteria which degrade biphenyls and polychlorinated biphenyls accumulate the corresponding chlorobenzoic acids [3,4], and thus chlorobenzoate degradation can be the rate-limiting step in the biodegradation process of these biphenyl compounds [5].

4-Chlorobenzoate (4-CBA) is water-soluble and has low toxicity, which enables many researchers to use it as the model compound in studying the biodegradation of halogenated aromatic compounds. Organisms that have been reported to be capable of degrading 4-CBA belong to several genera, including *Arthrobacter*, *Pseudomonas*, *Alcaligenes*, and *Nocardia* [6–9]. A microbial consortium is also described for degradation of 4-CBA under anoxic conditions [10]. 4-CBA is metabolized either via formation of 4-hydroxybenzoate (4-HBA) by dechlorination or via 4-chlorocatechol (4-CC) by benzoate dioxygenation [11]. Furthermore, the enzymes and genes involved in the dechlorination have been studied and described [12–14].

Much of the previous studies on 4-CBA-degrading bacteria, however, have mainly focused on degradative enzymes and pathways, and thus little information is available on the phenotypic and genetic properties of 4-CBA-utilizing isolates from the environments. In addition, the inconsistencies in strain identification and the heterogeneity of the strain sources in previous works make it difficult to suggest a conclusion on their ecological abundance and phylogenetic diversity.

In this study, we isolated 20 numerically dominant 4-CBA-degrading bacteria from agricultural soils of various locations in South Korea. We investigated their species diversity using fatty acid profiles in cell membrane and REP-PCR patterns of chromosome, and described their

* Corresponding author. Fax: (0331) 294-8533;
E-mail: kagil@plaza.snu.ac.kr

0168-6496/00/$20.00 © 2000 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.
PII: S0168-6496(99)00081-1

Downloaded from https://academic.oup.com/femsec/article/31/1/53/532138 by Pennsylvania Hospital Interlibrary Loan on 03 September 2020

COR-AAD1-00001223

physiological and genetic properties on 4-CBA biodegradation.

2. Materials and methods

2.1. Isolation of bacterial strains

Agricultural soil samples were taken from upland fields and rice fields of various locations in South Korea. The agricultural fields selected in this study have been under normal agricultural practices for more than 10 years. Samples from the top 15 cm of field soil were taken and kept at 4°C prior to use. A 10-g soil sample from each site was homogenized with 95 ml of a sterilized 0.85% saline solution by shaking the preparation on a rotary shaker (200 rpm). Samples (0.1 ml) of appropriate 10-fold dilutions were inoculated into most-probable-number tubes containing 3 ml of SMB (standard mineral base) mineral medium [15] supplemented with 300 μg ml^{-1} of 4-CBA (4-CBA mineral medium). The tubes were incubated at 30°C for 3 weeks, and degradation of 4-CBA was analyzed by spectrophotometry and reverse-phase HPLC on a μBondapack C18 column (3.9 by 300 mm: Waters, Milford, MA, USA) with an acetic acid (pH 5.0, 6 mM), methanol eluant (60% organic, 40% aqueous phase) at a flow rate of 1.0 ml min^{-1}. The culture of the terminal positive tube showing 4-CBA degradation was enriched by two additional transfers into fresh medium. Each enriched culture was streaked onto PTYG (peptone-tryptone-yeast extract-glucose) agar medium, and single colonies were then tested for 4-CBA degradation in fresh 4-CBA mineral medium before strain purification.

2.2. Media and culture conditions

All isolates were maintained on SMB mineral medium containing 4-CBA at a concentration of 300 ppm (μg ml^{-1}). PTYG agar medium containing (per l) 0.25 g of peptone (Difco), 0.25 g of tryptone (Difco), 0.5 g of yeast extract (Difco), 0.5 g of glucose, 0.03 g of magnesium sulfate, 0.003 g of calcium chloride, and 15 g of agar was used for strain purification and colony production for the REP-PCR.

2.3. Chemicals

4-Chlorobenzoate, 4-bromobenzoate, 4-iodobenzoate, 3-chlorobenzoate, benzoate, 4-hydroxybenzoate, and 4-chlorocatechol were obtained from Aldrich Chemical Co., and analytical grade 2,4-dichlorophenoxyacetic acid (2,4-D) was obtained from Sigma Chemical Co.

2.4. FAME analysis

The isolates were cultured on tryptic soy agar medium at 28°C for 48 to 72 h, then the cells were harvested from the plates by scraping with a sterile glass loop and used for FAME analysis. Saponification, methylation, and extraction were performed by using the procedure described in the MIDI manual (Microbial Identification, Inc.) [16].

2.5. Colony REP-PCR

The colony REP-PCR was performed using BOXA1R as a primer as described previously [17]. Each isolate was grown on the PTYG agar medium for 24 to 48 h, then a small amount of the cells were resuspended in 25 μl of PCR mixture. After PCR amplification, 10-μl samples of the REP-PCR products were separated by electrophoresis on horizontal 1% agarose gels.

2.6. Axenic culture experiment

After growth in 4-CBA or acetate mineral medium to produce cells induced or not induced for 4-CBA metabolism, cells were harvested, washed twice with an equal volume of 15 mM phosphate buffer (pH 7.0), and resuspended in the same buffer. Aliquots of the suspended cells were inoculated into duplicate flasks containing 200 ml of 4-CBA (300 ppm) mineral medium at a final density of ca. $OD_{550} = 0.005$. All cultures were incubated at 30°C and were aerated by shaking at 200 rpm on a rotary shaker. Aliquots of the cultures were regularly removed to determine cell growth and degradation of 4-CBA.

Table 1
Identification of the 4-chlorobenzoate-degrading isolates by FAME analysis

Isolates[a]	Identity determined by FAME analysis
HR1	*Micrococcus luteus*
HR2	*Micrococcus luteus*
HR3	*Oerskovia* sp.
HR4	*Micrococcus* sp.
HR5	*Cellulomonas flavigena*
HR6	*Micrococcus varians*
HR7	*Pseudomonas* sp.
HR8	*Micrococcus kristinae*
HR9	*Micrococcus varians*
HR10	*Micrococcus luteus*
HR11	*Oerskovia* sp.
HR12	*Micrococcus halobius*
HR13	*Micrococcus varians*
HR14	*Micrococcus halobius*
HR15	*Arthrobacter* sp.
HR16	Unidentifiable[b]
HR17	*Arthrobacter protophormiae*
HR18	*Micrococcus luteus*
HR19	*Arthrobacter globiformis*
HR20	Unidentifiable

[a]Isolates were obtained from agricultural soils.
[b]The isolate could not be identified because of poor matches with profiles in the MIDI library.

Downloaded from https://academic.oup.com/femsec/article/31/1/53/532168 by Pearson Hi-Ered International user on 03 September 2020

COR-AAD1-00001224

2.7. Degradation phenotype analysis

Each strain was cultured in PTYG medium. Cells were then harvested, washed, and prepared in sodium phosphate buffer as described above. Aliquots of the suspended cells were inoculated into culture tubes, each of which contained SMB mineral medium supplemented with one of the structural analogs at a concentration of 300 ppm. After 3-week incubation, the cultures were centrifuged to remove the cellular material, and the UV absorption was measured to determine the degradation of chloroorganic compounds.

2.8. Plasmid detection and conjugation

For detection of plasmid DNA, cells were lysed using a modified form [18] of the procedure of Kado and Liu [19]. To analyze the transferability of the 4-CBA degradation phenotype of the isolates, matings were performed on membrane filters as described by Willetts [20]. Transconjugants were selected on 4-CBA mineral medium containing 300 ppm 4-CBA, appropriate antibiotics, and 1.5% Noble agar.

2.9. Analysis of 4-CBA degradation pathway

Strains were cultivated in mineral medium containing 4-CBA or acetate to produce cells induced or not induced for 4-CBA metabolism as described above. Cells were then inoculated at a final concentration of ca. $OD_{550} = 0.5$ into duplicate SMB mineral medium supplemented with 4-hydroxybenzoate or 4-chlorocatechol at a concentration of 100 ppm. The disappearance of the intermediate was monitored by spectrophotometry.

3. Results and discussion

3.1. Isolation of 4-CBA-utilizing bacteria

Twenty 4-CBA-degrading bacteria, which were capable of utilizing 4-CBA as a sole source of carbon and energy, were isolated through enrichment processes from upland and rice field soils of various locations in South Korea (Table 1). 4-CBA-degrading bacteria seemed to be relatively ubiquitous in agricultural soils, since 48% (20 soils) of the 42 soil samples tested in this study showed positive degradation of 4-CBA, producing a single isolate from each of the positive soils. Among the isolates, nine strains were obtained from upland soils and the other 11 strains from rice soils, suggesting that 4-CBA degraders are relatively evenly distributed among upland fields and rice fields. 4-CBA in mineral medium was completely degraded by the isolates, and no dead-end products were accumulated during biodegradation of 4-CBA when analyzed with high-performance liquid chromatography. Fig. 1 shows typical growth and degradation curves by strain HR4 on 4-CBA medium.

3.2. FAME and taxonomic analysis

The results of the FAME analysis grouped the 20 isolates into 10 groups of strains at Euclidian distances of less than 10 (Fig. 2); it has been suggested that this distance approximates the species level [16]. All of the isolates clustered at a Euclidian distance of 80.2. The isolates that could be reasonably identified (similarity index > 0.6) by the FAME results are shown in Table 1. 65% (13 strains) of the 20 isolates were identified to the species level and 55% of the isolates constituted a dominant group of *Mi-*

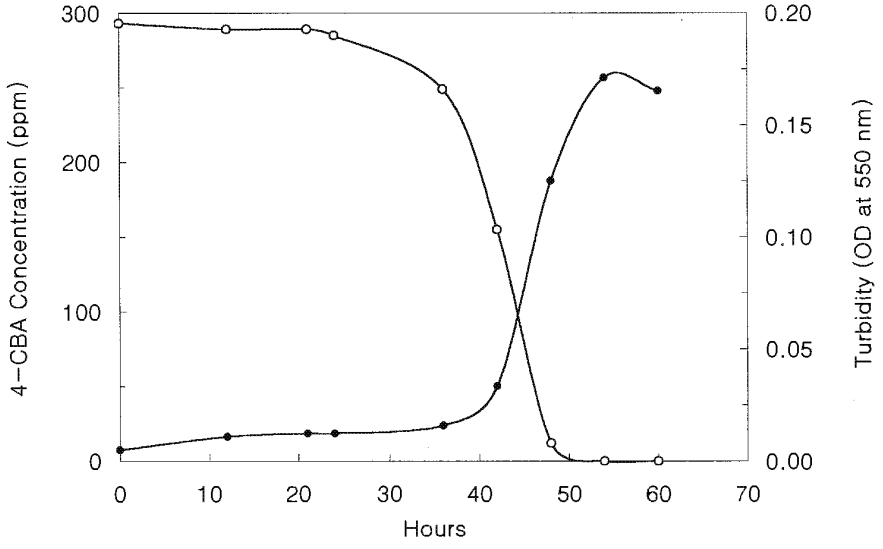


Fig. 1. Disappearance of 4-chlorobenzoate (○) and growth of bacteria (●) during degradation of 4-chlorobenzoate by strain HR4. Each point is the mean of two replicate liquid cultures. OD, optical density.

Downloaded from https://academic.oup.com/femsec/article/31/1/53/532188 by Pearson Hi-Ered International user on 03 September 2020

COR-AAD1-00001225

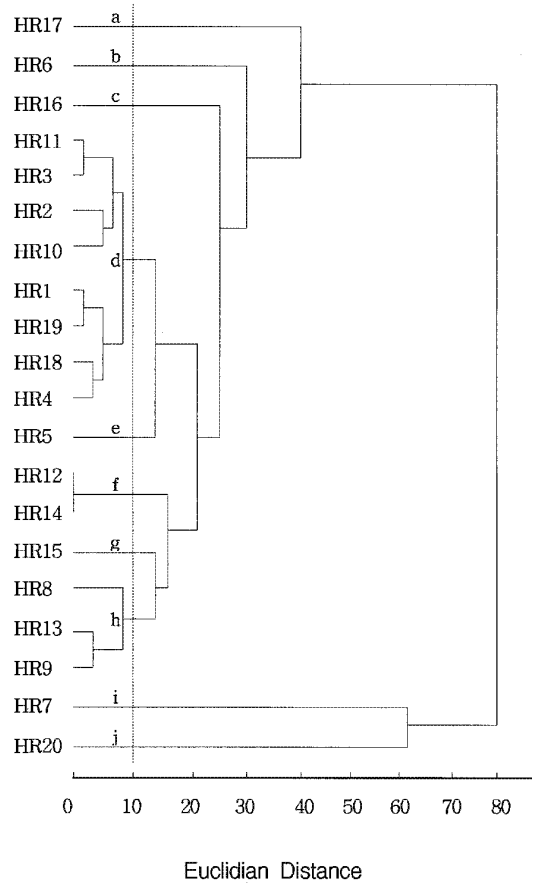


Fig. 2. Dendrogram based on the fatty acid profiles of the isolates. FAME groups a through j were defined at a Euclidean distance of less than 10.

crococcus species, which has not been reported as 4-CBA-degrading bacteria in the previous studies. The other strains belonged to *Arthrobacter*, *Oerskovia*, *Cellulomonas*, and *Pseudomonas* species. It is interesting to note that all of the *Micrococcus* spp., except strain HR9, were isolated from rice field soils. Unlike upland fields, the wet rice fields are characterized by submergence during part or all of the cropping period. Flooding of rice fields affects soil pH, redox state, and nutrient availability, and thus results in the differentiation of diverse microhabitats that can support the growth of a variety of microorganisms. Our results suggest that a new class of microorganisms, *Micrococcus* spp., are predominant among 4-CBA-degrading microorganisms in wetland rice fields, which are microbiologically and chemically unique from upland fields. The two strains HR16 and HR20 could not be identified due to their poor matches with profiles in the MIDI library.

3.3. REP-PCR analysis

A colony REP-PCR experiment was performed to study the genomic relatedness among the isolates belonging to the same or closely related FAME groups. The analysis of REP-PCR band patterns revealed that the 20 isolates produced 19 different DNA fingerprint patterns (Fig. 3), suggesting that most of them were very diverse in their chromosome structures even among the isolates identified as the same species. Of the isolates, strains HR12 and HR14, which were identified as *Micrococcus halobius* by the FAME method, exhibited identical REP-PCR DNA band patterns, though they were isolated from different locations. Identical patterns indicate that the two strains are very closely related or siblings.

3.4. Growth of 4-CBA-degrading bacteria in axenic cultures

To understand axenic growth patterns of the 4-CBA degraders, the strains were inoculated into 4-CBA mineral medium under both induced and uninduced conditions. In general, the isolates grew quickly on 4-CBA medium, showing doubling times ranging from 10.3 h (strain HR3) to 91.6 h (strain HR20) during the exponential growth phase. The growth curves of some representative strains are shown in Fig. 4. Under uninduced condition, strain HR3 exhibited short lag period (ca. 48 h) and thereafter began to grow exponentially. Strains HR14 and HR18 exhibited relatively longer lag periods (ca. 72 h), and strain HR20 showed the longest lag period (ca. 96 h). By contrast, when adapted to 4-CBA metabolism, their lag phases were significantly reduced, suggesting that the 4-CBA degradative enzymes are inducible by the presence of 4-CBA. The inducibility of 4-CBA degradative enzymes of the isolates was confirmed in resting cell experiment conducted with cells adapted or not adapted to 4-CBA metabolism. The cells adapted to 4-CBA metabolism min-

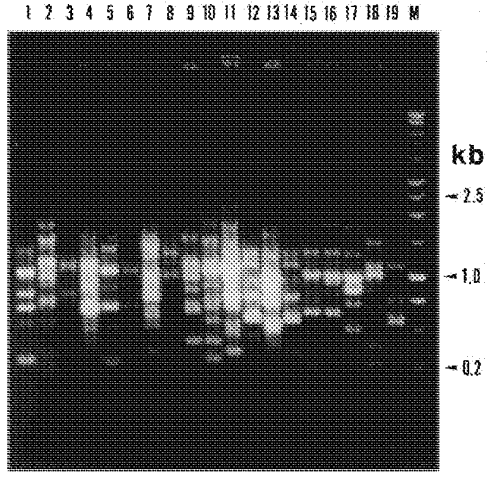


Fig. 3. REP-PCR band patterns of the isolates. Lanes: 1, HR1; 2, HR2; 3, HR3; 4, HR4; 5, HR5; 6, HR6; 7, HR7; 8, HR8; 9, HR9; 10, HR10; 11, HR11; 12, HR12; 13, HR13; 14, HR14; 15, HR15; 16, HR16; 17, HR17; 18, HR18; 19, HR19. M, 1 kb DNA ladder.

Downloaded from https://academic.oup.com/femsec/article/31/1/53/532188 by Pearson Hi-Ed International user on 03 September 2020

COR-AAD1-00001226

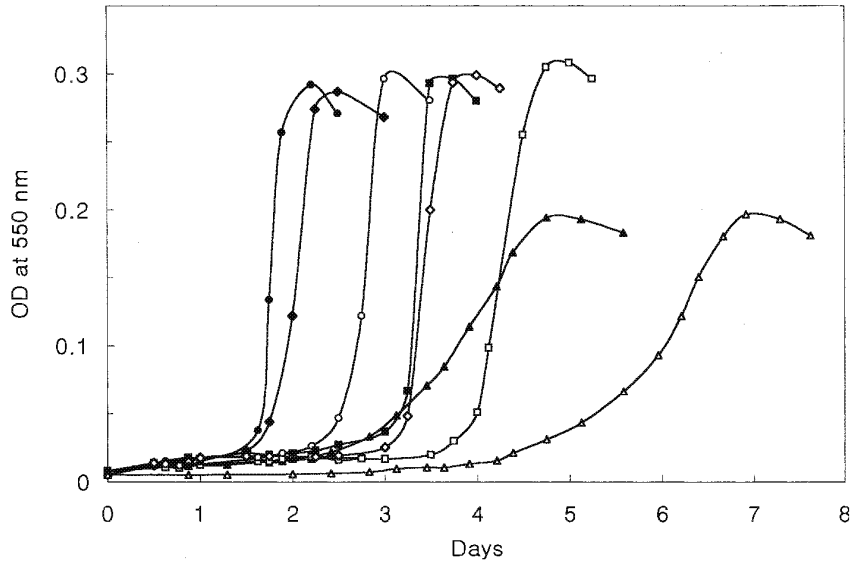


Fig. 4. Growth patterns of 4-chlorobenzoate-degrading bacteria in axenic cultures. Symbols: ● and ○, strain HR3; ■ and □, strain HR18; ◆ and ◇, strain HR14; ▲ and △, strain HR20. The inoculated bacteria were either adapted (solid symbols) or not adapted (open symbols) to 4-chlorobenzoate metabolism. Each point is the mean of two replicate liquid cultures. OD, optical density.

Table 2
Patterns of utilization[a] of substrates by the isolates

Isolate	Utilization of substrate					
	4-CBA	4-IBA	4-BrBA	3-CBA	BA	2,4-D
HR1	++	−	++	−	−	−
HR2	++	+	−	−	++	−
HR3	++	−	−	−	++	−
HR4	++	−	++	−	−	−
HR5	++	+	++	−	++	−
HR6	++	−	−	−	−	−
HR7	++	+	++	−	++	−
HR8	++	−	++	−	−	−
HR9	++	+	++	−	−	−
HR10	++	−	−	−	++	−
HR11	++	+	−	−	++	−
HR12	++	−	++	−	−	−
HR13	++	−	−	−	++	−
HR14	++	−	++	−	−	−
HR15	++	−	++	−	++	−
HR16	++	−	++	−	++	−
HR17	++	−	−	−	++	−
HR18	++	−	−	−	−	−
HR19	++	+	++	−	−	−
HR20	++	−	−	−	−	++

[a]The bacteria were grown on PTYG (peptone-tryptone-yeast extract-glucose) medium and then tested for the ability to use substrate. ++, >80% reduction in peak height from UV scanning and substantial growth (optical density at 550 nm, >0.13); +, 40 to 60% reduction in pak and moderate growth (optical density at 550 nm, >0.08); −, <15% reduction in peak and very scant growth (optical density at 550 nm, <0.01). 4-CBA, 4-chlorobenzoic acid; 4-IBA, 4-iodobenzoic acid; 4-BrBA, 4-bromobenzoic acid; 3-CBA, 3-chlorobenzoic acid; BA, benzoic acid; 2,4-D, 2,4-dichlorophenoxyacetic acid.

eralized completely 300 ppm of 4-CBA within 5 h, while the unadapted cells did not begin to degrade it during the incubation period of 11 h.

3.5. Degradative diversity analysis

The isolates were grown on PTYG medium, then examined for their ability to degrade other compounds related to 4-CBA. The isolates were generally restricted in their substrate utilization abilities (Table 2). Many of the isolates were able to utilize 4-bromobenzoate or benzoate, in addition to 4-CBA, but none of them could completely mineralize 4-iodobenzoate, probably due to the lack of uptake system for the compound. Of the isolates, strains HR6 and HR18 were the most restricted in their substrate use, with being unable to degrade any of the tested substrates, except 4-CBA. Since no isolates could utilize 3-chlorobenzoate, the oxidation specificity of 4-CBA and 3-chlorobenzoate appears to be very strict in the isolates.

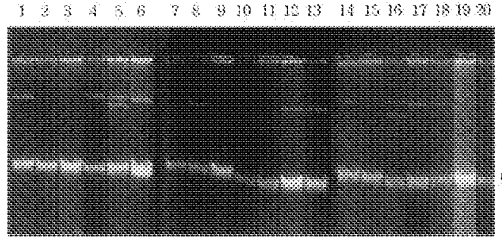


Fig. 5. Plasmid profiles of the isolates. Lanes: 1, HR3; 2, HR2; 3, HR1; 4, HR6; 5, HR7; 6, HR20; 7, HR9; 8, HR8; 9, HR19; 10, HR18; 11, HR17; 12, HR16; 13, HR15; 14, HR4; 15, HR5; 16, HR14; 17, HR13; 18, HR12; 19, HR11; 20, HR10.

Downloaded from https://academic.oup.com/femsec/article/31/1/53/532188 by Pearson Hi-Ered International user on 03 September 2020

COR-AAD1-00001227

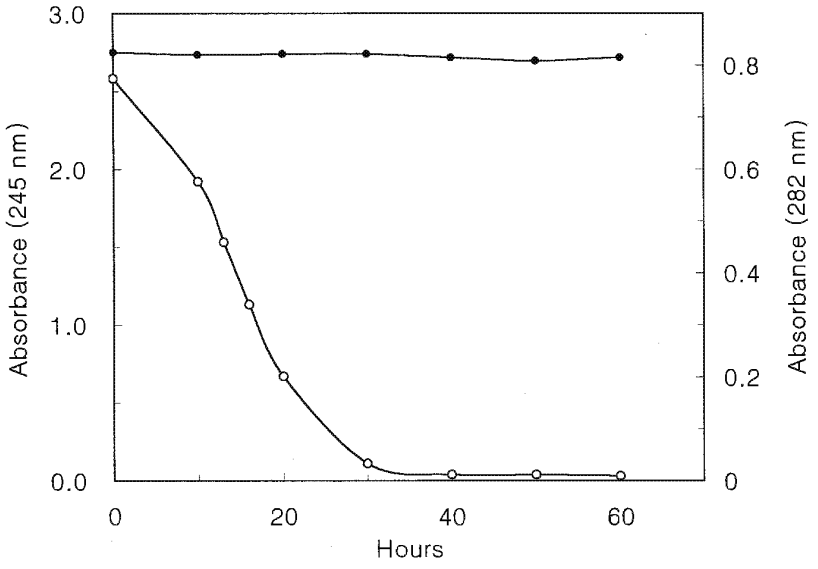


Fig. 6. Changes of UV absorption spectrum during degradation of 4-hydroxybenzoate and 4-chlorocatechol by strain HR4 grown on 4-chlorobenzoate. Symbols represent UV absorbance at 245 nm for 4-hydroxybenzoate (○) and 282 nm for 4-chlorocatechol (●).

3.6. *Plasmid detection and transferability of 4-CBA degradation phenotype*

When the isolates were subjected to Kado's plasmid detection procedure, 55% (11 strains) of the isolates exhibited one or two plasmid DNA bands (Fig. 5). Among the isolates, HR7 and HR14, which were identified as *Pseudomonas* sp. and *M. halobius*, respectively, exhibited two plasmid bands. While there is no clear relationship between the occurrence of plasmid and the FAME identifications, any of the four isolates identified as *Micrococcus luteus* did not contain plasmid DNA. Interestingly, the occurrence of plasmid was very rare in FAME group d, in which only strain HR3 of the eight isolates was observed to have plasmid DNA.

To investigate whether the 4-CBA degradative genes are on the plasmid and transmissible to other bacteria, the isolates containing plasmid were mated with antibiotic-resistant recipients. The 4-CBA degradation phenotype was not transferred at a detectable frequency ($< 10^{-9}$) in most of the isolates. However, in one strain, HR7, which contained two plasmids and was identified as *Pseudomonas* sp.

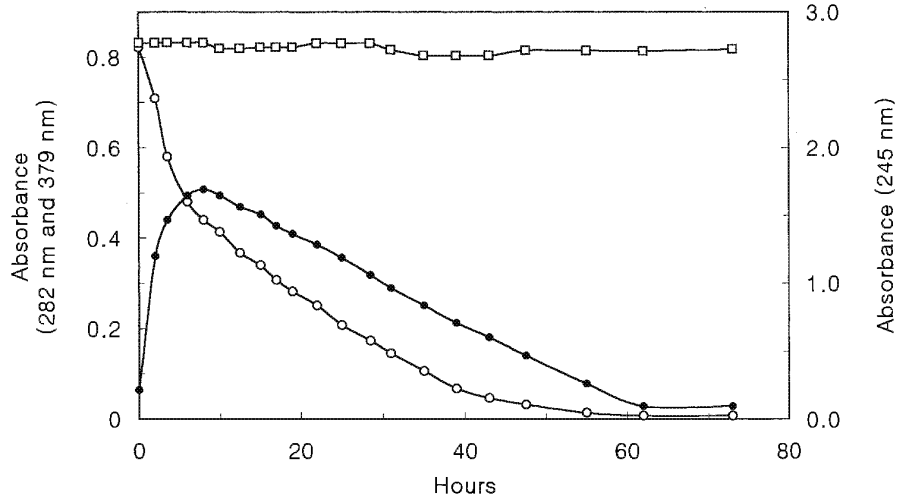


Fig. 7. Changes of UV absorption spectra during degradation of 4-chlorocatechol and 4-hydroxybenzoate by strain HR20 grown on 4-chlorobenzoate. Symbols represent UV absorbance at 282 nm for 4-chlorocatechol (○), 379 nm for an intermediate (●), and 245 nm for 4-hydroxybenzoate (□). Each point is the mean of two replicate liquid cultures.

Downloaded from https://academic.oup.com/femsec/article/31/1/53/532198 by Pennsylvania State University user on 03 September 2020

COR-AAD1-00001228

by FAME analysis, the 4-CBA degradation phenotype was transferred to the recipients (*Burkholderia* sp.) at an average frequency of 9.4×10^{-4}. Its two plasmids were transferred to the recipients, judging from the plasmid bands exhibiting identical electrophoretic mobilities in agarose gels containing the transconjugant and its respective donor HR7 (data not shown).

While the genes for 4-CBA dehalogenation were shown to be on the chromosome in some strains [13], the 4-CBA degradative genes have rarely been found on the plasmid DNA in the environmental isolates. Recently, there has been a report that indicated the presence of the dehalogenation gene on the plasmid pSS50 [21], but it was not clear whether the complete genes necessary for mineralization of 4-CBA were on the plasmid and whether the plasmid was self-transmissible. To our knowledge, it is the first report that self-transmissible plasmids from the environmental isolate contain the 4-CBA degradative genes. Further investigations are underway on the 4-CBA genes of the plasmids.

3.7. 4-CBA degradation pathway

To investigate the distribution of 4-CBA degradation pathway among the isolates, each strain was grown in 4-CBA or acetate mineral medium to produce cells adapted or not adapted to 4-CBA metabolism, and then washed with phosphate buffer and inoculated into 4-hydroxybenzoate or 4-chlorocatechol mineral medium. All the adapted isolates, except HR20, could mineralize 4-hydroxybenzoate very quickly, but they could not degrade 4-chlorocatechol. On the other hand, strain HR20 could degrade 4-chlorocatechol quickly, but could not attack 4-hydroxybenzoate (Fig. 6). During biodegradation of 4-CBA by strain HR20, the absorbance peak at 379 nm was sharply increased in 10 h and then gradually decreased (Fig. 7). It was previously reported that 5-chloro-2-hydroxymuconic semialdehyde was the intermediate corresponding to this kind of peak change during 4-CBA biodegradation via *meta*-cleavage pathway [22,23]. From the above results, most of our isolates appeared to degrade 4-CBA through 4-hydroxybenzoate, while one isolate, HR20, mineralized 4-CBA via formation of 4-chlorocatechol. While the two degradative pathways are described among other isolates of the previous works [6,13,22,23], the majority of organisms able to grow on 4-CBA appear to mineralize it through 4-hydroxybenzoate, which is consistent with the suggestion that dehalogenation by hydroxylation is very specific for the *para* position of halobenzoates in the microbial metabolism of halogenated benzoic acid [24]. It is interesting that the 3-chlorobenzoate-utilizing organisms degrade 3-chlorobenzoate mainly via chlorocatechol [25,26], while our 4-CBA-degrading isolates mineralize 4-CBA mainly via the corresponding hydroxybenzoate.

Acknowledgements

This study was supported by the Korea Science and Engineering Foundation (KOSEF) through the Research Center for Molecular Microbiology at Seoul National University.

References

[1] Furukawa, K., Tomizuka, N. and Kamibayashi, A. (1979) Effect of chlorine substitution on the bacterial metabolism of various polychlorinated biphenyls. Appl. Environ. Microbiol. 38, 301–310.
[2] Horavath, R.S. (1971) Cometabolism of the herbicide 2,3,6-trichlorobenzoate. J. Agric. Food Chem. 19, 291–293.
[3] Furukawa, K. and Chakrabarty, A.M. (1982) Involvement of plasmids in total degradation of chlorinated biphenyls. Appl. Environ. Microbiol. 44, 619–626.
[4] Masse, R., Messier, F., Peloquin, L., Ayotte, C. and Sylvestre, M. (1984) Microbial biodegradation of 4-chlorobiphenyl, a model compound of chlorinated biphenyls. Appl. Environ. Microbiol. 47, 947–951.
[5] Adriaens, P. and Focht, D.D. (1990) Continuous coculture degradation of selected polychlorinated biphenyl congeners by acinetobacter spp. in an aerobic reactor system. Environ. Sci. Technol. 24, 1042–1049.
[6] Marks, T.S., Smith, A.R.W. and Quirk, A.V. (1984) Degradation of 4-chlorobenzoic acid by *Arthrobacter* sp.. Appl. Environ. Microbiol. 48, 1020–1025.
[7] Keil, H., Klages, U. and Lingens, F. (1981) Degradation of 4-chlorobenzoate by *Pseudomonas* sp. CBS3 induction of catabolic enzymes. FEMS Microbiol. Lett. 10, 213–215.
[8] van den Tweel, W.J.J., Kok, J.B. and de Bont, J.A.M. (1987) Reductive dechlorination of 2,4-dichlorobenzoate to 4-chlorobenzoate and hydrolytic dechlorination of 4-chloro-, 4-bromo-, and 4-iodobenzoate by *Alcaligenes denitrificans* NTB-1. Appl. Environ. Microbiol. 53, 810–815.
[9] Klages, U. and Lingens, F. (1979) Degradation of 4-chlorobenzoic acid by a *Nocardia* sp.. FEMS Microbiol. Lett. 6, 201–203.
[10] Horowitz, A., Suflita, J.M. and Tiedje, J.M. (1983) Reductive dehalogenations of halobenzoates by anaerobic lake sediment microorganisms. Appl. Environ. Microbiol. 45, 1459–1465.
[11] Haggblom, M.M. (1992) Microbial breakdown of halogenated aromatic pesticides and related compounds. FEMS Microbiol. Rev. 103, 29–72.
[12] Chang, K.-H., Liang, P.-H., Beck, W., Scholten, J.D. and Dunaway-Mariano, D. (1992) Isolation and characterization of the three polypeptide components of 4-chlorobenzoate dehalogenase from *Pseudomonas* sp. strain CBS-3. Biochemistry 31, 5605–5610.
[13] Savard, P., Peloquin, L. and Sylvestre, M. (1986) Cloning of *Pseudomonas* sp. strain CBS3 gene specifying dehalogenation of 4-chlorobenzoate. J. Bacteriol. 168, 81–85.
[14] Schmitz, A., Gartemann, K.-H., Fiedler, J., Grund, E. and Eichenlaub, R. (1992) Cloning and sequence analysis of genes for dehalogenation of 4-chlorobenzoate from *Arthrobacter* sp. strain SU. Appl. Environ. Microbiol. 58, 4068–4071.
[15] Stanier, R.Y., Palleroni, N.J. and Doudoroff, M. (1966) The aerobic pseudomonads: a taxonomic study. J. Gen. Microbiol. 43, 159–271.
[16] Sasser, M. (1990) Technical note 102: Tracking a strain using the Microbial Identification System. MIS Inc., North Newark, DE.
[17] Versalovic, J., Schneider, M., de Bruijn, F.J. and Lupski, J.R. (1994) Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. Mol. Cell. Biol. 5, 25–40.
[18] Ka, J.O. and Tiedje, J.M. (1994) Integration and excision of a 2,4-dichloro-phenoxyacetic acid-degradative plasmid in *Alcaligenes para-*

Downloaded from https://academic.oup.com/femsec/article/31/1/53/532138 by Pennsylvania Hospital Library user on 03 September 2020

COR-AAD1-00001229

doxus and evidence of its natural intergeneric transfer. J. Bacteriol. 176, 5284–5289.
[19] Kado, C.I. and Liu, S.T. (1981) Rapid procedure for detection and isolation of large and small plasmids. J. Bacteriol. 145, 1365–1373.
[20] Willetts, N.S. (1988) Conjug. Methods Microbiol. 21, 49–77.
[21] Layton, A.C., Sanseverino, J., Wallace, W., Corcoran, C. and Sayler, G.S. (1992) Evidence for 4-chlorobenzoic acid dehalogenation mediated by plasmids related to pSS50. Appl. Environ. Microbiol. 58, 399–402.
[22] Seo, D.I., Lim, J.Y., Kim, Y.C., Min, K.H. and Kim, C.K. (1997) Isolation of *Pseudomonas* sp. S-47 and its degradation of 4-chlorobenzoic acid. J. Microbiol. 35, 188–192.
[23] Arensdorf, J.J. and Focht, D.D. (1995) A *meta* cleavage pathway for 4-chlorobenzoate, an intermediate in the metabolism of 4-chlorobiphenyl by *Pseudomonas cepacia* P166. Appl. Environ. Microbiol. 61, 443–447.
[24] Commandeur, L.C.M. and Parsons, J.R. (1990) Degradation of halogenated aromatic compounds. Biodegradation 1, 207–220.
[25] Dorn, E., Hellig, M., Reineke, W. and Knackmuss, H.-J. (1974) Isolation and characterization of a 3-chlorobenzoate degrading *Pseudomonas*. Arch. Microbiol. 99, 61–70.
[26] Chatterjee, D.K., Kellogg, S.T., Hamada, S. and Chakrabarty, A.M. (1981) Plasmid specifying total degradation of 3-chlorobenzoate by a modified *ortho* pathway. J. Bacteriol. 146, 639–646.

Downloaded from https://academic.oup.com/femsec/article/21/1/53/532168 by Pennsylvania Hospital Library user on 03 September 2020

COR-AAD1-00001230

In Vitro Cell. Dev. Biol.—Plant 37:300–304, March–April 2001
DOI: 10.1079/IVP2000165
© 2001 Society for In Vitro Biology
1054-5476/01 $10.00+0.00

IN VITRO ASSAY FOR 2,4-D RESISTANCE IN TRANSGENIC COTTON

BAO-HONG ZHANG*, HONG-MEI WANG, FANG LIU, YUN-HAI LI, AND ZHENG-DE LIU

Key Laboratory of Cotton Genetic Improvement of the Ministry of Agriculture, Cotton Research Institute, Chinese Academy of Agricultural Sciences, Anyang Henan 455112, P. R. China

(Received 22 May 2000; accepted 27 November 2000; T.A. Thorpe)

SUMMARY

2,4-Dichlorophenoxyacetic acid (2,4-D) resistant plants of transgenic cotton (*Gossypium hirsutum* L.) were produced using *Agrobacterium tumefaciens* containing a plasmid carrying the neomycin phosphotransferase II (*npt* II) and 2,4-D monooxygenase (*tfd* A) genes. An *in vitro* assay was performed to determine the sensitivity of seed germination, and the growth of seedlings of transgenic and non-transgenic cotton to various concentrations of kanamycin and 2,4-D. The results indicated that kanamycin caused the cotyledons of non-transgenic plants to turn white, but transgenic plants grew normally. Seed germination and seedling growth of non-transgenic plants were strongly inhibited by 2,4-D, but only slightly for transgenic plants. Transgenic plants and non-transgenic plants can be clearly distinguished by the use of 2 mg l^{-1} 2,4-D in seed germination medium. There was a high correlation between the response of seed germination and the growth of seedlings to kanamycin or 2,4-D, based on the germination ratio, albino ratio, dry weight or fresh weight. On this basis, we developed a rapid method for identifying transgenic plants that has been verified in the field. These findings will allow identification of cotton transformants at an early stage of plant development, saving time and improving cultivars containing the 2,4-D resistance trait.

Key words: Gossypium hirsutum L; transgenic cotton; *in vitro* assay; kanamycin; 2,4-D.

INTRODUCTION

Cotton is the most important economic fiber crop and is widely cultivated in many countries. There are over 180 million people worldwide depending on this crop for their livelihood (International Advisory Committee, 1996). Thus, genetic improvement of cotton is very important. Conventional breeding programs have produced steady improvements in many agronomic traits, but the great potential benefits of plant genetic engineering on cotton genetic improvement have been slow in coming. Since Umbeck et al. (1987) obtained the first transgenic cotton plants, many laboratories have obtained insect-resistant (Perlak et al., 1990; Cousins et al., 1991; Xie et al., 1991; Thomas et al., 1995; Jenkins et al., 1997; Li et al., 1998; Ni et al., 1998) or herbicide-resistant (Bayley et al., 1992; Lyon et al., 1993; Chen et al., 1994; Rajaseharam et al., 1996; Keller et al., 1997) transgenic cotton plants. Cotton is among the first transgenic crops to be commercialized, and has been widely cultivated in the USA, China and other countries (John, 1997; Zhang et al., 2000a).

2,4-Dichlorophenoxyacetic acid (2,4-D) is a herbicide commonly used to control broad-leaved weeds in the field. Because cotton has a high sensitivity to 2,4-D, 1 mg l^{-1} 2,4-D can damage cotton plants. The application of this herbicide in controlling weeds in cotton and adjacent fields has been limited.

We are interested in improving cotton by genetic engineering with a 2,4-D degradation enzyme gene to increase its resistance to the herbicide, 2,4-D. Genetically engineered broadleaf herbicide protection for 2,4-D in cotton was attempted by incorporating an *Alcaligenes eutrophus* gene encoding 2,4-D monooxygenase (*tfd* A). Transformants containing *tfd* A exhibited 50- to 100-fold more tolerance to 2,4-D compared with non-transformed plants (Bayley et al., 1992; Chen et al., 1994). We have developed reliable somatic embryogenesis and a plant regeneration system for cotton (Zhang et al., 1996a,b, 1999, 2000b). Using this system, we have obtained transgenic cotton containing the *tfd* A gene and transformed it into commercial cultivars *via* cross-breeding.

One of the requirements for transforming cotton is a suitable and rapid assay and identification method for screening authentic transformants. One of the most commonly used selective marker genes is neomycin phosphotransferase II (*npt* II) which encodes a phosphotransferase capable of phosphorylating aminoglycoside antibiotics such as kanamycin, neomycin, geneticin (G418) and paromomycin (Kapaun and Cheng, 1999). The marker gene has been commonly used in cotton genetic transformation for obtaining insect-resistant or disease-resistant transgenic cotton (John, 1997; Zhang and Zhao, 1997).

Although several methods have been developed to identify authentic transformants by assaying *npt* II and *tfd* A activities, these methods have required several elaborate *in vitro* procedures with

*Author to whom correspondence should be addressed. Present address: 1632 West 6th Street Apt F, Austin, TX 78703, USA; Email zbh68@hotmail.com

COR-AAD1-00001231

radioactively labeled substrates. These methods are expensive and have safety issues for humans. Thus, it is very important to successfully develop an efficient, rapid, simple and safe method for identification of authentic transformants.

In this paper, we determined the sensitivity of non-transgenic 2,4-D-resistant cotton and conventional cotton to kanamycin and the herbicide, 2,4-D. We have developed a non-destructive biotest for assay and selection of cotton transformants tolerant to kanamycin and 2,4-D in their early stages of plant development. These findings will save breeding time, cost, and speed up the application of plant genetic engineering to cotton breeding and production.

Materials and Methods

Plant materials. 2,4-D-resistant transgenic cotton plants, which were produced by transformed explants of cotton cv. Jimian 12 using *Agrobacterium tumefaciens* LBA 4404 with the plasmid pBI 121 carrying *npt* II and *tfd* A genes, were obtained from the Cotton Research Institute of Shanxi Academy of Agricultural Science, Shanxi, China. The plants were improved *via* cross-breeding and named transgenic cotton line 508. The non-transgenic cotton variety CRI 34 used in this study was obtained from the Cotton Research Institute of the Chinese Academy of Agricultural Sciences, Anyang, Henan, China.

Sterilization of cotton seeds and germination. Mature seeds without seed coats (kernels) were sterilized as follows: kernels were surface-sterilized with 0.1% $HgCl_2$ for 7 min and rinsed three times with sterile distilled water. The kernels were then placed on 0.7% agar-solidified medium (Agar from Beijing Biochemistry Company), pH 5.8, containing half-strength Murashige and Skoog (MS) inorganic salts (Murashige and Skoog, 1962), B_5 (Gamborg et al., 1968) vitamins, and 2% sucrose. The cultures were illuminated for 14 h d^{-1} with a light intensity of 3000 lux at 28 ± 2°C.

Sensitivity of cotton to kanamycin. Seeds without seed coats were inoculated in 250 × 30 mm test tubes containing the medium described as above with various concentrations of kanamycin. Kanamycin was filter-sterilized using a ProX 0.22 μm membrane filter and was then added to the medium at the following mg l^{-1} concentrations: 0, 75, 150 and 300. The experiment was arranged as a randomized complete block with 30 replicates and repeated three times respectively for non-transgenic variety CRI 34 and the transgenic line 508.

Sensitivity of cotton to herbicide 2,4-D. Seeds without seed coats were inoculated in 250 × 30 mm test tubes containing medium described as above with various concentrations of 2,4-D. 2,4-D was added to the medium at the following mg l^{-1} concentrations: 0, 2, 4, 6, 8 and 10 before sterilization at 121°C for 15 min. The experiment was arranged as a randomized complete block with 50–60 replicates and repeated three times respectively for non-transgenic variety CRI 34 and transgenic line 508.

Experimental records and data analysis. After 2 d of culture on the medium with various concentrations of kanamycin or 2,4-D, the number of germinated seeds of the conventional non-transgenic cotton variety CRI 34 and transgenic cotton line 508 were recorded, respectively, and the germination ratio was calculated. After 7 d of culture, the growth of roots, hypocotyls and cotyledons, and the number of seeds converted into plants were recorded. After 14 d of culture, the seedlings were measured for seedling height, fresh weights of roots, stems, cotyledons and whole plants. Roots, stems and cotyledons were then dried for 24 h at 121°C, and dry weights were measured.

Analysis of variance was used to determine if there were interactions between kanamycin or 2,4-D concentrations, so that we could find a suitable concentration for the identification of transgenic cotton resistant to kanamycin or 2,4-D. Data were transformed to percentage inhibition compared to a 0 mg l^{-1} kanamycin or 2,4-D control and were used to determine the concentration of 2,4-D which caused a 50% and 100% reduction in growth.

Identification of transgenic 2,4-D-resistant cotton in the field. Field identification was conducted in 1998 at the Cotton Research Institute, Chinese Academy of Agricultural Sciences near Anyang, Henan, P.R. China. Transgenic cotton bred line 508 and non-transgenic cotton cultivar CRI 34 were planted on April 25, 1998 at a rate of approximately 15 seeds m^{-1} of row with a 1.0 m row space. Final plant population was 9–10 plants m^{-2}. About 6 wk after planting, 200 mg l^{-1} 2,4-D was sprayed on the transgenic line 508 and CRI 34. About 2 wk after spraying, damage to the cotton leaves was determined. The plants whose leaves were twisted and damaged were non-resistant plants. However, the plants whose leaves were normal came from resistant plants. The numbers of resistant and non-resistant plants were recorded.

Results

Effects of kanamycin on the germination of seeds and the growth of seedlings. Cotton seeds were not sensitive to kanamycin for both transgenic seeds and non-transgenic seeds. The germination ratio varied only slightly when the concentrations of kanamycin increased from 0 to 300 mg l^{-1}. Both transgenic and non-transgenic seeds germinated and converted into plants. However, the effects of kanamycin on the growth of seedlings were different between transgenic and non-transgenic cotton. On the medium with kanamycin, although the growth rate of both transgenic and non-transgenic cotton seedlings decreased with the increasing kanamycin concentrations, the non-transgenic cotton grew more slowly than the transgenic cotton (Table 1).

The most obvious character that was affected by kanamycin was the change in cotyledon color. Low concentrations of kanamycin resulted in the cotyledons of non-transgenic seedlings losing their green color and initiated dying. In contrast, the seedlings of transgenic cotton grew normally. The percentage of non-transgenic cotton with albino cotyledons increased with increasing kanamycin concentrations up to 150 mg l^{-1}; thereafter, all cotyledons of non-transgenic seedlings

TABLE 1

SENSITIVITY OF TRANSGENIC LINE 508 AND NON-TRANSGENIC VARIETY CRI 34 TO KANAMYCIN, BASED ON THE GERMINATION OF SEEDS AND THE GROWTH RATE OF COTTON SEEDLINGS (28 d)

Cotton lines	Kanamycin concentration (mg l^{-1})	No. of seeds tested	No. of seeds germinated	Germination (%)	No. of seedlings established	% Seedling establishment	Height of seedling (cm)	No of albino cotyledons	% Albino cotyledons
508	0	30	26	86.7	22	73.3	9.73	0	0
	75	30	25	83.3	22	73.3	6.20	6	27.3
	150	30	25	83.3	22	73.3	3.94	7	31.8
	300	30	25	83.3	23	76.7	3.85	7	30.4
CRI 34	0	30	28	93.3	27	90.0	9.36	0	0
	75	30	27	90.0	27	90.0	4.82	23	85.2
	150	30	28	93.3	26	86.7	2.86	26	100.0
	300	30	28	93.3	28	93.3	3.48	28	100.0

COR-AAD1-00001232

TABLE 2

EFFECTS OF 2,4-D ON THE GERMINATION OF TRANSGENIC COTTON AND CONVENTIONAL NON-TRANSGENIC COTTON

2,4-D concentration (mg l^{-1})	Transgenic line 508				Non-transgenic variety CRI 34			
	No. of seeds inoculated	No. of seeds germinated	Germination (%)	Inhibition (%)	No. of seeds inoculated	No. of seeds germinated	Germination (%)	Inhibition (%)
0	50	39	78.0	0.0	60	53	88.3	0.0
2	50	36	72.0	7.7	56	2	3.6	95.9
4	49	35	71.4	8.5	58	1	1.7	98.1
6	48	34	69.4	11.0	60	1	1.7	98.1
8	56	37	66.1	15.3	60	0	0.0	100.0
10	46	27	58.7	24.8	60	0	0.0	100.0

lost their green color and started dying. Cotyledons of transgenic cotton were insensitive to kanamycin. The cotyledons of transgenic cotton plants grew normally on the medium supplemented with 150 or 300 mg l^{-1} kanamycin. These same concentrations caused the cotyledons of non-transgenic cotton to bleach (Table 1). It is clear that transgenic seedlings can be identified using 150 mg l^{-1} kanamycin.

Effects of 2,4-D on the germination of seeds and the growth of seedlings. Unlike kanamycin, cotton seeds were sensitive to 2,4-D, but the sensitivity of seeds to 2,4-D was different between transgenic and non-transgenic cotton (Table 2). Non-transgenic cotton was highly sensitive to 2,4-D. Low concentrations of 2,4-D inhibited the germination of non-transgenic cotton seeds. 2,4-D at 2 mg l^{-1} caused a 95% germination inhibition, and 8 mg l^{-1} 2,4-D caused 100% germination inhibition. 2,4-D at 2 and 8 mg l^{-1} only inhibited 7.7% and 15.3% of transgenic cotton seed germination and conversion to plants, respectively. In order to measure the inhibition of 2,4-D on transgenic cotton seed germination, we also added 10–150 mg l^{-1} 2,4-D in the seed germination medium. The results indicated that the transgenic seeds could germinate on the medium with 100 mg l^{-1} 2,4-D, but the germination ratio was low (data not shown). It showed that there was remarkable difference between transgenic and non-transgenic cotton in the germination ratio on the medium with 2,4-D (Table 2). Transgenic and non-transgenic cotton could be easily distinguished when they were placed on the germination medium supplemented with 2 mg l^{-1} 2,4-D.

In order to evaluate the sensitivity of transgenic and non-transgenic cotton seedling to 2,4-D, we also measured the plant height of 7- and 14-d-old seedlings, and fresh and dry weights of cotyledon, hypocotyl, root and whole plant of 14-d-old seedlings. The results indicated that 2,4-D greatly inhibited the growth of cotton seedlings. 2,4-D completely inhibited growth of non-transgenic seedlings at 2 mg l^{-1}. On the medium supplemented with 2 mg l^{-1} 2,4-D, the height of the non-transgenic cotton seedlings was only 0.49 cm after 14 d of culture, and it was only 6.0% of the control (without 2,4-D). The growth of hypocotyls of non-transgenic seedlings was inhibited completely on the media supplemented with 2,4-D over 2 mg l^{-1}. There were fewer effects of 2,4-D on transgenic seedlings than non-transgenic seedlings. Transgenic seedlings grew normally and the height of seedlings reached 6.75 cm after 14 d of culture on medium supplemented

TABLE 3

EFFECTS OF 2,4-D ON SEEDLING GROWTH OF TRANSGENIC COTTON LINE 508 AND THE CONVENTIONAL COTTON VARIETY CRI 34

Cotton line	2,4-D (mg l^{-1})	Height at 7 d[a] (cm)	Height at 14 d		Fresh weight at 14 d (g)[b]					Dry weight at 14 d (mg)[b]				
			cm[b]	Inhibition (%)	Cotyledon	Hypocotyl	Root	Total	Inhibition (%)	Cotyledon	Hypocotyl	Root	Total	Inhibition (%)
Transgenic line 508	0	4.3	7.59 a	0.0	303 a	206 a	88 a	597 a	0.0	40 a	13 a	7 a	60 a	0.0
	2	2.4	6.75 b	11.1	269 a	104 b	65 h	438 b	26.6	40 a	10 b	6 a	56 ab	6.7
	4	1.7	3.25 c	57.2	203 b	71 c	38 c	312 c	47.7	39 a	7 c	4 b	50 bc	16.7
	6	2.0	2.86 cd	62.3	186 b	50 c	34 c	280 c	52.1	39 a	7 c	4 b	50 c	16.7
	8	1.7	2.50 d	67.1	173 b	55 c	33 c	270 c	54.8	37 a	7 c	4 b	48 c	20.2
	10	2.1	2.48 d	67.3	190 b	59 c	36 c	261 c	56.3	34 a	10 b	5 b	49 c	18.3
Non-transgenic line CRI 34	0	6.0	8.21 a	0.0				732 a	0.0				64 a	0.0
	2	<1.0	0.49 b	94.0				314 b	57.1				52 bc	18.8
	4	<1.0		–				265 b	63.8				49 c	23.4
	6	<1.0		–				293 b	60.0				55 bc	14.1
	8	<1.0		–				307 b	58.1				55 bc	14.1
	10	<1.0		–				273 b	62.7				55 bc	14.1

[a] Data represent the average of items tested in a given treatment after 7 or 14 d of culture. For each treatment 45 seedlings were recorded in three replicates of 15 seedlings.

[b] Means followed by the same letter(s) within each column are not significantly different from each other at the 5% level by LSD test. Data represent the average of items tested in a given treatment after 14 d of culture. For each treatment, 45 seedlings were recorded in three replicates of 15 seedlings.

COR-AAD1-00001233

with 2 mg l^{-1} 2,4-D, while the growth of non-transgenic seedlings was completely inhibited (Table 3). Thus, 2,4-D could be considered as a suitable selective agent for identifying the transformants.

Formation and development of the root system was greatly affected by the herbicide. 2,4-D completely inhibited the growth and development of the radicle in non-transgenic seedlings, and there was no secondary root formation in the medium supplemented with 2,4-D. In contrast, the transgenic seedlings could produce a normal radicle and many secondary roots on 2,4-D media up to 10 mg l^{-1} of 2,4-D while the control cultivar, CRI 34, died.

Unlike plant height and the growth of roots, there was a slight influence of 2,4-D on the fresh and dry weight of non-transgenic seedlings. The reason is that there was little effect of 2,4-D on weight of cotyledons of either transgenic or non-transgenic cotton plants.

The sensitivity of cotton seedlings to 2,4-D was measured in the terms of the predicted concentration of 2,4-D (in mg l^{-1}) that caused a 50% growth inhibition (I_{50}) in fresh and dry weight. For transgenic seedlings, the I_{50} of fresh weight was about 6 mg l^{-1} 2,4-D. For non-transgenic seedlings, the I_{50} of fresh weight was less than 2 mg l^{-1} 2,4-D. Effects of 2,4-D on the dry weight of either transgenic or non-transgenic seedlings were less than on the fresh weight. These results indicated that non-transgenic seedlings were more sensitive than transgenic seedlings. The sensitivity of transgenic seedlings to 2,4-D decreased about fivefold. The moisture content of cells (data not shown) was only slightly affected by 2,4-D.

2,4-D also influenced the growth of the non-transgenic seedling cotyledon. Chlorophyll content was reduced by 2,4-D at concentrations of 2 mg l^{-1} or higher, with complete inhibition at 8 mg l^{-1}. In contrast, all seedlings of transgenic cotton remained green on the medium supplemented with the highest concentration tested (10 mg l^{-1}).

Preliminary establishment of an in vitro *identification method and field analysis of resistant 2,4-D transgenic cotton containing* npt *II and* tfd *A genes.* On the basis of analysis of kanamycin effects and the herbicide 2,4-D on the germination of seeds and the growth of seedlings of transgenic and non-transgenic cotton, we developed a rapid *in vitro* method for identifying and screening transgenic cotton materials. To examine the applicability of the method described above in identifying and screening for kanamycin and 2,4-D resistance in transgenic cotton, we chose seeds of cotton breeding line 508 for both laboratory and field evaluation in 1998. The breeding line 508 came from the hybrid progenies of commercial cultivar CRI 34 and a transgenic cotton breeding line containing the *tfd* A gene, which was unstable. The results indicated that 72.0%, 68.9% and 71.6% of the resistant plants were identified with 2 mg l^{-1} 2,4-D and 100 mg l^{-1} kanamycin in the laboratory, and 200 mg l^{-1} 2,4-D in the field, respectively. The results were very similar (Table 4). These results clearly demonstrated the reliability and efficiency of this non-destructive assay for 2,4-D resistance to identify and screen transgenic plants.

Discussion

Successful development of an efficient genetic transformation system and transgenic breeding for cotton requires an effective identification and selection protocol. With this effective identification method, it is easy to identify and screen transgenic plants from unstable populations.

The *npt* II gene that encodes neomycin phosphotransferase II confers resistance on its host cells to a wide range of aminoglycosides such as kanamycin, neomycin and G418. These antibiotics inhibit growth of plant cells by binding to the 30S ribosomal subunit, thereby inhibiting initiation of plastid translation (Wilmink and Dons, 1993; Kapaun and Cheng, 1999). Cotton cells transformed with the *npt* II gene can detoxify the antibiotic in the selection medium. Thus, transgenic cells, tissues, plants and seeds can grow on the medium with kanamycin. In contrast, non-transformed cotton cells could not live and grow on the medium with kanamycin because these cotton cells were sensitive to kanamycin. These results have been confirmed by Zhang et al. (2001). Although both transgenic and non-transgenic cotton seeds could germinate on MS medium with kanamycin. The cotyledons of non-transgenic seedlings lost green color and started dying after 20 d while transgenic seedlings grew normally. After 20 d of culture, transgenic and non-transgenic seedlings could be distinguished on the medium supplemented with 100 mg l^{-1} kanamycin.

Because cotton is highly sensitive to 2,4-D, low concentrations of 2,4-D can damage cotton and result in a loss of yield. Thus, we are interested in improving cotton by genetic engineering with a 2,4-D degradation gene (*tfd* A gene). The field trials indicated that transgenic cotton containing *tfd* A gene exhibited 50- to 100-fold more tolerance to the herbicide, 2,4-D, compared with non-transformed cotton (Bayley et al., 1992; Lyon et al., 1993; Chen et al., 1994). Our experiments indicate that cotton seeds and seedlings were highly sensitive to 2,4-D, and 2 mg l^{-1} 2,4-D could completely inhibit the growth of non-transgenic seedlings. Histological observations showed that 2,4-D inhibited the elongation of cotton cells (data not shown). There were also some effects of 2,4-D on transgenic cotton containing the *tfd* A gene, because this gene can express and produce 2,4-D monooxygenase which can degrade 2,4-D into a non-toxic form. Thus, transgenic cotton and non-transgenic cotton could be easily distinguished on the medium with only 2 mg l^{-1} 2,4-D.

Here we have described a rapid, simple, high-efficiency and safe method for identifying transgenic cotton. The method has been confirmed by field tests. The method described here will facilitate

TABLE 4

COMPARISON OF THE IDENTIFIED RESULTS OF TRANSGENIC LINE 508 BETWEEN IN THE LABORATORY WITH KANAMYCIN OR 2,4-D, AND IN THE FIELD

Identification methods	No. of plants	No. of resistant plants	No. of sensitive plants	Resistant plants (%)
Identified by 2 mg l^{-1} 2,4-D in the laboratory	50	14	36	72.0
Identified hy 100 mg l^{-1} kanamycin in the lahoratory	45	14	31	68.9
Identified by spraying 200 mg l^{-1} 2,4-D in the field	405	115	290	71.6

COR-AAD1-00001234

the application of gene transformation in cotton genetic improvement due to its rapidity and efficiency.

Acknowledgment

The authors would like to thank Dr. R. Malcolm Brown, Jr. of the University of Texas at Austin for critical reading of the manuscript.

References

Bayley, C.; Trolinder, N. L.; Ray, C.; Morgan, M. M.; Quisenherry, J. E.; Ow, D. W. Engineering 2,4-D resistance into cotton. Theor. Appl. Genet. 83:645–649; 1992.

Chen, Z. X.; Liewellyn, D. J.; Fan, Y. L.; Li, S. J.; Guo, S. D.; Jiao, G. L.; Zhao, J. X. 2,4-D resistant transgenic cotton plants produced hy *Agrobacterium*-mediated gene transfer. Sci. Agri. Sin. 27(4):31–37; 1994.

Cousins, Y. L.; Lyon, B. R.; Llewellyn, D. J. Transformation of an Australian cotton cultivar: prospects for cotton improvement through genetic engineering. Aust. J. Plant Physiol. 18:481–494; 1991.

Gamhorg, O.; Miller, R.; Ojima, K. Nutrient requirements of suspension cultures of soyhean root cells. Exp. Cell Res. 50:151–158; 1968.

International Advisory Committee. Cotton: review of world situation. Washington, DC: Monogram by International Advisory Committee; 1996.

Jenkins, J. N.; McCarty, J. C.; Buehler, R. E.; Kiser, J.; Williams, C.; Wofford, T. Resistance of cotton with δ-endotoxin genes from *Bacillus thuringiensis* var. *kurstaki* on selected *Lepidopteran* insects. Agron. J. 89:768–780; 1997.

John, M. E. Cotton crop improvement through genetic engineering. Crit. Rev. Biotechnol. 17:185–209; 1997.

Kapaun, J. A.; Cheng, Z. M. Aminoglycoside antibiotics inhibit shoot regeneration from Siberian elm leaf explants. Hortsciences 34:727–729; 1999.

Keller, G.; Spatola, L.; McCabe, D.; Martinell, B.; Swain, W.; John, M. E. Transgenic cotton resistant to herbicide hialaphos. Transgen. Res. 6:385–392; 1997.

Li, Y. E.; Chen, Z. X.; Wu, X.; Li, S. J.; Jiao, G. L.; Wu, J. H.; Fan, X. P.; Meng, J. H.; Zhu, Z.; Wang, W.; Zhu, Y.; Xu, H. L.; Xiao, G. F.; Li, X. H. Ohtaining transgenic cotton plants with Cowpea trypsin inhibitor gene. Acta Gossypii Sinica 10:237–243; 1998.

Lyon, B. R.; Cousins, Y. L.; Llewellyn, D. J.; Dennis, E. S. Cotton plants transformed with a bacterial-degradation gene are protected from accidental spray drift damage by the herhicide 2,4-dichorophenoxyacetic acid. Transgen. Res. 2:162–169; 1993.

Murashige, T.; Skoog, F. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant. 15:473–479; 1962.

Ni, W. C.; Zhang, Z. L.; Guo, S. D. Development of transgenic insect-resistant cotton plants. Sci. Agri. Sin. 31(2):8–13; 1998.

Perlak, F. J.; Deaton, R. W.; Armstrong, R. L.; Fuchs, R. L.; Sims, S. R.; Greenplate, J. T.; Fischhoff, D. A. Insect resistant cotton plants. Bio/Technology 8:939–943; 1990.

Rajaseharam, K.; Grulam, J. W.; Hudspeth, R. L. Herbicide-resistant Acale and Coker cottons transformation with a native gene encoding mutant forms of acetohydroxyacid synthase. Mol. Breed. 2:307–319; 1996.

Thomas, J. C.; Adams, D. G.; Keppenne, V. D.; Wasmann, C. C.; Brown, J. K.; Kanost, M. R.; Bohnert, H. J. Protease inhibitors of *Manduce sexta* expressed in transgenic cotton. Plant Cell Rep. 14:758–762; 1995.

Umbeck, P.; Johnson, G.; Barton, K. Genetically transformed cotton (*Gossypium hirsutum* L.) plants. Bio/Technology 5:235–266; 1987.

Wilmink, A.; Dons, J. J. M. Selective agents and marker genes for use in transformation of monocotyledonous plants. Plant Mol. Biol. Rep. 11:165–185; 1993.

Xie, D. X.; Fan, Y. L.; Ni, W. C.; Huang, J. Q. Transformed *Bacillus thuringiensis* crystal protein gene into cotton plants. China Sci. (B) 4:367–373; 1991.

Zhang, B. H.; Feng, R.; Li, X. L.; Li, F. L. Anther culture and plant regeneration of cotton (*Gossypium klotzschianum* Anderss). Chinese Sci. Bull. 41:145–148; 1996h.

Zhang, B. H.; Feng, R.; Liu, F.; Yao, C. B. Direct induction of cotton somatic embryogenesis. Chinese Sci. Bull. 44:766–767; 1999.

Zhang, B. H.; Liu, F.; Liu, Z. H.; Wang, H. M.; Yao, C. B. Effects of kanamycin on tissue culture and somatic embryogenesis in cotton. Plant Growth Regul. (in press); 2001.

Zhang, B. H.; Liu, F.; Yao, C. B. Plant regeneration *via* somatic embryogenesis in cotton (*Gossypium hirsutum* L.). Plant Cell Tiss. Organ Cult. 60:89–94; 2000b.

Zhang, B. H.; Liu, F.; Yao, C. B.; Wang, K. B. Recent progress in cotton biotechnology and genetic engineering in China. Curr. Sci. 79:37–44; 2000a.

Zhang, B. H.; Wang, Q. L.; Feng, R.; Li, F. L.; Li, F. G.; Li, X. L. Somatic embryony patterns and plant regeneration in *Gossypium hirsutum* L. J. Agr. Biotechnol. 4(1):44–50; 1996a.

Zhang, B. H.; Zhao, B. S. Cotton biotechnology and its application. Beijing: China Agricultural Press; 1997:1–201.

COR-AAD1-00001235

© Indian Academy of Sciences

Inheritance and segregation of exogenous genes in transgenic cotton

BAO-HONG ZHANG[1]*, TENG-LONG GUO[2] and QING-LIAN WANG[3]

[1]*Key Laboratory of Cotton Genetic Improvement of the Ministry of Agriculture, Cotton Research Institute, Chinese Academy of Agricultural Sciences, Anyang Henan 455 112, People's Republic of China*
[2]*Institute of Heze Agricultural Sciences, Heze, Shandong 273 000, People's Republic of China*
[3]*Henan Vocational Technical Teacher's College, Xinxiang Henan 453 003, People's Republic of China*

Abstract

Three transgenic cotton varieties (lines) were chosen for the study of inheritance and segregation of foreign Bt (*Bacillus thuringiensis* toxin) and *tfdA* genes in cotton. The transformed cotton varieties CCRI 30 and NewCott 33B expressing the Bt *cryIA* gene, and cotton line TFD expressing the *tfdA* gene were crossed with CCRI 19, CCRI 12 and Lumian 6. The results confirm inheritance and segregation of (i) the exogenous Bt gene in transgenic CCRI 30 and NewCott 33B, governing resistance to bollworm, and (ii) the exogenous *tfdA* gene in transgenic TFD, governing resistance to the herbicide 2,4-D. Both resistance characters were governed by a single dominant nuclear gene, and were not affected by cytoplasm. Our data support the conclusion that foreign traits encoded by single genes are inherited and expressed in Mendelian fashion in cotton. Our results also indicate that a practical backcross breeding program could be used to develop cotton cultivars combining one or more resistance traits from foreign and native gene sources.

[Zhang B. H., Guo T.-L. and Wang Q.-L. 2000 Inheritance and segregation of exogenous genes in transgenic cotton. *J. Genet.* **79**, 71–75]

Introduction

Cotton is an important economic and fibre crop worldwide. In 1987 Umbeck *et al.* first obtained transgenic cotton plants. These transgenic cotton plants carried two marker genes, encoding neomycin phosphotransferase II (NPT-II) and chloramphenicol acetyltransferase (CAT), which specify resistance to the antibiotics kanamycin and chloramphenicol, respectively. Since then, much progress has been made in genetic engineering of cotton (Zhang and Zhao 1997). Many laboratories have obtained insect-resistant (Zhang and Feng 1998; Zhang and Zhang 1998) or herbicide-resistant transgenic cotton plants via *Agrobacterium*-mediated transformation (Umbeck *et al.* 1987; Rajasekaran *et al.* 1996), particle bombardment (Finer and McMullen 1990; McCabe and Martinell 1993), or pollen tube pathway (Zhang *et al.* 2000). Up to now, at least seven foreign genes have been introduced into the cotton genome. Of these, Bt (Perlak *et al.* 1990; Cousins *et al.* 1991; Xie *et al.* 1991; Jenkins *et al.* 1997; Ni *et al.* 1998), CpTI (Li *et al.* 1998a), API (Thomas *et al.* 1995) and STKI (Wang *et al.* 1998) genes provide resistance to insect pests; *tfdA* (Bayley *et al.* 1992; Lyon *et al.* 1993; Chen *et al.* 1994), AHAS (Rajasekaran *et al.* 1996) and *bar* (Keller *et al.* 1997) genes provide resistance to herbicides. Other genes, for disease resistance, drought resistance, salt resistance, cold resistance and fibre improvement, are being sought for introduction into cotton.

Cotton is among the first transformed crops to be commercialized, yet the inheritance and segregation of exogenous genes in commercial transgenic cotton varieties have not been reported. Although several recent studies have reported the inheritance and expression of some marker genes (NPT-II gene and CAT gene) and Bt *cryIA*, and environmental effects on expression of these genes (Sachs *et al.* 1998; Umbeck *et al.* 1989; Li *et al.* 1998), these studies were limited to the transgenic regenerative plants. Up to now, all of the commercial transgenic cotton varieties are not transgenic regenerative plants; they were bred by cross and/or backcross between transgenic regenerative plants and commercial varieties.

*For correspondence: 1632 West 6th Street Apt F, Austin, TX 78703, USA. E-mail: zbh68@hotmail.com, zbh68@yahoo.com.

Keywords. transgenic cotton; Bt gene; *tfdA* gene; inheritance; dominant gene.

COR-AAD1-00001236

It is very important to understand the inheritance, segregation and stability of exogenous genes for a given transgenic event. We conducted this study to plan an efficient plant breeding strategy and ensure the stability of expression in different genetic backgrounds. The objective was to characterize the inheritance, segregation and stability of two foreign genes (Bt and *tfdA*) in commercial transgenic cotton varieties.

Materials and methods

Plant material: Two commercial transgenic insect-resistant cotton varieties, CCRI 30 and NewCott 33B, and one transgenic herbicide-resistant cotton line, TFD, were selected for this study. CCRI 30 and NewCott 33B carry a Bt (*Bacillus thuringiensis*) gene and express the CryIA insecticidal proteins for resistance to *Heliothis virescens* and *Helicoverpa zea* (Lepidoptera: Noctuidae). TFD carries the *Alcaligenes eutrophus tfdA* gene which encodes 2,4-D monooxygenase, which confers resistance to the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D). CCRI 30 was provided by the Cotton Research Institute of the Chinese Academy of Agricultural Sciences, NewCott 33B comes from the commercial variety bred by Monsanto and Delta and Pine Land Company. CCRI 30 was passed for wide planting after a national examination and approval in 1998. NewCott 33B was imported by the Jida Company and was permitted for planting in 1997, and passed the biosafety evaluation in 1998 by the Chinese government. TFD comes from the Cotton Research Institute in Shanxi Province, and then was purified by selfcrossing at the Cotton Research Institute. The nontransgenic varieties used, CCRI 12, CCRI 19 and Lumian 6, come from the Cotton Research Institute. All materials were purified by continuous two-generation selfcross, to ensure the purity of resistance character genes.

Experimental design: Transgenic and nontransgenic cotton were planted by conventional method. At luxuriant blooming, CCRI 30, NewCott 33B and TFD were selfcrossed or crossed with conventional nontransgenic cotton cv. CCRI 12, CCRI 19 and Lumian 6. More than 100 selfcrossed or crossed bolls were harvested for the next experiments.

In the next year, the selfcrossed populations of transgenic materials and the F_1 seeds were planted in the field, and were examined for resistance to bollworm or 2,4-D by the methods described below. At the same time, more than 50 F_1 plants were selfcrossed or backcrossed with recurrent parents to produce F_2 and BC_1 populations.

In the third year, the F_2 and BC_1 seeds were planted in the field, and resistance was checked by the same method. The BC_1 plants were backcrossed with recurrent parents to produce BC_2 populations.

Identification of insect resistance: At six-to-eight-leaf stage, three to five bollworms were placed on the shoot of all test plants. Five days later, survival of bollworm and damage to the cotton plant were examined. Plants without live bollworms and no damage were classified as insect-resistant plants. Plants with live bollworm(s) and much damage were classified as not insect resistant.

Identification of herbicide resistance: The identification of resistance to 2,4-D was conducted in the field. At four-leaf stage, test plants were sprayed with 200 mg/L 2,4-D. After two weeks, growth of the plants was investigated. The plants without damage were resistant plants and the damaged plants were nonresistant.

Data analysis: Chi-square goodness-of-fit tests were performed on data from the F_2, F_3, BC_1 and BC_2 populations derived from crosses between CCRI 30, NewCott 33B or TFD and the nontransgenic conventional varieties CCRI 12, CCRI 19 or Lumian 6 to determine if the observed segregation ratios of Bt-positive or *tfdA*-positive plants to negative plants fit the expected Mendelian 3:1, 7:5 or 1:1 phenotypic ratio, respectively.

Results and analysis

Inheritance and segregation of exogenous Bt gene in transgenic cotton

After five days of exposure to bollworms, conventional cotton variety CCRI 19 was damaged badly in leaves and shoots. The leaves showed many holes and the shoots had been bitten off. The transgenic insect-resistant Bt-cotton, both CCRI 30 and NewCott 33B, grew well, and there was almost no trace of damage by bollworms.

Transgenic cotton varieties CCRI 30 and NewCott 33B were crossed with conventional cotton variety CCRI 19 to produce F_1 populations. The F_1 plants grew healthily and were not damaged by bollworms. The results (table 1) indicate that they are all insect-resistant plants. The results confirm that the insect-resistance character controlled by the exogenous Bt gene is a dominant character. It did not matter whether the transgenic insect-resistant cotton cultivars CCRI 30 and NewCott 33B were male parent or female parent in the cross with conventional cotton cultivar CCRI 19; all of the F_1 plants resisted bollworms (table 1). This indicates that the insect-resistance characteristic controlled by the Bt gene was not affected by cytoplasmic factors.

Table 1. Resistance of Bt-cotton F_1 populations to bollworm.

Cross	Resistant	Susceptible	Total
NewCott 33B × CCRI 19	89	0	89
CCRI 19 × NewCott 33B	107	0	107
CCRI 30 × CCRI 19	63	0	63
CCRI 19 × CCRI 30	97	0	97

COR-AAD1-00001237

Table 2. Segregation of backcross and F_2 populations for resistance to bollworm.

Cross	Total	Resistant	Susceptible	Ratio	Test ratio	χ^2	P
NewCott 33B × CCRI 19 F_2	243	183	60	3.05:1	3:1	0.012	0.50 ~ 0.95
CCRI 19 × NewCott 33B F_2	357	266	91	2.92:1	3:1	0.031	0.50 ~ 0.95
CCRI 30 × CCRI 19 F_2	277	208	69	3.01:1	3:1	0.001	0.95 ~ 0.99
CCRI 19 × CCRI 30 F_2	433	326	107	3.05:1	3:1	0.019	0.50 ~ 0.95
(NewCott 33B × CCRI 19) × NewCott 33B BC_1	58	58	0				
(CCRI 19 × NewCott 33B) × NewCott 33B BC_1	67	67	0				
(CCRI 30 × CCRI 19) × CCRI 30 BC_1	63	63	0				
(CCRI 19 × CCRI 30) × CCRI 30 BC_1	63	63	0				
(NewCott 33B × CCRI 19) × CCRI 19 BC_1	140	69	71	0.97:1	1:1	0.029	0.50 ~ 0.95
(CCRI 19 × NewCott 33B) × CCRI19 BC_1	204	102	102	1:1	1:1	0	>0.99
(CCRI 30 × CCRI 19) × CCRI 19 BC_1	265	135	130	1.04:1	1:1	0.094	0.50 ~ 0.95
(CCRI 19 × CCRI 30) × CCRI 19 BC_1	237	118	119	0.99:1	1:1	0.004	0.95

To determine how many Bt gene copies were present in the cotton genome, F_1 plants of Bt-cotton and conventional cotton were selfcrossed to produce F_2 seeds and backcrossed with recurrent parent cultivars to produce BC_1 seeds. The phenotypic segregation ratios of the insect-resistance character and Chi-square values are presented in table 2. Phenotypic segregation ratios in F_2 populations are 2.92–3.05 : 1, and the Chi-square test indicated that segregation in F_2 populations fits the Mendelian 3:1 monogenic ratio. The results indicate that the insect-resistance characteristic conferred by the Bt gene is controlled by one pair of dominant genes. Phenotypic segregation ratios in BC_1 and their Chi-square goodness-of-fit tests also support this conclusion (table 2). In addition, the test results of BC_1 populations also indicate that the insect-resistance character provided by the Bt gene was not affected by cytoplasmic factors.

Inheritance and segregation of exogenous tfdA gene in transgenic cotton

Cotton is highly sensitive to 2,4-D. Even low concentrations of 2,4-D make cotton leaves to twist and fall off, cause great damage to the crop, and result in yield loss. Use of 2,4-D is prohibited in cotton fields and adjacent fields. Genetically engineered broadleaf herbicide protection for 2,4-D in cotton was obtained by incorporating an *Alcaligenes eutrophus* gene encoding 2,4-D monooxygenase (*tfdA*). Transformants carrying *tfdA* exhibited 50-fold to 100-fold more tolerance to 2,4-D compared with nontransformed plants (Bayley *et al.* 1992; Chen *et al.* 1994). We have bred transgenic cotton carrying the *tfdA* gene, and obtained elite transgenic 2,4-D-resistant cotton line TFD. Field examination indicated that TFD is highly resistant to 2,4-D. TFD plants could grow well when 1000 mg/L of 2,4-D was sprayed on their leaves. In contrast, 50 mg/L of 2,4-D could greatly damage nontransgenic cotton leaves and plants.

It is very important to research the inheritance and segregation of exogenous *tfdA* gene in transgenic cotton for breeding and application of transgenic 2,4-D-resistant cotton cultivars. We found that when TFD was crossed with CCRI 12, CCRI 19 or Lumian 6, irrespective of whether TFD was the male or female parent, all of the F_1 plants resisted 2,4-D (table 3). The result confirms that the herbicide-resistance characteristic provided by the exogenous *tfdA* gene is a dominant phenotype and is not affected by cytoplasmic factors.

Table 3. Resistance of tfdA-cotton F_1 progeny to 2,4-D.

Cross	Resistant	Susceptible	Total
TFD × CCRI 19	36	0	36
CCRI 19 × TFD	30	0	30
TFD × CCRI 12	27	0	27
CCRI 12 × TFD	32	0	32
TFD × Lumian 6	53	0	53
Lumian 6 × TFD	28	0	28

Segregation for resistance was observed in F_2 populations of crosses of TFD and the conventional cultivars, and some of the F_2 plants were highly resistant to 2,4-D and others were not (table 4). The phenotypic segregation ratios (resistant : susceptible) were in the range 2.79–3.54 : 1. By the Chi-square test, the results fit the theoretical Mendelian segregation ratio of 3:1. This indicates that the herbicide-resistance character provided by the exogenous *tfdA* gene is a dominant character and is controlled by a pair of genes. It also indicates that one copy of *tfdA* was inserted in the cotton genome when cotton explants were transformed by the *Agrobacterium*-mediated method.

Segregation for resistance in the F_3 populations derived from TFD and CCRI 19 is also shown in table 4. The results indicate that segregation in the F_3 populations generally fit the Mendelian 7:5 monogenic ratio. These results support the conclusions from the F_2 populations.

The segregation ratios of BC_1 populations (table 5) also confirm that the herbicide-resistance character is governed by a single dominant nuclear gene pair. When conventional cultivars CCRI 12, CCRI 19 or Lumian 6 were the recurrent parent, the resistance segregation ratios in BC_1 populations were 0.84–1.26 : 1, which fit the Mendelian 1:1 monogenic ratio. When TFD was used as the recurrent parent, all BC_1 plants were herbicide resistant.

COR-AAD1-00001238

Table 4. Segregation of F_2 and F_3 populations for resistance to 2,4-D.

Cross	Total	Resistant	Susceptible	Ratio	Test ratio	χ^2	*P*
TFD × CCRI 19 F_2	125	92	33	2.79:1	3:1	0.131	0.50 ~ 0.95
CCRI 19 × TFD F_2	118	92	26	3.54:1	3:1	0.554	0.30 ~ 0.50
TFD × CCRI 12 F_2	131	99	32	3.09:1	3:1	0.023	0.50 ~ 0.95
CCRI 12 × TFD F_2	131	96	35	2.74:1	3:1	0.206	0.50 ~ 0.95
TFD × Lumian 6 F_2	127	96	31	3.10:1	3:1	0.024	0.50 ~ 0.95
Lumian 6 × TFD F_2	98	73	25	2.92:1	3:1	0.014	0.50 ~ 0.95
TFD × CCRI 19 F_3	128	78	50	7.80:5	7:5	0.356	0.50 ~ 0.95
CCRI 19 × TFD F_3	109	67	42	7.98:5	7:5	0.442	0.50 ~ 0.95

Table 5. Segregation of BC_1 populations for resistance to 2,4-D.

Cross	Total	Resistant	Susceptible	Ratio	Test ratio	χ^2	*P*
(TFD × CCRI 19) × CCRI 19 BC_1	97	48	49	0.98:1	1:1	0.010	0.50 ~ 0.95
(CCRI 19 × TFD) × CCRI 19 BC_1	86	48	38	1.26:1	1:1	1.163	0.20 ~ 0.30
(TFD × CCRI 12) × CCRI 12 BC_1	88	44	44	1:1	1:1	0	>0.99
(CCRI 12 × TFD) × CCRI 12 BC_1	101	47	54	0.87:1	1:1	0.485	0.30 ~ 0.50
(TFD × Lumian 6) × Lumian 6 BC_1	76	39	37	1.05:1	1:1	0.053	0.30 ~ 0.50
(Lumian 6 × TFD) × Lumian 6 BC_1	59	27	32	0.84:1	1:1	0.424	0.50 ~ 0.95
(TFD × CCRI 19) × TFD BC_1	62	62	0				
(CCRI 19 × TFD) × TFD BC_1	31	31	0				
(TFD × CCRI 12) × TFD BC_1	59	59	0				
(CCRI 12 × TFD) × TFD BC_1	58	58	0				

Discussion

This study was undertaken to determine the inheritance and segregation of two exogenous genes (Bt insecticidal protein gene and *tfdA* herbicide resistance gene) in commercial cotton cultivars. Knowledge of the inheritance and segregation pattern of the inserted gene would provide the opportunity to develop transgenic germplasm, strategies to maximize the efficiency of developing improved germplasm, and the ability to breed fine varieties carrying these foreign resistance genes.

The Bt and TfdA phenotypes segregated as simple dominant Mendelian traits. This is consistent with the earlier reports by Perlak *et al.* (1990) and Sachs *et al.* (1998). These two exogenous genes can thus be inherited steadily into progeny, which indicates that we can obtain new fine varieties via crossing and/or backcrossing conventional varieties with the transgenic variety.

Non-Mendelian segregation of exogenous Bt gene was observed in F_2 populations derived from transgenic regenerative plants and conventional cotton lines (Sachs *et al.* 1998). This inheritance pattern was not observed in commercial transgenic cotton cultivars. The reason may be that some of the exogenous genes were inserted into non-nuclear genomes during transformation. These genes could be lost in crosses with commercial cotton cultivars.

Transformation is considered to be a random event, with each inserted gene at a unique location (Kohel *et al.* 2000). Thus, each transgenic plant is the result of a separate insertion event, and this event could translate to different degrees of success in attempts at backcross improvement, depending on associated linkages (Kohel *et al.* 2000). The number and sites of gene insertions into the plant genome have been found to influence gene expression in transgenic regenerative cotton plants. Several researchers have found differences in injury by tobacco budworm (*Heliothis virescens*) and in the behaviour, growth and survival of the budworm (Benedict *et al.* 1992, 1996). These differences were attributed to positional effects on Bt gene expression and/or somaclonal variation (Benedict *et al.* 1996).

Conclusions

Our results have demonstrated the Mendelian inheritance of exogenous genes in cotton. Exogenous Bt and *tfdA* genes can be steadily inherited and expressed in commercial transgenic cotton cultivars, free of cytoplasmic effects. The insect-resistance and herbicide-resistance characters are governed by a single dominant nuclear Bt or *tfdA* gene pair, respectively. Classical cross and backcross can introgress the exogenous resistance genes of the transformant into conventional cotton varieties, and one can breed new cotton varieties carrying desired characters such as insect resistance, disease resistance and so on.

References

Bayley C., Trolinder N., Ray C., Morgan M. M., Quisenberry J. E. and Ow D. W. 1992 Engineering 2,4-D resistance into cotton. *Theor. Appl. Genet.* **83**, 645–649.

Benedict J. H., Altman D. W., Umbeck P. F. and Ring D. R. 1992 Behavior, growth, survival, and plant injury for *Heliothis virescens* (F.) (Lepidoptera: Noctuidae) on transgenic Bt cottons. *J. Econ. Entomol.* **85**, 589–593.

COR-AAD1-00001239

Benedict J. H., Sachs E. S., Altman D. W., Deaton W. R., Kohel R. J., Ring D. R. and Berberich S. A. 1996 Field performance of cottons expressing transgenic Cry IA insecticidal proteins for resistance to *Heliothis virescens* and *Helicoverpa zea* (Lepidoptera: Noctuidae) on transgenic Bt cottons. *J. Econ. Entomol.* **89**, 230–238.

Chen Z. X., Llewellyn D. J., Fan Y. L., Li S. J., Guo S. D., Jiao G. L. and Zhao J. X. 1994 2,4-D resistant transgenic cotton plants produced by *Agrobacterium*-mediated gene transfer. *Sci. Agric. Sinica* **27**, 31–37.

Cousins Y. L., Lyon B. R. and Llewellyn D. J. 1991 Transformation of an Australian cotton cultivar: prospects for cotton improvement through genetic engineering. *Aus. J. Plant Physiol.* **18**, 481–494.

Finer J. J. and McMullen M. D. 1990 Transformation of cotton (*Gossypium hirsutum* L.) via particle bombardment. *Plant Cell Rep.* **8**, 586–589.

Gasser C. S. and Fraley R. T. 1989 Genetically engineering plants for crop improvement. *Science* **244**, 1293–1299.

Jenkins J. N., McCarty J. C., Buehler R. E., Kiser J., Williams C. and Wofford T. 1997 Resistance of cotton with δ-endotoxin genes from *Bacillus thuringiensis* var. kurstaki on selected Lepidopteran insects. *Agron. J.* **89**, 768–780.

Keller G., Spatola L., McCabe D., Martinell B., Swain W. and John M. E. 1997 Transgenic cotton resistant to herbicide bialaphos. *Transgenic Res.* **6**, 385–392.

Kohel R. J., Quisenberry J. E., Cartwright G. and Yu J. 2000 Linkage analysis of transgenes inserted into cotton via Agrobacterium tumefaciens transformation. *J. Cotton Sci.* **4**, 65–69.

Li Y. E., Chen Z. X., Wu X., Li S. J., Jiao G. L., Wu J. H., Fan X. P., Meng J. H., Zhu Z., Wang W., Zhu Y., Xu H. L., Xiao G. F. and Li X. H. 1998a Obtaining transgenic cotton plants with cowpea trypsin inhibitor gene. *Acta Gossypii Sinica* **10**, 237–243.

Li Y. E., Jiao G. L., Li S. J., Fan X. P., Wu X. and Chen Z. X. 1998b Obtaining of homozygous transgenic cotton lines and their genetic analysis. *Sci. Agric. Sinica* **31**, 96–98.

Lyon B. R., Cousins Y. L., Llewellyn D. J. and Dennis E. S. 1993 Cotton plants transformed with a bacterial degradation gene are protected from accidental spray drift damage by the herbicide 2,4-dichlorophenoxyacetic acid. *Transgenic Res.* **2**, 162–169.

McCabe D. E. and Martinell B. J. 1993 Transformation of elite cotton cultivars via particle bombardment of meristems. *Bio/Technology* **11**, 596–598.

Ni W. C., Zhang Z. L. and Guo S. D. 1998 Development of transgenic insect-resistant cotton plants. *Sci. Agric. Sinica* **31**, 8–13.

Perlak F. J., Deaton R. W., Armstrong R. L., Fuchs R. L., Sims S. R., Greenplate J. T. and Fischhoff D. A. 1990 Insect resistant cotton plants. *Bio/Technology* **8**, 939–943.

Rajasekaran K., Grula J. W. and Hudspeth R. L. 1996 Herbicide-resistant Acale and Coker cottons transformation with a native gene encoding mutant forms of acetohydroxyacid synthase. *Mol. Breeding* **2**, 307–319.

Sachs E. S., Benedict J. H., Stelly D. M., Taylor J. F., Altman D. W., Berberich S. A. and Davis S. K. 1998 Expression and segregation of genes encoding CryIA insecticidal proteins in cotton. *Crop Sci.* **38**, 1–11.

Thomas J. C., Adams D. G., Keppenne V. D., Wasmann C. C., Brown J. K., Kanost M. R. and Bohnert H. J. 1995 Protease inhibitors of *Manduca sexta* expressed in transgenic cotton. *Plant Cell Rep.* **14**, 758–762.

Umbeck P., Johnson G. and Barton K. 1987 Genetically transformed cotton (*Gossypium hirsutum* L.) plants. *Bio/Technology* **5**, 236–266.

Umbeck P., Swain W. and Yang N. S. 1989 Inheritance and expression of genes for kanamycin and chloramphenicol resistance in transgenic cotton plant. *Crop Sci.* **29**, 196–201.

Wang W., Zhu Z. and Deng C. Y. 1998 Obtaining of pest resistant cotton by transforming mediated Agrobacterium. In *New frontiers in cotton research, World Cotton Research Conference-2*, p. 119. Athens, Greece.

Xie D. X., Fan Y. L., Ni W. C. and Huang J. Q. 1991 Transformed *Bacillus thuringiensis* crystal protein gene into cotton plants. *China Sci.* **B4**, 367–373.

Zhang B.-H. and Feng R. 1998 Achievement, problems and strategies of transgenic insect-resistant cotton. *Acta Agron. Sinica* **24**, 246–248.

Zhang B.-H. and Zhang L. Z. 1998 *Pest-resistant cotton and its cultivation*. Shangdong Science and Technology Press, Jinan.

Zhang B.-H. and Zhao B. S. 1997 *Cotton biotechnology and its application*. Chinese Agricultural Press, Beijing.

Zhang B.-H., Liu F., Yao C. B. and Wang K. B. 2000 Recent progress in cotton biotechnology and genetic engineering in China. *Curr. Sci.* **79**, 37–44.

Received 8 March 2000; in revised form 14 June 2000

COR-AAD1-00001240

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Dec. 1996, p. 4318–4322 Vol. 62, No. 12
0099-2240/96/$04.00+0
Copyright © 1996, American Society for Microbiology

Complete Microbial Degradation of Both Enantiomers of the Chiral Herbicide Mecoprop [(*RS*)-2-(4-Chloro-2-Methylphenoxy)propionic Acid] in an Enantioselective Manner by *Sphingomonas herbicidovorans* sp. nov.

CHRISTIAN ZIPPER, KATHRIN NICKEL, WERNER ANGST, AND HANS-PETER E. KOHLER*

Swiss Federal Institute for Environmental Science and Technology (EAWAG) and Swiss Federal Institute of Technology (ETH), CH-8600 Dübendorf, Switzerland

Received 5 April 1996/Accepted 24 September 1996

***Sphingomonas herbicidovorans* MH (previously designated *Flavobacterium* sp. strain MH) was able to utilize the chiral herbicide (*RS*)-2-(4-chloro-2-methylphenoxy)propionic acid (mecoprop) as the sole carbon and energy source. When strain MH was offered racemic mecoprop as the growth substrate, it could degrade both the (*R*) and the (*S*) enantiomer to completion, as shown by biomass formation, substrate consumption, and stoichiometric chloride release. However, the (*S*) enantiomer disappeared much faster from the culture medium than the (*R*) enantiomer. These results suggest the involvement of specific enzymes for the degradation of each enantiomer. This view was substantiated by the fact that resting cells of strain MH grown on (*S*)-mecoprop were able to degrade the (*S*) but not the (*R*) enantiomer of mecoprop. Accordingly, resting cells of strain MH grown on (*R*)-mecoprop preferentially metabolized the (*R*) enantiomer. Nevertheless, such cells could transform (*S*)-mecoprop at low rates. Oxygen uptake rates with resting cells confirmed the above view, as oxygen consumption was strongly dependent on the growth substrate. Cells grown on (*R*)-mecoprop showed oxygen uptake rates more than two times higher upon incubation with the (*R*) than upon incubation with the (*S*) enantiomer and vice versa.**

Chiral (*RS*)-2-(4-chloro-2-methylphenoxy)propionic acid (mecoprop) (Fig. 1) is a systemic, postemergence herbicide developed in the 1950s for use in the control of broad-leaved weeds in cereal crops and lawns, and commercially it is often used in formulations together with other herbicides (22). Already in 1953, studies about the stereochemistry of plant growth regulators revealed that the herbicidal activity was associated only with the (*R*) enantiomer of mecoprop (11). Enantiomerically pure (*R*)-2-(4-chloro-2-methylphenoxy)propionic acid was made available commercially, beginning in 1987, as mecoprop-P, yet many commercial formulations contain mecoprop—the racemic mixture—and not mecoprop-P—the pure (*R*) enantiomer (22). The importance of considering the stereochemistry of pesticides, crop protectants, detergents, and drugs—not only for the investigation of their biological effects, but also for studying their environmental fate—has recently been emphasized (1, 2, 15). Nevertheless, such considerations are missing in many studies about the environmental fate and the microbial degradation of mecoprop (4, 7, 9, 12), and this might be one of the reasons why only limited progress has been made in unravelling the exact mechanism of microbial mecoprop degradation.

While the microbial degradation pathway of 2,4-D (2,4-dichlorophenoxyacetic acid) has been investigated in detail (3, 10, 13, 14, 20), the microbial metabolism of mecoprop remains to be elucidated. Mecoprop is known to degrade in natural environments by the action of microorganisms, and disappearance rates in soil ranging from 5 to 11 days have been reported (16, 17). Although mecoprop degradation most likely occurs in the top soil, the rather high mobility of the pollutant constitutes a risk of contamination for underlying aquifers and groundwater resources (5). Mecoprop vanished when added at low concentrations to an aerobic aquifer under laboratory conditions. In this case, a lag phase was observed after initial degradation of approximately 50% of the compound, which was followed by complete degradation. This phenomenon was attributed to microbial discrimination between the two stereoisomers of mecoprop (5).

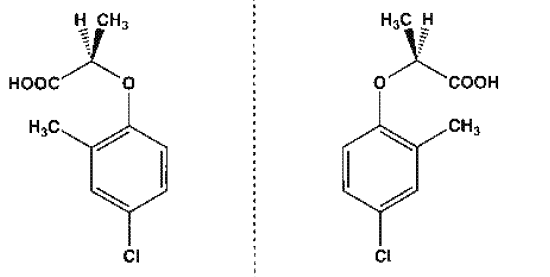
The failure of being able to isolate pure cultures of microorganisms with the ability to use mecoprop as the sole carbon and energy source has led researchers to postulate that in natural environments the complete degradation of mecoprop may only occur through the synergistic activities of a consortium of microorganisms (19). However, two pure bacterial cultures with the ability to utilize mecoprop as a growth substrate have been recently described (6, 19). Both strains—*Sphingomonas herbicidovorans* MH (previously designated *Flavobacterium* sp. strain MH [6]) and *Alcaligenes denitrificans*—were able to use mecoprop as the sole carbon and energy source. Experiments with the pure enantiomers of mecoprop as growth substrates revealed that *A. denitrificans* exclusively degraded the (*R*) enantiomer of mecoprop, while the (*S*) enantiomer remained unaffected (19). In the case of *S. herbicidovorans* MH, only mecoprop—the racemic mixture—was tested as a substrate in growth experiments and therefore no conclusions about enantiomer-specific processes could be drawn (6).

In this study, we investigated the degradation of mecoprop by *S. herbicidovorans* MH by using enantiomer-specific analytical tools and the pure enantiomers as well as the racemic mixture of 2-(4-chloro-2-methylphenoxy)propionic acid as growth substrates. We were able to unequivocally demonstrate for the first time that a pure bacterial culture is able to degrade

* Corresponding author. Mailing address: Department of Microbiology, EAWAG, Ueberlandstrasse 133, CH-8600 Dübendorf, Switzerland. Phone: 41 1 823 5521. Fax: 41 1 823 5547. Electronic mail address: kohler@eawag.ch.

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001241



(*R*)-2-(4-Chloro-2-methylphenoxy)-propionic acid | (*S*)-2-(4-Chloro-2-methylphenoxy)-propionic acid

FIG. 1. Structural formulas of the (*R*) and the (*S*) enantiomer of 2-(4-chloro-2-methylphenoxy)propionic acid (mecoprop).

both enantiomers of mecoprop to completion. The data presented suggest the involvement of specific enzymes for the initial attack on each enantiomer. We intend to stress the necessity of considering stereochemistry when studying the degradation of chiral compounds.

MATERIALS AND METHODS

Bacterial strain and culture conditions. *S. herbicidovorans* MH (previously designated *Flavobacterium* sp. strain MH) was a gift from F. Streichsbier, Vienna, Austria. It was isolated from a soil column with (*RS*)-2-(2,4-dichlorophenoxy) propionic acid as the carbon and energy source (6). Strain MH was identified by the Deutsche Stammsammlung von Mikroorganismen und Zellkulturen (DSM), Braunschweig, Germany, as a new species belonging to the *Sphingomonas* group within the α subgroup of the proteobacteria.

(i) Mineral medium. The chloride-free mineral salts medium used for growth of strain MH consisted of 20 mM phosphate buffer (KH_2PO_4-Na_2HPO_4, pH 6.5) containing $(NH_4)_2SO_4$ (0.12 g/liter), $MgSO_4 \cdot 7H_2O$ (73 mg/liter), and $Ca(NO_3)_2 \cdot 4H_2O$ (1 mg/liter). The following trace elements were supplied at 0.2 ml/liter from an acidified (H_3BO_3, 0.1 g/liter; H_2SO_4, 5 ml/liter) stock solution containing $FeSO_4 \cdot 7H_2O$ (2.5 g/liter), $MnSO_4 \cdot H_2O$ (0.75 g/liter), $ZnSO_4 \cdot 7H_2O$ (1.3 g/liter), $CuSO_4 \cdot 5H_2O$ (0.25 g/liter), $Co(NO_3)_2 \cdot 6H_2O$ (0.3 g/liter), $Na_2MoO_4 \cdot 2H_2O$ (0.15 g/liter), and $NiSO_4 \cdot 7H_2O$ (0.1 g/liter). After autoclaving, 1 ml of a vitamin solution containing pyridoxin-HCl (0.1 g/liter), thiamine-HCl (0.05 g/liter), riboflavin (0.05 g/liter), nicotinic acid (0.05 g/liter), Ca-pantothenate (0.05 g/liter), 4-aminobenzoic acid (0.05 g/liter), nicotinamide (0.05 g/liter), vitamin B_{12} (0.05 g/liter), biotin (0.02 g/liter), and folic acid (0.02 g/liter) was added by sterile filtration (0.2 μm) per liter of medium. The medium was supplemented with 20 mg of peptone (Biolife, Milan, Italy) per liter. The carbon sources were added to the mineral medium before autoclaving, as no losses could be observed upon heating to 121°C for 30 min.

(ii) Complex medium. The chloride-free complex medium consisted of 6.0 g of peptone per liter, 1.2 g of yeast extract per liter, and 0.4 g of glucose per liter in the mineral medium described above.

Chemicals. When not explicitly stated, chemicals were purchased from Fluka Chemicals, Buchs, Switzerland. Analytical standards of (*R*)-mecoprop were obtained from Riedel-de Haën, Seelze, Germany. Technical grade (*RS*)-mecoprop (92%) was donated by BASF Chemicals, Limburgerhof, Germany. It was recrystallized twice, first from 80% formic acid and then from petroleum (bp, 60 to 80°C), to form small, white crystals. The (*S*) and the (*R*) enantiomer of mecoprop were resolved from the racemic mixture by letting the recrystallized technical grade (*RS*)-mecoprop react with (*R*)- and (*S*)-phenylethylamine, respectively, to form the phenylethylamine diastereoisomeric salts by a procedure described by Matell (11). The enantiomer excess (ee) of the (*S*)- and the (*R*)-mecoprop was determined by enantiomer-specific high-performance liquid chromatography (HPLC) as 99 and 92%, respectively. The authenticity of the (*S*)- and the (*R*)-mecoprop was confirmed by ^{1}H- and ^{13}C-nuclear magnetic resonance spectroscopy.

Growth experiments. Precultures used for the inoculation of the media of the growth experiments were grown in chloride-free complex medium in 300-ml Erlenmeyer flasks at a temperature of 30°C. In order to remove the residual complex carbon substrate prior to the growth experiments, the precultures were harvested by centrifugation (15 min, 7,500 × *g*, 4°C) during the midexponential growth phase, washed, and resuspended in mineral medium in such a way that an optical density of 1.3 absorbance units at 546 nm (0.3 g [dry weight] per liter) was achieved. This cell suspension was used for inoculation (2% [vol/vol]). Growth experiments were performed in 1-liter Erlenmeyer flasks equipped with magnetic stirring bars. The flasks were equipped with a silicon tube that allowed easy sterile sampling. The experiments were carried out in a water bath at a temperature of 30°C with stirring held constant at 370 rpm. The initial mecoprop concentration was 0.46 mM (100 mg/liter).

Determination of growth yields. Dry weight was measured by filtering defined volumes of the culture broth through tared 0.20-μm-pore-size polycarbonate filters dried at 100°C (Nucleopore, Tübingen, Germany). Yields were calculated after determination of residual substrate and dissolved organic carbon (DOC).

Substrate consumption of resting cells. Cells were grown on (*R*)- or (*S*)-mecoprop, centrifuged, and washed. Each cell suspension (67 mg [dry weight] per liter for (*R*)-mecoprop-grown cells and 86 mg [dry weight] per liter for (*S*)-mecoprop-grown cells) was divided into three parts and the assay substrates (*R*)-, (*S*)-, and (*RS*)-mecoprop were added at a concentration of 0.2 mM in the final assay volume of 20 ml. The cell suspensions were transferred to 100-ml Erlenmeyer flasks and incubated at 30°C on a rotary shaker operated at 125 rpm.

Oxygen uptake of resting cells. Oxygen uptake was measured polarographically with a Clark type oxygen electrode (Rank Brothers, Cambridge, United Kingdom). The electrode was mounted to a reaction vessel (3 ml) whose temperature was held constant at 30°C. The assay was carried out with a substrate concentration of 1 mM and biomass concentrations in the range of 40 to 130 mg (dry weight) per liter. Specific oxygen uptake rates were corrected for endogenous rates.

Analytical procedures. (i) DOC. Samples were centrifuged (12,000 × *g* for 15 min), acidified with HCl (pH 2), freed from dissolved CO_2 by purging with N_2 for 10 min, and finally analyzed in a Tocor 2 carbon analyzer (Maihak, Hamburg, Germany).

(ii) HPLC. HPLC analyses were performed on a Waters 625 LC system with a WISP 700 autosampler and a 901 photodiodearray detector (Waters Millipore Corp., Milford, Mass.). (*R*)- and (*S*)-mecoprop were separated on a Nucleodex-α-PM column (200 by 4.0 mm) with permethylated α-cyclodextrin as the stationary phase (Macherey-Nagel, Düren, Germany) and detected at a wavelength of 230 nm. The system was operated isocratically with an eluent consisting of 70% methanol and 30% NaH_2PO_4 (50 mM, pH 3.0) at a flow rate of 0.7 ml/min. Retention times were typically 8.5 min for (*R*)- and 11.5 min for (*S*)-mecoprop. The detection limit for mecoprop was 4 μM.

(iii) Ion chromatography. Chloride concentration was measured with an ion chromatograph equipped with a conductivity detector (Dionex, Olten, Switzerland) after separation on an anion-exchange column (Dionex IonPac AS11; alkanol quaternary ammonium as functional group) with a gradient from 5 to 12 mM NaOH as the mobile phase at a flow rate of 1 ml/min. Subsequent elution of other anions was obtained with 25 mM NaOH. The detection limit for chloride was 20 μM.

(iv) ^{1}H- and ^{13}C-NMR. ^{1}H- and ^{13}C-NMR spectra were obtained with a 400-MHz Bruker AMX-400 NMR spectrometer (Spectrospin, Fällanden, Switzerland). Samples were dissolved in $CDCl_3$.

(v) UV-VIS spectroscopy. UV-visible light (VIS) spectra were obtained with an UVIKON 860 spectrophotometer (Kontron, Zürich, Switzerland).

RESULTS AND DISCUSSION

Taxonomical characteristics of strain MH. Strain MH is a yellow-pigmented, gram-negative, oxidase- and catalase-positive rod with a length of 1.5 to 4.0 μm and a width of 0.6 to 0.8 μm. It is aminopeptidase positive and ADH and urease negative, lyses only esculin but not gelatin or DNA, cannot reduce nitrate, is not able to grow anaerobically, and does not form acids from sugars. Analysis of the cellular fatty acids showed that the strain belongs to the *Sphingomonas* group within the α subgroup of the proteobacteria. Partial sequencing of the gene for the 16S rRNA allowed an unequivocal classification of strain MH as a member of this group. The highest degree of similarity (99%) was obtained with the gene for the 16S rRNA of *Rhizomonas suberifaciens* and *S. yanoikuyae*. DNA-DNA hybridization experiments between strain MH and the type strains of these two species (37.4% identity to *R. suberifaciens* DSM 7465 and 33.8% identity to *S. yanoikuyae* DSM 7462) showed that strain MH belongs to a new species within the group. Although it has been suggested that *S. yanoikuyae* should be placed into the genus *Rhizomonas* (18, 21), we decided to name strain MH *S. herbicidovorans*, because—according to opinion 14 of the *International Code of Nomenclature of Bacteria* (8)—the genus name *Rhizomonas* is a nomen rejiciendum and therefore usage of this name is in violation of rule 28a of the code. *S. herbicidovorans* MH has been deposited as

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001242

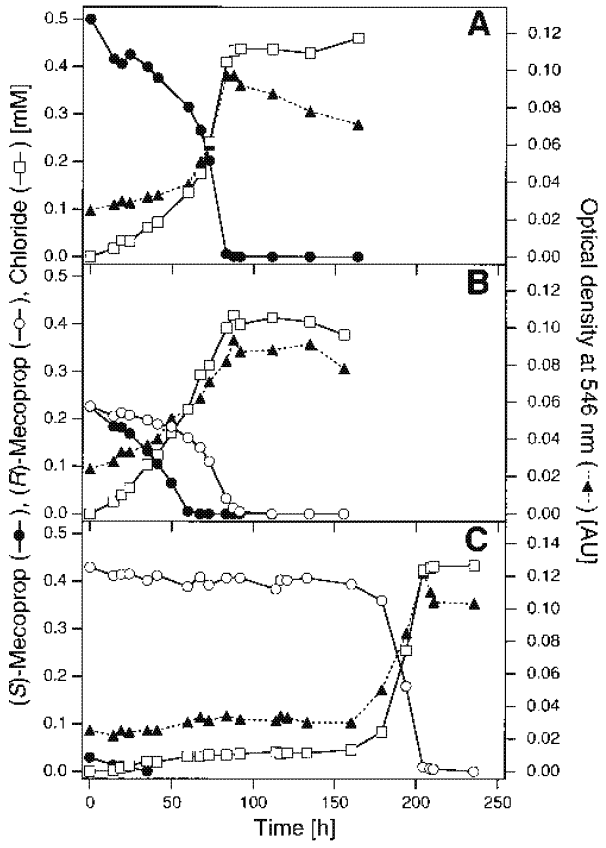


FIG. 2. Growth experiments with *S. herbicidovorans* MH incubated with (*R*)-2-(4-chloro-2-methylphenoxy)propionic acid, (*S*)-2-(4-chloro-2-methylphenoxy) propionic acid, and racemic (*RS*)-2-(4-chloro-2-methylphenoxy)propionic acid (mecoprop). (A) Growth on the pure (*S*) enantiomer. (B) Growth on the racemic mixture. (C) Growth on the pure (*R*) enantiomer. The inocula for these experiments were noninduced cells grown on complex media.

the type strain with the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig (DSM 11019).

It took at least 3 days at 30°C before circular, brightly yellow colonies with a smooth surface could be observed on nutrient agar plates. The shape of the colonies changed to a flower-like appearance when plates were incubated over longer time periods (up to 7 days). Cells secreted a brownish pigment into the medium when grown on solid or liquid nutrient broth. Cells taken from the exponential growth phase were rod shaped and often sticked together forming a chain, whereas cells taken from the stationary phase were almost of a coccoid shape. In mineral salts medium with racemic mecoprop as the sole source of carbon and energy, strain MH needed supplements of vitamins and peptone (20 mg/liter) for growth. Minimal agar plates had to be incubated up to 7 days until colonies appeared. These colonies were smaller than the colonies grown on nutrient agar and not as intensely colored, but upon examination with a stereoscopic microscope, both colony types looked identical.

Growth characteristics. Growth experiments with the single pure enantiomers as well as with the racemic mixture of mecoprop as growth substrates revealed that strain MH could completely degrade both enantiomers of mecoprop (Fig. 2A to C), albeit in an enantioselective manner. The inocula for all the growth experiments were noninduced cells grown on complex media. Chloride release was stoichiometric in all growth experiments, a strong hint for the destruction of the aromatic ring. Growth on (*S*)-mecoprop was completed after about 3.5 days irrespective of whether the (*R*) enantiomer was present. When racemic mecoprop was the growth substrate, degradation of the enantiomers was sequential, with (*S*)-mecoprop being degraded first. The (*R*) enantiomer was only slowly degraded during the first 50 h of such experiments, but the degradation rate for (*R*)-mecoprop markedly increased after (*S*)-mecoprop was completely consumed (Fig. 2B). Cells started to grow after a lag phase of about 7 days, when pure (*R*)-mecoprop was the sole carbon and energy source (Fig. 2C). Once growth commenced, cells grew as fast as when (*S*)-mecoprop was the growth substrate. In all growth experiments, strain MH did not excrete any metabolites that could be detected by HPLC into the medium. The fact that *S. herbicidovorans* MH was able to utilize both enantiomers of the chiral herbicide (*RS*)-2-(4-chloro-2-methylphenoxy)propionic acid (mecoprop) as the sole carbon and energy source is in contrast to findings with *A. denitrificans*, which—besides strain MH—is the only strain maintained in pure culture that has been described to metabolize mecoprop. *A. denitrificans* could only degrade the (*R*) enantiomer of mecoprop, while the (*S*) enantiomer remained untouched (19). The enantiomer-specific degradation pattern (Fig. 2) suggests that there is more than one enzyme system involved in the degradation of (*S*)- and the (*R*)-mecoprop in *S. herbicidovorans* MH.

The growth rates of strain MH growing on (*S*)-, (*R*)-, and (*RS*)-mecoprop (0.5 mM) were 0.039, 0.032, and 0.016 h^{-1} with yields of 0.35, 0.30, and 0.30 g (dry weight) per g of substrate, respectively. Hence, both enantiomers are equally well suited as a carbon and energy source for this bacterium. The values of the growth yields are equal to the one obtained with *Alcaligenes eutrophus* JMP134 growing on 2-methylphenoxyacetic acid and about twice the ones for growth of *A. eutrophus* on 2,4-dichlorophenoxyacetic acid and 4-chloro-2-methylphenoxyacetic acid, respectively (14). DOC levels measured at the beginning of a growth experiment and after depletion of the substrates corresponded to a complete conversion of mecoprop into biomass and presumably carbon dioxide. The decreased growth rate of 0.016 h^{-1} with racemic mecoprop as the growth substrate indicates that in this case the growth rate was affected by the concentration, which for each enantiomer was only half the value of what it was in the experiments with the pure enantiomers. Additional growth experiments showed that increasing the concentration of the racemic mixture caused an increase of the growth rate (data not shown). Apart from (*R*)-, (*S*)- and (*RS*)-mecoprop, strain MH was able to grow on (*R*)- and (*RS*)-dichlorprop [(*R*)- and (*RS*)-(2,4-dichlorophenoxy)propionic acid], 2,4-dichlorophenoxyacetic acid, 4-chloro-2-methylphenoxyacetic acid, and pyruvate.

Evidence for enantioselectivity in oxygen uptake and substrate consumption experiments with washed cell suspensions. The oxygen uptake by washed cells of strain MH grown on (*R*)-mecoprop was much higher for the (*R*) enantiomer than for the (*S*) enantiomer (Table 1). Accordingly, cells grown with (*S*)-mecoprop as the sole carbon and energy source were adapted to utilize the (*S*) enantiomer at much higher rates. In both cases, the rate of oxygen consumption due to the addition of the racemic mixture of mecoprop was in the same range as the one due to the addition of the pure enantiomer that was used as the growth substrate. Cells were also grown with the racemic mixture as the carbon and energy source. In this case, oxygen consumption due to the addition of the racemic mixture was higher than the one due to the addition of the single

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001243

TABLE 1. Oxygen uptake rates with (*R*)- and (*S*)-mecoprop and the racemic mixture by resting cells of *S. herbicidovorans* MH grown on these compounds[a]

Assay substrate	Absolute oxygen uptake rates (μmol of O_2/g [dry wt] per min) by cells grown on[b]:		
	(*R*)-Mecoprop	(*S*)-Mecoprop	(*RS*)-Mecoprop
(*R*)-Mecoprop	59.1 ± 3.3	23.5 ± 3.0	39.7 ± 2.0
(*S*)-Mecoprop	26.9 ± 1.9	89.9 ± 0.8	47.4 ± 3.3
(*RS*)-Mecoprop	65.3 ± 1.4	76.2 ± 12.4	56.9 ± 2.6

[a] All rates for cells grown on complex medium were below detection limit (<5 μmol of O_2 per g [dry weight] per min). Other substrates tested for oxygen uptake are mentioned in the text.
[b] Values are means of triplicates ± standard deviations.

pure enantiomers. A reasonable explanation for these findings is the cumulative effect of two enzyme systems acting each on one of the enantiomers. Cells grown on complex carbon substrates were not adapted to the utilization of (*R*)- or (*S*)-mecoprop (Table 1). The same was true for cells grown on pyruvate. When (*R*)- and (*RS*)-dichlorprop were used as the assay substrates, the oxygen uptake pattern was nearly identical to the one obtained with (*R*)- and (*RS*)-mecoprop. Strain MH was also grown with (*R*)- and (*RS*)-dichlorprop as the sole carbon and energy source. The oxygen uptake pattern with such cells and all the test substrates was nearly identical to the one with (*R*)- and (*RS*)-mecoprop as the growth substrates. Experiments with the pure (*S*)-dichlorprop could not be carried out because this enantiomer was not at our disposal. The specific oxygen uptake rates by resting cell suspensions support the existence of enantioselective degradation.

The results presented in Table 2 clearly show that cells of *S. herbicidovorans* MH grown on (*S*)-mecoprop were not able to transform the (*R*) enantiomer. Washed-cell suspensions of such cells effected disappearance of (*S*)-mecoprop and concomitant formation of chloride upon incubation with the pure (*S*) enantiomer as well as with the racemic mixture. The results of experiments with washed cells of strain MH grown on (*R*)-mecoprop (Table 2) were almost congruent to the results with cells grown on (*S*)-mecoprop—the cells preferentially transformed the enantiomer that was used as the growth substrate. Again, these data suggest the involvement of two independent enzymes for, at least, the initial transformation of each enantiomer. However, a slight uptake of (*S*)-mecoprop with concomitant chloride release by cells grown on the (*R*) enantiomer was evident (Table 2). This finding indicates that the enzymes responsible for the degradation of (*S*)-mecoprop may be constitutively expressed at low levels or that the (*R*)-mecoprop-degrading enzymes are able to slowly turn over (*S*)-mecoprop. A pathway based on racemization of (*R*)- to (*S*)-mecoprop and further degradation of the latter one would have been a third explanation, but in such a case much higher consumption rates for (*S*)-mecoprop by (*R*)-mecoprop-grown cells would have been expected. Yet, the possibility remains that the absence of a transport molecule for (*S*)-mecoprop in such cells prevented degradation. It seems unlikely, though, that the enzymes responsible for breakdown of (*S*)-mecoprop were induced without the concomitant induction of a potential transport molecule. Although not completely excluded, a degradation pathway based on racemization is not favored by the data from the substrate consumption experiments.

TABLE 2. Rates of consumption of (*R*)- and (S)-mecoprop and rates of formation of chloride by resting cells of *S. herbicidovorans* MH grown on (S)- and (*R*)-mecoprop[a]

Compound assayed	Rate of consumption or formation (μmol/g [dry wt]/min) with cells grown on[a]:	
	(*S*)-Mecoprop	(*R*)-Mecoprop
Incubations with (*S*)-mecoprop		
(*S*)-Mecoprop	−8.1 ± 0.3	−2.2 ± 0.9
Chloride	7.6 ± 0.8	1.9 ± 0.1
Incubations with (*R*)-mecoprop		
(*R*)-Mecoprop	0	−14.7 ± 1.6
Chloride	0	13.6 ± 0.3

[a] Values are parameter estimates ± standard deviations from linear regression analysis.

Conclusions. The data presented clearly demonstrate the complete microbial degradation of both enantiomers of the herbicide mecoprop in an enantioselective manner in *S. herbicidovorans* MH. This study proves the importance of employing enantiomer-specific analytical tools and the necessity of treating enantiomers as distinct molecules when studying the degradation of chiral compounds. The apparent neglect of stereochemistry has hampered progress in the elucidation of the mecoprop degradation pathway. However, we believe that progress can be made and further work on the biochemistry of mecoprop degradation is being carried out in our laboratory.

ACKNOWLEDGMENTS

We thank Roland Hany for running the NMR measurements and Hans-Rudolf Buser and Markus D. Müller for stimulating discussions and their help with enantiomer-specific analysis. We are grateful to Marc Suter and Carlo Kanz for critical reading of the manuscript.

REFERENCES

1. **Ariëns, E. J.** 1989. Racemates—an impediment in the use of drugs and agrochemicals, p. 31–80. *In* A. M. Krstulovic (ed.), Chiral separations by HPLC. Ellis Horwood Limited, Chichester, United Kingdom.
2. **Armstrong, D. W., G. L. Reid III, M. L. Hilton, and C. D. Chang.** 1993. Relevance of enantiomeric separations in environmental science. Environ. Pollut. **79:**51–58.
3. **Fukumori, F., and R. P. Hausinger.** 1993. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetate monooxygenase" is an α-ketoglutarate-dependent dioxygenase. J. Bacteriol. **175:**2083–2086.
4. **Gintautas, P. A., S. R. Daniel, and D. L. Macalady.** 1992. Phenoxyalkanoic acid herbicides in municipal landfill leachates. Environ. Sci. Technol. **26:** 517–521.
5. **Heron, G., and T. H. Christensen.** 1992. Degradation of the herbicide mecoprop in an aerobic aquifer determined by laboratory batch studies. Chemosphere **24:**547–557.
6. **Horvath, M., G. Ditzelmüller, M. Loidl, and F. Streichsbier.** 1990. Isolation and characterization of a 2-(2,4-dichlorophenoxy)propionic acid-degrading soil bacterium. Appl. Microbiol. Biotechnol. **33:**213–216.
7. **Kilpi, S.** 1980. Degradation of some phenoxy acid herbicides by mixed cultures of bacteria isolated from soil treated with 2-(2-methyl-4-chloro)phenoxypropionic acid. Microb. Ecol. **6:**261–270.
8. **Lapage, S. P., P. H. A. Sneath, E. F. Lessel, V. B. D. Skerman, H. P. R. Seeliger, and W. A. Clark (ed.).** 1992. International code of nomenclature of bacteria (1990 revision). American Society for Microbiology, Washington, D.C.
9. **Lappin, H. M., M. P. Greaves, and J. H. Slater.** 1985. Degradation of the herbicide mecoprop [2-(2-methyl-4-chlorophenoxy)propionic acid] by a synergistic microbial community. Appl. Environ. Microbiol. **49:**429–433.
10. **Loos, M. A., R. N. Roberts, and M. Alexander.** 1967. Phenols as intermediates in the decomposition of phenoxyacetates by an *Arthrobacter* species. Can. J. Microbiol. **13:**679–690.
11. **Matell, M.** 1953. Stereochemical studies on plant growth regulators. VII. Optically active α-(2-methyl-4-chlorophenoxy)propionic acid and α-(2,4-dichlorophenoxy)-n-butyric acid and their steric relations. Ark. Kemi. **6:**365–373.
12. **Oh, K.-H., and O. H. Tuovinen.** 1994. Microbial treatment of fertilizer solid waste containing phenoxyalkanoic herbicides 2,4-D and MCPP. J. Chem. Tech. Biotechnol. **61:**299–305.
13. **Pieper, D. H., K.-H. Engesser, and H.-J. Knackmuss.** 1989. Regulation of catabolic pathways of phenoxyacetic acids and phenols in *Alcaligenes eutrophus* JMP 134. Arch. Microbiol. **151:**365–371.

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001244

14. **Pieper, D. H., W. Reineke, K.-H. Engesser, and H.-J. Knackmuss.** 1988. Metabolism of 2,4-dichlorophenoxyacetic acid, 4-chloro-2-methylphenoxyacetic acid and 2-methylphenoxyacetic acid by *Alcaligenes eutrophus* JMP 134. Arch. Microbiol. **150:**95–102.
15. **Simoni, S., S. Klinke, C. Zipper, W. Angst, and H.-P. E. Kohler.** 1996. Enantioselective metabolism of chiral 3-phenylbutyric acid, an intermediate of linear alkylbenzene degradation, by *Rhodococcus rhodochrous* PB1. Appl. Environ. Microbiol. **62:**749–755.
16. **Smith, A. E., and A. J. Aubin.** 1991. Effects of long-term 2,4-D and MCPA field applications on the soil breakdown of 2,4-D, MCPA, mecoprop and 2,4,5-T. J. Environ. Qual. **20:**436–438.
17. **Smith, A. E., and B. J. Hayden.** 1981. Relative persistence of MCPA, MCPB and mecoprop in Saskatchewan soils, and the identification of MCPA in MCPB-treated soils. Weed Res. **21:**179–183.
18. **Takeuchi, M., H. Sawada, H. Oyaizu, and A. Yokota.** 1994. Phylogenetic evidence for *Sphingomonas* and *Rhizomonas* as nonphotosynthetic members of the alpha-4 subclass of the *Proteobacteria*. Int. J. Syst. Bacteriol. **44:**308–314.
19. **Tett, V. A., A. J. Willets, and H. M. Lappin-Scott.** 1994. Enantioselective degradation of the herbicide mecoprop [2-(2-methyl-4-chlorophenoxy)propionic acid] by mixed and pure bacterial cultures. FEMS Microbiol. Ecol. **14:**191–200.
20. **Tiedje, J. M., J. M. Duxbury, M. Alexander, and J. E. Dawson.** 1969. 2,4-D metabolism: pathway of degradation of chlorocatechols by *Arthrobacter* sp. J. Agric. Food Chem. **17:**1021–1026.
21. **van Bruggen, A. H. C., K. N. Jochimsen, E. M. Steinberger, P. Segers, and M. Gillis.** 1993. Classification of *Rhizomonas suberifaciens*, unnamed *Rhizomonas* species, and *Sphingomonas* spp. in rRNA superfamily IV. Int. J. Syst. Bacteriol. **43:**1–17.
22. **Worthing, C. R., and R. J. Hance.** 1991. The pesticide manual—a world compendium, ninth ed. The British Crop Protection Council. Unwin Brothers Limited, Old Woking, England.

Downloaded from http://aem.asm.org/ on September 3, 2020 at Corteva Agriscience

COR-AAD1-00001245

JOURNAL OF BACTERIOLOGY, July 1998, p. 3368–3374 Vol. 180, No. 13
0021-9193/98/$04.00+0
Copyright © 1998, American Society for Microbiology. All Rights Reserved.

Enantioselective Uptake and Degradation of the Chiral Herbicide Dichlorprop [(*RS*)-2-(2,4-Dichlorophenoxy)propanoic acid] by *Sphingomonas herbicidovorans* MH

CHRISTIAN ZIPPER,† MONIKA BUNK, ALEXANDER J. B. ZEHNDER, AND HANS-PETER E. KOHLER*

Swiss Federal Institute for Environmental Science and Technology, CH-8600 Dübendorf, and Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland

Received 3 February 1998/Accepted 22 April 1998

Sphingomonas herbicidovorans **MH was able to completely degrade both enantiomers of the chiral herbicide dichlorprop [(*RS*)-2-(2,4-dichlorophenoxy)propanoic acid], with preferential degradation of the (*S*) enantiomer over the (*R*) enantiomer. These results are in agreement with the recently reported enantioselective degradation of mecoprop [(*RS*)-2-(4-chloro-2-methylphenoxy)propanoic acid] by this bacterium (C. Zipper, K. Nickel, W. Angst, and H.-P. E. Kohler, Appl. Environ. Microbiol. 62:4318–4322, 1996). Uptake of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D (2,4-dichlorophenoxyacetic acid) was inducible. Initial uptake rates of cells grown on the respective substrate showed substrate saturation kinetics with apparent affinity constants (K_t) of 108, 93, and 117 μM and maximal velocities (V_{max}) of 19, 10, and 21 nmol min^{-1} mg of $protein^{-1}$ for (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D, respectively. Transport of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D was completely inhibited by various uncouplers and by nigericin but was only marginally inhibited by valinomycin and by the ATPase inhibitor *N,N'*-dicyclohexylcarbodiimine. Experiments on the substrate specificity of the putative transport systems revealed that (*R*)-dichlorprop uptake was inhibited by (*R*)-mecoprop but not by (*S*)-mecoprop, (*S*)-dichlorprop, or 2,4-D. On the other hand, the (*S*)-dichlorprop transport was inhibited by (*S*)-mecoprop but not by (*R*)-mecoprop, (*R*)-dichlorprop, or 2,4-D. These results provide evidence that the first step in the degradation of dichlorprop, mecoprop, and 2,4-D by *S. herbicidovorans* is active transport and that three inducible, proton gradient-driven uptake systems exist: one for (*R*)-dichlorprop and (*R*)-mecoprop, another for (*S*)-dichlorprop and (*S*)-mecoprop, and a third for 2,4-D.**

Research on the biodegradation of chiral xenobiotics can provide a better understanding of the processes that govern enantioselectivity in the microbial degradation and environmental fate of stereoisomers. We have examined the microbial degradation of the widely used chiral herbicide mecoprop [(*RS*)-2-(4-chloro-2-methylphenoxy)propanoic acid] by a pure bacterial strain that was isolated from soil. This strain, *Sphingomonas herbicidovorans* MH, can utilize both enantiomers of mecoprop as sole carbon and energy sources for growth (24). Growth experiments with the pure enantiomers as well as with racemic mecoprop revealed that the (*S*) enantiomer disappeared much faster from the culture medium than the (*R*) enantiomer. Based on these results, we concluded that specific catabolic enzymes must be involved in the degradation of each enantiomer of mecoprop (24). This view was confirmed by the finding that two α-ketoglutarate-dependent dioxygenases are involved in the enantioselective degradation of mecoprop in *S. herbicidovorans* MH, one that is specific for (*R*)-mecoprop and one that is specific for (*S*)-mecoprop (17). The (*S*) enantiomer-specific dioxygenase activity is constitutively expressed, whereas the (*R*) enantiomer-specific enzyme activity is present only when cells grow on the (*R*) enantiomer. Both enzymes are present in cells grown on the racemic mixture. Extracts of cells grown on complex medium and extracts of cells grown on (*S*)-mecoprop contain equally high activities of the (*S*) enantiomer-specific dioxygenase (17). However, intact cells grown on complex medium show very low oxygen uptake rates with (*S*)-mecoprop, whereas (*S*)-mecoprop-grown cells show high oxygen consumption with (*S*)-mecoprop (24). These findings suggested the existence of specific uptake systems and led us to examine the uptake of phenoxyalkanoic acids by *S. herbicidovorans* after growth on various substrates.

Many aromatic compounds are taken up by bacteria through energy-dependent transport systems (4, 5, 10, 11, 15, 16). The uptake of 4-chlorobenzoate by the coryneform bacterium NTB-1 is inducible and coupled to the proton motive force (4). The uptake of 4-toluene sulfonate by *Comamonas testosteroni* T-2 is also inducible and shows substrate saturation kinetics (11). Transport of 4-toluene sulfonate is inhibited completely by uncouplers but only marginally by ATPase inhibitors, and the authors proposed that 4-toluene sulfonate is taken up by an inducible secondary proton symport system (11). The uptake of 4-hydroxybenzoate by *Pseudomonas putida* is driven by the proton motive force, and the gene that encodes a 4-hydroxybenzoate transporter was identified, cloned, and expressed in *Escherichia coli*. Furthermore, it was found that 4-hydroxybenzoate transport is an integral feature of the β-ketoadipate pathway in *P. putida* (5, 15, 16). The uptake of the herbicide 2,4-D (2,4-dichlorophenoxyacetic acid) by *Ralstonia eutropha* JMP134 is inducible, is sensitive to metabolic inhibitors, and shows substrate saturation kinetics (10). Moreover, the authors identified a gene, designated *tfdK*, whose translation product shows resemblance to transport proteins of the major facilitator superfamily, and they suggested that *tfdK* encodes the 2,4-D transporter of *R. eutropha* JMP134.

* Corresponding author. Mailing address: Department of Microbiology, EAWAG, Ueberlandstrasse 133, CH-8600 Duebendorf, Switzerland. Phone: 41 1 823 5521. Fax: 41 1 823 5547. E-mail: kohler@eawag.ch.

† Present address: BMG Engineering, CH-8952 Schlieren, Switzerland.

Downloaded from http://jb.asm.org/ on September 4, 2020 at Corteva Agriscience

COR-AAD1-00001246

Our results indicate that *S. herbicidovorans* MH has two distinct uptake systems for chiral phenoxypropionic acid herbicides, one for the (*R*) enantiomer and another one for the (*S*) enantiomer of dichlorprop and mecoprop. Furthermore, evidence was obtained that 2,4-D was taken up by a third transport system.

MATERIALS AND METHODS

Organism and growth conditions. *S. herbicidovorans* MH (DSM 11019) was cultured and maintained as described before (24), unless stated otherwise. Growth experiments were performed in 1-liter Erlenmeyer flasks equipped with magnetic stirring bars. The flasks were equipped with a silicon tube that allowed easy sterile sampling. The experiments were carried out in a water bath at a temperature of 30°C, with stirring held constant at 370 rpm (24).

Uptake assays. We performed the transport experiments with (*R*)- and (*S*)-[^{14}C]dichlorprop synthesized from ^{14}C-labeled 2,4-dichlorophenol (23). Radiolabeled (*R*)- and (*S*)-mecoprop were not commercially available. Uptake of ^{14}C-labeled compounds was routinely assayed by a filtration method (18). Cells were harvested at the mid-exponential growth phase by centrifugation (15 min, 7,500 × *g*, 4°C), washed twice with ice-cold minimal medium (24), and resuspended in a phosphate buffer (20 mM Na_2HPO_4-KH_2PO_4 [pH 6.5]) at concentrations of 0.15 ± 0.03 mg of protein per ml. Cell suspensions could be kept on ice and aerated for 2 h without significant loss of transport activity. In some experiments, the cell suspension was preincubated for 5 min with potential uptake inhibitors. The cell suspension was acclimated at 30 ± 2°C for 5 min before the addition of the ^{14}C-labeled compound. At intervals of 15 s, uptake was terminated by filtering a 200-μl sample through a 0.45-μm-pore-size cellulose-nitrate filter (Sartorius AG, Göttingen, Germany) that had been prewashed with 2 ml of phosphate buffer (washing buffer) containing 50 μM of unlabeled 2,4-D or (*R*)- or (*S*)-dichlorprop. After filtration, the filters were washed once with 5 ml of the washing buffer to remove extracellular ^{14}C-labeled compounds and then transferred to scintillation vials. After addition of the scintillation cocktail (Filter-Count; Packard Instrument Company, Meriden, Conn.), the radioactivity was determined in a liquid scintillation counter (Kontron Analytical, Zürich, Switzerland). The kinetic uptake parameters K_s (apparent affinity constant) and V_{max} (maximum velocity) were calculated with a computer program (IGOR Pro, WaveMetrics Inc., Lake Oswego, Oreg.). The calculations were based on weighted nonlinear regression analysis of the Michaelis-Menten model. Error estimates for the biological variance were obtained from the data of four independent assays with different cell batches at one specific substrate concentration and by making the assumption that the error was heteroscedastically distributed with a constant relative standard deviation.

Estimation of intracellular volume. The specific intracellular volume of *S. herbicidovorans* MH was estimated from the dimensions of the cells (average length of the rods, 2.8 μm; average width, 0.7 μm [24]), the protein content (0.15 ± 0.03 mg ml^{-1} at an optical density at 546 nm of 1), and the cell number ([1.71 ± 0.13] × 10^9 ml^{-1} at an optical density at 546 nm of 1). From these data, a specific internal cell volume of 1.2 μl/mg of protein was calculated. This value is in the same range as the volumes of other gram-negative rods (11, 13) and was used for all calculations.

Analytical procedures. (i) HPLC analysis. High-pressure liquid chromatography (HPLC) analyses were performed on a Gynkotek HPLC system with a M480G pump, a Gina 50T autosampler, and a UVD340S photodiode array detector (Gynkotek GmbH, Germering, Germany). The system was operated isocratically with an eluent consisting of 70% (vol/vol) methanol and 30% (vol/vol) NaH_2PO_4 (50 mM, pH 3.0) at a flow rate of 0.7 ml/min. Forty microliters of the samples was injected, and the eluting compounds were detected at a wavelength of 230 nm. The analytes were separated on a Nucleodex-α-PM column (200 by 4.0 mm) with permethylated α-cyclodextrin as the chiral stationary phase (Macherey-Nagel, Düren, Germany). Retention times were typically 6.0, 6.7, and 8.6 min for 2,4-D, (*R*)-dichlorprop, and (*S*)-dichlorprop, respectively. The detection limits for the dichlorprop enantiomers and 2,4-D were 0.1 and 0.05 μM, and the relative standard deviations (n = 3) were ±0.8% and ±0.7%, respectively.

(ii) Determination of ATP content. ATP was determined by means of the firefly luciferase assay with an ATP bioluminescence assay kit (HS II; Boehringer, Mannheim, Germany).

(iii) Protein determination. Protein contents were determined by the method of Lowry et al. (12), with bovine serum albumin as the standard.

(iv) Ion chromatography. Chloride concentrations were determined with an ion chromatograph equipped with a conductivity detector (Dionex, Olten, Switzerland) after separation on an anion-exchange column (IonPac AS11; Dionex) (24).

Chemicals. [*ring*-U-^{14}C]2,4-D (≥98%; 807 GBq/mol) was obtained from Sigma Chemical Co. (St. Louis, Mo.). [*ring*-U-^{14}C](*R*)-dichlorprop (≥98.3%; enantiomeric purity, 99.4%; 126 GBq/mol) and [*ring*-U-^{14}C](*S*)-dichlorprop (≥98.4%; enantiomeric purity, >99.4%; 166 GBq/mol) were prepared by Amersham, Buckinghamshire, England, as described elsewhere (23). All other chemicals were obtained from Fluka, Buchs, Switzerland, or Merck, Dietikon, Switzerland.

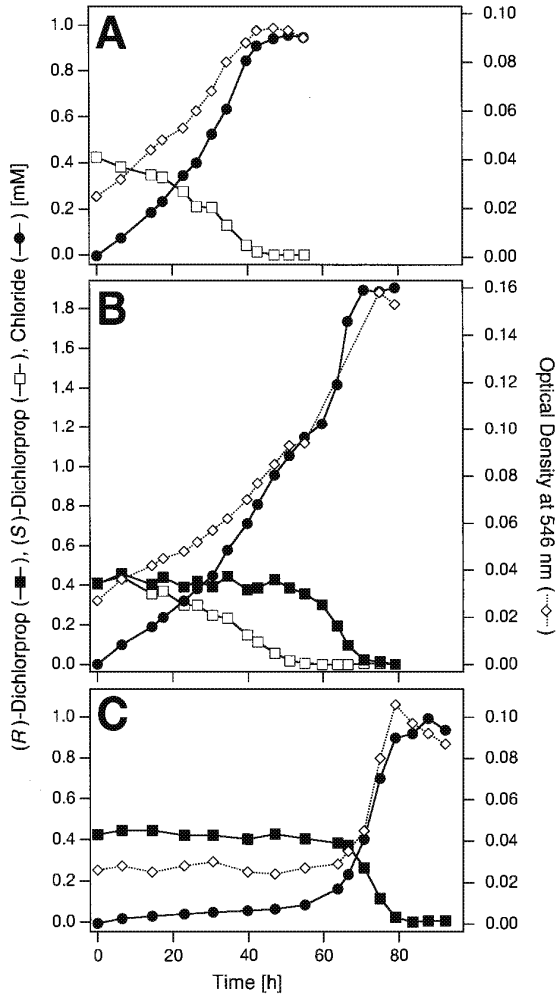


FIG. 1. Growth experiments with *S. herbicidovorans* MH incubated with (*R*)-dichlorprop, (*S*)-dichlorprop, and racemic (*RS*)-dichlorprop. Precultures were grown in chloride-free complex medium for 12 h at 30°C. (A) Growth on the pure (*S*) enantiomer; (B) growth on the racemic mixture; (C) growth on the pure (*R*) enantiomer. An A_{546} of 0.1 corresponds to 20.8 mg (dry weight) per liter.

RESULTS

Growth on dichlorprop and 2,4-D. *S. herbicidovorans* MH utilized both enantiomers of dichlorprop as sole carbon and energy sources for growth (Fig. 1A and C). However, the (*S*)-dichlorprop was completely removed from the culture medium after 43 h of incubation (Fig. 1A), while in experiments with (*R*)-dichlorprop as the growth substrate, 82 h passed until the degradation was complete (Fig. 1C). When strain MH was offered racemic dichlorprop as the growth substrate, it could degrade both enantiomers, with the (*S*)-dichlorprop being preferentially degraded (Fig. 1B). Chloride release was stoichiometric with (*R*)-, (*S*)-, and (*RS*)-dichlorprop as the growth substrate (Fig. 1), which is a strong indication for the mineralization of the aromatic ring. In all growth experiments, strain MH did not excrete any metabolites into the medium that could be detected by HPLC, and the levels of dissolved organic carbon at the beginning and at the end of a growth experiment (data not shown) corresponded to a complete conversion of dichlorprop

Downloaded from http://jb.asm.org/ on September 4, 2020 at Corteva Agriscience

COR-AAD1-00001247

TABLE 1. Growth characteristics of *S. herbicidovorans* MH grown on (*R*)-dichlorprop, (*S*)-dichlorprop, (*RS*)-dichlorprop, and 2,4-D and of the strain grown on (*R*)-mecoprop, (*S*)-mecoprop, and (*RS*)-mecoprop[a]

Growth substrate	Mean growth rate (h^{-1}) ± SD	Specific degradation rate (mmol h^{-1} g [dry wt]$^{-1}$)	Mean yield (g [dry wt] g of substrate^{-1}) ± SD
(*R*)-Dichlorprop	0.082 ± 0.006	2.3	0.37 ± 0.06
(*R*)-Mecoprop	0.032	ND[b]	0.30 ± 0.15
(*S*)-Dichlorprop	0.036 ± 0.001	1.4	0.35 ± 0.02
(*S*)-Mecoprop	0.039	ND	0.35 ± 0.05
Racemic mixture of dichlorprop			0.29 ± 0.03
(*R*)-Dichlorprop	0.043 ± 0.003	2.4	
(*S*)-Dichlorprop	0.029 ± 0.004	1.1	
Racemic mixture of mecoprop	0.016[c]	ND	0.30 ± 0.04
2,4-D	0.133 ± 0.003	4.4	0.19 ± 0.05

[a] Values for growth on mecoprop were taken from reference 24.
[b] ND, not determined.
[c] In this case, the growth rate was affected by the concentration, which for each enantiomer was only half of what it was in the experiments with the pure mecoprop enantiomers. Additional growth experiments showed increased growth rates in experiments with higher (*RS*)-mecoprop concentrations (24).

into biomass and presumably carbon dioxide. 2,4-D was also completely degraded by strain MH, as shown by biomass formation, substrate consumption, and stoichiometric chloride release (data not shown). Typically, growth on 2,4-D started without a lag phase and was complete after 15 h of incubation. The growth rates and the growth yields obtained are listed in Table 1.

Induction of uptake. Cells of *S. herbicidovorans* MH grown on (*R*)-dichlorprop were induced for the uptake of (*R*)-dichlorprop and 2,4-D, but such cells did not take up (*S*)-dichlorprop (Fig. 2A). When *S. herbicidovorans* MH grew on (*S*)-dichlorprop, uptake of (*S*)-dichlorprop and 2,4-D was induced but uptake of (*R*)-dichlorprop was not (Fig. 2B). Interestingly, cells of *S. herbicidovorans* MH grown on 2,4-D were able to take up 2,4-D but not the dichlorprop enantiomers (Fig. 2C). Control experiments with succinate-grown cells revealed that these cells took up (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D, but with very low rates (Fig. 2D). However, these rates were slightly higher than the rates for uptake by noninduced cells in the experiments described above (Fig. 2A to C).

Kinetics of uptake. The kinetic parameters of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D uptake were estimated from the initial uptake rates (15 to 60 s) at concentrations from 0.7 to 450 μM. Uptake of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D showed monophasic saturation kinetics (Fig. 3), with K_t values of 108 ± 21, 93 ± 19, and 117 ± 25 μM and V_{max} values of 19 ± 2, 10 ± 1, and 21 ± 3 nmol min^{-1} mg of $protein^{-1}$, respectively. Succinate-grown cells (Fig. 3C) took up 2,4-D with much lower rates than did 2,4-D-grown cells, and the relationship between 2,4-D concentration and 2,4-D uptake in such cells could be modeled with a linear function.

Effect of structurally related compounds on uptake. To study the substrate specificity of the (*R*)- and (*S*)-dichlorprop uptake, we examined how the addition of unlabeled (*R*)-dichlorprop, (*S*)-dichlorprop, (*R*)-mecoprop, (*S*)-mecoprop, and 2,4-D affected transport. The unlabeled compounds were added at concentrations almost threefold higher than those of the labeled compounds. The presence of (*S*)-dichlorprop, (*S*)-mecoprop, and 2,4-D exerted only a small influence on (*R*)-dichlorprop uptake (Table 2). As expected, (*R*)-dichlorprop uptake was reduced by unlabeled (*R*)-mecoprop and (*R*)-dichlorprop. Similarly, (*R*)-mecoprop, (*R*)-dichlorprop, and 2,4-D had only a small effect on (*S*)-dichlorprop uptake, whereas the (*S*) enantiomers of mecoprop and dichlorprop reduced the uptake of (*S*)-dichlorprop (Table 2).

Substrate accumulation. To measure substrate accumulation, cells were grown on succinate, 2,4-D, (*R*)-dichlorprop, and (*S*)-dichlorprop. Cells from each batch were incubated in 60 μM [^{14}C]2,4-D, [^{14}C](*R*)-dichlorprop, and [^{14}C](*S*)-dichlorprop solutions for 60 s, and the accumulated radioactivity was determined. Apparent intracellular substrate concentrations were calculated with a cell volume of 1.2 μl/mg of protein. Division of these values by the external substrate concentration yielded the apparent accumulation of the substrates within the cells (Table 3).

Effect of metabolic inhibitors on uptake. The uncouplers carbonylcyanide *m*-chlorophenylhydrazone (CCCP), carbonylcyanide-4-trifluoromethoxyphenylhydrazone (CCFP), and 2,4-dinitrophenol (DNP), which can dissipate the proton motive force Δp (i.e., both the electrical potential Δψ and the pH gradient ΔpH) (9), inhibited the uptake of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D almost completely (Table 4). Almost complete inhibition of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D uptake was also observed in the presence of nigericin (Table 4), a K^+ ionophore that catalyzes an electroneutral

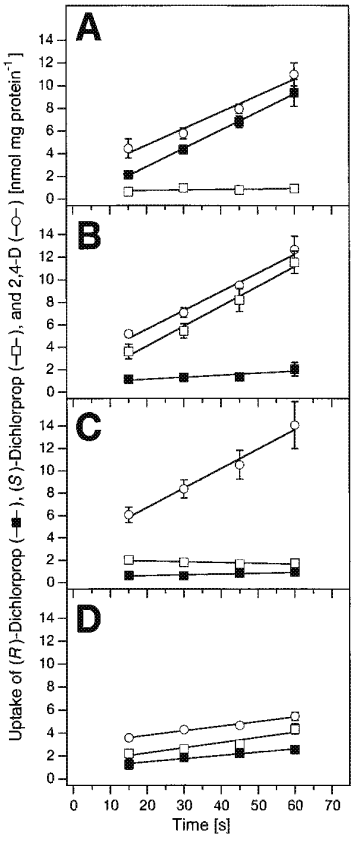


FIG. 2. Induction experiments with *S. herbicidovorans* MH grown on (*R*)-dichlorprop (A), (*S*)-dichlorprop (B), 2,4-D (C), and succinate (D). Cells were washed and resuspended in 20 mM phosphate buffer, pH 6.5 (0.15 mg of protein per ml), and uptake was started by the addition of 60 μM ^{14}C-labeled (*R*)-dichlorprop, (*S*)-dichlorprop, or 2,4-D. Values represent means of three replicates; standard deviations are indicated by the error bars.

Downloaded from http://jb.asm.org/ on September 4, 2020 at Corteva Agriscience

COR-AAD1-00001248

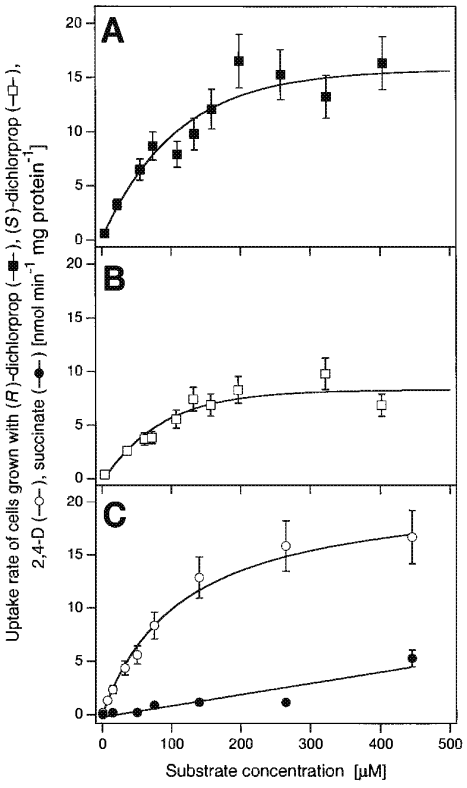


FIG. 3. Effect of substrate concentration on uptake of (*R*)-dichlorprop (A), (*S*)-dichlorprop (B), and 2,4-D (C) by *S. herbicidovorans* MH grown on (*R*)-dichlorprop (A), (*S*)-dichlorprop (B), and succinate and 2,4-D (C). The initial uptake rates (15 s to 2 min) are plotted versus the external concentration of the substrate in the assay (0.7 to 450 μM). Analysis of the data gave K_t values of 108, 93, and 117 μM and V_{max} values of 19, 10, and 21 nmol min^{-1} mg of $protein^{-1}$ for (*R*)- and (*S*)-dichlorprop and 2,4-D, respectively. Error bars represent relative standard deviations (15%) (see Materials and Methods).

K^+-H^+ exchange and thus selectively dissipates the ΔpH in the presence of high K^+ concentrations (9). Valinomycin, another potassium ionophore which can selectively dissipate the Δψ in the presence of high concentrations of K^+ (9), showed little effect on uptake of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D. On the other hand, valinomycin and nigericin in combination reduced the transport of all three substrates to values from 0.4 to 10.9% of the control values (Table 4). In combination, these two ionophores function as an uncoupler (9). Although the ATPase inhibitor *N,N'*-dicyclohexylcarbodiimine (DCCD) reduced uptake of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D to some extent, no correlation between uptake activity and internal ATP concentration was found. For instance, the ATP content of (*S*)-dichlorprop-grown cells treated with nigericin was 83% of the control, whereas uptake of (*S*)-dichlorprop was almost completely inhibited (Table 4). Moreover, transport of (*S*)-dichlorprop was completely inhibited by the uncouplers CCCP, CCFP and DNP, whereas the ATP concentrations were still in the range of 50 to 73% of the control (Table 4).

DISCUSSION

Growth characteristics. Racemic mixtures of the chiral herbicide dichlorprop were degraded in an enantioselective manner by *S. herbicidovorans* MH (Fig. 1B). The growth patterns of strain MH incubated with (*R*)-, (*S*)-, and (*RS*)-mecoprop, a structurally related compound, were essentially similar, with preferential degradation of the (*S*) enantiomer (24). The growth rates with dichlorprop were in the same range, and the growth yields were nearly identical to those obtained with mecoprop (Table 1). Hence, the two enantiomers of these compounds were equally well suited as a carbon and energy source for this bacterium. These results are in contrast to findings with *Alcaligenes denitrificans*, which exclusively degraded (*R*)-mecoprop (19, 20), and with laboratory soil systems with low concentrations of mecoprop, where extensive enantiomerization was observed (2, 14). Strain MH was also able to degrade the (achiral) 2,4-D to completion, with maximum growth rates that were higher than those with the enantiomers of dichlorprop and mecoprop (Table 1).

Induction and specificity of uptake. Our data clearly indicate that the uptake of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D was inducible. The data presented in Fig. 2 clearly demonstrate that the uptake of dichlorprop was enantioselective and that there must be a separate uptake system for each enantiomer. These results substantiate the view that uptake of both enantiomers was carrier mediated, as (chiral) transport proteins can be expected to bind enantiomers in a selective manner (6). D- and L-xylose were transported at different rates by the yeast *Rhodotorula gracilis*, a fact that was used to illustrate the substrate specificity of carrier-mediated transport systems (6, 7). Because cells grown on 2,4-D could not take up (*R*)-dichlorprop and (*S*)-dichlorprop and because uptake of (*R*)-dichlorprop and (*S*)-dichlorprop by (*R*)- and (*S*)-dichlorprop-grown cells, respectively, was only slightly reduced by 2,4-D

TABLE 2. Effect of structurally related compounds on uptake of (*R*)- and (*S*)-dichlorprop by *S. herbicidovorans* MH[a]

Addition	Uptake of (*R*)-dichlorprop[b] ± SD (% of control)	Uptake of (*S*)-dichlorprop[c] ± SD (% of control)
None (control)	100 ± 3.4	100 ± 18.9
2,4-D	86 ± 0.9	87 ± 10.8
(*R*)-Dichlorprop	57	95 ± 14.0
(*R*)-Mecoprop	50 ± 2.3	84 ± 4.5
(*S*)-Dichlorprop	89 ± 0.8	54 ± 14.9
(*S*)-Mecoprop	96 ± 1.8	25 ± 10.4

[a] Cells grown on (*R*)-dichlorprop or (*S*)-dichlorprop were washed and resuspended in 20 mM phosphate buffer (pH 6.5). Initial uptake rates [with 30 μM ^{14}C-labeled (*R*)- and (*S*)-dichlorprop] were determined in the presence or absence of an unlabeled, structurally related compound (80 μM). Values represent the means of duplicate determinations.

[b] Uptake by (*R*)-dichlorprop-grown cells; 100% corresponds to an uptake rate of 2.5 ± 0.2 nmol min^{-1} mg of $protein^{-1}$.

[c] Uptake by (*S*)-dichlorprop-grown cells; 100% corresponds to an uptake rate of 1.5 ± 0.3 nmol min^{-1} mg of $protein^{-1}$.

TABLE 3. Accumulation of radioactivity by cells of *S. herbicidovorans* MH grown on succinate, 2,4-D, (*R*)-dichlorprop, and (*S*)-dichlorprop

Assay substrate	Mean accumulation of radioactivity (c'_i/c_e)[a] ± SD of cells grown on:			
	Succinate	2,4-D	(*R*)-Dichlorprop	(*S*)-Dichlorprop
2,4-D	20 ± 1	134 ± 27	108 ± 14	103 ± 15
(*R*)-Dichlorprop	15 ± 8	3 ± 3	114 ± 17	5 ± 1
(*S*)-Dichlorprop	25 ± 6	0	0	116 ± 13

[a] c'_i is the apparent intracellular substrate concentration, which was obtained in the presence of catabolic activity and consequently represents the sum of the intracellular concentrations of the substrate and all metabolites. c_e is the external substrate concentration in the transport assay (60 μM).

Downloaded from http://jb.asm.org/ on September 4, 2020 at Corteva Agriscience

COR-AAD1-00001249

TABLE 4. Effects of inhibitors on the uptake of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D and on the intracellular ATP concentration in *S. herbicidovorans* MH[a]

Addition	% of control					
	Cells grown on (*R*)-dichlorprop		Cells grown on (*S*)-dichlorprop		Cells grown on 2,4-D	
	Uptake of (*R*)-dichlorprop	Internal ATP concn	Uptake of (*S*)-dichlorprop	Internal ATP concn	Uptake of 2,4-D	Internal ATP concn
CCCP (50 μM)	7.4	47	0.0	50	3.8	61
CCFP (50 μM)	0.8	56	0.0	64	1.4	60
DNP (500 μM)	11.9	44	4.3	73	7.1	61
Valinomycin (100 μM)	39.3	49	77.6	88	92.5	97
Nigericin (100 μM)	0.0	61	1.5	83	6.8	47
Valinomycin plus nigericin (100 μM each)	7.0	73	10.9	78	0.4	65
DCCD (50 μM)	21.0	34	9.3	19	37.5	29

[a] Cells grown on (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D were washed and resuspended in 20 mM phosphate buffer (pH 6.5). After incubation for 5 min at 30°C in the presence or absence of an inhibitor, initial uptake rates {with 60 μM [^{14}C](*R*)-dichlorprop, 115 μM [^{14}C](*S*)-dichlorprop and 30 μM [^{14}C]2,4-D} and intracellular ATP concentrations were determined. In the absence of an inhibitor (control), the initial uptake rates were 7.1 ± 0.5, 5.5 ± 0.1, and 4.2 ± 0.5 nmol min^{-1} mg of protein^{-1} for (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D, respectively, and the internal ATP concentration was 1.4 ± 0.3 mM for all substrates; these values were taken as 100%.

(Table 2), we suggest that there is a specific transport system for 2,4-D. The first common metabolite in the degradation of dichlorprop and 2,4-D by strain MH is 2,4-dichlorophenol (17), which is further metabolized via the modified *ortho* pathway (8). Regulation of transport systems by downstream metabolites of the *ortho* pathway was reported by Nichols and Harwood (16). β-Ketoadipate, a common metabolite of benzoate and 4-hydroxybenzoate degradation by *P. putida*, induces the 4-hydroxybenzoate transporter, whereas benzoate represses it. These findings made clear why *P. putida* cells preferentially degrade benzoate over 4-hydroxybenzoate in a mixture of both compounds (16). Filer and Harker demonstrated that dichloromuconate is the inducing agent of the 2,4-D pathway in *R. eutropha* JMP134 (3). The model for induction of the uptake of phenoxyalkanoic acids is presented in Fig. 4. The suggestion that 2,4-D uptake of strain MH was induced by a common metabolite of 2,4-D and dichlorprop degradation agrees well with our data but needs more elaborate experiments for definitive proof. On the other hand, the inducing agents of the (*R*)-dichlorprop and (*S*)-dichlorprop carriers must be (*R*)-dichlorprop and (*S*)-dichlorprop, respectively. This is a necessary consequence of the induction pattern of the (*R*)-dichlorprop and (*S*)-dichlorprop carriers (Fig. 2A and B), because, except for the first step, (*R*)- and (*S*)-dichlorprop are degraded identically (17). From the data presented in Table 2, we conclude that (*R*)-mecoprop was transported by the (*R*)-dichlorprop carrier and that (*S*)-mecoprop was transported by the (*S*)-dichlorprop carrier. This shows that the transport systems were specific

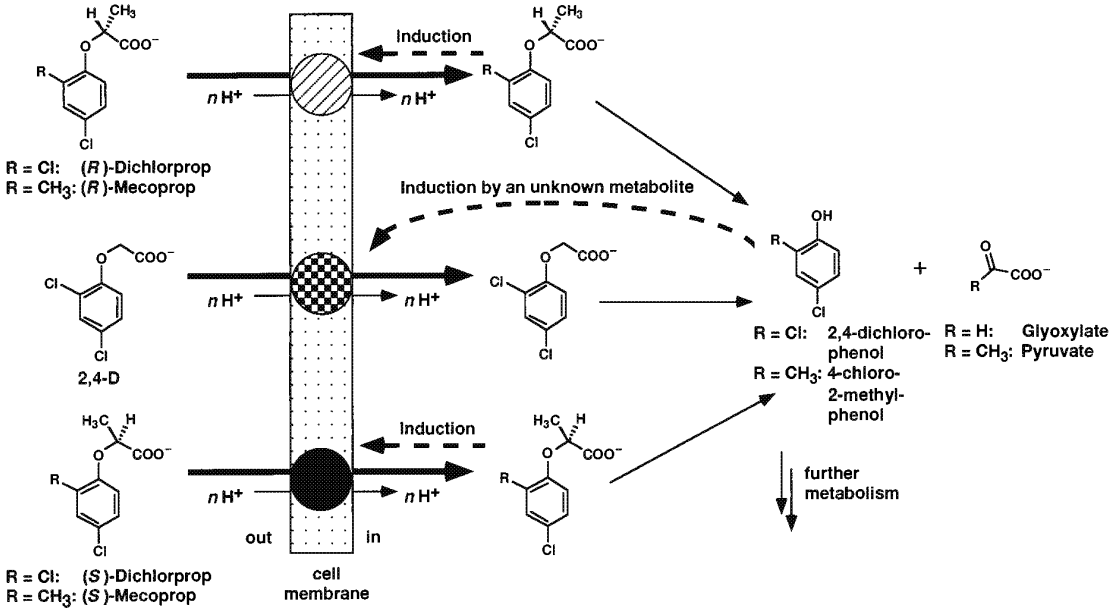


FIG. 4. Model proposed for the uptake of (*R*)- and (*S*)-dichlorprop, (*R*)- and (*S*)-mecoprop, and 2,4-D by *S. herbicidovorans* MH. Evidence is provided that strain MH has three inducible uptake systems: one for (*R*)-dichlorprop and (*R*)-mecoprop, one for (*S*)-dichlorprop and (*S*)-mecoprop, and a third for 2,4-D. Studies with inhibitors of energy metabolism indicate that the proton gradient is the driving force and that the uptake of these phenoxyalkanoic acids proceeds in symport together with protons. The number of protons (*n*) transported together with the phenoxyalkanoic acids was equal or larger than 1.

Downloaded from http://jb.asm.org/ on September 4, 2020 at Corteva Agriscience

COR-AAD1-00001250

with respect to the substitution and the configuration at the stereogenic center (C-2 atom of the alkanoic side chain) and rather unspecific with respect to the aromatic moiety of the phenoxyalkanoic acids.

Kinetics of uptake. The rates of uptake of 2,4-D-grown cells in response to increasing 2,4-D concentrations were in accordance with saturation kinetics (Fig. 3C). Saturation kinetics was also observed for the transport of (*R*)-dichlorprop and (*S*)-dichlorprop into (*R*)-dichlorprop- and (*S*)-dichlorprop-grown cells, respectively, in response to increasing substrate concentrations (Fig. 3A and B). Saturation kinetics is often interpreted as an indicator for carrier-mediated transport for which the uptake rate increases with increasing substrate concentration up to a maximal velocity, at which the carrier is saturated (6). The apparent affinity constants and the maximal velocities for transport of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D, are similar to those reported for uptake of 4-toluene sulfonate by *C. testosteroni* T-2 (11) and 1 to 2 orders of magnitude higher than those reported for the transport of benzoic acid into *Alcaligenes denitrificans* (13) and 4-hydroxyphenylacetic acid into *Klebsiella pneumoniae* (1). Saturation kinetics was also reported for adsorption of 2,4-D to the cell wall of *Pseudomonas fluorescens*, a strain that does not metabolize 2,4-D (21). But as adsorption is neither inducible nor sensitive to metabolic inhibitors, we excluded it as an important contributor to the observed saturation kinetics. Biological membranes are good diffusion barriers for charged molecules, but they are not completely impermeable. It is generally assumed that the diffusion of deprotonated (negatively charged) organic acids is low compared with diffusion of the protonated (neutral) acid (6). The uptake assays with strain MH were performed at pH 6.5, at which more than 99.9% of the 2,4-D molecules are deprotonated (pK_a of 2,4-D is 2.6 [22]). Therefore, we assume that diffusion of the protonated species of 2,4-D was negligible. To assess the impact of unspecific transport of phenoxyalkanoic acids, we examined the 2,4-D uptake by succinate-grown cells. A linear relationship between 2,4-D concentration and uptake rate was found for these cells (Fig. 3C), indicating that 2,4-D uptake of succinate-grown cells might be due to diffusion of 2,4-D into the cells. But it cannot be excluded that succinate-grown cells are partially induced for uptake of 2,4-D, (*R*)-dichlorprop, and (*S*)-dichlorprop. In induction experiments (Fig. 2), it was evident that (*R*)- and (*S*)-dichlorprop uptake was higher with succinate-grown cells (Fig. 2D) than with 2,4-D-grown cells (Fig. 2C). Therefore, we report apparent uptake rates and did not correct the values for these influences. Furthermore, these influences were smaller than the standard deviation of the measurements in the concentration range that was used for most experiments (0.7 to 300 μM). We assume that diffusion of dichlorprop was also of minor importance, as dichlorprop and 2,4-D have similar pK_a values (pK_a of dichlorprop is 3.0 [22]).

Carrier-mediated transport. The apparent accumulation factors were obtained in the presence of catabolic activity (Table 3). Therefore, they cannot be used as an argument for active transport. However, as laid out below, they are proof for the existence of carrier-mediated transport for 2,4-D, (*R*)-dichlorprop, and (*S*)-dichlorprop. Cells grown on complex medium and on succinate constitutively express dioxygenase activities for the degradation of 2,4-D and (*S*)-dichlorprop (17). But succinate-grown cells accumulated significantly smaller amounts of 2,4-D or (*S*)-dichlorprop than (*S*)-dichlorprop-grown cells. Furthermore, 2,4-D-grown cells, which contain dioxygenase activity toward (*S*)-dichlorprop, did not accumulate radioactivity from this latter chemical. Growth on 2,4-D and (*S*)-dichlorprop triggers much higher accumulation of 2,4-D and (*S*)-dichlorprop, respectively, than growth on other substrates (Table 3). These results clearly demonstrate that metabolic substrate depletion together with passive diffusion does not account for the accumulation of radioactivity in strain MH.

Facilitated diffusion or active transport? Two types of carrier-mediated transport systems can be distinguished: facilitated diffusion and active transport (6). The accumulation of compounds against the concentration gradient and the dependence on metabolic energy (i.e., ATP or proton motive force) are characteristic for active transport systems. To study if transport of 2,4-D, (*R*)-dichlorprop, and (*S*)-dichlorprop depended on metabolic energy, we tested the effect of different inhibitors of energy metabolism on the uptake rate (Table 4). Evidence in favor of an ATP-driven transport mechanism could not be obtained. In fact, the inhibitor studies and the lack of correlation between transport activity and ATP pools argue against such a mechanism. On the other hand, uptake of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D was strongly inhibited by all uncouplers that were tested and by a combination of valinomycin and nigericin (Table 4). These results suggest that the proton motive force Δp drives the uptake of 2,4-D, (*R*)-dichlorprop, and (*S*)-dichlorprop. Inhibition experiments with valinomycin and nigericin gave evidence that the ΔpH rather than the $\Delta\psi$ was necessary for transport, as nigericin completely inhibited uptake of 2,4-D, (*R*)-dichlorprop, and (*S*)-dichlorprop. In contrast, valinomycin only had a small effect. These results support the conclusion that the uptake of all three compounds was driven by the proton gradient rather than by ATP or by the electrical potential. Therefore, we propose that the uptake of phenoxyalkanoic acids by *S. herbicidovorans* MH proceeded in symport with one or more protons (Fig. 4). More detailed investigations are needed for a full understanding of these uptake systems.

By combining the results of previous studies (17, 24) with those in this paper, we demonstrate that the selective degradation of the enantiomers of mecoprop (24) and dichlorprop by the soil isolate *S. herbicidovorans* MH was due not only to enantioselective metabolism (17) but also to enantioselective uptake. To the best of our knowledge, this is the first report of enantioselective uptake of environmentally relevant chiral xenobiotics by bacteria.

ACKNOWLEDGMENTS

We are grateful to Kathrin Nickel, Hauke Harms, and Sol Resnick for critical reading of the manuscript.

REFERENCES

1. **Allende, J. L., A. Gibello, M. Martin, and A. Garrido-Pertierra.** 1992. Transport of 4-hydroxyphenylacetic acid in *Klebsiella pneumoniae*. Arch. Biochem. Biophys. **292:**583–588.
2. **Buser, H.-R., and M. D. Müller.** 1997. Conversion reactions of various phenoxyalkanoic acid herbicides in soil. 2. Elucidation of the enantiomerization process of chiral phenoxy acids from incubation in a D_2O/soil system. Environ. Sci. Technol. **31:**1960–1967.
3. **Filer, K., and A. R. Harker.** 1997. Identification of the inducing agent of the 2,4-dichlorophenoxyacetic acid pathway encoded by plasmid pJP4. Appl. Environ. Microbiol. **63:**317–320.
4. **Groenewegen, P. E. J., A. J. M. Driessen, W. N. Konings, and J. A. M. de Bont.** 1990. Energy-dependent uptake of 4-chlorobenzoate in the coryneform bacterium NTB-1. J. Bacteriol. **172:**419–423.
5. **Harwood, C. S., N. N. Nichols, M.-K. Kim, J. L. Ditty, and R. E. Parales.** 1994. Identification of the *pcaRKF* gene cluster from *Pseudomonas putida*: involvement in chemotaxis, biodegradation, and transport of 4-hydroxybenzoate. J. Bacteriol. **176:**6479–6488.
6. **Höfer, M., and J. G. Boggett.** 1981. Transport across biological membranes. Pitman Publishing Ltd., Boston, Mass.
7. **Höfer, M., and A. Kotyk.** 1968. Tight coupling of monosaccharide transport and metabolism in *Rhodotorula gracilis*. Folia Microbiol. **13:**197–204.
8. **Horvath, M., G. Ditzelmüller, M. Loidl, and F. Streichsbier.** 1990. Isolation and characterization of a 2-(2,4-dichlorophenoxy)propionic acid-degrading soil bacterium. Appl. Microbiol. Biotechnol. **33:**213–216.

Downloaded from http://jb.asm.org/ on September 4, 2020 at Corteva Agriscience

COR-AAD1-00001251

9. **Konings, W. N., K. J. Hellingwerf, and G. T. Robillard.** 1981. Transport across bacterial membranes, p. 257–283. *In* S. L. Bonting and J. J. H. H. M. de Pont (ed.), Membrane transport, vol. 2. Elsevier/North-Holland Biomedical Press, Amsterdam, The Netherlands.
10. **Leveau, J. H. J., A. J. B. Zehnder, and J. R. van der Meer.** 1998. The *tfdK* gene product facilitates the uptake of 2,4-dichlorophenoxyacetate by *Ralstonia eutropha* JMP134(pJP4). J. Bacteriol. **180:**2237–2243.
11. **Locher, H. H., B. Poolman, A. M. Cook, and W. N. Konings.** 1993. Uptake of 4-toluene sulfonate by *Comamonas testosteroni* T-2. J. Bacteriol. **175:**1075–1080.
12. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193:**265–275.
13. **Miguez, C. B., C. W. Greer, J. M. Ingram, and R. A. MacLeod.** 1995. Uptake of benzoic acid and chloro-substituted benzoic acids by *Alcaligenes denitrificans* BRI 3010 and BRI 6011. Appl. Environ. Microbiol. **61:**4152–4159.
14. **Müller, M. D., and H.-R. Buser.** 1997. Conversion reactions of various phenoxyalkanoic acid herbicides in soil. 1. Enantiomerization and enantioselective degradation of the chiral 2-phenoxypropionic acid herbicides. Environ. Sci. Technol. **31:**1953–1959.
15. **Nichols, N. N., and C. S. Harwood.** 1997. PcaK, a high-affinity permease for the aromatic compounds 4-hydroxybenzoate and protocatechuate from *Pseudomonas putida*. J. Bacteriol. **179:**5056–5061.
16. **Nichols, N. N., and C. S. Harwood.** 1995. Repression of 4-hydroxybenzoate transport and degradation by benzoate: a new layer of regulatory control in the *Pseudomonas putida* β-ketoadipate pathway. J. Bacteriol. **177:**7033–7040.
17. **Nickel, K., M. J.-F. Suter, and H.-P. E. Kohler.** 1997. Involvement of two α-ketoglutarate-dependent dioxygenases in enantioselective degradation of (*R*)- and (*S*)-mecoprop by *Sphingomonas herbicidovorans* MH. J. Bacteriol. **179:**6674–6679.
18. **Poolman, B., E. J. Smid, and W. N. Konings.** 1987. Kinetic properties of a phosphate-bond-driven glutamate-glutamine transport system in *Streptococcus lactis* and *Streptococcus cremoris*. J. Bacteriol. **169:**2755–2761.
19. **Tett, V. A., A. J. Willetts, and H. M. Lappin-Scott.** 1994. Enantioselective degradation of the herbicide mecoprop [2-(2-methyl-4-chlorophenoxy)propionic acid] by mixed and pure bacterial cultures. FEMS Microbiol. Ecol. **14:**191–200.
20. **Tett, V. A., A. J. Willetts, and H. M. Lappin-Scott.** 1997. Biodegradation of the chlorophenoxy herbicide (*R*)-(+)-mecoprop by *Alcaligenes denitrificans*. Biodegradation **8:**43–52.
21. **Wedemeyer, G.** 1966. Uptake of 2,4-dichlorophenoxyacetic acid by *Pseudomonas fluorescens*. Appl. Microbiol. **14:**486–491.
22. **Worthing, C. R., and R. J. Hance.** 1991. The pesticide manual—a world compendium. 9th ed. The British Crop Protection Council, Farnham, Great Britain.
23. **Zipper, C.** 1998. Ph.D. dissertation. Swiss Federal Institute of Technology, Zürich, Switzerland.
24. **Zipper, C., K. Nickel, W. Angst, and H.-P. E. Kohler.** 1996. Complete microbial degradation of both enantiomers of the chiral herbicide mecoprop [(*RS*)-2-(4-chloro-2-methylphenoxy)propionic acid] in an enantioselective manner by *Sphingomonas herbicidovorans* sp. nov. Appl. Environ. Microbiol. **62:**4318–4322.

Downloaded from http://jb.asm.org/ on September 4, 2020 at Corteva Agriscience

COR-AAD1-00001252

Isolation and Characterization of 2-Methyl-4-Chlorophenoxyacetic Acid-Degrading Bacteria from Agricultural Soils

Seung-Hee Cho and Jong-Ok Ka*

Division of Applied Biology and Chemistry, College of Agricultural and Life Science, Seoul National University, Suwon 441-744, Korea

Received March 3, 1999

Seven numerically dominant 2-methyl-4-chlorophenoxyacetic acid (MCPA)-degrading bacteria were isolated from agricultural soils. The isolates utilized the herbicide MCPA as a sole carbon source, producing significant biomass in MCPA mineral medium. They exhibited diverse herbicide degradation capabilities, but most of them grew very slowly in mineral medium containing herbicide. The chromosomal DNA patterns of the isolates obtained by polymerase chain reaction amplification of repetitive extragenic palindromic sequences were distinct from each other. One isolate, SH3, which was identified as *Sphingomonas* species by fatty acid methyl ester analysis, was able to degrade 5 different phenoxyacetic acid herbicides within 4 days. This strain contains two plasmids, and the smaller one has a crucial role in herbicide degradation. MCPA treated into agricultural soils without indigenous MCPA-degraders persisted for a long time, but the application of the isolate SH3 resulted in rapid decline of MCPA concentration in the soil.

Key words : *biodegradation, 2-methyl-4-dichlorophenoxyacetic acid, bacteria.*

Large amounts of man-made chlorinated organic chemicals have been used in agriculture as herbicides and pesticides. Among them, 2-methyl-4-chlorophenoxyacetic acid (MCPA) is a selective hormone-type herbicide widely used in post-emergence control of broad-leaved weeds in cereals, grassland, and on non-crop land. Unlike many of the synthetic compounds released into the environment, MCPA is known to be easily degraded in agricultural soils within 1 to 4 months.[1]

There have been sporadic reports on microbial degradation of MCPA in soil and isolation of MCPA-utilizing microorganisms.[2-6] The persistence of MCPA in soil is longer than that of 2,4-dichlorophenoxyacetic acid and 4-chloro phenoxyacetic acid,[7] and the degradation rate of MCPA in soil is enhanced by amendment with fertilizers and organic materials.[8] Repeated amendments with phenoxy herbicides, such as MCPA or 2,4-D, accelerate the degradation rate of the herbicide in soil and proliferate the herbicide-degrading microorganisms.[9] Microorganisms that have been reported to be capable of degrading MCPA belong to several genera, including *Arthrobacter*, *Pseudomonas*, and *Alcaligenes*.[4,10] It is also reported that MCPA is metabolized mainly via formation of 3-chloro-5-methylcatechol using ortho- or meta-cleavage pathway.[6]

However, the previous studies on MCPA-degrading isolates have mainly focused on degradative enzyme systems and pathways, thus giving little information on the phenotypic and genetic diversity of MCPA-degrading microorganisms. Furthermore, since many of the previously described MCPA-degrading microorganisms were initially isolated for their ability to metabolize other phenoxy herbicides, e.g. 2,4-D or mecoprop, few studies have been available on physiological and ecological properties of MCPA-degrading microorganisms isolated from MCPA selection.

In this study, we isolated seven numerically dominant MCPA-degrading bacteria from ninety three agricultural soils amended with MCPA, and investigated their diversity, physiological and genetic properties. In addition, soil microcosms were used to study MCPA degradation patterns in natural field soils with and without the inoculated MCPA-degrading isolates.

Materials and Methods

Media and culture conditions. All isolates were maintained on SMB mineral medium[11] containing MCPA at a concentration of 300 μg/ml. PTYG medium containing peptone 0.25 g, tryptone 0.25 g, yeast extract 0.5 g, glucose 0.5 g, magnesium sulfate 0.03 g, and calcium chloride 0.03 g in 1 liter of sterile distilled water was used for strain purification and colony production for the REP-PCR.

*Corresponding author
Phone: 82-331-290-2475; Fax: 82-331-294-8513
E-mail: kagol@plaza.snu.ac.kr

Abbreviations: 4-CPA, 4-chlorophenoxyacetic acid; 2,4-D, 2,4-dichlorophenoxyacetic acid; 2,4-DB, 2,4-dichlorophenoxybutyric acid; 2,4-DP, 2-(2,4-dichlorophenoxy)propionic acid; FAME, fatty acid methyl ester; HPLC, high performance liquid chromatography; MCPA, 2-methyl-4-chlorophenoxyacetic acid; MCPB, 4-(2-methyl-4-chlorophenoxy)butyric acid; MCPP, 2-(2-methyl-4-chlorophenoxy)propionic acid; MPN, most probable number; PTYG, peptone-tryptone-yeast extract-glucose; REP-PCR, repetitive extragenic palindromic sequence-polymerase chain reaction; SMB, standard mineral base.

COR-AAD1-00001253

Chemicals. Analytical grade MCPA and 2,4-D were obtained from Sigma Chemical Co., and 4-CPA, 2,4-DP, MCPP, 2,4-DB, and MCPB were obtained from Aldrich Chemical Co.

Isolation of bacterial strains. Agricultural field soil samples were taken from diverse countrywide sites in Korea. Samples from the top 15 cm of field soil were taken and kept at 4°C prior to use. Each soil was treated with MCPA at a concentration of 300 μg/g and incubated at room temperature for one month. A 10 g of soil sample was homogenized in 95 ml of a sterilized saline solution (0.85%) by shaking on a rotary shaker (200 rpm). Samples (0.1 ml) of appro-priate 10-fold dilutions were inoculated into MPN tubes containing 3 ml of MCPA mineral medium (SMB mineral medium containing 300 μg/ml of MCPA). The tubes were incubated at 30°C for three weeks and degradation of MCPA was analyzed by spectrophotometry. The culture of the ter-minal positive tube showing MCPA degradation was enriched by two additional transfers into fresh medium. Each enriched culture was streaked onto PTYG agar medium, and single colonies were then tested for MCPA degradation in fresh MCPA mineral medium before strain purification.

FAME analysis. The isolates were cultured on tryptic soy agar at 28°C for 48 to 72 h, then the cells were harvested from the plates by scraping with a sterile glass loop and used for FAME analysis. Saponification, methylation, and extraction were performed according to the procedure previously described in the MIDI manual (Microbial Identification, Inc.).[18]

Colony REP-PCR. The colony REP-PCR was performed using BOXA1R as a primer as described previously.[11,19] Each isolate was grown on the PTYG agar for 24 to 48 h, then a small amount of the cells were resuspended in 25 μl of PCR mixture. After PCR amplification, 10 μl samples of the REP-PCR products were separated by electrophoresis on horizontal 1% agarose gels.

Axenic culture experiment. After growth in MCPA mineral medium or PTYG medium, cells were harvested, washed twice with an equal volume of 15 mM phosphate buffer (pH 7.0), and resuspended in the same buffer. Aliquots of the suspended cells were inoculated into duplicate flasks containing 200 ml of MCPA (300 μg/ml) mineral medium at a final density of OD_{600}=0.005 for the axenic growth experiment or OD_{600}=0.5 for the resting cell experiment. All cultures were incubated at 30°C and were aerated by shaking at 200 rpm on a rotary shaker. Aliquots of the cultures were regularly removed to determine cell growth and degradation of MCPA.

Degradative phenotype analysis. Each strain was cultured in PTYG medium. Cells were then harvested, washed, and prepared in sodium phosphate buffer as described above. Aliquots of the suspended cells were inoculated into culture tubes, each of which contained SMB mineral medium supplemented with one of the structural analogs at a concentration of 300 μg/ml. After three week incubation, the cultures were centrifuged to remove the cellular material, and the UV absorption was measured to determine the degradation of chloroorganic compounds.

Table 1. MCPA-degrading bacteria and their isolation soil sites.

Isolates	Soil sites[a]
SH1	Chunan, Chungchongnam-Do
SH2	Yangsan, Kyungsangnam-Do
SH3	Chunan, Chungchongnam-Do
SH4	Namcheju, Cheju-Do
SH5	Yagok, Chungchongbuk-Do
SH6	Yangpyoung, Kyungki-Do
SH7	Chungwon, Chungchongbuk-Do

[a]All soils were from agriculture fields.

Plasmid detection and conjugation. For detection of plasmid DNA, cells were lysed using a modified form[16] of the procedure of Kado and Liu.[17] To analyze the transferability of the MCPA degradative phenotype of the isolates, matings were performed on membrane filters as described by Willets.[16] Transconjugants were selected on 2,4-D mineral medium containing 300 μg/ml MCPA, appropriate antibiotics, and 1.5% Noble agar.

Degradation of MCPA in soil. Two soil samples with or without the indigenous MCPA-degrading microorganisms were obtained from agricultural fields. The soil was stored at field moisture levels at 4°C until used. The isolate SH3 showing the highest degradation rate of MCPA among the isolates (Table 1) was chosen as the inoculum strain. The strain SH3 was grown at 30°C in PTYG medium, harvested, washed, and prepared in sodium phosphate buffer. Each of the soil samples was inoculated with the strain SH3 at a density of 1.0×10^8 cells/g soil. The soil was thoroughly mixed, and 300 g was transferred to each of duplicate beakers; two other control replicates were not inoculated with strain SH3. Inoculated and uninoculated soils were treated with MCPA dissolved in 0.1 M NaH_2PO_4 buffer (pH 7.0) to a concentration of 300 μg/g and thoroughly mixed. The disappearance of MCPA from soil was monitored by HPLC,[19] and the soils were respiked with MCPA (300 μg/g) after it was removed until a total of 3 cycles of degradation had been completed.

Results and Discussion

Distribution of MCPA-degrading microorganisms. The numbers of MCPA-degrading microorganisms in diverse agricultural soils were estimated with MPN procedure. The MCPA-degrading populations generally ranged from 0 to 3.4×10^7 cells/g soil in ninety three different agricultural soils, and 86% of the soil samples apparently did not show any positive degradation of MCPA during the incubation time. The population values for MCPA-degraders in the tested soils are relatively lower than the previous result,

COR-AAD1-00001254

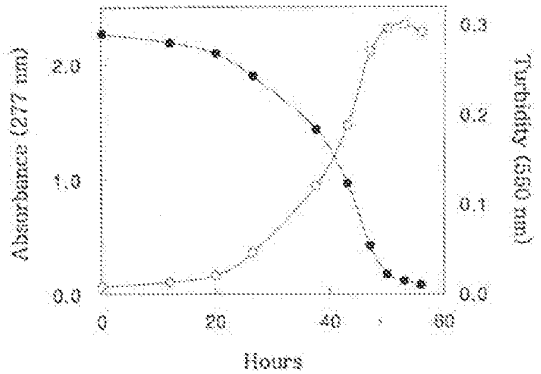


Fig. 1. Disappearance of MCPA (●—●) and growth of bacterial strain SH3 (○—○) during biodegradation of MCPA. Each point is the mean for two replicate liquid cultures.

0.34 to 1377 cells/g,[20] and the observation that only 14% of the tested soils showed positive degradation of MCPA suggests that the MCPA-degrading microorganisms are not widely distributed in our agricultural soils.

Degradation of MCPA. MCPA in mineral medium was completely degraded by the isolates, and no intermediate products were detected by HPLC analysis. Fig. 1 shows typical growth and degradation curves of strain SH3 on MCPA medium. The isolates were able to utilize MCPA as a sole carbon and energy source, as indicated by significant biomass increases in MCPA mineral medium.

FAME and REP-PCR analysis. FAME method was used to analyze the phylogenetic properties of the isolates, but only the SH3 strain of the isolates was reasonably identified as a *Sphingomonas* species. The other strains could not be identified due to their poor growth on laboratory media.

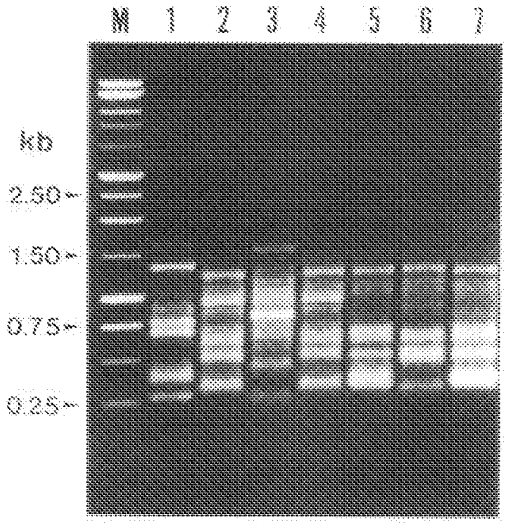


Fig. 2. REP-PCR patterns of strains. Lanes: 1, SH1; 2, SH2; 3, SH3; 4, SH4; 5, SH5; 6, SH6; 7, SH7; and M, 1 kb ladder DNA size marker.

A colony REP-PCR experiment was performed to study the genomic relatedness of the isolates (Fig. 2). The analysis of REP-PCR band patterns with densitometer revealed that the isolates produced 6 different DNA fingerprint patterns. Strains SH2 and SH4 isolated from different locations exhibited the same band pattern, suggesting that they were very closely related in their chromosome structure. Strains SH5 and SH7 also showed a similar band pattern.

Growth properties of the isolates. To understand axenic growth patterns of the MCPA degraders, the isolates were inoculated into MCPA mineral medium under induced and uninduced conditions. One of the most unique features of the isolates was that most of them are slow-growing bacteria compared to previously reported MCPA degraders.[3] Fig. 3 shows typical growth patterns of strains SH1 and SH3 on MCPA medium. Under the uninduced condition, strain SH3 exhibited short lag period (ca. 1 day) and thereafter began to grow exponentially. Strain SH1 exhibited relatively longer lag period (ca. 3 days). On the other hand, the rest of the isolates took about 30 days to completely mineralize 300 μg/ml of MCPA. These slow-growing isolates had long lag periods (10 to 20 days) and their doubling times during the exponential growth phase were about 20 times longer than that of strain SH3. The previously described microorganisms able to degrade phenoxy herbicides grow well in herbicide medium, and it usually takes several days to completely mineralize the herbicides.[3] Recently, Kamagata *et al.* reported on the isolation of slow-growing, oligotrophic bacteria, but their isolates still completely mineralize the herbicide within 1 week.[2] Our strains are not oligotrophic, since they grow in mineral medium containing more than 1000 μg/ml of MCPA.

Under induced condition, our isolates grew more quickly (Fig. 3), suggesting that their MCPA degradative enzymes were inducible by the presence of MCPA. The inducibility of MCPA degradative enzymes of the isolates was confirmed

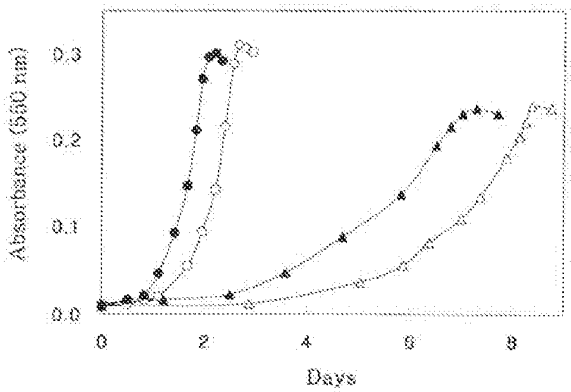


Fig. 3. Growth patterns of MCPA-degrading bacteria in axenic culture. Symbols: ●—● and ○—○, strain SH3; ▲—▲ and △—△, strain SH1. The bacteria were either adapted (solid symbols) or not adapted (open symbols) to MCPA metabolism. Each point is the mean for two replicate flask cultures.

COR-AAD1-00001255

Table 2. Ability of MCPA-degrading strains to degrade other phenoxy acid herbicides.

Herbicide	Strains[a]						
	SH1	SH2	SH3	SH4	SH5	SH6	SH7
MCPA	+	+	+	+	+	+	+
2,4-D	+	+	+	+	+	+	+
MCPB	-	+	-	+	+	-	-
2,4-DB	+	+	-	+	+	-	+
4CPA	+	+	+	+	-	-	+
MCPP	-	-	+	-	-	-	-
2,4-DP	-	-	+	-	-	-	-

[a]The bacterial strains were grown on PTYG and then tested for herbicide utilization capabilities. +, >90% reduction in peak height from UV scanning and substantial growth (OD_{600}>0.25); -, <15% reduction in peak and very scant growth (OD_{600}<0.01).

in resting cell experiments conducted with cells adapted or not adapted to MCPA metabolism. The cells adapted to MCPA metabolism mineralized MCPA quickly without any lag period, while the unadapted cells could not degrade it during the incubation period.

Degradative diversity analysis. The isolates were grown on PTYG medium, then examined for their ability to degrade other compounds related to MCPA (Table 2). The isolates exhibited diverse patterns in their herbicide utilization abilities. Strains SH1, SH2, SH4, and SH7 were able to utilize MCPA, 2,4-D, 2,4-DB and 4-CPA as the sole carbon sources, as indicated by complete disappearance of absorbance peak and by substantial cell growth. Strain SH6 was restricted in its herbicide degradation spectrum. Of the isolates, strain SH3 shows the most interesting degradation capability, since it rapidly degraded five different herbicides, including MCPP and 2,4-DP which have a three-carbon side chain.

Plasmid profile of the isolates. When the isolates were subjected to Kado's plasmid detection procedure, only strain SH3 exhibited two plasmid bands and the other strains contained no plasmids (Fig. 4). To investigate whether the MCPA degradative genes were transmissible to other bacteria, the isolates were mated with antibiotics-resistant recipients, *Pseudomonas cepacia* and *Alcaligenes* sp. In all of the isolates, the MCPA degradation phenotype was not transferred at a detectable frequency ($<10^{-9}$) into the two recipients, suggesting that their MCPA degradation genes were located in the chromosomal DNA or on the untransmissible plasmid. Through a curing process using sodium dodecyl sulfate, a cured strain was obtained from strain SH3. This cured strain did not have the smaller one of the two plasmids and concomitantly lost the ability to degrade MCPP and 2,4-DP. Moreover, its MCPA and 2,4-D degradation rates were significantly decreased, taking 20 to 30 days to completely mineralize 300 μg/ml of MCPA. This result suggests that the smaller plasmid of strain SH3 has very important genes for metabolism of herbicides. In future studies, more attention needs to be placed on understanding the role of this new degradative plasmid in herbicide metabolism.

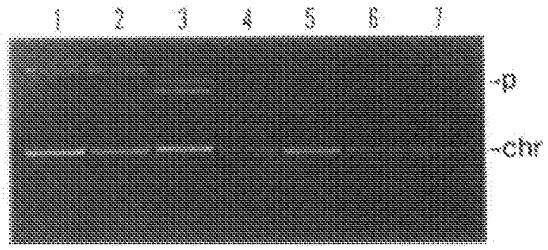


Fig. 4. Plasmid profiles of MCPA-degraders. Lanes: 1, SH1; 2, SH2; 3, SH3; 4, SH4; 5, SH5; 6, SH6; and 7, SH7.

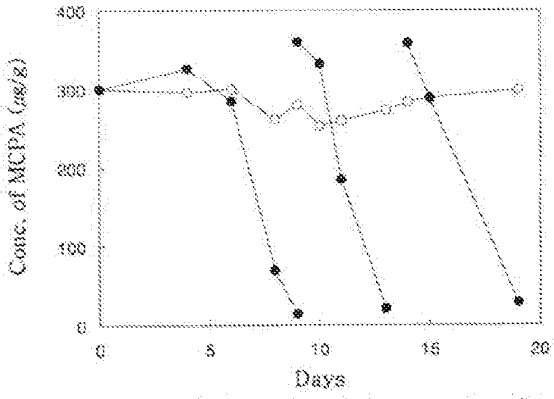


Fig. 5. Degradation of MCPA in soil microcosms inoculated with strain SH3 (●—●) or with only indigenous microbial population (○—○). Each point is the mean for two replicate microcosms.

Degradation of MCPA in soil. The patterns of degradation of MCPA in natural field soils were analyzed with and without MCPA degraders. MCPA was observed to be quickly degraded after short lag periods in field soils with indigenous MCPA degraders. The first dose of MCPA (300 μg/g) in this soil was degraded within two weeks, and it took only about one week for the second and third doses of MCPA to be completely degraded, respectively. However, MCPA was persistent for quite a long time in natural field soils without any indigenous MCPA-degrading microorganisms (Fig. 5). The application of one of the isolates into this natural field soil at a density of 10^7 cells/g soil resulted in quick degradation of MCPA, taking about 1 week for the complete removal of 300 μg/g of MCPA in adapted soils. The result suggested that MCPA used in agricultural practices could persist in the soils without MCPA degraders and that the MCPA-degrading strains isolated in this study thus could help in removing the persistent MCPA residues through in situ bioremediation.

Acknowledgments. This study was supported by Korean Ministry of Education through Research Fund (1997).

COR-AAD1-00001256

References

1. Smith, A. and Hayden, B. J. (1981) Relative persistence of MCPA, MCPB and mecoprop in Saskatchewan soils, and the identification of MCPA in MCPB-treated soils. *Weed Res.* **21**, 179-183.
2. Helweg, A. (1987) Degradation and adsorption of ^{14}C-MCPA in soil-influence of concentration, temperature and moisture content on degradation. *Weed Res.* **27**, 287-296.
3. Smith, A. E. and Aubin, A. J. (1991) Effects of long-term 2,4-D and MCPA field applications on the soil breakdown of 2,4-D, MCPA, Mecoprop, and 2,4,5-T. *J. Environ. Qual.* **20**, 436-438.
4. Smith, A. E., Mortensen, K., Aubin, A. J. and Molloy, M. M. (1994) Degradation of MCPA, 2,4-D, and other phenoxyalkanoic acid herbicides using an isolated soil bacterium. *J. Agric. Food Chem.* **42**, 401-405.
5. Alexander, M. and Aleen, M. I. H. (1961) Effect of chemical structure on microbial decomposition of aromatic herbicides. *J. Agr. Food Chem.* **9**, 44-47.
6. Stephen, D. Y. and Kuwatsuka, S. (1980) Effect of organic matter and chemical fertilizers on the degradation of benthiocarb and MCPA herbicides in the soil. *Soil Sci. Plant Nutr.* **26**, 541-549.
7. Roeth, F. W. (1986) Enhanced herbicide degradation in soil with repeat application. *Rev. Weed Sci.* **2**, 45-65.
8. Loos, M. A., Roling, J. M. and Alexander, M. (1967) Phenoxyacetate herbicide detoxication by bacterial enzymes. *J. Agr. Food Chem.* **15**, 858-860.
9. Kilpi, S., Backstrom, V. and Korhola, M. (1980) Degradation of 2-methyl-4-chlorophenoxyacetic acid (MCPA), 2,4-dichlorophenoxyacetic acid (2,4-D), benzoic acid and salicylic acid by *Pseudomonas* sp. HV3. *Microbiol. Lett.* **8**, 177-182.
10. Pieper, D. H., Reineke, W., Engesser, K. H. and Knackmuss, H. J. (1988) Metabolism of 2,4-dichlorophenoxyacetic acid, 4-chloro-2-methylphenoxyacetic acid and 2-methylphenoxyacetic acid by *Alcaligenes eutrophus* JMP 134. *Arch. Microbiol.* **150**, 95-102.
11. Ka, J. O., Holben, W. E. and Tiedje, J. M. (1994) Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D-treated field soils. *Appl. Environ. Microbiol.* **60**, 1106-1115.
12. Stanier, R. Y., Palleroni, N. J. and Doudoroff, M. (1966) The aerobic pseudomonads: A taxonomic study. *J. Gen. Microbiol.* **43**, 159-271.
13. Sasser, M. (1990) Technical Note 102: Tracking a Strain Using the Microbial Identification System. MIS Inc., North Newark, DE.
14. de Bruijn, F. J. (1992) Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. *Appl. Environ. Microbiol.* **58**, 2180-2187.
15. Versalovic, J., Schneider, M., de Bruijn, F. J. and Lupski, J. R. (1994) Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol. Cell. Biol.* **5**, 25-40.
16. Ka, J. O. and Tiedje, J. M. (1994) Integration and excision of a 2,4-dichlorophenoxyacetic acid-degradative plasmid in *Alcaligenes paradoxus* and evidence of its natural intergeneric transfer. *J. Bacteriol.* **176**, 5284-5289.
17. Kado, C. I. and Liu, S. T. (1981) Rapid procedure for detection and isolation of large and small plasmids. *J. Bacteriol.* **145**, 1365-1373.
18. Willetts, N. S. (1988) Conjugation. *Methods Microbiol.* **21**, 49-77.
19. Ka, J. O., Holben, W. E. and Tiedje, J. M. (1994) Use of gene probes to aid in recovery and identification of functionally dominant 2,4-dichlorophenoxyacetic acid-degrading populations in soil. *Appl. Environ. Microbiol.* **60**, 1116-1120.
20. Loos, M. A., Schlosser, J. F. and Mapham, W. R. (1979) Phenoxy herbicide degradation in soils: quantitative studies of 2,4-D- and MCPA-degrading microbial populations. *Soil Biol. Biochem.* **11**, 377-385.
21. Kamagata, Y., Fulthorpe, R. A., Tamura, K., Takami, H., Forney, L. J. and Tiedje, J. M. (1997) Pristine environments harbor a new group of oligotrophic 2,4-dichlorophenoxyacetic acid-degrading bacteria. *Appl. Environ. Microbiol.* **63**, 2266-2272.

COR-AAD1-00001257

The Journal of Microbiology, June 2004, p.87-93 Vol. 42, No. 2
Copyright ⓒ 2004, The Microbiological Society of Korea

Genetic and Phenotypic Diversity of (*R/S*)-Mecoprop [2-(2-Methyl-4-Chlorophenoxy)Propionic Acid]-Degrading Bacteria Isolated from Soils

Jong-Sung Lim, Mee-Kum Jung, Mi-Soon Kim, Jae-Hyung Ahn, and Jong-Ok Ka*

School of Agricultural Biotechnology, Seoul National University, Seoul 151-742, South Korea

(Received January 29, 2004 / Accepted April 2, 2004)

Twelve mecoprop-degrading bacteria were isolated from soil samples, and their genetic and phenotypic characteristics were investigated. Analysis of 16S rDNA sequences indicated that the isolates were related to members of the genus *Sphingomonas*. Ten different chromosomal DNA patterns were obtained by polymerase-chain-reaction (PCR) amplification of repetitive extragenic palindromic (REP) sequences from the 12 isolates. The isolates were found to be able to utilize the chiral herbicide mecoprop as a sole source of carbon and energy. While seven of the isolates were able to degrade both (*R*)- and (*S*)-mecoprop, four isolates exhibited enantioselective degradation of the (*S*)-type and one isolate could degrade only the (*R*)-enantiomer. All of the isolates were observed to possess plasmid DNAs. When certain plasmids were removed from isolates MP11, MP15, and MP23, those strains could no longer degrade mecoprop. This compelling result suggests that plasmid DNAs, in this case, conferred the ability to degrade the herbicide. The isolates MP13, MP15, and MP24 were identified as the same strain; however, they exhibited different plasmid profiles. This indicates that these isolates acquired different mecoprop-degradative plasmids in different soils through natural gene transfer.

Key words: mecoprop-degrading bacteria, biodegradation, mecoprop, bacterial diversity

Phenoxyalkanoic acid herbicides have enormous agricultural utility, and thus many studies have focused on their persistence and fate in natural environments. (Torstensson *et al.*, 1975; Smith and Hayden, 1981; Smith and Aubin, 1991). 2-(4-chloro-2-methylphenoxy) propionic acid (mecoprop, MCPP) is a chlorinated phenoxyalkanoic herbicide, and is used to control broad-leaved weeds in cereals and grassland. While phenoxyacetic acids, such as 2,4-dichlorophenoxyacetic acid (2,4-D) and 4-chloro-2-methylphenoxyacetic acid (MCPA), are achiral, phenoxypropionic acids, including mecoprop and 2-(2,4-chlorophenoxy)propionic acid (dichlorprop, 2,4-DP), are chiral due to an asymmetrically substituted carbon atom. Most commercial formulations of mecoprop are racemic mixtures, containing both the *R* and *S* enantiomers of mecoprop, although only the *R* enantiomer has been reported to be herbicidally active (Aberg, 1956).

Among the phenoxyalkanoic acids, 2,4-D has been most extensively studied regarding its environmental fate, metabolizing microorganisms, degradation pathways, and degradative genes and enzymes (Loos *et al.*, 1967; Don *et al.*, 1985; Pieper *et al.*, 1988; Ka *et al.*, 1994; Chung and Ka, 1998). In contrast, relatively little information is available on the microorganisms able to degrade the chiral phenoxyalkanoic acid herbicide mecoprop (Horvath *et al.*, 1990; Tett *et al.*, 1994). The main reason for this dearth of information is that, unlike the 2,4-D-degrading organisms, mecoprop-degrading microorganisms are fairly rare and thus not easily isolated in pure cultures on laboratory media. Mecoprop is considered more recalcitrant than 2,4-D due to the propionic acid side chain linked to the oxygen atom of the phenoxyacid moiety (Alexander and Aleem, 1961). It requires long and continuous cultivation to obtain degradation-positive cultures of the herbicide, and the prolonged enrichment often results in multi-membered mixed cultures (Lappin *et al.*, 1985; Tett *et al.*, 1994). The difficulty of obtaining pure cultures implies that complete mineralization of the herbicides in nature can only occur through the syntrophism of a consortium of microorganisms (Tett *et al.*, 1994).

In this work, we studied the occurrence of mecoprop-degrading microorganisms in two different types of natural soils, i.e., agricultural soils with no history of mecoprop treatment, and golf course soils, which had been routinely treated with the herbicide for several years. We were able to isolate twelve mecoprop-degrading bacteria from golf course soil samples, but none from agricultural soils. We investigated species diversity by 16S rDNA sequence analysis and REP-PCR patterns of chromosomes, and described physiological and genetic properties of the isolates on mecoprop biodegradation.

* To whom correspondence should be addressed.
(Tel) 82-2-880-4673; (Fax) 82-2-871-2095
(E-mail) kagil@plaza.snu.ac.kr

COR-AAD1-00001258

Materials and Methods

Media and culture conditions

All isolates were cultivated on mineral medium (Park and Ka, 2003) containing mecoprop at a concentration of 300 μg/ml. Peptone-tryptone-yeast extract-glucose (PTYG) medium containing (per liter) 0.25 g of peptone (Difco, USA), 0.25 g tryptone (Difco, USA), 0.5 g of yeast extract (Difco, USA), 0.5 g of glucose, 0.03 g of magnesium sulfate, and 0.003 g of calcium chloride was used for strain purification and colony production for repetitive extragenic palindromic PCR (REP-PCR). All cultures were incubated at 30°C and liquid cultures were aerated by shaking at 200 rpm on a rotary shaker (Vision Co., Korea).

Chemicals

2-(2-methyl-4-chlorophenoxy)propionic acid (MCPP, mecoprop), 2-(2,4-dichlorophenoxy) propionic acid (2,4-DP, dichlorprop), 4-(2,4-dichlorophenoxy)butyric acid (2,4-DB), and 4-(2-methyl-4-chlorophenoxy)butyric acid (MCPB) were obtained from Aldrich Chemical Co., Milwaukee, USA, and analytical grade 2,4-dichlorophenoxyacetic acid (2,4-D), 2-methyl-4-chlorophenoxyacetic acid (MCPA), and (*RS*)-enantiomers of mecoprop were purchased from Sigma Chemical Co. (USA).

Analysis of phenoxyalkanoic acids

HPLC analyses were performed on a Shimadzu LC-10 series high-performance liquid chromatography (Shimadzu Co., Tokyo, Japan). A Nucleodex-α-PM column (200×4 mm) with permethylated α-cyclodextrin (Machcrey-Nagel, Düren, Germany) was used to separate mecoprop enantiomers, and a UV detector was set at 230 nm. The eluent system consisted of 70% methanol and 30% NaH_2PO_4 (50 mM, pH 3.0) at a flow rate of 0.7 ml/min. Retention times were typically 8.5 min for (*R*)-mecoprop, and 11.5 min for (*S*)-mecoprop. UV spectra were recorded with a UVIKON 860 spectrophotometer (Kontron Ins., Milan, Italy) at wavelengths from 340 to 220 nm.

Isolation of mecoprop-degraders

Soil samples were taken from South Korean golf courses and agricultural fields in various locations. Golf courses selected in this study had been treated with racemic mecoprop for several years, while the upland fields had undergone normal agricultural practices growing dry crops, such as legumes and barley for more than 10 years, and had no history of mecoprop treatment. Samples from the top 15 cm of soil were taken, sifted through a 2-mm-pore sieve, and kept at 4 °C prior to use. 100 g of each soil sample was transferred to each 250-ml sterile beaker, treated with racemic mecoprop dissolved in a 0.1 M NaH_2PO_4 buffer (pH 7.0) to a concentration of 300 ppm, and thoroughly mixed. Four weeks after mecoprop application, a 10-g soil sample from each beaker was homogenized with 95 ml of a sterile 0.85% saline solution, by shaking the preparation at 200 rpm on a rotary shaker. After homogenization, 0.1 ml samples of appropriate 10-fold dilutions were inoculated into test tubes containing 3 ml of mecoprop medium (300 μg/ml). The tubes were incubated at 30°C for 4 weeks and the amount of mecoprop degradation was ascertained by spectrophotometry and HPLC. The culture of the terminal positive tube (showing substantial cell growth and less than 50% of the mecoprop remaining) was enriched by two additional transfers into fresh medium. Each enriched culture was streaked onto PTYG agar medium, and single colonies were then tested for mecoprop degradation in fresh mecoprop medium, before strain purification.

16S rDNA sequence analysis

Total genomic DNA was extracted from the isolates and PCR amplification of 16S rRNA genes was performed with 27f and 1492r as previously described (Lane, 1991). The amplified 16S rRNA genes were sequenced using an ABI Prism BigDye Terminator Cycle Sequencing Ready Kit according to manufacturers instructions (Perkin-Elmer) with the sequencing primers 27f and 519r (Hugenholtz *et al.*, 1998). Approximately 400 unambiguous nucleotide positions were used for comparison to the data in GenBank using the Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990). Sequences from nearest relatives were identified from the Ribosome Database Project (RDP) using the SIMILARITY-RANK program of the RDP (Maidak *et al.*, 2000).

Nucleotide sequence accession numbers

The partial 16S rDNA sequences of the 12 isolates have been deposited in the GenBank under the following accession numbers: MP1, AY521007; MP9, AY521008; MP11, AY521009; MP12, AY521010; MP13, AY521011; MP15, AY521012; MP18, AY521013; MP19, AY521014; MP20, AY521015; MP23, AY521016; MP24, AY521017; MP27, AY521018.

Colony REP-PCR

Colony REP-PCR was performed using BOXA1R as a primer, as described previously (de Bruijn, 1992; Versalovic *et al.*,1994). Each isolate was grown on the PTYG agar medium for 24 to 48 h, and a small amount of cells was resuspended in 25 μl of PCR mixture. After PCR amplification, 10 μl samples of the REP-PCR products were separated by electrophoresis on horizontal 1.5% agarose gels.

Degradative phenotype analysis

Each strain was grown on PTYG medium. Cells were harvested by centrifugation at 10,000×g for 10 min at 4°C, washed twice with an equal volume of 15 mM phosphate buffer (pH 7.0), and resuspended in the same buffer. Ali-

COR-AAD1-00001259

quots of suspended cells were inoculated into culture tubes, each of which contained mineral medium supplemented with a 300 μg/ml concentration of one of the structural analogs. After 4 weeks of incubation, the cultures were centrifuged to remove cellular material, and UV spectra were measured from 340 to 220 nm to determine the degree of phenoxyacetate degradation.

Axenic culture experiment

After growth in PTYG medium, cells were harvested, washed, and resuspended in sodium phosphate buffer. Aliquots of suspended cells were inoculated into duplicate flasks containing 200 ml of racemic mecoprop medium (300 μg/ml) at a final density of OD_{550}=0.005. All cultures were incubated at 30°C, and were aerated by shaking at 200 rpm on a rotary shaker. Aliquots of the cultures were regularly removed to determine cell growth and mecoprop degradation.

Plasmid detection and conjugation

In order to detect plasmid DNA, cells were lysed using a modified form (Ka and Tiedje, 1994) of the previously published procedure (Kado and Liu, 1981). Plasmid curing was performed by the sodium dodecyl sulfate method (Crosa *et al.*, 1994). To analyze transferability of the mecoprop degradative phenotype in the isolates, matings were performed on membrane filters, as described by Willetts (Willetts, 1988), using *Alcaligenes eutrophus* JMP228 and *Pseudomonas cepacia* DBO1 as recipients. Transconjugants were selected on mecoprop mineral media containing either rifampicin (100 μg/ml) for *Alcaligenes eutrophus* JMP228, or kanamycin (75 μg/ml) for *Pseudomonas cepacia* DBO1, 1.5% Noble agar (Difco, USA), and 300 μg/ml of mecoprop.

PCR amplification of the* tfd *genes

Partial gene sequences specific to the 2,4-D degradation pathway were amplified by PCR with specific primers targeted for the *tfdA, B,* and *C* genes of the 2,4-D degradative plasmid pJP4 (Don *et al.*, 1985). PCR cycles and primers for the *tfd* genes were previously reported (Berthelet *et al.*, 1996; Vallaeys *et al.*, 1996; Kamagata *et al.*, 1997). Amplification of the *tfdA, B,* and *C* genes with corresponding primers was expected to produce 362 bp, 205 bp, and 361 bp DNA fragments, respectively.

Results and Discussion

Isolation of mecoprop-degrading bacteria

Both agricultural and golf course soil samples were subjected to a simple enrichment procedure to isolate mecoprop-degrading microorganisms. However, none of the 67 agricultural soil samples showed any detectable mecoprop degradation during 4 weeks of incubation. Thus, no mecoprop-degrading organisms were recognized in any of the agricultural samples. By way of contrast, 12 of the 23 golf course soils, which had been treated with mecoprop for several years, exhibited positive degradation of the herbicide. Twelve mecoprop-degrading bacteria were isolated and purified from the positive tubes of 12 different soils (Table 1).

It usually takes a long time to enrich microbial cultures which are able to degrade mecoprop or dichlorprop, and many of the researches have had to use soil-column techniques and chemostat cultures to enrich them (Lappin *et al.*, 1985; Horvath *et al.*, 1990). This difficulty suggested that microorganisms which possess the complete degradative pathway required to mineralize the herbicides are very rare in natural environments (Tett *et al.*, 1994). This was true for the agricultural soils tested in this work, as we observed no positive degradation of mecoprop in any of the 67 different agricultural soil samples. In contrast, about 52% of the golf course soil samples exhibited positive degradation of the herbicide, and twelve strains able to independently mineralize (*R*) and/or (*S*)-type mecoprops were readily isolated from these soils with a simple enrichment procedure. The results suggest that mecoprop-degrading microorganisms are ubiquitous in soils which

Table 1. Nearest relatives of the mecoprop-degrading isolates based upon 16S rDNA sequence analysis

Isolate	GenBank accession no.	Soil sites	Nearest relative	Similarity(%)[a]
MP 1	AY521007	Seongnam, Gyeonggi-do	*Sphingomonas wittichii*	95
MP 9	AY521008	Seongnam, Gyeonggi-do	*Sphingomonas agrestis*	99
MP11	AY521009	Yongin, Gyeonggi-do	*Sphingomonas suberfaciens*	97
MP12	AY521010	Incheon, Gyeonggi-do	*Sphingomonas* sp. VM506	98
MP13	AY521011	Incheon, Gyeonggi-do	*Sphingomonas agrestis*	99
MP15	AY521012	Yangsan, Gyeongnam	*Sphingomonas agrestis*	99
MP18	AY521013	Jinhae, Gyeongnam	*Sphingomonas* sp. VM506	98
MP19	AY521014	Gyeongsan, Gyeongbuk	*Sphingomonas* sp. JS1	98
MP20	AY521015	Gyeongsan, Gyeongbuk	*Sphingomonas* sp. JS1	97
MP23	AY521016	Jeju-si, Jeju-do	*Sphingomonas agrestis*	98
MP24	AY521017	Namjeju-gun, Jeju-do	*Sphingomonas agrestis*	99
MP27	AY521018	Gyeongsan, Gyeongbuk	*Sphingomonas* sp. JS1	97

[a]Based upon approximately 400 nucleotide positions at the 5 end of the 16S rRNA gene

COR-AAD1-00001260

are routinely treated with mecoprop, but are, at best, very rare in environments with no history of mecoprop treatment.

16S rDNA sequence and REP-PCR analyses

16S rDNA sequence analysis confirmed that all isolates were related to members of the genus *Sphingomonas* (Table 1). The isolates were all Gram-negative, and belonged to the alpha subgroup of Proteobacteria. Although the mecoprop-degrading strains were isolated from different soils, some isolates were most closely related to the same species, notably, *S. agrestis* and *Sphingomonas* sp. JS1. REP-PCR experiments were performed in order to study genomic relatedness among the isolates that appeared most closely related. It was revealed that the 12 isolates produced 10 different DNA fingerprint patterns (Fig. 1). This suggests that, although the isolates all belonged to the genus *Sphingomonas*, they possessed radically different genomic fingerprints. Strains MP13, MP15, and MP24, however, which were classified as *S. agrestis*, exhibited identical genomic DNA fingerprints. This indicated that the three strains are either siblings, or at least very closely related. Since all of the strains were isolated from different soil samples, their detection frequencies reflect their ubiquity in the soils examined.

Enantioselective degradation of (RS)-mecoprop in axenic cultures

To understand the preferential degradation patterns of the enantiomers of racemic mixtures by the mecoprop degraders, each strain was grown on PTYG, and inoculated into racemic mecoprop medium. The growth and mecoprop degradation curves of representative strains are shown in Fig. 2. Isolates were classified into five groups, according to their growth characteristics on racemic mecoprop. Group I contained strain MP11, which degraded both enantiomers almost simultaneously and grew rapidly after an initial lag period of about 1 day (Fig. 2A). Group II included strains MP13, MP19, MP20, and MP23, which also could degrade both enantiomers, but degraded (*S*)-mecoprop preferentially (Fig. 2B). Group III contained strains MP12 and MP15, which both degraded mecoprop in an enantioselective manner, but exhibited a long lag phase of about 4 days before the strains began to grow on (*R*)-mecoprop (Fig. 2C). Group IV consisted of strains MP1, MP9, MP18, and MP24, which could degrade (*S*)-mecoprop, but were unable to degrade (*R*)-mecoprop (Fig. 2D). Group V contained strain MP27, which could not degrade (*S*)-mecoprop, but degraded (*R*)-mecoprop after

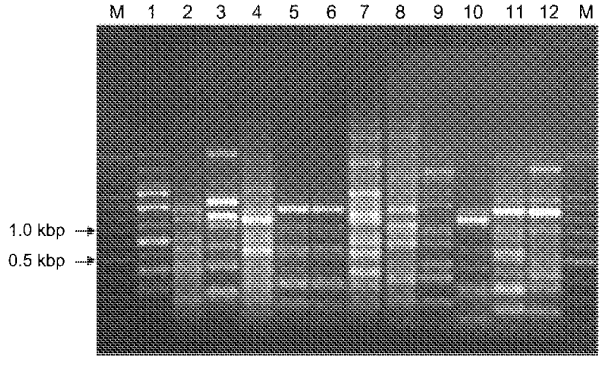


Fig. 1. REP-PCR band patterns of the isolates. Lanes : 1, MP1; 2, MP9; 3, MP11; 4, MP12; 5, MP13; 6, MP15; 7, MP18; 8, MP19; 9, MP20; 10, MP23; 11, MP24; 12, MP27; M, DNA size marker.

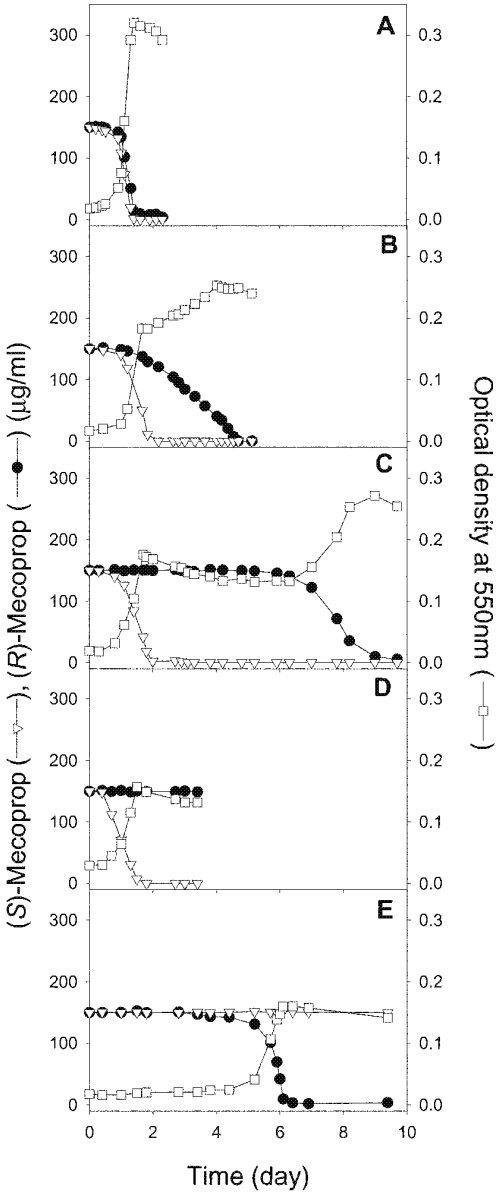


Fig. 2. Growth patterns of mecoprop-degrading bacteria on racemic mecoprop. Disappearance of mecoprop enantiomers and growth by strain MP11 (A), strain MP19 (B), strain MP15 (C), strain MP1 (D), and strain MP27 (E). Each point is the mean for two replicate liquid cultures.

COR-AAD1-00001261

Table 2. Substrate utilization patterns by the mecoprop-degrading isolates.

Isolate[a]	Substrate[b]					
	Mecoprop	Dichlorprop	2,4-D	MCPA	2,4-DB	MCPB
MP 1	+(S)	+(S)	+	++	-	-
MP 9	+(S)	+(S)	+	++	-	-
MP11	++	++	+	++	-	-
MP12	++	++	+	++	-	-
MP13	++	++	+	++	-	-
MP15	++	++	+	++	-	-
MP18	+(S)	+(S)	+	++	-	-
MP19	++	++	+	++	-	-
MP20	++	++	-	-	-	-
MP23	++	++	+	++	-	-
MP24	+(S)	+(S)	+	++	-	-
MP27	+(R)	+(R)	+	++	-	-

[a]The isolates were grown on PTYG before the test of substrate utilization.
[b]Test for the utilization of substrates. ++, >95% reduction in peak height as determined by UV scanning and substantial growth (optical density at 550 nm >0.23); +, 40 to 60% reduction in peak height and moderate growth (optical density at 550 nm >0.14); –, <10% reduction in peak height and very scant growth (optical density at 550 nm <0.01); +(S), utilization of only (*S*)-type enantiomer; +(R), utilization of only (*R*)-type enantiomer. mecoprop, 2-(2-methyl-4-chlorophenoxy)propionic acid; dichlorprop, 2-(2,4-dichlorophenoxy)propionic acid; 2,4-D, 2,4-dichlorophenoxyacetic acid; MCPA, 2-methyl-4-chlorophenoxyacetic acid; 2,4-DB, 4-(2,4-dichlorophenoxy)butyric acid; MCPB, 4-(2-methyl-4-chlorophenoxy)butyric acid.

an initial lag period of about 5 days (Fig. 2E). The final culture absorbances of groups I, II, and III, which degraded both enantiomers, were about twice as high as those of groups IV and V, which degraded only one enantiomer. Observed growth characteristics on racemic mecoprop indicate that (*RS*)-mecoprop degradation patterns in the isolates are much more diverse than those of the enantioselective mecoprop-degraders, *Alcaligenes denitrificans* and *Sphingomonas herbicidovorans* MH (Tett *et al*., 1994; Zipper *et al*., 1996). *A. denitrificans* could degrade only the (*S*) enantiomer of mecoprop, like our group IV strains, and *S. herbicidovorans* MH showed a sequential degradation pattern, similar to our group II strains. The strains belonging to groups I, III, and V of this study exhibited new growth characteristics on mecoprop, indicating that racemic mecoprop biodegradation mechanisms are very diverse among mecoprop-degrading microbial populations in nature. The diverse enantiomer-specific degradation patterns of the isolates support the previous observation that separate enzyme systems are required for the degradation of each mecoprop enantiomer (Zipper *et al*., 1996).

Degradative diversity analysis

Isolates were grown on PTYG medium, and examined for their ability to degrade other mecoprop-related compounds. All mecoprop-degrading isolates, except strain MP20, were also able to degrade related herbicides, such as (*RS*)-dichlorprop, 2,4-D, and MCPA (Table 2). When group IV strains, which could not degrade (*R*)-type mecoprop, were inoculated into minimal medium containing pure (*R*)-type dichlorprop as the sole carbon source, they were unable to degrade and grow on the (*R*)-type dichlorprop. In contrast, strain MP27 of group V, which degraded only (*R*)-mecoprop, was able to degrade only (*R*)-dichlorprop. The other isolates, which could degrade (*RS*)-mecoprop, also degraded (*RS*)-dichlorprop. These results suggested that degradation enzyme systems for mecoprop enantiomers in the isolates are closely related to those for dichlorprop enantiomers. All isolates degraded 2,4-D to completion, but growth yields were lower than with (*RS*)-mecoprop. It is of note that most of the isolates could utilize (*S*)-mecoprop as a carbon and energy source, while relatively limited populations degraded the (*R*)-mecoprop. These results are in agreement with the observation that the herbicidally active (*R*)-mecoprop was more slowly degraded than the inactive (*S*)-type enantiomer in environmental samples (Heron and Christensen, 1992; Müller and Buser, 1997).

Plasmid detection and its relationship to mecoprop degradation phenotype

When the isolates were subjected to Kado's plasmid detection procedure (Kado and Liu, 1981), all isolates exhibited one to five plasmid DNA bands (Fig. 3). These plasmids were stably maintained in cells cultivated for two months,

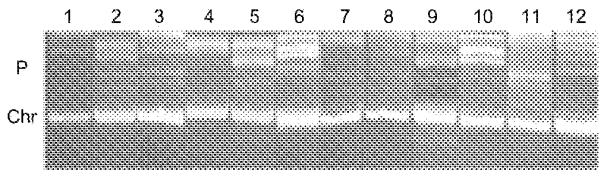


Fig. 3. Plasmid profiles of the isolates. Lanes : 1, MP1; 2, MP9; 3, MP11; 4, MP12; 5, MP13; 6, MP15; 7, MP18; 8, MP19; 9, MP20; 10, MP23; 11, MP24; 12, MP27. P, plasmid band; Chr, chromosomal band.

COR-AAD1-00001262

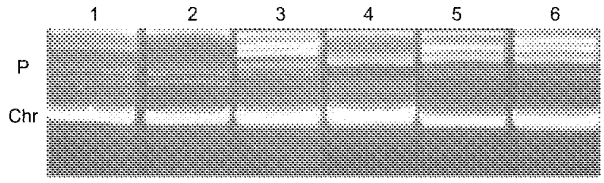


Fig. 4. Plasmid profiles of the representative isolates and their cured strains. Lanes : 1, MP11; 2, cured strain of MP11; 3, MP15; 4, cured strain of MP15; 5, MP23; 6, cured strain of MP23. P, plasmid band; Chr, chromosomal band.

with repeated transfers into Luria-Bertani medium. To investigate the relationship between mecoprop degradation phenotype and these plasmids, some representative isolates were subjected to sodium dodecyl sulfate plasmid curing procedures (Fig. 4). The cured strain of MP11 lost the ability to degrade (*RS*)-mecoprop when it lost the second plasmid from the top (Fig. 4, lane 2). Initially, strain MP15 was able to degrade both (*RS*)-mecoprop and (*RS*)-dichlorprop. However, when the second to fifth plasmids from the top were lost, producing smaller plasmids (Fig. 4, lane 4), the cured strain of MP15 could no longer degrade any of the enantiomers. In the case of strain MP23, part of the first plasmid was eliminated through the curing procedure, producing a smaller plasmid (Fig. 4, lane 6), and the resultant cured strain could not degrade (*RS*)-mecoprop, (*RS*)-dichlorprop, or 2,4-D. These results suggested that plasmids contained the essential mecoprop-degradation genes. When the plasmid-containing isolates were mated with two antibiotic-resistant recipients, *Pseudomonas cepacia* DBO1 and *Alcaligenes eutrophus* JMP228, the mecoprop degradation phenotype was not transferred to either of the two recipients at a detectable frequency ($<10^{-9}$).

Interestingly, strains MP13, MP15, and MP24 exhibited different plasmid profiles, although they were identified as the same strain, *Sphingomonas agrestis*, based on their colony morphologies, genomic DNA fingerprints, and 16S rDNA sequences. These three isolates showed different growth characteristics on racemic mecoprop, each strain belonging to three different groups (Fig. 2). The substrate utilization pattern of strain MP24 was also different from those of strains MP13 and MP15 (Table 2). Since the three isolates identified as the same strain had the same genomic DNA fingerprint, their differences in mecoprop degradation properties were clearly due to their different plasmid profiles. Most of the previous mecoprop-degrading microorganisms were isolated through long periods of subcultivation from multi-membered mixed cultures, suggesting that the mecoprop-degradative pathway in a pure culture resulted from the exchange of genetic information between several organisms in the mixed culture (Horvath *et al.*, 1990; Tett *et al.*, 1994). Genes specifying pesticide-biodegradative ability are often contained on plasmids, which can facilitate the transfer of genetic information among microbial populations (Pemberton *et al.*, 1979). Our results indicated that the three isolates acquired different catabolic plasmid DNAs from their neighboring microbial communities, while they were living in different soils under mecoprop treatment. The different plasmid profiles leading to distinct catabolic activities in the same isolates demonstrate how microbial populations adapt and evolve in natural environments amended with herbicides. In our laboratory, we are currently investigating how each plasmid of the three strains is involved in the degradation of (*RS*)-mecoprop, and related phenoxy herbicides.

PCR amplification of the tfd *genes*

Since all of the pure isolates and the syntrophic pairs were able to degrade 2,4-D as well as mecoprop, we investigated whether the isolates have any *tfd*-like genes of the 2,4-D degradative plasmid pJP4. When PCR amplification was performed for every isolate using specific primers selected from the internal sequences of the *tfdA*, *B*, and *C* genes, none of the isolates exhibited any PCR DNA bands (data not shown). The results indicate that the genes involved in 2,4-D degradation in these isolates do not have any significant homology with the well-known *tfd* genes of the 2,4-D degradative plasmid pJP4 (Don *et al.*, 1985).

Acknowledgment

This study was supported by the BioGreen 21 Project.

References

Aberg, B. 1956. On the effects of para-substitution in some plant growth regulators with phenyl nuclei, pp. 93-116. *In* R.L. Wain and F. Wightman (Eds.), The Chemistry and Mode of Action of Plant Growth Substances. Butterworth Scientific Publications, London.

Alexander, M. and M.I.H. Aleem. 1961. Effect of chemical structure on microbial decomposition of aromatic herbicides. *J. Agric. Food Chem.* 9, 44-47.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment tool. *J. Mol. Biol.* 215, 403-410.

Berthelet, M., L.G. Whyte, and C.W. Greer. 1996. Rapid, direct extraction of DNA from soils for PCR analysis using polyvinylpolypyrrolidone spin columns. *FEMS Microbiol. Lett.* 138, 17-22.

Chung, M.J. and J.O. Ka. 1998. Isolation and characterization of 2,4-dichlorophenoxyacetic acid-degrading bacteria from paddy soils. *J. Microbiol.* 36(4), 256-261.

Crosa, J.H., M.E. Tolmasky, L.A. Actis, and S. Falkow. 1994. Plasmids, p. 365-386. *In* P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg (eds.), Methods for General and Molecular Bacteriology. Am. Soc. Microbiol., Washington.

de Bruijn, F.J. 1992. Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria.

COR-AAD1-00001263

Appl. Environ. Microbiol. 58, 2180-2187.

Don, R.H., A.J. Weightman, H.-J. Knackmuss, and K.N. Timmis. 1985. Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4). *J. Bacteriol.* 161, 85-90.

Heron, G. and T.H. Christensen. 1992. Degradation of the herbicide mecoprop in an aerobic aquifer determined by laboratory batch studies. *Chemosphere* 24, 547-557.

Horvath, M., G. Ditzelmüller, M. Loidl, and F. Streichsbier. 1990. Isolation and characterization of a 2-(2,4-dichlorophenoxy)propionic acid-degrading soil bacterium. *Appl. Microbiol. Biotechnol.* 33, 213-216.

Hugenholtz, P., C. Pitulle, K.L. Hershberger, and N.R. Pace. 1998. Novel division level bacterial diversity in a Yellowstone hot spring. *J. Bacteriol.* 180, 366-376.

Ka, J.O., W.E. Holben, and J.M. Tiedje. 1994. Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D-treated field soils. *Appl. Environ. Microbiol.* 60, 1106-1115.

Ka, J.O. and J.M. Tiedje. 1994. Integration and excision of a 2,4-dichlorophenoxyacetic acid-degradative plasmid in *Alcaligenes paradoxus* and evidence of its natural intergeneric transfer. *J. Bacteriol.* 176, 5284-5289.

Kado, C.I. and S.T. Liu. 1981. Rapid procedure for detection and isolation of large and small plasmids. *J. Bacteriol.* 145, 1365-1373.

Kamagata, Y., R.A. Fulthorpe, K. Tamura, H. Takami, L.J. Forney, and J.M. Tiedje. 1997. Pristine environments harbor a new group of oligotrophic 2,4-dichlorophenoxyacetic acid-degrading bacteria. *Appl. Environ. Microbiol.* 63, 2266-2272.

Lane, D.J. 1991. 16S/23S rRNA sequencing, pp. 115-148. *In* E. Stackebrandt and M. Goodfellow (Eds.), Nucleic Acid Techniques in Bacterial Systematics. John Wiley and Sons, England.

Lappin, H.M., M.P. Greaves, and J.H. Slater. 1985. Degradation of the herbicide mecoprop [2-(2-methyl-4-chlorophenoxy)propionic acid] by a synergistic microbial community. *Appl. Environ. Microbiol.* 49, 429-433.

Loos, M.A., R.N. Roberts, and M. Alexander. 1967. Phenols as intermediates in the decomposition of phenoxyacetates by an *Arthrobacter* species. *Can. J. Microbiol.* 13, 679-690.

Maidak, B.L., J.R. Cole, T.G. Lilburn, C.T. Parker Jr, P.R. Saxman, J.M. Stredwick, G.M. Garrity, B. Li, G.J. Olsen, S. Pramanik, T.M. Schmidt, and J.M. Tiedje. 2000. The RDP (Ribosomal Database Project) continues. *Nucleic Acids Res.* 28(1), 173-174.

Müller, M.D. and H. Buser. 1997. Conversion reactions of various phenoxyalkanoic acid herbicides in soil. 1. Enantiomerization and enantioselective degradation of the chiral 2-phenoxypropionic acid herbicides. *Environ. Sci. Technol.* 31, 1953-1959.

Park, H.D. and J.O. Ka. 2003. Genetic and phenotypic diversity of dichlorprop-degrading bacteria isolated from soils. *J. Microbiol.* 41(1), 7-15.

Pemberton, J.M., B. Corney, and R.H. Don. 1979. Evolution and spread of pesticide degrading ability among soil microorganisms, pp. 287-299. *In* K.N. Timmis and A. Puhler (Eds.), Plasmids of Medical, Environmental and Commercial Importance. Elsevier/North-Holland Biomedical Press, Amsterdam.

Pieper, D.H., W. Reineke, K.-H. Engesser, and H.-J. Knackmuss. 1988. Metabolism of 2,4-dichlorophenoxyacetic acid, 4-chloro-2-methylphenoxyacetic acid and 2-methylphenoxyacetic acid by *Alcaligenes eutrophus* JMP134. *Arch. Microbiol.* 150, 95-102.

Smith, A.E. and A.J. Aubin. 1991. Effects of long-term 2,4-D and MCPA field applications on the soil breakdown of 2,4-D, MCPA, mecoprop, and 2,4,5-T. *J. Environ. Qual.* 20, 436-438.

Smith, A.E. and B.J. Hayden. 1981. Relative persistence of MCPA, MCPB and mecoprop in Saskatchewan soils, and the identification of MCPA in MCPB-treated soils. *Weed Res.* 21, 179-183.

Tett, V.A., A.J. Willetts, and H.M. Lappin-Scott. 1994. Enantioselective degradation of the herbicide mecoprop [2-(2-methyl-4-chlorophenoxy)propionic acid] by mixed and pure bacterial cultures. *FEMS Microbiol. Ecol.* 14, 191-200.

Torstensson, N.T.L., J. Stark, and B. Goransson. 1975.The effect of repeated applications of 2,4-D and MCPA on their breakdown in soil. *Weed Res.*15,159-164.

Vallaeys,T., R.R. Fulthorpe, A.M. Wright, and G. Soulas. 1996. The metabolic pathway of 2,4-dichlorophenoxyacetic acid degradation involves different families of *tfdA* and *tfdB* genes according to PCR-RFLP analysis. *FEMS Microbiol. Ecol.* 20, 163-172.

Versalovic, J., M. Schneider, F.J. de Bruijn, and J.R. Lupski. 1994. Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol. Cell Biol.* 5, 25-40.

Willetts, N.S. 1988. Conjugation. *Methods Microbiol.* 21, 49-77.

Zipper, C., K. Nickel, W. Angst, and H.-P.E. Kohler. 1996. Complete microbial degradation of both enantiomers of the chiral herbicide mecoprop ((*RS*)-2-(4-chloro-2-methylphenoxy) propionic acid) in an enantioselective manner by *Sphingomonas herbicidovorans* sp. nov.. *Appl. Environ. Microbiol.* 62, 4318-4322.

COR-AAD1-00001264

学会賞受賞論文

（奨励賞）

農薬の微生物代謝とその分解酵素の進化的解析

田村勝徳*

東京大学分子細胞生物学研究所

Microbial Pesticide Degradations and Evolutionary Analysis of Degrading Enzymes

Katsunori TAMURA

Institute of Molecular and Cellular Biosciences, University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo 113-0032

はじめに

地球上に誕生した微生物は30数億年の間進化し続け，天然に存在する広範な芳香族化合物を基質として分解し利用してきた．その間に微生物間での遺伝子交換や代謝遺伝子群の再編成などにより，代謝経路やその制御機構のレンジを広げてきた．物質文明の基礎となった各種産業の発展は，こうした微生物を含む自然界の物質循環系にとって未経験の新しい構造を持つ多くの化合物を産出してきた．特に，農業の分野では夥しい数のハロゲン化有機化合物が人工的に合成され，食料生産の安定確保のために直接的あるは間接的に圃場に投与され続けている．それらの難分解性化合物を利用し生育できる微生物は非常に限られており，分解能を示す微生物の単離および特徴付けや，分解経路の起源とその進化過程を解き明かすことは，難分解性化合物に対する環境保全を目的とした微生物機能の開発といった観点からも極めて重要かつ興味深い課題である．

ペンタクロロニトロベンゼン（PCNB）は，難防除病害であるアブラナ科作物の根こぶ病などに卓効を示す土壌殺菌剤としてかつて圃場に大量散布されていた．本薬剤は，物理化学的に極めて安定で，一般に微生物分解に対しても耐性であることから，土壌中で長期間残留することが知られている[1,2]．このため，ケミカルハザードとして圃場を取り巻く生態系へ何らかの影響を及ぼすことが懸念されていた．

一方，ホルモン系除草剤2,4-ジクロロフェノキシ酢酸（2,4-D）は広葉雑草の防除に広く用いられているが，それらの圃場の土壌中では速やかに代謝分解される．分解微生物やその遺伝子に関してはこれまでに多くの研究報告があるが，自然界に類縁化合物が存在しないこと，代謝酵素の中で最初のステップを触媒する2,4-Dジオキシゲナーゼは唯一2,4-Dの代謝に特異的な酵素であること[3]から，本酵素を含む分解経路の起源に関心がもたれた．

こうした背景から，本研究は2種類の芳香族塩化化合物を例に取り，PCNB剤の土壌中における微生物代謝機構の解明および2,4-D分解経路の進化的な解析を行ったものである．

1. PCNB分解微生物の分布と特性

PCNBは，難防除病害であるアブラナ科作物の根こぶ病，バレイショのそうか病，ナス，トマト，キュウリ，ハクサイ等の苗立枯病などの*Rhizoctonia*属の病原菌に卓効を示すこと[4]から，土壌殺菌剤として毎年圃場に大量散布されていた．本薬剤は，物理化学的に極めて安定で，一般に微生物分解に対しても耐性であることから，多年に渡って圃場に残留することが知られている[1,2]．一方，毎年繰り返し散布されるために，分解微生物の集積等による根こぶ病菌に対する殺菌力の低下が問題となっていた[5]．また，本剤の残留物や代謝分解産物が圃場を取り巻く土壌環境等の生態系に及ぼす影響が危惧されていた．PCNBの土壌中での分解について，微生物の関与が報告されているが[1,2,6-8]，それら分解能を持った微生物の単離や特徴付けに関する報告はなされていなかった．そこで，広範な環境中から採集した試料を用いてPCNB分解活性を調査すると共に，分解微生物の分布と特性を検討した．

* 〒113-0032 東京都文京区弥生1-1-1

COR-AAD1-00001265

Table 1 PCNB degrading activity of soil samples.

Place	No. of samples[a]	No. of samples with PCNB degradability[b]
Koumi Town*	22	22 (100)
Minakami Town	4	3 (75)
Wako City	14	12 (86)
Shiorin-Koen Area	10	7 (70)
Kasukawa City	6	3 (50)
Nasu-Kogen Area	12	7 (58)
Furano Town	12	11 (92)
Yakushima Area	10	5 (50)
Tsumagoi Village*	24	22 (92)
Total	114	92 (81)

a) Samples were collected from diverse positions of each place.
b) PCNB degradation means that more than 20 % of PCNB was degraded.
* These places are cultivated fields where PCNB has been continually applied. All cultures were incubated under still conditions.

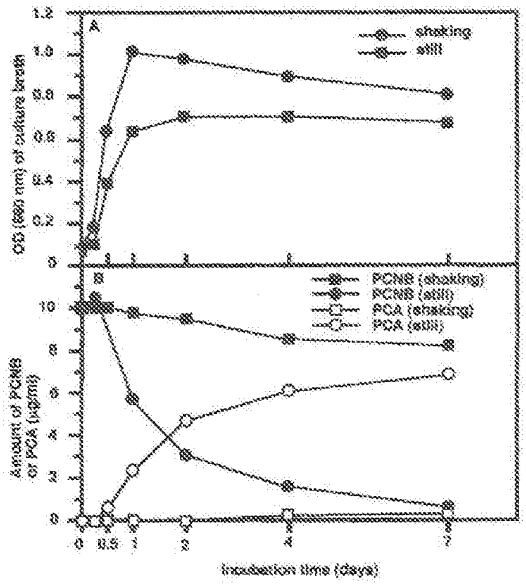


Fig. 1 Time course of bacterial population (A) and metabolism of PCNB to PCA by *P. aeruginosa* I-41 (B) under different culture conditions.
(A) The optical density (660 nm) of culture broth under shaking condition or still conditions. (B) Initial concentration of PCNB was 10 μg/ml, and residual PCNB and produced PCA were extracted with *n*-hexane and analyzed by GC.

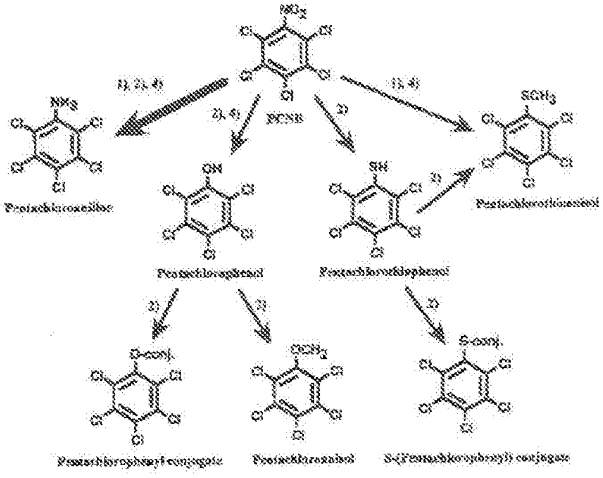


Fig. 2 Metabolic pathway of PCNB.
Each chemical is metabolized in 1) plant, 2) fish, 3) mammal, or 4) soil.

まず，PCNB 剤を毎年繰り返し散布している圃場から土壌を採集したほか，山間地域，海浜地域，湖沼地域，火山帯地域などの PCNB に曝されていないと考えられる種々の環境中から試料を採集し，それらの PCNB 分解能を試験した．培地は 0.1%の酵母エキスを含む最少塩培地を用い，PCNB を 20 μg/ml の濃度で添加した．5 ml の培地に約 1 g の試料を加え，28°C で 1 週間，暗黒下で振とう条件または静置条件で培養した後，残留した PCNB および代謝産物を *n*-ヘキサンで抽出し，その一定量を用いてガスクロマトグラフィー/マススペクトロメトリー (GC/MS) により定量分析した．その結果，PCNB の散布および汚染の有無に関わらず，供試した計 114 の試料中 92 の試料(約 80%)で 20%以上の PCNB の分解が認められた (Table 1)．

次に，PCNB の分解活性が認められた試料から主として細菌を単離し，さらに PCNB 分解能を検定したところ，群馬県水上町の山岳地帯から採取した土壌から，嫌気的条件下で効率的に PCNB を PCA に変換する I-41 株を選抜した．I-41 株は静置培養条件下で 10 μg/ml の PCNB を 7 日間でほぼ完全に分解したが，振とう条件下では 20%以下の分解に留まった．経時的には，本細菌の増殖は約 2 日間で定常期に達し，その後 PCNB の代謝量は徐々に増加した．7 日目には，添加した PCNB の約 70%に対応する PCA が検出された (Fig. 1)．また，I-41 株は細菌学的諸性状を調査したところ，*Pseudomonas aeruginosa* と同定された．

これらの結果から，PCNB を分解する微生物は PCNB の散布の有無に関わらず広範な土壌中に生息しているものと判断された．このことは，PCNB 分解菌の中でも強分解菌として単離された I-41 株が山岳地帯からの土壌に由来すること，本細菌が様々な土壌や水圏環境等に幅広く存在する *P. aeruginosa*[9] と同定されたことからも支持される．さらに，1) PCNB の分解活性を示した殆どの土壌サンプルにおいて PCA が検出されたこと，2) I-41 株による PCNB の主要な代謝産物が PCA であったこと，3) 著者らの予備実験や他の研究グループの報告[10,11] において繰り返し PCNB を散布した土壌における主要な残留物は PCA であったこ

COR-AAD1-00001266

Table 2 Metabolism of PCNB and its related compounds by washed cells[a] of *P. aeruginosa* I-41.

Compounds (10 μg/ml)	Rate of metabolism (%)
Pentachloronitrobenzene	100
2,3,5,6-Tetrachloronitrobenzene	34
2,4,5-Trichloronitrobenzene	51
Pentachloroaniline	46
Pentachlorothioanisole	0
Pentachlorophenol	0
Hexachlorobenzene	0
Tetrachloro-isophthalonitrile	0

a) Bacterial cells of isolate I-41 were grown for 24 hr in MSY medium containing PCNB (1 μg/ml), collected by centrifugation, washed twice with sterilized water and suspended in MSY medium (OD660 = 1.0), and PCNB or its related compounds were added (10 μg/ml). After 3 days incubation, each compound was analyzed.

とから，土壌環境中における PCNB の主要な微生物代謝経路は PCA への還元的変換であると結論付けられた[14,15] (Fig. 2)．

また，I-41 株は PCNB の他，2,3,5,6-テトラクロロニトロベンゼン（2,3,5,6-TeCNB）および 2,4,5-トリクロロニトロベンゼン（2,4,5-TCNB）を各々PCNB の場合の 34%および 51%の効率で代謝し，主な代謝産物として各々2,3,5,6-テトラクロロアニリン (2,3,5,6-TeCA) および 2,4,5-トリクロロアニリン (2,4,5-TCA) を生じたのに加え，PCA を 46%の効率で代謝した．これに対して，関連化合物であるペンタクロロチオアニソール（PCTA），ペンタクロロフェノール（PCP），ヘキサクロロベンゼン（HCB）およびテトラクロロイソフタロニトリル（TPN）を全く代謝しなかった (Table 2)．

I-41 株から無細胞抽出液を調製し，PCNB から PCA への変換が還元酵素反応によることを確認した．過去のニトロアリル還元酵素に関する報告[16]にあるのと同様に，この PCNB を還元代謝する酵素はニトロアリル化合物に対してさほど高い基質特異性を持たないと考えられた．したがって，I-41 株による PCNB の還元反応はこの化合物に特異的であるわけではなくむしろ付随的な反応であるものと推察された．さらに，PCNB の PCA への変換は，糸状菌[5,16]や植物[11,15–18]，動物[5,19]においても観察されており，生物学的に非特異的な反応であると考えられた[11]．

土壌微生物の増殖に及ぼす PCNB および PCA の影響を調べるために，I-41 株を用いて，培地に添加する PCNB の濃度が I-41 株の生育または PCNB の代謝に及ぼす効果および各種濃度の PCA が I-41 株の生育に及ぼす効果を検討した．その結果，培地に添加した 200 μg/ml 以上の PCNB は I-41 株による PCNB の代謝そのものは阻害しないが，I-41 株の増殖を著しく抑制することが明らかとなった (Fig. 3)．また，200 μg/ml 以上の PCA によって I-41 株の

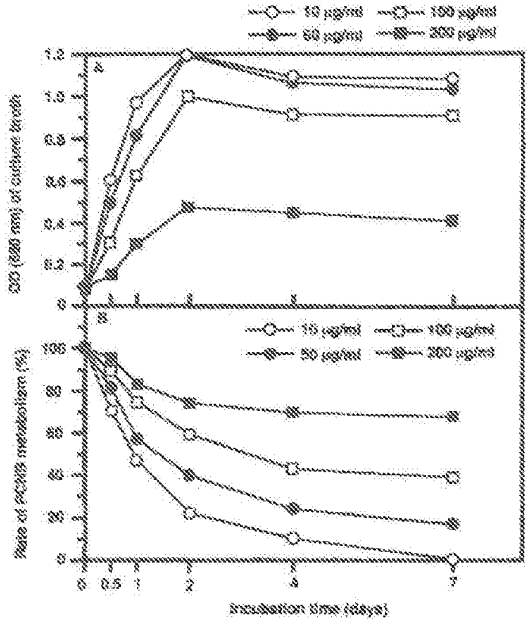


Fig. 3 Effect of PCNB concentrations on growth rate (A) and PCNB metabolism (B) by *P. aeruginosa* I-41.
(A) The optical density (660 nm) of culture broth under still conditions. Initial concentrations of PCNB are indicated.

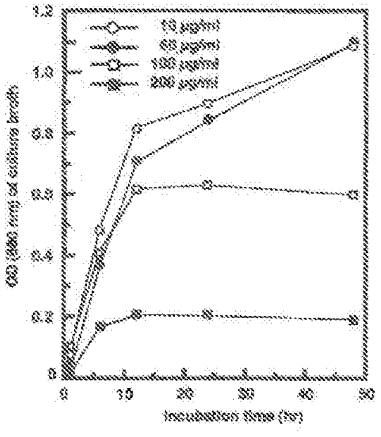


Fig. 4 Effect of PCA concentrations on growth rate of *P. aeruginosa* I-41.
The optical density (660 nm) of culture broth under still conditions. Initial concentrations of PCNB are indicated.

生育は著しく抑制された (Fig. 4)．PCA による I-41 株の生育阻害効果の機構は不明であるが，Ko & Farly[17]は様々な微生物に対して PCA は生育阻害活性を持つことを報告しており，Marthy & Kaufman[20]は嫌気的条件で PCA は土壌中の呼吸作用を強く阻害することを示している．これらの事実から，毎年大量に土壌散布されている圃場では，高濃度の PCNB やその代謝産物である PCA の蓄積が，本来の

COR-AAD1-00001267

防除対照である植物病原菌以外に土壌に生息する微生物の群集密度に何らかの影響を与える可能性が考えられた[32,33]。

一方，PCNB 剤を毎年繰り返し散布している長野県のキャベツ栽培圃場由来の土壌試料から，好気的条件下で PCNB を代謝する *Pseudomonas* 属細菌 N9002 株を単離した．本菌は振とう培養条件で 30 μg/ml の PCNB を 7 日間で 80%以上分解したが，培養液中に PCA の蓄積や予想される代謝産物である PCTA および PCP は検出されなかった．また，本菌は PCNB の他，PCA および PCTA を分解したが，PCP の分解能は認められなかった．本菌から調製した無細胞抽出液は PCNB，PCA および PCTA を代謝したことから，代謝経路に必要な複数の酵素が含まれていることが確認された．N9002 株による PCNB の分解経路に脱塩素反応が含まれるかどうかを検討するために，水銀チオシアネート発色法により培養液中に遊離した塩素イオンの検出を行ったところ，経時的な PCNB の減少に伴って遊離した塩素イオン量の増加が認められた (Fig. 5)．以上の結果から，N9002 株が新規の PCNB 代謝経路を有する可能性が示唆され，その詳細な解明が今後期待される．

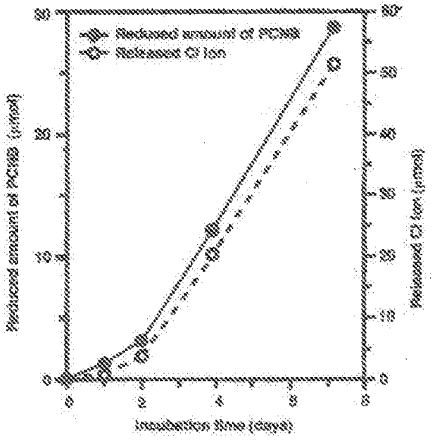


Fig. 5 Degradation of PCNB and release of chloric ion by *Pseudomonas* sp. N9002.
Initial concentration of PCNB was 50 μmol, and residual PCNB was extracted with *n*-hexane and analyzed by GC. The amount of chloric ion in the culture medium at each incubation time was measured by using a mercuric thiocyanate colorimetric method.

2. 2,4-D 分解経路の進化的解析

2,4-D はホルモン系合成除草剤で，主に広葉雑草を防除するために世界的に広く用いられている．本剤は土壌中で速やかに分解されるが，そのような土壌からはこれまでに多くの 2,4-D 分解菌が単離され，それらの系統的分類，代謝遺伝子群の解析や土壌中での個体群動態などの研究が行われてきた[34–38]．中でも，*Ralstonia eutropha* JMP134 株は 2,4-D 分解菌研究のモデルとして最も研究の進んだ細菌の一つで，2,4-D 分解に必要な代謝遺伝子群 (*tfdA*～*tfdF*) および制御遺伝子 (*tfdS*，*tfdR*) を伝達性プラスミド pJP4 上に有することが知られている[39–39]．さらに，このような伝達性プラスミドが 2,4-D 分解遺伝子の水平伝播に関与することが示されている[40]．

2,4-D 代謝の最初のステップを担う *tfdA* 遺伝子は，2,4-D から 2,4-ジクロロフェノールへの変換を触媒するジオキシゲナーゼをコードする遺伝子で[38,39]，その後の代謝経路を担う他の *tfd* 遺伝子群とは異なり，進化的に関連のある遺伝子が他の代謝系からは見い出されていない．また，2,4-D が世に生まれて半世紀余であるが，現在までのところ自然界には 2,4-D に類縁な化合物の存在は知られていない．こうした背景から，2,4-D 分解菌が如何にして代謝経路を獲得してきたかを知ることはとりわけ興味深い．そこで 2,4-D 分解遺伝子の起源を探る目的で，かつて 2,4-D に全く曝されていないと考えられる原始土壌を試料に用いて 2,4-D 分解菌の探索を行い，分離された菌の性状を生理学的および遺伝学的に調査するとともに，2,4-D 代謝遺伝子の伝達性について検討した．

世界各地から採集された天然土壌の中から注意深く 30 試料を選び，まず 2,4-D 分解活性を調査した．そのうち 7 試料に分解活性が認められ，その中で最も強い活性を示したハワイ火山性土壌から分解菌を単離することにした．土壌試料を 2,4-D を含む培地で 10 回以上集積培養した後，2,4-D の分解に伴う pH 指示薬の変色あるいは ^{14}C-2,4-D の取り込みを指標にして分解菌を単離した．採取場所の異なる 3 つの試料からは REP-PCR 法により各々遺伝的背景が異なることが判明した 3 種類の細菌 (HW1，HWK12，

Table 3 2,4-D degrading isolates from Hawaiian pristine soil.

Strain	Origin	Morphology	REP-PCR* Type	*tfdA*	Phylogenetic position based on 16S rDNA
HW1	HW1	Short rod	I	YES	*Variovorax*
HW13	HW13	Long bent rod	II	NO	*Bradyrhizobium* or *Nitrobacter*
HWK12	HWK12	Long bent rod	III	NO	*Bradyrhizobium* or *Nitrobacter*

*REP-PCR: Repetitive Extragenic Palindromic Polymerase Chain Reaction

COR-AAD1-00001268

HW13）が分離された[8]（Table 3）．菌の形態はHW1株が短桿状，HWK12株およびHW13株が湾曲した桿状で，*tfdA*に特異的なプライマーを用いたPCRではHW1株からのみ期待された約300 bpの断片が増幅された[8]．なお，土壌試料を用いた同様のPCRではHW1株が分離された土壌からのみ増幅断片が検出されたことから，土壌を採取した当初から試料中に*tfdA*遺伝子が存在していたことを確認している．これら3種細菌はリボソームRNAの部分塩基配列に基づいて，各々HW1株が*Variovorax*属，HWK12株およびHW13株が*Bradyrhizobium*属あるいは*Nitrobacter*属菌に分類された．さらに，HW1株から増幅された約300 bpの遺伝子断片の塩基配列はpJP4由来の*tfdA*のそれに対して98%の相同性を示したことから，両者がほぼ同一の遺伝子であることが判明した[8]．

次に，HW1株を用いてさらに2,4-D代謝遺伝子の解析を行った．まず，pJP4由来の*tfdA*～*tfdC*の遺伝子プローブを用いてサザン解析を行ったところ，HW1株から抽出したDNAは*tfdA*とはハイブリッドを形成したが，その後の代謝経路を担う*tfdB*および*tfdC*とはハイブリッドを形成しなかった．HW1株から*tfdA*を含む3.0 kbの*Sma*I断片を広宿主域プラスミドベクターにクローン化した後，*R. eutropha* JMP228（pBH501aE）株内で発現させ，2,4-Dの分解能を確認した．なお，JMP228株はpJP4を除去したJMP134株のリファンピシン耐性変異株で，プラスミドpBH501aEはpJP4の*tfdA*遺伝子領域を部分的にカナマイシン耐性遺伝子（*nptII*）カセットで置換することにより*tfdA*を不活化した変異体である[10]．また，HW1株が持つ*tfdA*遺伝子が他の細菌に転移し得るかどうかを明らかにするため，JMP228株またはJMP228（pBH501aE）株を受容菌として用いてプレート交配法により接合実験を行った．各々の受容菌については，JM228株は一連の*tfd*遺伝子群を相補するかどうか，JMP228（pBH501aE）株は*tfdA*遺伝子のみを相補するかどうかを同時に検討するために供試したものである．その結果，後者を受容菌に用いた場合にのみ，1 mMの2,4-Dおよびリファンピシンを含む無機塩培地上で5×10^{-7}，カナマイシンを添加した同培地上では4×10^{-6}の頻度で2,4-D分解能を獲得した接合体が得られた．この結果から，HW1株の*tfdA*遺伝子が伝達性のプラスミドなどを介した接合伝達によりJMP228（pBH501aE）株へ転移したものと考えられた．そこで，常法によりHW1株からプラスミドの検出を試みたが確認できなかったため，パルスフィールドゲル電気泳動を行ったところ，約620 kbの巨大プラスミド様のDNAを有していることが判明した．このプラスミド様DNAはpJP4の*tfdA*プローブとハイブリッドを形成したが，*tfdB*および*tfdC*プローブとはハイブリッドを形成しなかったことから，HW1株は*tfdA*遺伝子のみをプラスミド上に保有していることが明らかになった（Fig. 6）．さらに，2,4-D分解能を獲得した接合体を用いたサザン解析から，HW1株の巨大プラスミドは接合体内で維持されておらず，*tfdA*はpBH501aE上に位置していることが明らかとなった．また，これらのことを実証するために，HW1株からクローン化した*tfdA*を含むプラスミドpBSH1をエレクトロポレーション法でJMP228株（pBH501aE）に導入することによっても，2,4-D代謝能を持つ形質転換体が得られた．これら接合体および形質転換体は*tfdA*をpJP4のそれと全く同位置に持つことが確認された．HW1株の*tfdA*遺伝子ORFの上流と下流領域の塩基配列はpJP4におけるそれらに対して90%以上の相同性を示したことから，*tfdA*が相同組換えによりpBH501aEの*nptII*カセットと置換したものと考えられた．一方，*nptII*をプローブとしたサザン解析は，これら接合体および形質転換体はいずれもpBH501aEと同位置に*nptII*を持つプラスミドを同時に保有していることを示した．以上より，HW1株の遺伝子伝達には巨大プラスミドの接合伝達および引き続き起こるHW1株の*tfdA*遺伝子近傍の配列とpBH501aE中の配列の間での相同的組み換えが関与していることが示唆された．

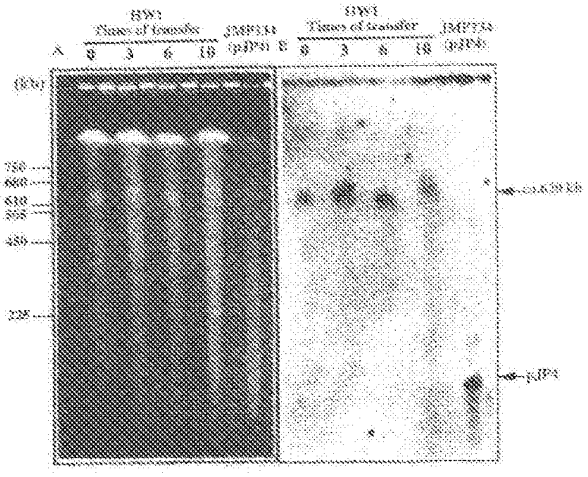


Fig. 6 (A) Detection of plasmid DNA from *Variovorax* sp. HW1 by pulsed field gel electrophoresis (PFGE) and (B) hybridization with *tfdA* probe.
A, The numbers of left side indicate the sizes of *Saccharomyces cerevisiae* DNA used as the molecular standard marker. HW1 culture was transferred for each indicated time, and was subjected to be analyzed by PFGE. The range of DNA separation was 100 kb~1.2 Mb. Hybridization was carried out under high stringency conditions.

かつて2,4-Dに曝されたことのないと考えられるハワイ大山性土壌から分離されたHW1株が，如何なる目的で2,4-D分解遺伝子である*tfdA*を有していたかは現在のところ全く不明であるが，2,4-Dが使用される以前からその分解遺伝子が既に自然界に存在していたことは確かであろ

COR-AAD1-00001269

う．さらに，原始土壌に生息する細菌が持つ *tfdA* が他の菌に伝達され得るという事実は，HW1 株のような細菌の 2,4-D 代謝遺伝子が現在圃場等に広く拡散し，これまでに多数分離されている 2,4-D 分解菌のそれの起源である可能性を強く示唆するものと考えられる[7]．

おわりに

実際に，圃場に散布された薬剤が土壌中で如何なる運命を辿るのか，あるいは残留農薬が土壌微生物の群集密度にどのような影響を及ぼすかを正確に把握するためには，圃場レベルでの緻密な実験検討が必要であるが，本研究で得られた結果は，今後少なくとも PCNB 剤の土壌微生物による代謝分解の全容解明を進める上で重要な指針を与えるものと考えられる．これまでの過去の研究成果により，2,4-D 分解微生物やそれらの代謝遺伝子に関して体系的な理解が可能になった観があるが，2,4-D に曝された歴史を持たない原始土壌にその分解菌および代謝遺伝子の起源を求めるという本研究での新たな切り口により，微生物進化の足跡を垣間見る機会が得られたものと思われる．これら PCNB の微生物分解機構に対する理解や，2,4-D 代謝経路の起源や伝播性に関する知見は，微生物等の働きを利用して難分解性化学物質を分解処理するバイオリメディエーション技術の応用開発に向けた学術的基盤の整備に少なからず資するものと期待される．

本研究を遂行するにあたり，理化学研究所主任研究員 山口勇先生には，農薬の分解微生物などに関する貴重なご指導，ご鞭撻を賜りました．また，同主任研究員 工藤俊章先生，東京工業大学名誉教授 磯越弘毅先生，長岡技術科学大学教授 福田雅夫先生，同教授 故矢野圭二先生，ミシガン州立大学教授 J. M. Tiedje 先生，同助教授 L. J. Forney 先生には，微生物学などに関する様々なご指導，ご教示を賜りました．衷心より深く感謝申し上げます．也通産省工業技術院主任研究員 鎌形洋一先生，海洋科学センター主任研究員 高見英人先生，現明治大学講師 前田晋久先生，現ペーデュー大学助教授 C. Nakatsu 先生，現トロント大学助教授 R. A. Fulthorpe 先生には共同研究を通して様々なご教示を頂きました．理化学研究所微生物制御研究室研究員 長谷川志氏には実際の研究に携わって頂きました．その他にも様々な方々から，ご援助ならびにご協力をして頂きました．厚く御礼申し上げます．

本研究は，旧科学技術庁新技術事業団微生物進化プロジェクトチームおよびミシガン州立大学微生物生態センターの共同研究プロジェクトとして行われたものです．ここに記して謝意の意を表します．

引用文献

1) W. H. Ko & J. D. Farley: *Phytopathology* **59**, 64 (1969)
2) T. Nakanishi: *Ann. Phytopathol. Soc. Jpn.* **38**, 249 (1972)
3) Y. Kamagata, R. R. Fulthorpe, K. Tamura, H. Takami, L. J. Forney & J. M. Tiedje: *Appl. Environ. Microbiol.* **63**, 2266 (1997)
4) T. Nakanishi & H. Oku: *Ann. Phytopathol. Soc. Jpn.* **36**, 67 (1970)
5) K. Suyama & H. Okazaki: Abstr., 19th Annual Meeting of Pestic. Sci. Soc. Jpn. D111, 1994 (in Japanese)
6) C. I. Chacko, J. L. Lockwood & M. Zabik: *Science* **154**, 893 (1966)
7) T. Nakanishi & H. Oku: *Ann. Phytopathol. Soc. Jpn.* **35**, 339 (1969)
8) T. Nakanishi & H. Oku: *Phytopathology* **59**, 1761(1969)
9) N. R. Krieg & J. G. Holt: "Bergey's Manual of Systematic Bacteriology," Vol. 1, 8th Ed., The Williams &Wilkins Company, London, p. 164, 1984
10) N. B. Murthy & D. D. Kaufman: *J. Agric. Food Chem.* **26**, 1151 (1978)
11) M. Ogiso & H. Tanabe: *J. Pesticide Sci.* **7**, 391 (1982)
12) K. Tamura, Y. Hasegawa, T. Kudo & I. Yamaguchi: *J. Pesticide Sci.* **20**, 145 (1995)
13) K. Tamura, Y. Hasegawa, T. Kudo & I. Yamaguchi: "Microbial Diversity and Genetics of Biodegradation," K. Horikoshi *et al.* Ed., Japan Scientific Societies Press, Tokyo, p.113, 1997
14) I. Yamashita: *Biochem. J.* **27**, 88 (1954)
15) S. Gorbach & U. Wagner: *J. Agric. Food Chem.* **15**, 654 (1967)
16) G. L. Lamoureux, J.-M. Gouot, D. G. Davis & D. G. Rusness: *J. Agric. Food Chem.* **29**, 996 (1981)
17) D. G. Rusness & G. L. Lamoureux: *J. Agric. Food Chem.* **28**, 1070(1980)
18) P. W. Aschbacher & V. J. Feil: *J. Agric. Food Chem.* **31**, 1150 (1983)
19) E. J. Kuchar, F. O. Geenty, W. P. Griffith & R. J. Thomas: *J. Agric. Food Chem.* **17**, 1237 (1969)
20) M. A. Bhat, M. Tsuda, K. Horiecke, M. Nozaki, C. S. Vaidyanathan & T. Nakazawa: *Appl. Environ. Microbiol.* **60**, 307 (1994)
21) G. R. Chaudry & G. H. Huang: *J. Bacteriol.* **170**, 3897 (1988)
22) J. O. Ka & J. M. Tiedje: *J. Bacteriol.* **145**, 1365 (1994)
23) J. O. Ka, W. E. Holben & J. M. Tiedje: *Appl. Environ. Microbiol.* **60**, 1106 (1994)
24) J. O. Ka, W. E. Holben & J. M. Tiedje: *Appl. Environ. Microbiol.* **60**, 1121 (1994)
25) A. A. Mač, R. O. Marits, N. R. Ausmees, V. M. Kõiv & A. L. Heinaru: *J. Gen. Microbiol.* **139**, 3165 (1993)
26) P. S. Amy, J. W. Schulke, L. M. Frazier & R. J. Seidler: *Appl. Environ. Microbiol.* **49**, 1237 (1985)
27) R. H. Don & J. M. Pemberton: *J. Bacteriol.* **161**, 466 (1985)
28) R. H. Don, A. J. Weightman, H.-J. Knackmuss & K. N. Timmis: *J. Bacteriol.* **145**, 681 (1985)
29) B. Kaphammer & R. H. Olsen: *J. Bacteriol.* **172**, 5856 (1990)
30) B. Kaphammer B., J. J. Kukor & R. H. Olsen: *J. Bacteriol.* **172**, 2280 (1990)
31) R. H. Don & J. M. Pemberton: *J. Bacteriol.* **145**, 681 (1981)
32) F. Fukumori & R. P. Hausinger: *J. Bacteriol.* **175**, 2083 (1993)
33) F. Fukumori & R. P. Hausinger: *J. Biol. Chem.* **268**, 24311 (1993)
34) W. E. Holben, B. M. Schroeter, V. G. Calabrese, R. H. Olsen, J. K. Kukor, V. O. Biederbeck, A. E. Smith & J. M. Tiedje: *Appl. Environ. Microbiol.* **58**, 3941 (1992)

COR-AAD1-00001270

Society Awards 2001

(on high prospectiveness)

Microbial Pesticide Degradations and Evolutionary Analysis of Degrading Enzymes

Introduction

Although the major metabolic pathways of microorganisms evolved millions of years ago, molecular evolution has not ceased. Developments in the chemical industries over the past several decades have resulted in the synthesis of numerous artificial compounds and their release into the natural environment. In agriculture, for example, halogenated organic compounds have been applied in abundance to cultivated land as pesticides in order to obtain stability in crop yields. The introduction of xenobiotic chemicals to nature has created opportunities for the evolution of metabolic pathways that allow microorganisms either to exploit new carbon sources or to detoxify harmful compounds.

The pesticide pentachloronitrobenzene (PCNB) has been widely used as a fungicide, especially for the control of clubroot disease of cruciferous vegetables. This compound is physicochemically stable and generally resistant to microbial degradation, and it is thought to remain in soil for many years after its application. Given that large quantities of this fungicide have been repeatedly applied to cultivated fields during the last few decades, residual PCNB or its metabolites may affect nontarget soil microorganisms and result in ecological hazards. In addition, the synthetic herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) has also been applied to crops over the past 5 decades for the control of broad-leaf weeds. Several microorganisms that are able to degrade 2,4-D have been isolated from soil with a history of exposure to this herbicide, and the catabolic genes of these isolates have been extensively investigated. Among the genes that contribute to 2,4-D catabolism, *tfdA*, which encodes an enzyme responsible for the first step of this process, is relatively specific to the 2,4-D degradation pathway. Given that no naturally occurring analogs of 2,4-D are known, it is of interest to determine the mechanism by which the ability to degrade 2,4-D developed in soil microorganisms.

In this study, the microbial degradation of PCNB in soil environments and the molecular evolution of the 2,4-D metabolic pathway were examined.

Distribution and Characterization of PCNB-degrading Microorganisms

Although several studies of the degradation of PCNB in soil have implicated soil microorganisms in this process, few such studies have described the characteristics of the PCNB-degrading microbes. We therefore screened for the presence of PCNB degraders in soil samples that were collected not only from cultivated fields exposed to PCNB but also from noncultivated areas including mountainous terrain, volcanic zones, and beach regions as well as industrial activated sludge. Each soil sample (1 g) was placed in 5 ml of mineral salt medium containing 0.1% yeast extract and PCNB (20 μg/ml), and was incubated for 7 days at 28°C with or without shaking. Under nonagitated conditions, 92 out of a total of 114 samples mediated an extent of PCNB degradation of > 20%, suggesting that PCNB-degrading microorganisms were present in most of the soil environments irrespective of the previous application of PCNB. From a soil sample with a high PCNB degradation rate, a bacterial strain (I-41) was isolated as the most efficient degrader of PCNB and identified as *Pseudomonas aeruginosa*. Incubation of I-41 under nonagitated conditions with PCNB at a concentration of 10 μg/ml resulted in the degradation of almost all of the pesticide within 1 week; in contrast, only 18% of the PCNB was degraded under shaking conditions. Characterization of the time course of PCNB degradation by this bacterium revealed that bacterial growth achieved stationary phase within 2 days and that the breakdown of PCNB correlated with the increase in the bacterial population. In addition, pentachloroaniline (PCA) was detected in most soil samples that mediated PCNB degradation, and the principal metabolite of PCNB in I-41 was confirmed to be PCA. Together with previous observations indicating that PCA is the major metabolite of PCNB in soil samples from fields repeatedly subjected to PCNB application, our results suggest that the main pathway for PCNB degradation in soil environments is the reductive conversion of PCNB to PCA.

I-41 mediated the reduction not only of PCNB to PCA but also of tetrachloronitrobenzen (TeCNB) to tetrachloroaniline (TeCA) and of trichloronitrobenzen (TCNB) to trichloroaniline (TCA). The observation that a cell-

COR-AAD1-00001271

free extract of I-41 also mediated the reductive degradation of PCNB demonstrated that this process is the result of an enzyme-catalyzed reaction. The enzyme system responsible for the reduction of PCNB thus appears to possess a relatively low substrate specificity among nitroaryl compounds. Indeed, nitroaryl reductase systems have been shown to exhibit low substrate specificity. The reduction of PCNB by I-41 might thus be an incidental property of this bacterium rather than reflect the existence of a dedicated PCNB degradation pathway. Moreover, the degradation of PCNB to PCA has been observed in some fungi, plants, and animals, suggesting that catalysis of this reaction is a nonspecific biological property.

PCNB is a potent fungicide and is especially effective against *Plasmodiophora brassicae* and *Rhizoctonia solani*. However, PCNB at concentrations of >200 μg/ml also markedly inhibited the growth of I-41, as did PCA at 200 μg/ml. These observations suggest that high concentrations of PCNB or PCA might affect the population of soil microorganisms in addition to the target plant pathogens in fields repeatedly exposed to PCNB in large amounts.

Another PCNB-degrading bacterium, designated *Pseudomonas* sp. N9002, was isolated from a cabbage field in which PCNB had been applied annually for at least 5 years. This organism mediated the degradation of virtually all of the fungicide within 1 week when incubated under shaking conditions with PCNB at a concentration of 20 μg/ml. Neither PCA, pentachlorothioanisol (PCTA), nor pentachlorophenol (PCP) was detected during the degradation of PCNB by N9002, whereas the release of free chlorine ions was observed with the use of the colorimetric mercury thiocyanate method. These results suggest that N9002 degrades PCNB through an unknown mechanism that includes dechlorination.

Molecular Evolution of a 2,4-D Metabolic Pathway

Several microorganisms able to degrade 2,4-D have been isolated to date from soils with a history of exposure to this herbicide, and their catabolic genes have been extensively investigated. *Ralstonia eutropha* JMP134 is one of the best characterized 2,4-D degraders and has been shown to harbor a plasmid, pJP4, on which the catabolic genes (*tfdA* through *tfdF*) and corresponding regulatory genes (*tfdR* and *tfdS*) responsible for 2,4-D degradation are located. The *tfdA* gene encodes α-ketoglutarate-dependent 2,4-D dioxygenase (TfdA), which catalyzes the first step of 2,4-D metabolism. Unlike the other *tfd* catabolic genes, *tfdA* is relatively specific to 2,4-D degradation, and other sequences evolutionarily related to *tfdA* have not been detected. Therefore, little is known of the origin and essential function of *tfdA* in natural environments.

Recently, 2,4-D-degrading activity has been detected in the soil of several pristine ecosystems that had not been exposed to 2,4-D or to other halogenated organic compounds. With the use of 2,4-D dye indicator medium or ^{14}C-labeled 2,4-D medium, a previously unidentified group of oligotrophic bacteria capable of degrading 2,4-D was thus isolated from enrichment cultures of volcanic soil derived from geographically remote regions of Hawaii. Phylogenetic analysis based on 16S rDNA sequencing and characterization of the physiological properties of these bacteria revealed that isolate HW1 belongs to the genus *Variovorax* and that isolates HWK12 and HW13 possess 16S rDNA sequences that exhibit a high degree of similarity to those of *Bradyrhizobium* species. All of the isolates grew slowly on 2,4-D medium, with mean generation times of 16 to 30 hr, which are substantially longer than those of previously identified 2,4-D degraders. Polymerase chain reaction-mediated amplification and Southern hybridization analysis revealed that HW1 harbors a *tfdA* gene almost identical to that of pJP4 but does not possess genes homologous to *tfdB* or to *tfdC*. Cloning of a 3.0-kb *Sma*I fragment containing *tfdA* of HW1 in a broad-host range plasmid vector and introduction of the resulting plasmid into *R. eutropha* JMP228 (pBH501aE) resulted in complementation of the inability of this strain to degrade 2,4-D; JMP228 is a rifampicin-resistant derivative of JMP134 that has been cured of pJP4, and pBH501aE is a *tfdA*$^-$ derivative of pJP4 in which the *tfdA* gene is disrupted by insertion of a kanamycin resistance gene cassette. Furthermore, the 2,4-D degradation phenotype (or, at least, TfdA activity) of HW1 was shown to be transmissible to JMP228 (pBH501aE), suggesting that the 2,4-D catabolic gene (or genes) is located on a conjugative plasmid like pJP4. Thus, HW1 appeared an appropriate bacterial strain for studies of the genetic origin of the 2,4-D degradation pathway and the assembly of 2,4-D catabolic genes.

Paised-field gel electrophoresis revealed that HW1 harbors a large plasmid (~620 kb), on which *tfdA* is located. Experimental mating of HW1 and JMP228 strains showed that transconjugants with the ability to degrade 2,4-D were obtained (at a frequency of 5×10^{-7}) only with JMP228 (pBH501aE) as the recipient. Hybridization analysis demonstrated that the large plasmid was not maintained in the transconjugants and that the *tfdA* gene of these cells was located on pBH501aE. Electroporation of the cloned *tfdA* gene of HW1 into JMP228 (pBH501aE) also resulted in the generation of transformants with the ability to degrade 2,4-D. Analysis of both transconjugants and transformants confirmed that the *tfdA* gene was located on the pBH501aE at the position corresponding to that of *tfdA* in pJP4. These results suggest that the HW1 plasmid is conjugatively

COR-AAD1-00001272

transferred to JMP228 (pBH501aE) and that the *tfdA* gene of HW1 subsequently undergoes homologous recombination with pBH501aE. Furthermore, the presence in a pristine environment of a 2,4-D-degrading bacterium that harbors the canonical *tfdA* gene on a transmissible plasmid indicates that *tfdA* in soil microorganisms such as HW1 might be the origin of the gene in 2,4-D degraders present in 2,4-D-treated cultivated fields, although the intrinsic function of the gene in untreated environments remains to be determined.

Perspectives

Although the behavior of PCNB in nature and the effects of residual pesticide on the population dynamics of soil microorganisms require further investigation in different types of fields, our results provide important insight into the microbial degradation of PCNB. Previous studies of the microbial degradation of 2,4-D have resulted in the identification of various 2,4-D degraders and their catabolic genes. Our strategy of characterizing 2,4-D-degrading microorganisms isolated from pristine environments may provide a rare opportunity to define recent events in the molecular evolution of new metabolic pathways for pesticide degradation. Furthermore, progress in our understanding both of the mechanisms of microbial degradation of PCNB in soil environments as well as of the molecular evolution of 2,4-D metabolic pathways should eventually lead to the development of bioremediation protocols for the treatment of soil or groundwater that is contaminated with these chlorinated aromatic compounds.

COR-AAD1-00001273

Biotechnology Letters, Vol 20, No 11, November 1998, pp. 1073–1076

Isolation and characterization of a stable 2,4-dichlorophenoxyacetic acid degrading bacterium, *Variovorax paradoxus*, using chemostat culture

Tatiana Vallaeys[1*], Lionel Albino[1], Guy Soulas[1], Alice D. Wright[2] and Andrew J. Weightman[3]

(1*) Institut National de la Recherche Agronomique. Microbiologie des Sols. BV1540. 21034 Dijon France
(2) Center for Microbial Ecology, Michigan State University, East-Lansing, Michigan U.S.A.
(3) School of Pure and Applied Biology. University of Wales, P.O. Box 915. Cardiff CF1 3TL Wales UK

A strain of *Variovorax paradoxus* degrading 2,4-dichlorophenoxyacetic acid (2,4-D) was isolated from the Dijon area (France) using continuous chemostat culture. This strain, designated TV1, grew on up to 5 mM 2,4-D and efficiently degraded the herbicide as sole carbon source as well as in presence of soil extracts. It also degraded phenol and 2-methyl, 4-chlorophenoxyacetic acid at 3 mM and 2,4-dichlorophenol at 1 mM. This organism contained a stable 200 kb plasmid, designated pTV1, which showed no similarity in its restriction pattern with the archetypal 2,4-D catabolic plasmid pJP4. However, pTV1 contained an 11 kb *Bam*HI fragment which hybridized at low stringency with the 2,4-D degradative genes *tfdA*, *tfdB* and *tfdR* from pJP4. PTV1 partial *tfdA* sequence showed 77 % similarity with the archetypal *tfdA* gene sequence from *Ralstonia eutropha* JMP134. Tn*5* mutagenesis confirmed the involvement of this gene in the 2,4-D catabolic pathway.

Keywords: *Variovorax paradoxus*, *tfdA*, α-ketoglutarate dioxygenase, 2,4-D

Introduction

The herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) has been widely used since 1945 to control dicotyledoneous weeds and is still used today as part of complex formulations. A number of bacterial strains able to degrade 2,4-D have been isolated (Don and Pemberton, 1981; Fulthorpe *et al.*, 1995; Suwa *et al.*, 1996), some of which have been genetically characterized. In most cases, the *tfd* genes have been localized on transmissible plasmids (Don and Pemberton, 1981; Mae *et al.*, 1993, Ka and Tiedje, 1994), some of which have been found unstable in their host strain (Mae *et al.*, 1993; Ka and Tiedje, 1994). Among these plasmids, the pJP4, isolated from the 2,4-D degrading bacterium *Ralstonia eutropha* (formerly *Alcaligenes eutrophus*) JMP134, has been extensively studied. The 2,4-D catabolic (*tfdABCDEF*) genes carried by this plasmid have been characterized and sequenced (Streber *et al.*, 1987; Perkins *et al.*, 1990). In this paper, we describe the isolation of a *Variovorax paradoxus* strain by chemostat enrichment and the characterization of its stable plasmid, pTV1.

Material and methods

Bacterial strains and media

Strains used in this study are listed in Table 1. Strain identification, based on phenotypic properties was performed by the Pasteur Institute (Paris, France). Loos' (1976) medium was used to grow 2,4-D degrading strains. Luria Broth agar (LB) (Sambrook *et al.*, 1989) was used to grow *E. coli* strains.

Culture in chemostat

Soil from the Dijon (France) INRA experimental station was used to prepare soil extract solution (Girard and Rougieux, 1967). The carbon concentration in the soil extract solution was determined using a Dohrmann DC 80 carbon analyzer and adjusted to 50 mg per liter with distilled water. Twenty ml KH_2PO_4/ Na_2HPO_4 buffer pH 6.8 and 1 ml of a mineral base (0.028 g l^{-1} $FeSO_4$ 7(H_20); 2 μg l^{-1} H_3BO_3; 1.8 μg l^{-1} $MnSO_4$; 0.2 μg l^{-1} $ZnSO_4$; 0.08 μg l^{-1} $CuSO_4$; 0.01 μg l^{-1} $NaMoO_4$) were added to a total 1 l of soil extract medium. The carbon/nitrogen ratio of the medium was then adjusted to 5:1 with NH_4NO_3. A 2-liter chemostat (Biolaffite) was inoculated with a 100 mg soil aggregate of the same soil. Samples were collected every 4 hours from the chemostat and 2,4-D concentration was determined from it absorbancy at 283 nm. After 4 months continuous culture, significant 2,4-D degradation was observed in the chemostat. In order to simplify the diversity of the chemostat microbial community, the carbon derived from the stock soil extract solu-

© 1998 Chapman & Hall

COR-AAD1-00001274

Table 1 Bacterial strains and plasmids used in this study.

Strain		Plasmid	Reference
Ralstonia eutropha	JMP134	pJP4	Don and Pemberton (1981)
	JMP222	–	Don and Pemberton (1981)
Variovorax paradoxus	TV1	pTV1	this study
	TV1	pTV1::Tn5	this study
Escherichia coli	S17-1::Tn5	–	Galimand *et al.* (1988)
	S17–1::Tn5	pTV5	this study
	C600	–	Appleyard (1954)
	C2013	pRK2013	Figurski and Helinski (1979)
	C600	pBR325–*mob*	N.C. McClure, unpublished
	C600	pTV5::Tn5	this study
	DHα	pUC::*tfdA*	Holben *et al.* (1992)
	DHα	pUC::*tfdB*	Holben *et al.* (1992)
	DHα	pUC::*tfdC*	Holben *et al.* (1992)
	DHα	pUC::*tfdD*	Holben *et al.* (1992)
	DHα	pUC::*tfdE*	Holben *et al.* (1992)
	DHα	pUC::*tfdF*	Holben *et al.* (1992)
	JM107	pKT231::*tfdR*	N.C. McClure, unpublished
Pseudomonas putida	UWC1	pD10	McClure *et al.* (1989)

tion was progressively replaced by 2,4-D up to 200 mg l^{-1} and suspension-dilutions of the culture medium were plated onto selective Loos medium. Positive isolates were tested for its ability to release $^{14}CO_2$ from ^{14}C- 2,4-D, labeled either on the side chain or on the ring (Vallaeys *et al.*, 1997). Ability to degrade alternative substrates (acetate, malate, succinate, phenol, benzoate, 3-chlorobenzoate, 2,4-dichlorobenzoate, 2,4-dichloropenol, m-fluorobenzoate, 2,4,5-T and MCPA) was tested using batch culture. Stability of the 2,4-D degrading potential of the strain was checked after 30 subculture cycles on LB medium by plating the strain on Loos (1976) medium.

Molecular methods, transposon Tn5 mutagenesis and *tfd* marker exchange by homologous recombination

Plasmids were extracted according to Wheatcroft and Williams (1981) (pJP4, pTV1) and Sambrook *et al.* (1989). Endonuclease restrictions and electrophoresis were carried out using standard method (Sambrook *et al.*, 1989). Southern transfer to nylon membranes (Hybond N, Amersham) was done with the Vacugene 2016 system (Pharmacia). Homologies between plasmids were investigated by Southern hybridization using the whole plasmid pJP4; the *Sst*I C fragment of pJP4 carrying the *tfdBCDEF* genes (Mc Clure *et al.*, 1989); the *Sst*I fragment carrying *tfdR* (Mc Clure, unpublished) and the *tfd* gene probes (Holben *et al.*, 1992). Hybridizations were performed according to Fulthorpe *et al.* (1995). Cloning of the pTV1 *Bam*HI fragment carrying *tfd*-like genes into the plasmid pBR325-mob (N.C. McClure, unpublished) and transformation of *E. coli* S17.1::Tn5 strain with the resulting plasmid pTV5 were performed using standard methods (Sambrook *et al.*, 1989). Transformant *E. coli* strains were selected on LB agar supplemented with ampicillin (50 mg l^{-1}) and kanamycin (50 mg l^{-1}). *E. coli* strain S17.1::Tn5 (pTV5) was used as donor in membrane matings with the recipient *E. coli* C600. LB agar supplemented with 50 mg l^{-1} kanamycin and rifampicin was used for the selection of recipients *E. coli* C600 (pTV5::Tn5). Transposon Tn5 insertions were localized by restriction endonuclease digestions. *E. coli* C600 (pTV5::Tn5) strains containing Tn5 inserted in the pTV1 *Bam*HI fragment were used as donors in a triparental matings with *E. coli* C600 carrying the helper plasmid pRK2013 (Figurski and Helinski, 1979) and recipient *V. paradoxus* TV1. As Tn5 encoded streptomycin resistance is expressed in non-enterobacteria but not in *E. coli* (Selvaraj and Iyer, 1984), Loos (1976) medium supplemented with 50 mg l^{-1} kanamycin and 100 mg l^{-1} streptomycin was used to select *V. paradoxus* TV1 (pTV1::Tn5) recombinants. PCR amplification of the *tfdA*-like gene was done as described by Vallaeys *et al.* (1996). Amplified fragment was purified using the Gene Clean kit (BIO 101, La Jolla). Sequencing of the purified product was done using the Applied Biosystems automatic sequencer (Perkin Elmer, Cetus). Alignments of the partial *tfdA*-like gene sequences were obtained using the CLUSTAL V software (Higgins and Sharp, 1988) after computer-assisted homology search on GenBank and EMBL databases.

Results

2,4-D degrading potential of the strain and radioisotopic 2,4-D assays

After 4 months of continuous-flow culture with 2,4-D as the sole carbon source, a strain of *Variovorax paradoxus* was isolated, which gave a positive reaction on Loos medium,

COR-AAD1-00001275

and rapidly converted both ring and chain labeled ^{14}C-2,4-D to $^{14}CO_2$. This strain, designated TV1, was still able to degrade 2,4-D after 30 subculture cycles on non-selective media. It was able to grow on up to 5 mM 2,4-D as sole carbon source, as well as in the presence of soil or yeast extracts, on phenol and 2-methyl, 4-chlorophenoxyacetic acid (3 mM) and on 2,4-dichlorophenol (1 mM).

Localization and cloning of the 2,4-D degrading gene of *V. paradoxus* TV1

V. paradoxus TV1 was shown to contain a plasmid, designated pTV1, which was about 200 kb in size and had a different restriction pattern from plasmid pJP4 (Fig.1).

An 11 kb-*Bam*HI fragment of pTV1 showed homology to a fragment of pJP4 carrying the genes *tfdA*, *tfdB*, and *tfdC* and to the *tfdR* gene probe (results not shown). This *Bam*HI fragment was cloned using plasmid pBR325-*mob* to give pTV5. After Tn*5*-mutagenesis of the *tfd*-like region in pTV5, marker exchange by homologous recombination was used to insert *tfd*::Tn*5* mutations into plasmid pTV1. White 2,4-D$^-$ mutants were detected on Loos medium. These were assayed for *tfdA* activity and some were found not to release $^{14}CO_2$ from ^{14}C-2,4-D indicating Tn*5* insertional inactivation of a *tfdA*-like gene. The presence of a *tfdA*-like gene on the cloned *Bam*HI fragment from pTV1 was confirmed by PCR. The sequence of the amplified fragment (Genbank accession number U65531) showed 77% similarity to the canonical *tfdA* gene from *R. eutropha* JMP134, 93% to the *tfdA*-like gene from *Burkholderia* sp. RASC and 99% to the *tfdA*-like gene from *Halomonas* sp. Amino-acid sequence alignment confirmed the existence of 3 types of *tfdA*-like genes (Fig.2).

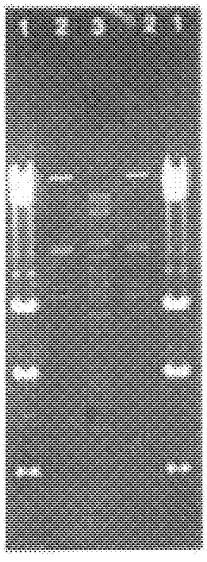


Figure 1 The 2 plasmids pJP4 from *R. eutropha* JMP134 and pTV1 from *V. paradoxus* showed distinct sizes and restriction patterns. Lanes contains as follows: (1)lambda/*Hin*dIII, (2): pJP4/*Eco*RI, (3): pTV1/*Eco*RI.

Discussion

This work shows the existence of a novel stable 2,4-D degradative plasmid, pTV1, carrying part or all of the *tfd*-like genes on an 11 kb *Bam*HI fragment, and showing low sequence similarity to isofunctional genes of *R. eutropha* JMP134. This plasmid provides its host strain, *Variovorax paradoxus* TV1 with the ability to efficiently degrade 2,4-D when the herbicide constitutes the only source of carbon and energy as well as in presence of either yeast- or soil-

```
          10         20         30         40         50         60
1   YDALPRDLQSELEGLRAEHYALNSRFLLGDTDYSEAQRNAMPPVNWPLVRTHAGSGRKFL
2   ..D..E.FKK..Q.........H...I....E...S.........S...I...........
3   ..D..E.FKK..Q.............I.........S.........S...............
4   ..D..E.FKK..Q.............I.........S.........S...............

C   YD+LP+D+++EL+GLRAEHYAL+SRF+LGDT+YSE+QRNAMPPV+WPL+RTHAGSGRKFL

          70         80         90        100
1   FIGAHASHVEGLPVAEGRMLLAELLEHATQREFVYRHRWNVGDLVM
2   ........I..R.....................PK.....S.K......
3   ........I..R.....................K.....S.K......
4   ......G.I..R.....................K.....S.K......

C   FIGAHA:H+EG+PVAEGRMLLAELLEHATQ++FVYRH+W+VGDLVM
```

Figure 2 Amino-acid sequence similarity between partial amplified sequences of the *tfdA* gene from (1) *R. eutropha* JMP134 (Streber *et al.*, 1987) and the *tfdA*-like gene from (2) *Burkholderia* sp. RASC (Suwa *et al.*, 1996), (3) *Halomonas* sp. I.18 (Maltseva *et al.*, 1996) and (4) *V. paradoxus* TV1 (U65531). (C) consensus. Dots indicate identity.

COR-AAD1-00001276

extract, both in pure and mixed culture conditions. Finally, this plasmid harbors a different type of *tfdA* gene than the one classically reported in *R. eutropha* JMP134 and various other 2,4-D degrading isolates. Interestingly, the sequence data reported here confirm a similarity of >99% between the *tfdA*-like genes from *V. paradoxus* strain TV1, *Rhodoferax fermentans* strains tfd31 and B6–9 (β-group of Proteobacteria) (Fulthorpe *et al.*, 1995), but also *Halomonas* sp. strain I-18 (γ-group of Proteobacteria) (Maltseva *et al.*, 1996), which have been isolated from other continents, using batch enrichment. Top *et al.* (1996) reported the presence of a pTV1 type *tfdA* gene on a plasmid encoding only the α-ketoglutarate dioxygenase and lacking the other genes of the 2,4-D catabolic pathway. Although pTV1 failed to transfer to various recipients (unpublished results), such results give evidence of catabolic gene transfer between different bacterial genera.

Although 2,4-D is an analogue of auxine, it does not occur naturally in the environment, unlike the intermediates generated throughout it degradation pathway. In addition, the α-ketoglutarate dependent dioxygenase encoded by the *tfdA* gene seems specific to the 2,4-D catabolic pathway: firstly, Pieper *et al* (1989) showed that this enzyme could not be induced by unsubstituted phenoxyacetates, secondly the *tfdA* sequence does not show significant similarity to any other known sequences (Fukumori and Hausinger, 1993). However, the *tfdA* gene sequences show significant divergence suggesting an ancient origin of the gene and possibly, substrate specialization previously to the use of the herbicide 2,4-D. Different efficiencies and stabilities of 2,4-D degrading capabilities have been reported (Vallaeys *et al.*, 1997). These could be linked to culture conditions, to strain species (Maltseva *et al.*, 1996) but also to location of the *tfd* genes on an unstable plasmid (Mae *et al.*, 1993) or to the *tfdA* gene type.

The study of the origin and phylogenies of degradative genes with respect to the organisms which contain them is a pre-requisite to a better understanding of mechanisms involved in the evolution and dissemination of xenobiotic-degradative capabilities amongst bacteria and also, to the optimization of degradative properties for bioremediation purposes.

Acknowledgments

We thank R. Fulthorpe for Advice and L. Courde for assistance with sequence alignment.

References

Appleyard, RK (1954). *Genetics* 39: 440

Don, RH and Pemberton, JM (1981). *J. Bacteriol.* 145: 681–686

Figurski, DH and Helinski, DR(1979). *Proc. Natl. Acad. Sci. USA* 76: 1648–1652

Fukumori, F and Hausinger, RP (1993). *J. Bacteriol.* 175: 2083–2086

Fulthorpe, R, McGowan, C, Maltseva, O, Holben, W and Tiedje, JM (1995). *Appl. Environ. Microbiol.* 61 :3274–3281

Galimand, M, Vieille, C, Perroud, B, Onyeocha, I and Elmerich, C (1988). Advances in genetics of *Azospirillum brasilense* Sp7: use of Tn5 mutagenesis for gene mapping and identification. In: *Azospirillum IV: Genetics, physiology, ecology*, W. Kling-müller ed pp 1–7, Berlin: Springer Verlag

Girard, H and Rougieux, R (1967). *Techniques de microbiologie agricole.* Paris: Dunod

Higgins, DG and Sharp, PM (1988). *Gene* 73: 237–244

Holben, WE, Schroeter, BM, Calabrese, VGM, Olsen, RH, Kukor, JK, Biederbeck, VO, Smith, AE, and Tiedje, JM (1992). *Appl. Microbiol. Environ.* 58:3941–3948.

Ka, JO and Tiedje, JM (1994). *J. Bacteriol.* 176: 5284–5289.

Loos, MA (1976). *Can. J. Microbiol.* 21: 104–107.

Mae, AA, Martis, RO, Ausmees, NR, Koiv, VM and Heinaru, AL (1993). *J. Gen. Microbiol.* 139: 3165–3170.

Maltseva, OV, McGowan, C, Fulthorpe, R and Oriel, P (1996). *Microbiology.* 142: 1115–1122.

Mc Clure, NC, Weightman, AJ and Fry, JC (1989). *Appl. Environ. Microbiol.* 55: 2627–2634.

Perkins, EJ, Gordon, MP, Caceres, O and Lurquin, PF (1990). *J. Bacteriol.* 172: 2351–2359.

Pieper, DH, Reineke, W, Engesser, K and Knackmuss, HJ (1989). *Arch. Microbiol.* 151:365–371.

Sambrook, J, Fritsch, EF and Maniatis, T (1989). *Molecular cloning, a laboratory manual*, 2nd edn, New-York: Cold Spring Harbour Laboratory Press

Selvaraj.G and Iyer, VN (1984). *J. Bacteriol.* 158:580–589.

Streber, W, Timmis, KN and Zenk, MH (1987). *J. Bacteriol.* 169: 2950–2955.

Suwa, Y, Wright, A, Fukimori, F, Nummy, KA, Hausinger, RP, Holben, WE and Forney, LJ (1996). *Appl. Environ. Microbiol.* 62: 2464–2469.

Top, EM, Maltseva, OM and Forney, LJ (1996). *Appl. Environ. Microbiol.* 62:1470–1476

Vallaeys, T, Fulthorpe, RR, Wright, A and Soulas, G (1996). *FEMS. Microbiol. Ecol.* 20:163–172.

Vallaeys, T, Persello-Cartieaux, F, Rouard, N, Lors, C, Laguerre, G and Soulas, G (1997). *FEMS Microb. Ecol.* 24:269–278.

Wheatcroft, R and Williams, PA (1981). *J. Gen. Microbiol.* 124 : 433–437.

Received: 4 September 1998
Revisions requested: 23 September 1998
Revisions received: 8 October 1998
Accepted: 8 October 1998

COR-AAD1-00001277

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1993 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 268, No. 32, Issue of November 15, pp. 24311-24317, 1993
Printed in U.S.A.

Purification and Characterization of 2,4-Dichlorophenoxyacetate/α-Ketoglutarate Dioxygenase*

(Received for publication, April 15, 1993, and in revised form, June 25, 1993)

Fumiyasu Fukumori‡§ and Robert P. Hausinger‡¶

From the ‡Center for Microbial Ecology and Departments of Microbiology and Biochemistry, Michigan State University, East Lansing, Michigan 48824 and the §Research Development Corporation of Japan, Chiyoda-ku, Tokyo 100, Japan

The *Alcaligenes eutrophus* 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase, encoded by the *tfdA* gene of plasmid pJP4, is an Fe(II)-dependent enzyme that catalyzes the conversion of 2,4-dichlorophenoxyacetate to 2,4-dichlorophenol and glyoxylate concomitant with the decomposition of α-ketoglutarate to form succinate and carbon dioxide (Fukumori, F., and Hausinger, R. P. (1993) *J. Bacteriol.* 175, 2083–2086). Using recombinant *Escherichia coli* cells that overexpress the *tfdA* gene, the thermolabile enzyme (stable only up to 30 °C) was purified to apparent homogeneity (specific activity of 16.9 μmol of substrate converted min^{-1} mg of $protein^{-1}$) by a simple two-step procedure. The native protein has an apparent M_r of 50,000 ± 2,500, consistent with a homodimeric structure. Ferrous ion is absolutely required for activity and cannot be replaced by several other divalent cations tested. Ascorbic acid stimulates dioxygenase activity and reduces the rate of enzyme inactivation by a metal ion-mediated process. The enzyme exhibits maximum activity at pH 6.5–7, however, it is stable over a pH range of 6.5–11. Although capable of hydroxylating a wide range of phenoxyacetates and related compounds, the enzyme exhibits the greatest affinity (K_m 17.5 ± 1.0 μM) and highest catalytic efficiency for 2,4-dichlorophenoxyacetate. Similarly, α-ketoglutarate is the preferred co-substrate (K_m 3.20 ± 0.54 μM) for the enzyme, but it can utilize a range of other α-ketoacids with lower efficiency. Results from chemical modification studies are consistent with the presence of multiple essential histidine residues in the enzyme.

The biodegradation of 2,4-dichlorophenoxyacetic acid (2,4-D),[1] a broadleaf herbicide, has been studied as a model for microbial decomposition of chloroaromatic compounds (1–3). *Alcaligenes eutrophus* JMP134 is the most extensively studied 2,4-D degrader and the pathway found in this microbe is shown in Fig. 1. The genes responsible for the first five steps in this pathway (*tfdA*, *-B*, *-C*, *-D*, and *-E*) have been localized to the transmissible plasmid, pJP4 (4). The nucleotide sequences of *tfdCDE* (5) exhibit extensive similarity to the well studied *clcABD* and *tcbCDE* chlorocatechol degradation genes (reviewed in Ref. 3) and similarly provide for the catabolism of dichlorocatechol. The sequence of the *tfdB* gene (5), encoding 2,4-DCP hydroxylase, also exhibits similarity to other genes involved in phenol hydroxylation (*e.g.* see Ref. 6). In contrast to the downstream genes, however, the sequence of the *tfdA* gene (7) does not exhibit obvious evolutionary relationships to any other known genes. The *tfdA* gene product (TfdA) generally has been referred to as 2,4-D monooxygenase; however, we recently demonstrated that the enzyme carries out a very different type of reaction (8) that is mechanistically related to reactions catalyzed by a group of α-KG-dependent dioxygenases (9). The reaction requires α-KG and Fe(II), in addition to 2,4-D and oxygen, and produces carbon dioxide, succinate, 2,4-DCP, and glyoxylate (Fig. 2).

Here, we detail the purification and describe the catalytic properties of 2,4-D/α-KG dioxygenase. In addition, we examine the substrate ranges of the enzyme for other phenoxyacetates and α-ketoacids. Furthermore, we describe evidence consistent with the presence of multiple essential histidine residues in the enzyme.

EXPERIMENTAL PROCEDURES

Chemicals—Organic chemicals were obtained from Sigma or Aldrich and were of the highest purity available. α-[1-^{14}C]KG (1.8 MBq/mmol) was obtained from DuPont NEN.

Media, Organisms, and Plasmid—Modified LB medium (10 g of Tryptone, 5 g of yeast extract, and 5 g of NaCl per liter (pH 7.2)) supplemented with 100 μg/ml of ampicillin was used to grow the host strain *Escherichia coli* JM109 (10). *A. eutrophus* JMP134 (4) was grown at 30 °C in 1/10 PTYG medium (0.25 g of peptone, 0.25 g of Tryptone, 0.5 g of yeast extract, 0.5 g of glucose, 0.03 g of $MgSO_4$, and 0.0035 g of $CaCl_2$ per liter (pH 7.2)) supplemented with 2.5 mM 2,4-D. Plasmid pUS311 (8) that encodes the *tfdA* gene had been generated from pKJS31 (7).

Purification of TfdA—Unless stated otherwise, all purification steps were done at 4 °C. Cultures (2 liters) of *E. coli* JM109(pUS311) were grown overnight at 30 °C with moderate aeration. Cells were harvested by centrifugation (6,000 × *g*) for 20 min and suspended in 25 ml of 20 mM Tris, 1 mM EDTA, 0.4 mg/liter leupeptin (pH 7.2). The cell suspensions were stored at −20 °C until use. Suspended cells were adjusted to 1 mM phenylmethylsulfonyl fluoride and passed through a French pressure cell at 120 MPa. Cell extracts were obtained after clarification of disrupted cells by centrifugation at 100,000 × *g* for 30 min. To the extract was added 20 mM Tris, 1 mM EDTA, 0.1 mg/liter leupeptin (pH 7.2) (Buffer A) containing 3.4 M $(NH_4)_2SO_4$ and the fraction precipitating between 1.0 and 1.7 M was collected. The precipitate was dissolved in 20 ml of Buffer A and dialyzed against 4 liters of the same buffer for 16 h. The solution was then chromatographed on a Mono Q 10/10 column (Pharmacia LKB Biotechnology Inc.) in Buffer A and eluted with a 100 ml of linear salt gradient to 0.2 M NaCl at room temperature (about 20 °C). TfdA was found to elute at approximately 0.05 M NaCl.

Enzyme Assays—2,4-D/α-KG dioxygenase was assayed at 30 °C. The typical assay mixture (total volume of 1 ml) contained 1 mM α-KG, 50 μM ascorbate, 50 μM $(NH_4)_2Fe(SO_4)_2$, and 1 mM 2,4-D in 10 mM imidazole buffer (pH 6.75). Aliquots of the reaction mixture were quenched by the addition of EDTA. The concentration of released phenol derivative was

* This work was supported by the National Science Foundation-Science and Technology Center Grant BIR 9120006 and the Michigan State University Agricultural Experiment Station. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed. Tel.: 517-353-9675; Fax: 517-353-8957.

[1] The abbreviations used are: 2,4-D, 2,4-dichlorophenoxyacetate; α-KG, α-ketoglutarate; 2,4-DCP, 2,4-dichlorophenol; DEP, diethylpyrocarbonate; TNBS, trinitrobenzenesulfonic acid; EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide; 2,4,5-T, 2,4,5-trichlorophenoxyacetate; Bis-tris, bis(2-hydoxyethyl)iminotris(hydroxymethyl)methane; MOPS, 3-(*N*-morpholino)propanesulfonic acid; MES, 2-(*N*-morpholino)ethanesulfonic acid; CAPS, 3-(cyclohexylamino)propanesulfonic acid; CHES, 2-(cyclohexylamino)ethanesulfonic acid.

The Journal of Biological Chemistry ASBMB

Downloaded from www.jbc.org by on August 5, 2009

COR-AAD1-00001278

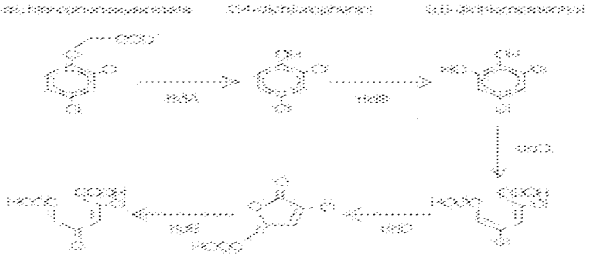


FIG. 1. **Pathway for degradation of 2,4-D in *A. eutrophus* JMP134.**

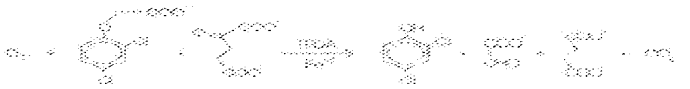


FIG. 2. **Overall reaction for 2,4-D/α-ketoglutarate dioxygenase.**

measured by a modification of the 4-aminoantipyrine method (11); to 1 ml of reaction mixture was added 100 μl of a pH 10 buffer solution (3.09 g of boric acid, 3.73 g of potassium chloride, and 44 ml of 1 N sodium hydroxide brought to 1 liter with distilled water), 10 μl of 2% 4-aminoantipyrine, and 10 μl of 8% potassium ferricyanide. The absorbance at 510 nm was determined after 20 min with a Gilford Response spectrophotometer. One unit of activity is defined as the amount of enzyme that forms 1 μmol of phenol derivative/min at 30 °C.

Examination of the kinetic parameters for 2,4-D and its analogues and for α-KG and its analogues involved holding the concentration of one substrate constant at 1 mM while varying the concentration of the substrate under consideration. For each of the phenoxyacetate derivatives the appropriate modified phenol compound was used as a standard in the spectrophotometric assay to account for their slightly different extinction coefficients. In contrast, 2,4-DCP served as the standard for kinetic evaluation of each of the α-ketoacids. Because of the extremely low K_m value for α-KG, special conditions were used with this substrate. Enzyme (0.1 μg) was incubated with α-[1-^{14}C]KG (145 kBq/mmol) under standard conditions for up to 1 min, the reaction was stopped by the addition of 1 N HCl, reaction mixtures were incubated for 1 h at room temperature with shaking to drive off the released carbon dioxide (α-KG is stable under the low pH conditions used in this experiment), and the remaining radioactivity in the solution was measured. Rates for α-KG utilization were calculated based on the released radioactivity. The capacity of the enzyme to react with several other non-phenoxyacetate substrates was assessed by monitoring spectroscopic changes or by high performance liquid chromatography (Hewlett-Packard HP-1050) methods using a C_{18} (octadecyl silane) reverse-phase column (4.0 × 250 mm) at a flow rate of 1.6 ml/min in various compositions of methanol, 0.1% H_3PO_4 in H_2O. For all compounds that served as substrate for the enzyme, apparent kinetic constants were obtained (12) from rates that were calculated by using linear regression analysis. Values of k_{cat} were based on a subunit M_r of 32,171 (7).

Analytical Methods—Protein concentration was measured (13) using bovine serum albumin as a standard. The native M_r of TfdA was estimated by gel filtration through a Superose 12 HR 10/30 column (Pharmacia; flow rate 15 ml/h) in 20 mM Tris, 1 mM EDTA, 100 mM NaCl (pH 7.2). The calibration proteins were thyroglobulin, M_r 670,000; bovine γ-globulin, M_r 158,000; chicken ovalbumin, M_r 44,000; equine myoglobin, M_r 17,000; and vitamin B-12, M_r 1,350 (Bio-Rad).

Gel Electrophoresis and Immunoblotting—Denaturing SDS-polyacrylamide gel electrophoresis was carried out on 12% polyacrylamide gels (14). Gels were stained with Coomassie Brilliant Blue. Standard proteins included phosphorylase *b*, M_r 92,500; bovine serum albumin, M_r 66,200; ovalbumin, M_r 45,000; carbonic anhydrase, M_r 31,000; soybean trypsin inhibitor, M_r 21,500; and lysozyme, M_r 14,400 (Bio-Rad).

Antibodies were generated against purified TfdA in a white New Zealand rabbit, and the IgG fraction was obtained from the serum (15). Immunoblots were prepared by blotting the proteins onto nitrocellulose, probing with anti-TfdA rabbit IgG, and developing with goat anti-rabbit IgG-alkaline phosphatase conjugates (16).

Chemical Modification Studies—DEP was freshly diluted with ethanol and its concentration was determined by using 10 mM imidazole (pH 7.0) based on the increase in absorbance at 240 nm (ε = 3000 M^{-1} cm^{-1}) (17). Reactions were carried out by incubating the enzyme (1.2 μM) with DEP and other additions as indicated in 10 mM MOPS buffer (pH 6.5) at room temperature (21–22 °C). Aliquots (25 μl) were withdrawn at various time intervals, the reactions were quenched with 785 μl of 16.3 mM imidazole buffer (pH 6.75), and residual activities were determined as in the standard procedure. Reactivation of the DEP-modified TfdA (treated for 30 min with 1 mM DEP) was examined after adjusting the solution to 10 mM imidazole and 0.5 M hydroxylamine (using neutralized reagent). TfdA activity at various times was assessed after removal of excess reagent by using a Bio-Rad 10 DG desalting column. Inactivation studies of the enzyme (15 nM) by various other reagents, dissolved in H_2O, were performed in 10 mM imidazole or 10 mM MOPS buffers (pH 6.75). Aliquots (2 μl) were withdrawn at various time intervals and residual activities were determined as described above.

RESULTS

Purification of TfdA—2,4-D/α-KG dioxygenase was purified to over 95% homogeneity from *E. coli* JM109(pUS811) with 56% yield by using the 2-step procedure summarized in Table I. Comparison of the isolated protein to cell extracts from the recombinant strain by denaturing gel electrophoresis (Fig. 3) demonstrated that the cells produced the *tfdA* gene product as more than 10% of the total protein. Gel filtration chromatography was used to estimate a M_r of 60,000 ± 2,500 for the native enzyme, consistent with a homodimeric structure (subunit M_r of 32,171; Ref. 7).

The presence of protease inhibitors during early stages of TfdA purification enhances the stability of the enzyme by pre-

TABLE I
Purification of 2,4-D/α-ketoglutarate dioxygenase

Purification step	Total protein	Total activity	Specific activity	Purification
	mg	*units (%)*	*units/mg*	*fold*
Cell extracts	750	1890 (100)	2.5	1
Ammonium sulfate precipitation	235	1400 (83)	6.2	2.8
Mono Q	58	960 (58)	16.9	7.7

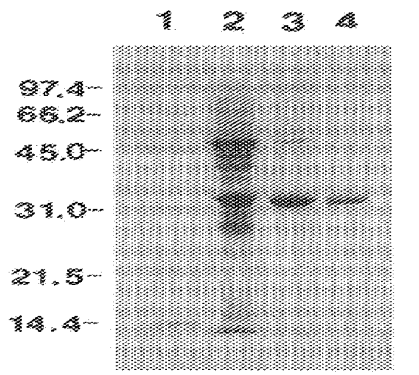


FIG. 3. **Purification of Alcaligenes eutrophus JMP134 2,4-D/α-KG dioxygenase from *E. coli* JM109(pUS811).** Protein samples were analyzed by denaturing gel electrophoresis on a 12% polyacrylamide gel and stained with Coomassie Brilliant Blue. *Lane 1*, molecular weight markers with the relative molecular mass values indicated in kilodaltons on the *left margin*; *lane 2*, supernatant of disrupted cells of *E. coli* JM109(pUS811); *lane 3*, fraction precipitated from 1–1.7 M ammonium sulfate; *lane 4*, fraction containing purified TfdA as eluted from Mono Q ion exchange column.

Downloaded from www.jbc.org by on August 5, 2009

COR-AAD1-00001279

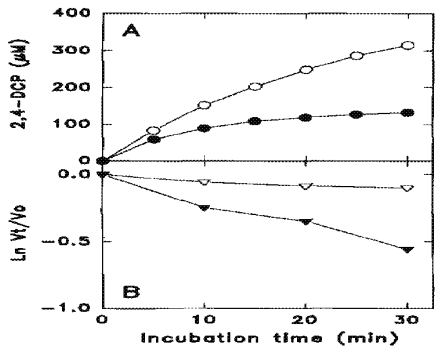


FIG. 4. **Time- and ferrous ion-dependent inactivation of 2,4-D/α-KG dioxygenase.** *A*, Progress curves of 2,4-DCP formation for TfdA (1 μg/ml) incubated in the standard assay with (○) and without (●) ascorbate. *B*, TfdA (2 μg/ml) was incubated at room temperature in 10 mM imidazole buffer (pH 6.75) with no additions (▽), and with 100 μM ferrous ion and 100 μM ascorbate (▼). Aliquots were withdrawn over time and assayed under standard conditions.

venting conversion of the subunit to an inactive TfdA fragment of apparent M_r 27,000. Whereas NH_2-terminal sequence analysis of the non-degraded subunit revealed the residues expected from DNA sequence analysis (S-V-V-A-N-; Ref. 7), the amino-terminal sequence of the proteolytic fragment of TfdA (F-L-Y-A-E-L-) was consistent with hydrolytic cleavage after arginine (residue 77) and loss of the NH_2-terminal fragment. By using anti-TfdA IgG in Western blot analysis of various samples, the conversion was found to occur after cell disruption rather than during the cell cultivation period. The same methods were used to demonstrate that the M_r of 32,000 form of the protein was present in cell extracts of *A. eutrophus* JMP134.

Maintenance of low temperatures during bacterial growth and enzyme isolation also is important for obtaining high levels of activity. As previously described (8), TfdA forms aggregates in *E. coli* JM109(pUS311) cells that are grown at 37 °C but remains soluble for cells grown at 30 °C. Indeed, we have observed that colonies grown at the higher temperature are turbid, whereas those grown at the more moderate temperature remain clear. Furthermore, the purified 2,4-D/α-KG dioxygenase was found to be very sensitive to thermal inactivation. At a protein concentration of 0.3 μM, TfdA remains stable only up to 30 °C and is completely inactivated within 5 min at 40 °C. In contrast, the purified enzyme was found to be stable at 4 °C for at least several weeks after purification.

Catalytic Properties of TfdA—As previously noted (8), activity of 2,4-D/α-KG dioxygenase requires ferrous ion and is enhanced by ascorbic acid. The ferrous ion requirement could not be met by other divalent metal ions including Co(II), Cu(II), Li(II), Mg(II), Mn(II), Ni(II), and Zn(II) when added as sulfate salts at a concentration of 100 μM. As shown by the time-dependent decrease in enzyme activity in Fig. 4*A*, however, ferrous ion alone is unable to sustain enzyme catalysis over long time periods. The rate of activity loss was greatly reduced (although not completely eliminated) by inclusion of ascorbic acid in the assay. The ascorbic acid serves to prevent inactivation rather than to reactivate the enzyme as shown by studies in which addition of ascorbate to partially active enzyme in the ferrous ion-alone treatment failed to restore full activity. Similarly, activity could not be recovered from inactive enzyme that was dialyzed against 1 mM EDTA and/or 1 mM dithiothreitol and then reassayed in the presence of ferrous ion plus ascorbate. Catalytic turnover of the enzyme was not required for inactivation as shown by the loss of activity during enzyme incubation with ferrous ion prior to addition of substrate (Fig. 4*B*). Although not completely characterized, the inactivation clearly results from a metal ion-mediated event as demonstrated by the retention of activity when the enzyme is stored in the absence of metal ions and the presence of EDTA. To minimize enzyme inactivation in the kinetic studies described below, reactions were initiated by addition of enzyme to assay mixtures and sufficiently short assay periods were used to calculate initial rates. The concentration dependences for ferrous ion and ascorbate on initial 2,4-D/α-KG dioxygenase rates are shown in Fig. 5.

The pH dependence of 2,4-D/α-KG dioxygenase was examined. The enzyme activity exhibits a narrow pH optimum of 6.5–7.0 (Fig. 6*A*). In contrast, TfdA is stable over a wide pH

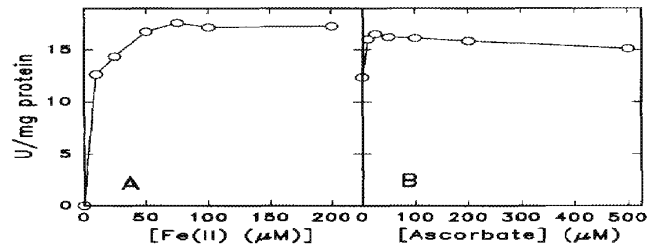


FIG. 5. **Effect of Fe(II) and ascorbic acid concentrations on the activity of 2,4-D/α-KG dioxygenase.** *A*, Fe(II) concentration dependence on 2,4-D/α-KG dioxygenase activity; *B*, ascorbic acid concentration dependence on enzyme activity.

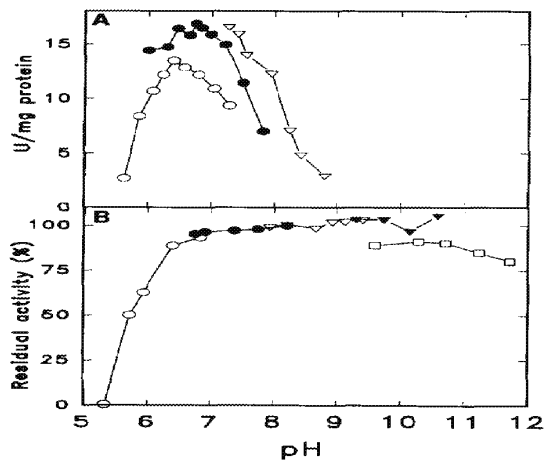


FIG. 6. **pH dependence of 2,4-D/α-KG dioxygenase.** *A*, enzyme activity. The enzyme assays were done in the following buffers at a concentration of 5 mM: Bis-Tris, pH 5.6–7.3 (○); imidazole, pH 6.0–7.6 (●); Tris, pH 7.3–8.8 (▽), all at 30 °C. *B*, enzyme stability. The residual enzyme activity after incubation at various pH values for 20 h at 4 °C was determined by using standard assay conditions. The enzyme was incubated in the following buffers at a concentration of 10 mM: MES, pH 5.3–6.8 (○); imidazole, pH 6.7–8.2 (●); Tris, pH 7.9–9.4 (▽); CHES, pH 9.3–10.6 (▼); CAPS, pH 9.6–11.7 (□).

Downloaded from www.jbc.org by on August 5, 2009

COR-AAD1-00001280

TABLE II
Phenoxyacetate substrate specificity and kinetic parameters for 2,4-D / α-ketoglutarate dioxygenase
All experiments were performed at 30 °C in 10 mM imidazole buffer (pH 6.75) containing 1 mM α-KG, 50 μM ascorbate, 50 μM $(NH_4)_2Fe(SO_4)_2$, and the indicated substrates over the indicated concentration ranges.

Substrate (concentration range studied)	K_m	k_{cat}	k_{cat}/K_m
	μM	min^{-1}	$min^{-1}\cdot mM^{-1}$ (%)
Phenoxyacetate (200–1000 μM)	460 ± 23	443 ± 10	960 (3)
2-Chlorophenoxyacetate (40–200 μM)	110 ± 6.5	380 ± 11	3,450 (11)
4-Chlorophenoxyacetate (40–200 μM)	117 ± 6.2	595 ± 15	5,090 (17)
2,3-Dichlorophenoxyacetate (50–200 μM)	102 ± 8.5	288 ± 12	2,820 (9)
2,4-Dichlorophenoxyacetate (10–50 μM)	17.5 ± 1.0	529 ± 16	30,200 (100)
3,4-Dichlorophenoxyacetate (50–500 μM)	219 ± 7.6	307 ± 5	1,400 (5)
2,4,5-Trichlorophenoxyacetate (20–150 μM)	59.6 ± 3.4	96 ± 3	1,610 (5)
4-Chloro-2-methylphenoxyacetate (20–150 μM)	89.0 ± 8.7	233 ± 12	2,620 (9)
DL-2-Phenoxypropionate (0.5–5 mM)	1170 ± 120	5.1 ± 0.2	4.4 (—)
2-(2,4-Dichlorophenoxy)propionate (0.05–1 mM)	191 ± 6.7	61 ± 3	320 (1)
3-Phenoxypropionate (5–25 mM)	12900 ± 3200	3.2 ± 0.4	0.25 (—)

range (although it is totally inactivated by acidic conditions (pH < 6.5)) as shown by measuring the remaining activity after incubation in various pH buffers at 4 °C for 20 h (Fig. 6*B*).

Substrate Range of TfdA—In addition to 2,4-D, the enzyme catalyzes the α-KG-dependent release of the expected phenol derivatives from several other phenoxyacetates (Table II). The enzyme exhibits greatest affinity and catalytic efficiency toward 2,4-D, however, many ring-substituted compounds are reasonable substrates with similar catalytic rates but lower affinity. Non-halogenated phenoxyacetate possesses a larger K_m value than do the halogenated substrates. Similarly, the K_m value for 2-phenoxypropionate is substantially larger than that of 2-(2,4-dichlorophenoxy)propionate. The additional methyl group in the side chain of these two compounds, however, greatly decreases their k_{cat} values for hydroxylation, perhaps due to the change from a secondary to a tertiary carbon atom. Although it is a very poor substrate 3-phenoxypropionate is hydroxylated by the enzyme, demonstrating that the substrate binding site can accommodate one extra methylene carbon in the side chain. In contrast, TfdA exhibits no activity toward analogues in which (i) the phenoxy group is two carbons atoms more distant from the carboxyl group (4-(2,4-dichlorophenoxy)-butyrate), (ii) the carboxyl and phenoxy groups are located at adjacent positions on an aromatic ring (2-phenoxybenzoate), (iii) the acid group is reduced (2-phenoxyethanol), or (iv) the phenoxy oxygen is replaced by a methylene group (hydrocinnamic acid). Similarly, the enzyme exhibits insignificant activity toward methyl esters of phenoxyacetate and 4-chlorophenoxyacetate. Although trace amounts of product phenols were observed (<0.5%), they may have arisen from uncatalyzed hydrolysis of the substrate, with subsequent enzymatic action on the phenoxyacetates, as the amount of product did not depend on the enzyme concentration. Finally, because other investigators have characterized a mechanistically related 4-hydroxyphenylpyruvate dioxygenase (17–19), we tested whether this compound could serve as a substrate for TfdA. We found that 4-hydroxyphenylpyruvate is not utilized by this enzyme.

The K_m and k_{cat} values for several α-ketoacids were determined to examine the specificity of the enzyme for α-KG (Table III). Although the enzyme exhibits a broad α-ketoacid substrate range, the highest affinity and greatest catalytic rate were observed for α-KG. The non-α-ketoacid carboxyl group is not required for recognition by the enzyme, but the K_m values for the two substrates possessing a second acidic group are significantly increased and the catalytic rates are greatly reduced compared to the other substrates. Addition of an extra methylene group between the α-ketoacid group and the free carboxyl group, as in α-ketoadipate, leads to small changes in the kinetic constants: a 6.4-fold increase in K_m and a 55% decrease in k_{cat} compared to α-KG. In contrast, removal of one of the methylene groups, as in oxalacetate, led to an ineffective substrate. Furthermore, using 25 μg of enzyme and 50-min incubation times, β-ketoglutarate, malonate, succinate, and glutarate (all up to 1 mM concentrations) were found to be incapable of supporting 2,4-D hydroxylation. Finally, in the absence of 2,4-D, no decomposition of α-KG was observed.

TABLE III
α-Ketoacid substrate specificity and kinetic parameters for 2,4-D / α-ketoglutarate dioxygenase
Enzyme reactions were performed at 30 °C in 10 mM imidazole buffer (pH 6.75) containing 1 mM 2,4-D,50 μM ascorbate, 50 μM $(NH_4)_2Fe(SO_4)_2$, and the indicated substrates over the indicated ranges of concentration.

Substrate (concentration range studied)	K_m	k_{cat}	k_{cat}/K_m
	μM	min^{-1}	$min^{-1}\cdot mM^{-1}$ (%)
α-KG (1.5–10 μM)	3.20 ± 0.54	643 ± 44	2.0×10^6 (100)
α-Ketoadipate (10–100 μM)	20.6 ± 1.1	290 ± 5	14100 (7.05)
Pyruvate (500–2500 μM)	1020 ± 86	58 ± 2	60 (0.03)
α-Ketobutyrate (150–800 μM)	464 ± 61	89 ± 6	190 (0.10)
α-Ketovalerate (300–1250 μM)	607 ± 47	404 ± 14	660 (0.33)
α-Ketocaproate (200–1000 μM)	583 ± 50	158 ± 7	270 (0.14)
α-Ketoisovalerate (200–1000 μM)	745 ± 36	16 ± 1	20 (0.01)

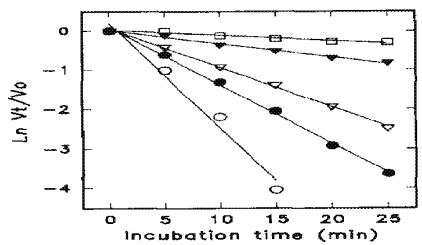


FIG. 7. **Kinetics of inactivation of 2,4-D/α-KG dioxygenase with DEP.** Pseudo-first-order plots for TfdA inactivation by DEP at concentrations of 0.13 mM (□), 0.27 mM (▼), 0.67 mM (▽), 1.0 mM (●) and 1.34 mM (○) in 10 mM MOPS buffer (pH 6.5, 21–22 °C). Ln (V_t/V_0) (where V_t is the velocity at time t and V_0 is the initial velocity) is shown as a function of time.

Chemical Modification of TfdA—DEP modification of TfdA in MOPS buffer (pH 6.75) resulted in the rapid loss of enzyme activity. A plot of the natural logarithm of remaining activity *versus* time at various concentrations of DEP exhibits pseudo first-order kinetics (Fig. 7). Replotting the apparent first-order

ASBMB The Journal of Biological Chemistry

Downloaded from www.jbc.org by on August 5, 2009

COR-AAD1-00001281

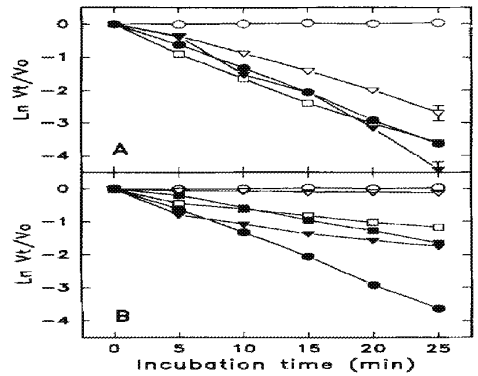


FIG. 8. **Substrate/ligand protection against DEP inactivation of 2,4-D/α-KG dioxygenase.** Residual activities are plotted against the incubation time. The enzyme was incubated with 1 mM DEP in 10 mM MOPS buffer (pH 6.5, 22 °C) in the presence of the indicated compounds. *A*, 1 mM 2,4-D (▽), 1 mM α-KG (▼), 0.5 mM Fe(II) and 0.5 mM ascorbate (□), and no addition (●). *B*, 1 mM 2,4-D, 0.5 mM Fe(II), and 0.5 mM ascorbate (▼), 1 mM α-KG, 0.5 mM of Fe(II), and 0.5 mM ascorbate (□), 1 mM 2,4-D and 1 mM α-KG (▽), 0.1 mM 2,4-D and 0.1 mM α-KG (■), and no addition (●). Activity was also measured in a control sample to which ethanol was added instead of DEP (○).

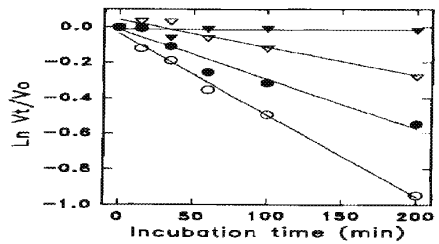


FIG. 9. **Kinetics of inactivation of 2,4-D/α-KG dioxygenase with TNBS.** Pseudo-first-order plots for inactivation of TfdA by TNBS at concentrations of 0 mM (▼), 0.5 mM (▽), 1 mM (●), and 2.5 mM (○) in 10 mM MOPS buffer (pH 6.75, 22 °C).

rate constants as a function of DEP concentration yields a straight line that provides a second-order rate constant of 6.14 M^{-1} s^{-1}. The ability of 2,4-D, α-KG, Fe(II) plus ascorbate, and combinations of these substances to protect the enzyme against DEP inactivation was examined. Whereas none of the individual compounds are able to significantly protect the enzyme from inactivation by DEP (Fig. 8*A*), the combinations of 2,4-D plus Fe(II) or α-KG with Fe(II) clearly decreases the inactivation rate (Fig. 8*B*). Furthermore, the combined presence of 2,4-D and α-KG is very effective in protecting the enzyme from inactivation by DEP.

Although commonly considered a histidine-selective reagent, DEP is known to react with tyrosine, lysine, and cysteine residues (17). We were unable to use the reported spectroscopic method (17) to determine whether a DEP-reactive tyrosine residue is present in TfdA because the DEP-modified protein forms a precipitate. This problem also prevented spectroscopic quantitation of the total DEP-reactive residues. Two approaches were used to assess whether the enzyme possesses a DEP-reactive lysine residue: hydroxylamine treatment of DEP-

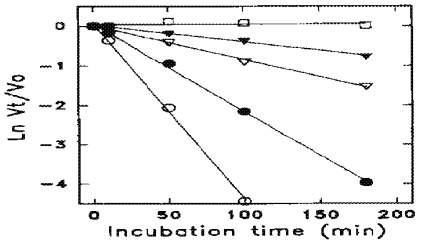


FIG. 10. **Kinetics of inactivation of 2,4-D/α-KG dioxygenase with EDC.** Pseudo-first-order plots for inactivation of TfdA by EDC at concentrations of 0 mM (□), 10 mM, (▼), 20 mM (▽), 50 mM (●), and 100 mM (○) in 10 mM MOPS buffer (pH 6.75, 22 °C).

modified protein and enzyme reactivity toward TNBS. Of the possible DEP-reactive residues usually only modified lysine or cysteine groups are stable to hydroxylamine treatment. Although we found that DEP-treated enzyme was incapable of being reactivated by 0.5 M hydroxylamine (not shown), it remains possible that one or more carbethoxyhistidine residues may be inaccessible to the reagent. The latter explanation is most consistent with our results using the lysine-selective reagent, TNBS. The enzyme was slowly inactivated by TNBS (Fig. 9), however, no protection from inactivation was observed when 1 mM α-KG, 1 mM 2,4-D, or 1 mM α-KG plus 1 mM 2,4-D was included in the reaction mixture. With regard to possible cysteine reactivity toward DEP, neither iodoacetamide (to 100 mM) nor *N*-ethylmaleimide (to 25 mM) inactivated the enzyme after 100 min of incubation. Similarly, 2,3-butanedione (to 100 mM) failed to inactivate the enzyme. In contrast, the carboxyl group-specific modifier, EDC, slowly inactivates the enzyme at high concentrations with pseudo-first-order kinetics (Fig. 10). A replot of the apparent rate constants *versus* EDC concentration is linear and allows calculation of the second-order rate constant of 0.018 M^{-1} s^{-1}. Analysis of substrate protection was impossible with this reagent since 2,4-D and α-KG both have carboxyl groups. These results are inconsistent with the presence of essential cysteine or arginine residues or of highly reactive lysine or carboxyl groups in the enzyme.

DISCUSSION

2,4-D/α-KG dioxygenase was purified to apparent homogeneity from recombinant *E. coli* cells by a simple two-step procedure and several properties of the enzyme were examined. This discussion will focus on comparison of its enzymic properties to that of other α-KG-dependent enzymes, the substrate specificity of the enzyme, and evidence consistent with the presence of multiple essential histidine residues in TfdA.

The specific requirements for α-KG and Fe(II), and the stimulation of activity by ascorbate, are typical characteristics of α-KG-dependent-dioxygenases (9). For example, prolyl 4-hydroxylase is relatively stable when assayed in the presence of ascorbate, whereas, it is rapidly inactivated in the absence of this reductant with 90% of the enzyme becoming inactive in approximately one minute corresponding to 15–30 turnovers (20, 21). In contrast, ascorbate-free 2,4-D/α-KG dioxygenase activity is much more stable with the enzyme carrying out approximately 4000 turnovers before becoming completely inactivated after 30–40 min. Inactivation of prolyl or lysyl hydroxylases is thought to involve a side reaction of these enzymes in which α-KG is decarboxylated in the absence of the peptide substrate leaving the metallocenter in an oxidized state that subsequently must be reduced by ascorbate (22, 23). TfdA,

The Journal of Biological Chemistry ASBMB
COR-AAD1-00001282
Downloaded from www.jbc.org by on August 5, 2009

however, does not appear to catalyze such an ascorbate-dependent decarboxylation of α-KG uncoupled from 2,4-D hydroxylation and enzyme inactivation occurs in the absence of either substrate. Rather than serving to reduce an oxidized enzyme-bound metal center, ascorbate appears to reduce the rate of some other metal ion-mediated enzyme inactivation process. The decrease in enzyme activity that is observed in assays containing both ferrous ion and ascorbate is not due to product inhibition as shown by the lack of any effect when 1 mM 2,4-DCP was included in assay mixtures containing various concentrations (10, 20, 50, and 100 μM) of 2,4-D.

Our results establish the broad substrate range for the hydroxylase reaction of the *tfdA*-encoded dioxygenase. Pieper *et al.* (24) reported that whole cells of *A. eutrophus* JMP134 that were deregulated for *tfdA* expression could degrade a series of 11 phenoxyacetates with chloro- and methyl-substitutions on the phenol ring (although only three of these compounds were growth substrates). Detailed kinetic studies of degradation rates, however, were impossible in their whole cell preparations because of substrate transport and other complications. Here, we were able to evaluate the kinetic constants for seven of the compounds examined by Pieper *et al.* (24) and four other substrates of TfdA. With regard to the observed substrate range of TfdA, it is important to emphasize that not all of these *in vitro* substrates are used by the *A. eutrophus* cells *in situ*. For example, we established that TfdA catalyzes conversion of 2,4,5-T to 2,4,5-trichlorophenol despite the reported inability of *A. eutrophus* JMP134 to catabolize it (25). These results suggest that either the strain can not transport 2,4,5-T or it lacks the ability to degrade 2,4,5-trichlorophenol. Whereas we were unable to detect utilization of 4-(2, 4-dichlorophenoxy)butyric acid by purified TfdA, a *Flavobacterium* species could degrade 4-(2,4-dichlorophenoxy)butyric acid as well as 2-(2,4-dichlorophenoxy)propionic acid, 2-(4-chloro-2-methylphenoxy)propionic acid, and 2,4-D (26). These results may indicate that the enzyme corresponding to TfdA in *Flavobacterium* has a different substrate range or that another enzyme is required; *e.g.* the four-carbon side chain of the substrate may be oxidized to remove two carbons and form 2,4-D. Similarly, we were unable to demonstrate utilization of the methyl ester of 2,4-D despite the known ability of numerous microorganisms to degrade this compound (27). It is likely that a separate enzyme synthesized by these microbes first hydrolyses the methyl ester to release 2,4-D that is subsequently metabolized by a TfdA-like enzyme. Examination of the kinetic values obtained for TfdA reveals that the rates of catalysis for the various substrates are generally quite similar with the exception of the sluggish rates observed with the propionates. Rather, the major differences in the catalytic efficiencies for these substrates involve their affinities for the enzyme. TfdA appears to prefer chlorinated substrates over the unsubstituted compounds. The position of the phenoxy groups at the C-2 carbon of an acid is very important for recognition, although one extra methylene group can be tolerated. The importance of the carboxyl group for substrate recognition is consistent with the presence of a positively charged residue in the 2,4-D binding site.

Although TfdA exhibits greatest efficiency with α-KG as the cosubstrate, many other α-ketoacids are able to be used and α-ketoadipate supports significant levels of hydroxylation. Similar utilization of α-ketoadipate has been reported in the reaction of some other α-KG-dependent dioxygenases such as prolyl 4-hydroxylase (28), but not in γ-butyrobetaine hydroxylase (29). The greatly enhanced affinity of α-ketoacids containing a second carboxyl group is consistent with the presence of a positively charged residue at the α-ketoacid binding site.

Chemical modification studies of TfdA were used to rule out the presence of essential and accessible cysteine or arginine residues and of highly reactive and essential lysine and carboxyl groups. In contrast, DEP rapidly inactivates the enzyme. Myllylä *et al.* (30) reported DEP-dependent inactivation of prolyl 4-hydroxylase and provided some evidence for substrate protection. In our more detailed analysis of substrate protection, neither 2,4-D, α-KG, nor Fe(II) alone was found to protect the enzyme from inactivation; yet, combinations of 2,4-D or α-KG with Fe(II) provide some protection, and a combination of 2,4-D and α-KG afforded nearly complete protection. These results are compatible with the presence of multiple essential histidine residues, as illustrated in Fig. 11, but do not rule out the possible DEP-reactivity of alternate residues. Consistent with the expected requirements for positive charges at the binding sites of 2,4-D and α-KG, we propose that essential histidine residues are present at the binding sites for both of these substrates. In addition, we propose that one or more additional histidine residues may be buried in the protein at the Fe(II) binding site, consistent with the known ability of histidine residues to function as metallocenter ligands. For example, spectroscopic analysis of isopenicillin N synthase, possibly a mechanistically related α-KG-dependent enzyme, suggested the presence of three histidine ligands to Fe(II) in that protein (31, 32). Binding of 2,4-D and α-KG would protect the histidine residues at the substrate binding sites and additionally may protect the Fe(II) ligands by steric constraints. Binding of either 2,4-D plus Fe(II) or α-KG and Fe(II) might lead to the observed reduced rate of inactivation. In contrast, addition of any one compound alone is unable to protect the three or more distinct sites of inactivation postulated in this model. Further studies clearly are required to assess the validity of this model.

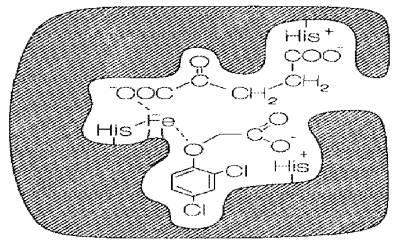


FIG. 11. **Model of the TfdA active site indicating potential roles of essential histidine residues.** The varied abilities of 2,4-D, α-KG, Fe(II), and mixtures of these compounds to protect against TfdA modification by DEP is interpretted in terms of a model that includes three distinct roles for histidine residues: assisting in the binding of 2,4-D, facilitating the binding of α-KG, and ligation of the metallocenter. The illustration is not meant to depict the Fe(II) coordination number or geometry, and the indicated Fe(II) coordination by the α-ketoacid group and the phenoxy oxygen is purely speculative at this time.

Acknowledgment—We thank Katherine Nummy for technical assistance.

REFERENCES

1. Chaudhry, G. R. & Chapalamadugu, S. (1991) *Microbiol. Rev.* **55,** 59–79
2. Häggblom, M. M (1992) *FEMS Microbiol. Rev.* **103,** 29–72
3. van der Meer, J. R., de Vos, W. M., Harayama, S. & Zehnder, A. J. B. (1992) *Microbiol. Rev.* **56,** 677–694
4. Don, R. H., Weightman, A. J., Knackmuss, H.-J. & Timmis, K. N. (1985) *J. Bacteriol.* **161,** 85–90
5. Perkins, E. J., Gordon, M. P., Caceres, O. & Lurquin, P. F. (1990) *J. Bacteriol.* **172,** 2351–2359
6. Nurk, A., Kasak, L. & Kivisaar, M. (1991) *Gene (Amst.)* **102,** 13–18
7. Streber, W. R., Timmis, K. N. & Zenk, M. H. (1987) *J. Bacteriol.* **169,** 2950–2955
8. Fukumori, F. & Hausinger, R. P. (1993) *J. Bacteriol.* **175,** 2083–2086
9. Abbott, M. T. & Udenfriend, S. (1974) in *Molecular Mechanisms of Oxygen Activation* (Hayaishi, O., ed) pp. 167–214, Academic Press, New York

Downloaded from www.jbc.org by on August 5, 2009

The Journal of Biological Chemistry ASBMB

COR-AAD1-00001283

10. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene (Amst.)* **33,** 103–119
11. King, R. J., Short, K. A. & Seidler, R. J. (1991) *Appl. Environ. Microbiol.* **57,** 1790–1792
12. Wilkinson, G. N. (1961) *Biochem. J.* **80,** 324–332
13. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem.* **193,** 265–275
14. Laemmli, U. K. (1970) *Nature* **227,** 680–685
15. McKinney, M. M. & Parkinson, A. (1987) *J. Immunol. Methods* **96,** 271–278
16. Harlow, E. & Lane, D. (1988) in *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
17. Miles, E. W. (1977) *Methods Enzymol.* **47,** 431–442
18. Lindblad, B., Lindstedt, G., Lindstedt, S. & Rundgren, M. (1977) *J. Biol. Chem.* **252,** 5073–5084
19. Bradley, F. C., Lindstedt, S., Lipscomb, J. D., Que, L., Jr., Roe, A. L. & Rundgren, M. (1986) *J. Biol. Chem.* **261,** 11693–11696
20. Myllylä, R., Kuutti-Savolainen, E.-R. & Kivirikko, K. I. (1978) *Biochem. Biophys. Res. Commun.* **83,** 441–448
21. de Jong, L., Albracht, S. P. J. & Kemp, A. (1982) *Biochim. Biophys. Acta* **704,** 326–332
22. Counts, D. F., Cardinale, G. J. & Udenfriend, S. (1978) *Proc. Natl. Acad. Sci. U. S. A.* **75,** 2145–2149
23. Myllylä, R., Majamaa, K., Günzler, V., Hanauske-Abel, H. M. & Kivirikko, K. I. (1984) *J. Biol. Chem.* **259,** 5403–5405
24. Pieper, D. H., Engesser, K.-H. & Knackmuss, H.-J. (1989) *Arch. Microbiol.* **151,** 365–371
25. Haugland, R. A., Schlemm, D. J., Lyons, R. P., III, Sferra, P. R. & Chakrabarty, A. M. (1990) *Appl. Environ. Microbiol.* **56,** 1357–1362
26. Horvath, M., Ditzelmüller, G., Loidl, M. & Streichsbier, F. (1990) *Appl. Microbiol. Biotech.* **33,** 213–216
27. Said, W. A. & Lewis, D. L. (1991) *Appl. Environ. Microbiol.* **57,** 1498–1503
28. Majamaa, K., Hanauske-Abel, M. H., Günzler, V. & Kivirikko, K. I. (1984) *Eur. J. Biochem.* **138,** 239–245
29. Ng, S.-F., Hanauske-Abel, H. M. & Englard, S. (1991) *J. Biol. Chem.* **266,** 1526–1533
30. Myllylä, R., Günzler, V., Kivirikko, K. I. & Kaska, D. D. (1992) *Biochem. J.* **286,** 923–927
31. Ming, L.-J., Que, L., Kriauciunas, A., Frolik, C. A. & Chen, V. J. (1990) *Inorg. Chem.* **29,** 1111–1112
32. Scott, R. A., Wang, S., Eidsness, M. K., Kriauciunas, A., Frolik, C. A. & Chen, V. J. (1992) *Biochemistry* **31,** 4596–4601

THE JOURNAL OF BIOLOGICAL CHEMISTRY ASBMB

Downloaded from www.jbc.org by on August 5, 2009

COR-AAD1-00001284

JOURNAL OF BACTERIOLOGY, Apr. 1993, p. 2083–2086
0021-9193/93/072083-04$02.00/0
Copyright © 1993, American Society for Microbiology
Vol. 175, No. 7

Alcaligenes eutrophus JMP134 "2,4-Dichlorophenoxyacetate Monooxygenase" Is an α-Ketoglutarate-Dependent Dioxygenase

FUMIYASU FUKUMORI[1,2] AND ROBERT P. HAUSINGER[1,3]*

Center for Microbial Ecology[1] and Departments of Microbiology and Biochemistry,[3] Michigan State University, East Lansing, Michigan 48824, and Research Development Corporation of Japan, Chiyoda-ku, Tokyo 100, Japan[2]

Received 2 November 1992/Accepted 20 January 1993

The *Alcaligenes eutrophus* JMP134 *tfdA* gene, encoding the enzyme responsible for the first step in 2,4-dichlorophenoxyacetic acid (2,4-D) biodegradation, was overexpressed in *Escherichia coli*, and several enzymatic properties of the partially purified gene product were examined. Although the *tfdA*-encoded enzyme is typically referred to as 2,4-D monooxygenase, we were unable to observe any reductant-dependent activity. Rather, we demonstrate that this enzyme is a ferrous ion-dependent dioxygenase that uses α-ketoglutarate as a cosubstrate. The α-ketoglutarate is converted to succinate concomitant with 2,4-D conversion to 2,4-dichlorophenol. By using [1-^{14}C]α-ketoglutarate, we established that carbon dioxide is the second product derived from α-ketoglutarate. Finally, we verified the proposal that glyoxylate is the second product derived from 2,4-D.

The 2,4-dichlorophenoxyacetic acid (2,4-D) biodegradation pathway is a paradigm for microbial metabolism of chloroaromatic compounds (10, 27). The scheme for mineralization of 2,4-D was first established in an *Arthrobacter* species (2, 3, 18, 32, 33) and was shown to involve side chain removal, hydroxylation of the resulting 2,4-dichlorophenol (2,4-DCP), opening of the 3,5-dichlorocatechol ring, and subsequent conversion of 2,4-dichloro-*cis*,*cis*-muconate to succinate, which is metabolized by the cells. This pathway also is operative in *Alcaligenes eutrophus* JMP134, in which several of the 2,4-D degradation genes (*tfdA*, *tfdB*, *tfdC*, *tfdD*, *tfdE*, and *tfdF*) have been localized to the transmissible plasmid pJP4 (7–9). Three of the downstream genes, *tfdCDE*, exhibit extensive nucleotide sequence similarity to the well-studied chlorocatechol degradation genes *clcABD* (9a) and *tcbCDE* (34), and the corresponding chlorocatechol dioxygenase, cycloisomerase, and hydrolase steps are biochemically well understood (9, 26). Similarly, the *A. eutrophus tfdB* gene (24) exhibits significant sequence similarity to the phenol monooxygenase gene *pheA* (22), and the nicotinamide-dependent, flavin-containing 2,4-DCP hydroxylase that it encodes has been purified and well characterized (17). In contrast to the other *tfd* genes, the *A. eutrophus tfdA* gene sequence (31) exhibits no obvious evolutionary relationship to other known sequences and no biochemical characterization of the *tfdA* gene product (TFDA) has been reported.

Here, we identify the complete enzymatic reaction catalyzed by the product of the *tfdA* gene. Although TFDA is generally referred to as 2,4-D monooxygenase (9, 11, 15, 23, 25, 31), we demonstrate that the enzyme is a ferrous ion-dependent and α-ketoglutarate-dependent dioxygenase that converts α-ketoglutarate to succinate and carbon dioxide concomitant with the conversion of 2,4-D to 2,4-DCP and glyoxylate.

* Corresponding author.

MATERIALS AND METHODS

Plasmid construction and cell growth. The 3.1-kbp *Sac*I fragment of plasmid pKJS31 (31), containing the *tfdA* gene, was subcloned into the *Sac*I site of plasmid pUC19, followed by elimination of a 1.5-kbp *Xba*I fragment and religation to yield plasmid pUS311. *Escherichia coli* JM109 (35) cells were transformed with pUS311, grown overnight at 30°C on modified LB medium (containing [per liter] 10 g of tryptone, 5 g of yeast extract, and 5 g of NaCl, pH 7.2) supplemented with 100 μg of ampicillin per ml, and harvested by centrifugation. All restriction enzyme digestions, ligations, transformations, and other common DNA manipulations were performed by standard procedures (29).

Preparation of cell extracts and partial purification of TFDA. Cultures (2 liters) of *E. coli* JM109(pUS311) were grown overnight at 30 or 37°C with moderate aeration. Cells were harvested by centrifugation (6,000 × *g*) for 20 min and then suspended in 25 ml of 20 mM Tris–1 mM EDTA–0.4 μg of leupeptin per ml (pH 7.2). Crude cell extracts were obtained by passing the cells twice through a precooled French pressure cell at 120 MPa (after adjusting the buffer to 1 mM phenylmethylsulfonyl fluoride) and clarifying the debris by centrifugation at 100,000 × *g* for 30 min at 4°C. TFDA was purified 19-fold from cell extracts (obtained from cultures that were grown at 30°C) by chromatography on DEAE-Sepharose (2.6-by-30-cm column, 2 ml/min) at 4°C, followed by Mono-Q 10/10 chromatography (2 ml/min) at room temperature in the presence of 1 mM EDTA and 20 mM Tris-HCl buffer (pH 7.7 and 7.2, respectively) and by using NaCl gradients. The enzyme elutes from the resins at approximately 0.1 and 0.05 M NaCl, respectively.

Polyacrylamide gel electrophoresis. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis was carried out by using the buffers of Laemmli (16) and included a 12% running gel and a 4.5% stacking gel. Gels were stained with Coomassie brilliant blue (5). Molecular weight markers included phosphorylase *b* (M_r, 92,500), bovine serum albumin (M_r, 66,200), ovalbumin (M_r, 45,000), carbonic anhydrase (M_r, 31,000), soybean trypsin inhibitor (M_r, 21,500), and lysozyme (M_r, 14,400).

COR-AAD1-00001285

Assays. The 2,4-D degrading activities of cell extracts and partially purified TFDA were determined by measuring 2,4-DCP formation after 10 min of incubation at 30°C in 1 mM 2,4-D–10 mM MOPS (morpholinepropanesulfonic acid) buffer (pH 6.5) containing the additions indicated in Results and Discussion. 2,4-DCP was assayed by a previously described 4-aminoantipyrine method (15). The A_{510} was determined after 20 min with a Gilford Response spectrophotometer (Gilford Instrument Laboratory, Inc., Oberlin, Ohio). Alternatively, 2,4-D disappearance and 2,4-DCP formation were monitored with an HP-1050 high-performance liquid chromatograph (HPLC) (Hewlett-Packard Co.) equipped with a C_{18} (octadecyl silane) reverse-phase column (4.0 by 250 mm) at a flow rate of 1.5 ml/min in various compositions of methanol–0.1% H_3PO_4 in H_2O with concomitant monitoring of the A_{230}. One unit of activity is defined as the amount of enzyme that forms 1 μmol of phenol derivative per min at 30°C. The protein concentration was measured by colorimetric methods (19) using bovine serum albumin as a standard.

Organic acid determinations. The succinate concentration was determined by using an Aminex HPX-78H column (Bio-Rad) with an eluent of 8 mM sulfuric acid at a flow rate of 0.5 ml/min with concomitant monitoring with a differential refractive index detector. By using the same organic acid separation column, glyoxylate and α-ketoglutarate were partially resolved from each other (elution times of ~11.7 and ~11.0 min, respectively) and well separated from succinate (elution time of ~15.2 min). The identities of succinate and glyoxylate in reaction mixtures were confirmed by well-established gas chromatography-mass spectrometry (GC-MS) methods (4).

Identification of CO_2 as a product of the reaction. Partially purified TFDA was incubated with 2.1 nmol of [1-^{14}C]α-ketoglutarate (1.8 GBq/mmol; New England Nuclear) in the presence of unlabeled α-ketoglutarate (99 μM), 2,4-D (100 μM), ascorbate (50 μM), and $(NH_4)_2Fe(SO_4)_2$ (100 μM) in 2 ml of 10 mM MOPS buffer (pH 6.5) at 30°C. The reaction was stopped by the addition of 1 ml of HCl, and the released $^{14}CO_2$ was trapped in phenethylamine-wetted filter paper by published procedures (5). The amount of radioactivity in the filter paper trap was assessed by scintillation counting.

RESULTS AND DISCUSSION

Preliminary cell extract analysis. The pJP4-derived *tfdA* gene was subcloned into pUC19 and transformed into *E. coli* JM109. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis verified that the recombinant cells synthesized high levels of a peptide with an M_r of 32,000, consistent with the reported properties of TFDA (31). Similar gel electrophoresis analysis of cell fractions revealed that TFDA formed insoluble aggregates in bacterial cells at 37°C, whereas the overexpressed protein was soluble when cells were grown at 30°C. Additional experiments focused on the cells that were grown at the lower temperature. Despite the apparent abundance of TFDA, the 2,4-D degradation activity in cell extracts of these samples remained very low, with typical specific activities of approximately 0.003 μmol of 2,4-D converted to 2,4-DCP per min per mg of protein. For comparison, whole cells of *A. eutrophus* JM134(pJP4) grown on 2,4-D degrade this substrate at a rate of 0.105 μmol/min/mg of protein, as shown by substrate disappearance (25). The trace level of activity in recombinant *E. coli* cell extracts was abolished by the addition of EDTA and other chelating agents or by gel filtration chromatography or dialysis. Addition of ferrous ions restored the low levels of activity to chelator-treated samples but not to protein fractions that were freed of low-molecular-weight components. Addition of NADPH or NADH had no effect on any of the samples. The latter results prompted us to question whether this gene product is indeed a monooxygenase, as is generally stated in the literature (9, 11, 15, 23, 25, 31).

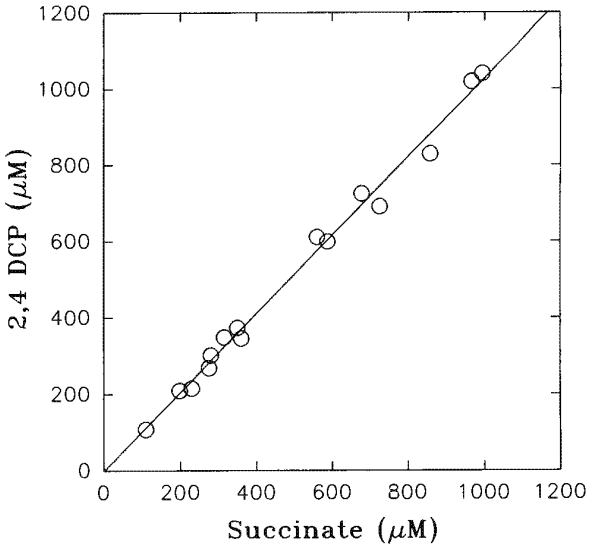


FIG. 1. Correlation between formation of 2,4-DCP and the concentration of succinate during 2,4-D degradation. The concentrations of 2,4-DCP and succinate were assessed by HPLC methods at various times after the addition of TFDA to an assay mixture containing 1 mM 2,4-D, 1 mM α-ketoglutarate, 50 μM ascorbate, 100 μM ferrous ammonium sulfate, and 10 mM MOPS buffer (pH 6.5).

Identification of TFDA as an α-ketoglutarate-dependent dioxygenase. Conversion of 2,4-D to 2,4-DCP by partially purified TFDA requires both ferrous ion and α-ketoglutarate. The activity of the partially purified enzyme was assayed in 10 mM MOPS buffer (pH 6.5) containing 1 mM 2,4-D with no additions and found to be 0.0 μmol/min/mg of protein. The activity was identical when ferrous ion alone (as ferrous ammonium sulfate) or α-ketoglutarate alone was added to the buffer mixture. When α-ketoglutarate (1 mM) and ferrous ion (100 μM) were both added, however, the activity of TFDA was 2.2 μmol/min/mg of protein. In addition, oxygen was confirmed to be required for this reaction (data not shown) and ascorbate (100 μM) was found to

TABLE 1. TFDA-catalyzed production of $^{14}CO_2$ from [1-^{14}C]α-ketoglutarate

Amt of TFDA in assay (μg)[a]	Reaction time (min)	Radioactivity trapped as CO_2	
		kcpm	%
10	10	89.5	58.5
	30	97.1	63.5
25	10	110.4	72.2
None	10	2.4	1.6
	30	3.6	2.4

[a] TFDA was incubated with 100 μM α-ketoglutarate (containing 1.53×10^5 cpm of radiolabeled material)–100 μM 2,4-D–50 μM ascorbate–100 μM ferrous ion in 2 ml of 10 mM MOPS buffer (pH 6.5) at 30°C. The released $^{14}CO_2$ was trapped as described in the text.

COR-AAD1-00001286

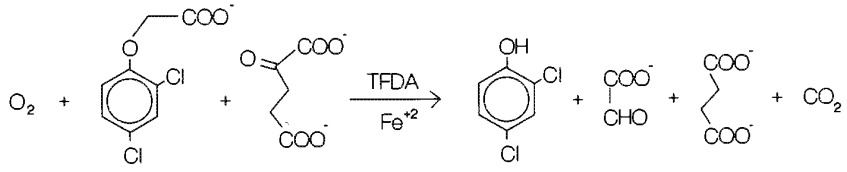


FIG. 2. Overall reaction for the activity of the enzyme encoded by the *tfdA* gene.

stimulate the activity when added along with ferrous ion and α-ketoglutarate (6.2 μmol/min/mg of protein). The α-ketoglutarate does not simply regulate the activity of the enzyme but rather is a cosubstrate in this reaction. The concentration of α-ketoglutarate decreases at the same rate as that observed for the disappearance of 2,4-D (data not shown), and a compound that coelutes with succinate is produced concomitantly with the release of 2,4-DCP (Fig. 1). Identification of succinate as a product of the reaction was confirmed by GC-MS analysis. Analysis of the enzyme kinetics demonstrated that the reaction was nearly linear, with an apparent inhibition or slow inactivation at long time points. By using the 10 min time points as reasonable estimates for initial rates, the K_ms for 2,4-D and α-ketoglutarate were determined. Over substrate concentration ranges of 10 to 50 μM for 2,4-D and 1.5 to 20 μM for α-ketoglutarate, the K_ms were estimated to be 21 ± 1.9 and 3.3 ± 0.4 μM, respectively.

Two products are generated by TFDA in addition to 2,4-DCP and succinate: glyoxylate and carbon dioxide. Early biochemical studies conducted with cell extracts of an *Arthrobacter* species provided compelling evidence that, at least in this microbe, 2,4-D is converted to 2,4-DCP and glyoxylate (32). Comparable studies have not been reported for *A. eutrophus*. Using partially purified *A. eutrophus* enzyme synthesized in *E. coli* and an organic acids-HPLC protocol, we found that a substance coeluting with glyoxylate is produced stoichiometrically with succinate. Identification of glyoxylate was verified by GC-MS analysis. By using a reaction mixture containing [1-^{14}C]α-ketoglutarate, the second product arising from α-ketoglutarate was demonstrated to be carbon dioxide (Table 1). We conclude that the overall reaction catalyzed by TFDA is that shown in Fig. 2 and propose the name 2,4-D/α-ketoglutarate dioxygenase for the product of the *tfdA* gene.

Comparison of TFDA to other α-ketoglutarate-dependent dioxygenases. The reaction carried out by 2,4-D/α-ketoglutarate dioxygenase is very similar to the reactions catalyzed by a class of hydroxylases that require ferrous ion and α-ketoglutarate (1). However, using the BESTFIT algorithm (6), we were unable to detect any significant sequence similarity between TFDA and several of these enzymes (prolyl hydroxylase [13], lysyl hydroxylase [21], hyoscyamine 6β-hydroxylase [20], aspartyl β-hydroxylase [14], and deacetoxycephalosporin synthase/hydroxylase [30]) or with the mechanistically related 4-hydroxyphenylpyruvate dioxygenase (28). Indeed, the TFDA sequence exhibited no significant similarity to any sequences available in Release 71.0 of GenBank by using the WORDSEARCH algorithm (6). Although the TFDA sequence is distinct from all other reported sequences, we propose that an enzyme analogous to TFDA functions in the catabolism of other phenoxyacetates, such as 2,4,5-trichlorophenoxyacetic acid (12), and perhaps in other biodegradative pathways. Experiments are in progress to purify and further characterize the 2,4-D/α-ketoglutarate dioxygenase.

ACKNOWLEDGMENTS

We thank Bill Holben for providing plasmid pKJS31 and John Breznak for use of the HPLC for organic acids determination.

This work was supported by NSF-Science and Technology Center Grant BIR9120006 and by the Research Development Corporation of Japan.

REFERENCES

1. **Abbott, M. T., and S. Udenfriend.** 1974. α-Ketoglutarate-coupled dioxygenases, p. 167–214. *In* O. Hayaishi (ed.), Molecular mechanisms of oxygen activation. Academic Press, New York.
2. **Bollag, J.-M., G. G. Briggs, J. E. Dawson, and M. Alexander.** 1968. 2,4-D metabolism. Enzymatic degradation of chlorocatechols. J. Agric. Food Chem. **16:**829–833.
3. **Bollag, J.-M., C. S. Helling, and M. Alexander.** 1968. 2,4-D metabolism. Enzymatic hydroxylation of chlorinated phenols. J. Agric. Food Chem. **16:**826–828.
4. **Chamberlin, B. A., and C. C. Sweeley.** 1987. Metabolic profiles of urinary organic acids recovered from absorbent filter paper. Clin. Chem. **33:**572–576.
5. **Cooper, T. G.** 1977. The tools of biochemistry, p. 102–103. John Wiley & Sons, New York.
6. **Devereux, J. P., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **10:**305–321.
7. **Don, R. H., and J. M. Pemberton.** 1981. Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J. Bacteriol. **145:**681–686.
8. **Don, R. H., and J. M. Pemberton.** 1985. Genetic and physical map of the 2,4-dichlorophenoxyacetic acid-degradative plasmid pJP4. J. Bacteriol. **161:**466–468.
9. **Don, R. H., A. J. Weightman, H.-J. Knackmuss, and K. N. Timmis.** 1985. Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4). J. Bacteriol. **161:**85–90.

9a. **Frantz, B., and A. M. Chakrabarty.** 1987. Organization and nucleotide sequence determination of a gene cluster involved in 3-chlorocatechol degradation. Proc. Natl. Acad. Sci. USA **84:** 4460–4464.

10. **Häggblom, M. M.** 1992. Microbial breakdown of halogenated aromatic pesticides and related compounds. FEMS Microbiol. Rev. **103:**29–72.
11. **Harker, A. R., R. H. Olsen, and R. J. Seidler.** 1989. Phenoxyacetic acid degradation by the 2,4-dichlorophenoxyacetic acid (TFD) pathway of plasmid pJP4: mapping and characterization of the TFD regulatory gene, *tfdR*. J. Bacteriol. **171:**314–320.
12. **Haugland, R. A., U. M. X. Sangodkar, P. R. Sferra, and A. M. Chakrabarty.** 1991. Cloning and characterization of a chromosomal DNA region required for growth on 2,4,5-T by *Pseudomonas cepacia* AC1100. Gene **100:**65–73.
13. **Helaakoski, T., K. Vuori, R. Myllylä, K. I. Kivirikko, and T. Pihlajaniemi.** 1989. Molecular cloning of the α-subunit of human prolyl 4-hydroxylase: the complete cDNA-derived amino acid sequence and evidence for alternative splicing of RNA transcripts. Proc. Natl. Acad. Sci. USA **86:**4392–4396.
14. **Jia, S., W. J. VanDusen, R. E. Diehl, N. E. Kohl, R. A. F. Dixon, D. O. Elliston, A. M. Stern, and P. A. Friedman.** 1992. cDNA cloning and expression of bovine aspartyl (asparaginyl) β-hydroxylase. J. Biol. Chem. **267:**14322–14327.
15. **King, R. J., K. A. Short, and R. J. Seidler.** 1991. Assay for detection and enumeration of genetically engineered microor-

COR-AAD1-00001287

ganisms which is based on the activity of a deregulated 2,4-dichlorophenoxyacetate monooxygenase. Appl. Environ. Microbiol. **57:**1790–1792.
16. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
17. **Liu, T., and P. J. Chapman.** 1984. Purification and properties of a plasmid-encoded 2,4-dichlorophenol hydroxylase. FEBS Lett. **173:**314–318.
18. **Loos, M. A., R. N. Roberts, and M. Alexander.** 1967. Phenols as intermediates in the decomposition of phenoxyacetates by an *Arthrobacter* species. Can. J. Microbiol. **13:**679–690.
19. **Lowry, O. H., N. J. Rosebrough, A. J. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193:**265–275.
20. **Matsuda, J., S. Okabe, T. Hashimoto, and Y. Yamada.** 1991. Molecular cloning of hyoscyamine 6β-hydroxylase, a 2-oxoglutarate-dependent dioxygenase, from cultured roots of *Hyoscyamus niger*. J. Biol. Chem. **266:**9460–9464.
21. **Myllylä, R., T. Pihlajaniemi, L. Pajunen, T. Turpeenniemi-Hujanen, and K. I. Kivirikko.** 1991. Molecular cloning of chick lysyl hydroxylase. Little homology in primary structure to the two types of subunit of prolyl 4-hydroxylase. J. Biol. Chem. **266:**2805–2810.
22. **Nurk, A., L. Kasak, and M. Kivisaar.** 1991. Sequence of the gene (*pheA*) encoding phenol monooxygenase from *Pseudomonas* sp. EST1001: expression in *Escherichia coli* and *Pseudomonas putida*. Gene **102:**13–18.
23. **Perkins, E. J., and P. F. Lurquin.** 1988. Duplication of a 2,4-dichlorophenoxyacetic acid monooxygenase gene in *Alcaligenes eutrophus* JMP134(pJP4). J. Bacteriol. **170:**5669–5672.
24. **Perkins, E. J., M. P. Gordon, O. Caceres, and P. F. Lurquin.** 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. **172:**2351–2359.
25. **Pieper, D. H., D.-H. Engesser, and H.-J. Knackmuss.** 1989. Regulation of catabolic pathways of phenoxyacetic acids and phenol in *Alcaligenes eutrophus* JMP 134. Arch. Microbiol. **151:**365–371.
26. **Pieper, D. H., A. E. Kuhn, K. Stadler-Fritzsche, P. Fischer, and H.-J. Knackmuss.** 1991. Metabolism of 3,5-dichlorocatechol by *Alcaligenes eutrophus* JMP 134. Arch. Microbiol. **156:**218–222.
27. **Rochkind-Dubinsky, M. L., G. S. Sayler, and J. W. Blackburn.** 1987. Chlorophenoxy and chlorophenyl herbicides. Microbiol. Ser. **18:**108–123.
28. **Rüetschi, •., B. Odelhög, S. Lindstedt, J. Barros-Söderling, B. Persson, and H. Jörnvall.** 1992. Characterization of 4-hydroxyphenylpyruvate dioxygenase. Primary structure of the *Pseudomonas* enzyme. Eur. J. Biochem. **205:**459–466.
29. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Samson, S. M., J. E. Dotzlaf, M. L. Slisz, F. W. Becker, R. M. Van Frank, L. E. Veal, W.-K. Yeh, J. R. Miller, S. W. Queener, and T. D. Ingolia.** 1987. Cloning and expression of the fungal expandase/hydroxylase gene involved in cephalosporin biosynthesis. Bio/Technology **5:**1207–1214.
31. **Streber, W. R., K. N. Timmis, and M. H. Zenk.** 1987. Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. **169:**2950–2955.
32. **Tiedje, J. M., and M. Alexander.** 1969. Enzymatic cleavage of the ether bond of 2,4-dichlorophenoxyacetate. J. Agric. Food Chem. **17:**1080–1084.
33. **Tiedje, J. M., J. M. Duxbury, M. Alexander, and J. E. Dawson.** 1969. 2,4-D metabolism. Pathway of degradation of chlorocatechols by *Arthrobacter* sp. J. Agric. Food Chem. **17:**1021–1026.
34. **Van der Meer, J. R., R. I. L. Eggen, A. J. B. Zehnder, and W. M. de Vos.** 1991. Sequence analysis of the *Pseudomonas* sp. strain P51 *tcb* gene cluster, which encodes metabolism of chlorinated catechols: evidence for specialization of catechol 1,2-dioxygenases for chlorinated substrates. J. Bacteriol. **173:**2425–2434.
35. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.

COR-AAD1-00001288

Acta Biotechnol. 17 (1997) 4, 351–356 Akademie Verlag

Short Communication

Isolation of Phenoxy Herbicide-Degrading *Rhodoferax* Species from Contaminated Building Material

EHRIG, A., MÜLLER, R. H., BABEL, W.

UFZ – Umweltforschungszentrum Leipzig–Halle GmbH
Sektion Umweltmikrobiologie
Permoserstraße 15
04318 Leipzig, Germany

Summary

Bacterial strains have been isolated from contaminated concrete debris which exhibit the metabolic capability to degrade 2,4-chlorinated and 4-chloro-2-methyl-substituted phenols and phenoxyalkanoic acids including phenoxyacetate and phenoxypropionate derivatives. These strains were taxonomically identified. Two of them were found to belong to the β-subgroup of the proteobacteria and showed strong similarity to *Rhodoferax fermentans*. Preliminary investigations by PCR amplification using respective primers revealed that the strains harbour *tfd*A-like gene sequences.

Introduction

In the search for alkaliphilic/alkalitolerant microorganisms capable of degrading chlorinated/methylated phenoxyalkanoic acids (which are widely applied as herbicides) we found strains with different substrate specificity in correspondence to the selective conditions applied. In each case, the source was contaminated building material from a demolished herbicide production plant. Using 2,4-D as a source of carbon and energy for growth, the enrichment resulted in a strain, *Comamonas acidovorans* P4 a, which was capable of degrading chlorinated/methylated phenoxyacetates productively, but did not attack the phenoxypropionic and phenoxybutyric derivatives [1]. With 2,4-dichlorophenoxybutyric acid (2,4-DB) as a source of carbon and energy, a consortium of strains was enriched, which functioned synergistically in the degradation of phenoxybutyric acid derivatives. First gram-positive isolates characterized as strains of *Aureobacterium testaceum* and *Rhodococcus erythropolis* (R. H. MÜLLER, H. MERTINGK and W. BABEL, submitted), etherolytically cleaved the substrates 2,4-DB and 4-chloro-2-methylphenoxybutyric acid (MCPB) resulting in the liberation of 2,4-dichlorophenol and 4-chloro-2-methylphenol, respectively. Other strains in this consortium, related to *Ochrobactrum anthropi*, were shown to be capable of completely degrading the latter compounds (MÜLLER *et al.*, submitted). No strain was available at this time with the capability to attack for instance 2,4-dichloro-

COR-AAD1-00001289

phenoxypropionate (2,4-DP) or 4-chloro-2-methylphenoxypropionate (MCPP) at alkaline pH values. Such surrounding conditions and these compounds are components in a herbicide mixture playing a role in a site currently being considered [2]. Consequently, strains were sought, which possess activities for the degradation of phenoxypropionate herbicides while tolerating an alkaline milieu. Their enrichment, isolation and preliminary description are shown in this paper.

Materials and Methods

Medium

Mineral salt solution was composed of [in mg/l]: $(NH_4)_2SO_4$, 3,760; KH_2PO_4, 680; $Ca(NO_3)_2$, 23.6; $MgSO_4 \times 7\ H_2O$, 71.2; $FeSO_4 \times 7\ H_2O$, 4.98; $CuSO_4 \times 5\ H_2O$, 0.785; $MnSO_4 \times 4\ H_2O$, 0.81; $ZnSO_4 \times 7\ H_2O$, 0.44; $Na_2MoO_4 \times 4\ H_2O$, 0.25. PYE (+F) medium contained: 3 g/l each of peptone and yeast extract and 10 mM fructose. For growth on solid medium, 2% agar (Oxoid, Agar No. 1) was used. The pH was adjusted to alkaline values (pH 8–10) with sodium hydroxide after autoclaving. Phenoxyalkanoic acid herbicides were applied as sodium salts (stock solution of 0.05 mM in 0.1 mM NaOH) after autoclaving.

Isolation

The strains were isolated from the contaminated concrete rubble of a demolished herbicide production plant. About 200 g of freshly collected debris from this site (diameter ≤ 2.5 mm) was treated by percolation with a mineral salt solution, supplied with 50 mg/l of 2,4-DP and 100 mg/l cycloheximide. This was performed in a column measuring 5 × 20 cm, under aerobic conditions, at room temperature and a pH of 9.5 ± 1; the debris material in the column was covered by a layer of glass wool. Percolate (suspension) from this fermenter was recirculated to the top of the column at a rate of 6 ml/h. The consumption of 2,4-DP was monitored daily, and the concentration readjusted to 50 mg/l as necessary. After a certain time, aliquots from the percolate and from the glass wool layer were streaked onto PYE (+F) plates, pH 9.5. The different morphotypes observed were isolated.

Degradation

Morphotypes, obtained after growth on PYE (+F) plates, were transferred into liquid PYE (+F) medium (pH 9.5) and grown overnight. Aliquots were transferred into mineral salt solution (pH 8.5 to 9) supplied with 2 g/l of sodium acetate and 0.1 g/l of yeast extract and grown to the stationary phase. Then 20 mg/l of the respective chlorinated phenols or phenoxyalkanoic acid was added, and the suspensions were incubated on a rotary shaker (200 rpm) at 30 °C. The consumption of the phenolic compounds was monitored by HPLC analysis [3].

Fatty Acid Analysis

The cultures were grown on Trypticase Soy Broth Agar (BECKTON DICKINSON, Cockeysville MD 21030, USA) at 28 °C for 24 h. Fatty acids were determined as the methyl esters by gas chromatography according to the technical notes of the MIDI Inc. automated microbial identification system (Barksdale Professional Centre, Newark DE 19711, USA), as indicated in [4].

Results and Discussion

Various strains with the metabolic capability to degrade phenoxy herbicides were isolated after percolation treatment using 2,4-DP as the sole source of carbon and

COR-AAD1-00001290

energy (Tab. 1). Strains P230, P2301, P212 a/2 showed a broad range of utilization of chlorinated phenols and phenoxyalkanoic acids including phenoxypropionates and phenoxyacetates. Utilization of certain substituted phenols is obligatory in strains with this metabolic disposition as these are intermediates in the degradation of the respective phenoxyalkanoates [5]. The broadest diversity in substrate utilization was observed with strain P230. Besides the dichlorinated compounds indicated in Tab. 1, this strain also degraded 4-chloro-2-methylphenol (MCP) within two days, 4-chloro-2-methylphenoxyacetate (MCPA) (6 days) and 4-chloro-2-methylphenoxypropionate (MCPP) (3 days). Strains P2303 and P211a showed predominant utilization of phenoxypropionates. Longer incubation with 2,4-D (two weeks) resulted in the utilization of this substrate also with strain P2303. None of the strains isolated was able to attack phenoxybutyrate derivates within 8 days of incubation. An additional set of strains only showed weak etherolytic cleavage of phenoxyalkanoic acids resulting in the liberation of the respective phenolic intermediates. Strain P212a/1 exhibited this property when tested with the three phenoxy acids, whereas the other two strains (P242 and P213a) only cleaved 2,4-dichlorophenoxybutyrate.

Tab. 1. Degradation of dichlorinated phenols and phenoxyalkanoic acids

Strain	DCP	2,4-D	2,4-DP	2,4-DB
P230	+ (1)[1)]	+ (1)	+ (1)	− (5)
P2301	+ (1)	+ (1)	+ (1)	− (3)
P212 a/2	+ (3)	+ (5)	+ (3)	− (8)
P2303	+ (2)	− (3)	+ (1)	− (2)
P211a	+ (2)	− (3)	+ (2)	− (3)
P212 a/1	− (8)	cl[2)] 0.2 (8)	cl 0.64 (8)	cl 4.5 (8)
P242	− (5)	− (5)	− (8)	cl 1.0 (8)
P213a	− (2)	− (2)	− (2)	cl 0.5 (2)

The individual substrates were supplied in doses of 20 mg/l.

1) Complete disappearance (+) or non-utilization (−) of the substrate after respective days (in brackets) of incubation.

cl[2)] The strains only cleaved the ether bond by forming DCP in an amount of [mg/l] and after the days indicated in brackets.

Strains P230 and P212 a/2 were further characterized according to microbiological and taxonomic aspects. The strains are gram-negative and exhibited positive oxidase reaction. They were motile. Strain P230 tended to form large aggregates and cell layers during growth in submerged culture; in emerged culture very rigid colonies were observed. Both aspects were not so pronounced with strain P212 a/2. The G+C content was 66.5% with strain P230 and 67.4% with strain P212 a/2 (P2301 and P2303 gave values of 68.7 and 69.0%, respectively). The fatty acid spectra of the various strains were very similar (Tab. 2). Slight differences were found with strain P242 which had a slightly higher content of C18:1 (Tab. 2). The strains used a broad spectrum of organic compounds for growth after three days of incubation including C5- and C6-carbohydrates and polyoles, (di)carboxylic acids and amino acids as tested with the Biotype 100 test kit by using Medium 2 (BIOMÉRIEUX, Marcy-l'Etoile, France). Strain P212a/2

COR-AAD1-00001291

Tab. 2. Fatty acid spectra

Strain	10:0 3OH	12:0	14:0	15:0	Sum3	Sum4	16:0	16:0m	17:0c	17:0	17:1	18:1	Sum7	19:0c
P230	3.0	3.9	1.0	3.8	4.6	23.2	31.0		20.1	0.8			6.65	
P2301	3.3	4.2	1.0	1.9	5.0	33.8	28.9		11.7				8.6	1.6
P212a/2	4.15	3.9	3.2	3.1	4.75	37.2	22.2		15.9				3.45	
P211a	3.9	3.8	0.7	1.0	5.1	37.3	29.5		9.2				9.0	0.6
P242	1.9			0.7		22.8	28.9	6.0		0.7	1.5	9.6		

c, cyclo; i, iso; m, 10-methyl.
Sum 3 (sum in Feature 3), 14:0 3OH/16:1 ISO I.
Sum 4 (sum in Feature 4), 16:1 w7c/15 iso 2OH.
Sum 7 (sum in Feature 7), 18:1 w7c/w9t/w12t.

COR-AAD1-00001292

showed a similar pattern, although slower growth was observed. Strains P230 and P212a/2 did not utilize methanol as a heterotrophic substrate. Growth of strain P230 as tested on a mixture of mannitol and succinate proceeded under aerobic and anaerobic conditions and by using both NH_4^+ and N_2 as the nitrogen source.

Partial sequence analysis of 16S rDNA by using the primer 530r revealed that strains P230 and P212a/2 belong to the β-branch of the proteobacteria. Concerning the isolate P230, a similarity of 96.6% to *Rhodoferax fermentans* strain FR2 and strain FR3 [6] was observed, as was a similarity of 95.4% to both reference strains with isolate P212a/2 (C. SPRÖER, DSMZ Braunschweig, personal communication). Preliminary investigations as to the pathway of phenoxy herbicide degradation showed that strains P230 and P212a/2 resulted in PCR amplification products by using primers derived from conserved regions of the *tfd*A gene of *Alcaligenes eutrophus* [7] (now *Ralstonia eutropha*, see [8]) equivalent in size to this reference strain. In a survey on the distribution of the *tfd*-like genes, the presence of *tfd*A has recently been shown in further strains of *Rhodoferax* [9].

Strains P230 and P212a/2 exhibit a wider substrate specificity with respect to phenoxy herbicides than strains often described in literature [5, 10]. A restricted utilization profile was shown for the most intensively investigated example, *A. eutrophus* JMP 134, which is limited to the degradation of phenoxyacetate herbicides [11, 12]. Only a few examples are known with a wider range of substrate utilization. This holds for a species of *Flavobacterium* [13] (now *Sphingomonas herbicidovorans* [14]) and of *Xanthobacter* [15] both of which degrade various phenoxyalkanoic acids with different specificities. The present strains extend this palette to alkaline conditions in particular. Investigations are underway to characterize the metabolic potential of these strains regarding the utilization of the phenolic compounds in the alkaline milieu from which they were derived.

Abbreviations

2,4-D – 2,4-Dichlorophenoxyacetate
2,4-DB – 2,4-Dichlorophenoxybutyrate
2,4-DP – 2,4-Dichlorophenoxypropionate
DCP – 2,4-Dichlorophenol
MCP – 4-Chloro-2-methylphenol
MCPP – 4-Chloro-2-methylphenoxypropionate
MCPA – 4-Chloro-2-methylphenoxyacetate
MCPB – 4-Chloro-2-methylphenoxybytyrate

Acknowledgement

The authors are indebted to Dr. C. HÄRTIG for determining the fatty acid spectra.

Received 22 September 1997
Received in revised form 22 October 1997
Accepted 28 October 1997

COR-AAD1-00001293

References

[1] HOFFMANN, D., MÜLLER, R. H., KIESEL, B., BABEL, W.: Isolation and characterization of an alkaliphilic bacterium capable of growing on 2,4-dichlorophenoxyacetic acid and 4-chloro-2-methylphenoxyacetic acid. Acta Biotechnol. **16** (1996), 121–131.

[2] MÜLLER, R. H., BABEL, W.: Biodegradative properties of alkaliphilic microorganisms from building materials of herbicide production plants. In: IChem^{E}, 2nd UK Congress of Biotechnology, Brighton, 4-6 July 1994, 73–75.

[3] OH, K. H., TUOVINEN, O. H.: Degradation of 2,4-dichlorophenoxyacetic acid by mixed cultures of bacteria. J. Ind. Microbiol. **6** (1990), 275–278.

[4] LOFFHAGEN, N., HÄRTIG, C., BABEL, W.: Fatty acid patterns of *Acinetobacter calcoaceticus* 69-V indicate sensitivity against xenobiotics. Appl. Microbiol. Biotechnol. **44** (1995), 526–531.

[5] SINTON, G. L., FAN, L. T., ERICKSON, L. E., LEE, S. M.: Biodegradation of 2,4-D and related compounds. Enzyme Microb. Technol. **8** (1986), 395–403.

[6] HIRAISHI, A., HOSHINO, Y., SATOH, T.: *Rhodoferax fermentans* gen. nov., sp. nov., a phototrophic purple nonsulphur bacterium previously referred to as the "*Rhodocyclus gelatinosus*-like" group. Arch. Microbiol. **155** (1991), 330–336.

[7] VALLAEYS, T., FULTHORPE, R. R., WRIGHT, A. M., SOULAS, G.: The metabolism of 2,4-dichlorophenoxyacetic acid degradation involves different families of *tfd*A and *tfd*B genes according to PCR-RFLP analysis. FEMS Microbiol. Ecol. **20** (1996), 163–172.

[8] YABUUCHI, E., KOSAKO, Y., YANO, I., HOTTA, H., NISHIUCHI, Y.: Transfer of two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. nov.: Proposal of *Ralstonia pickettii* (RALSTON, PALLERONI and DUODOROFF, 1973) comb. nov., *Ralstonia solanacearum* (SMITH, 1896) comb. nov. and *Ralstonia eutropha* (DAVIS, 1969) comb. nov. Microbiol. Immun. **39** (1995), 897–904.

[9] FULTHORPE, R. R., MCGOWAN, C., MALTSEVA, O. V., HOLBEN, W. E., TIEDJE, J. M.: 2,4-Dichlorophenoxyacetic acid-degrading bacteria contain mosaics of catabolic genes. Appl. Environ. Microbiol. **61** (1995), 3274–3281.

[10] HÄGGBLOM, M. M.: Microbial breakdown of halogenated aromatic pesticides and related compounds. FEMS Microbiol. Rev. **103** (1992), 29–72.

[11] DON, R. H., PEMBERTON, J. M.: Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J. Bacteriol. **145** (1981), 681–686.

[12] PIEPER, D. H., REINEKE, W., ENGESSER, K. H., KNACKMUSS, H. J.: Metabolism of 2,4-dichlorophenoxyacetic acid, 4-chloro-2-methylphenoxyacetic acid and 2-methylphenoxyacetic acid by *Alcaligenes eutrophus* JMP 134. Arch. Microbiol. **150** (1988), 95–102.

[13] HORVATH, M., DIETZELMÜLLER, G., LOIDL, M., STREICHSBIER F.: Isolation and characterization of a 2-(2,4-dichlorophenoxy)propionic acid-degrading soil bacterium. Appl. Microbiol. Biotechnol. **33** (1990), 213–216.

[14] ZIPPER, C., NICKEL, K., ANGST, W., KOHLER, H. P. E.: Complete microbial degradation of both enantiomers of the chiral herbicide mecoprop [(*RS*)-2-(4-chloro-2-methylphenoxy)propionic acid] in an enatioselective manner by *Sphingomonas herbicidovorans* sp. nov. Appl. Environ. Microbiol. **62** (1996), 4318–4322.

[15] DIETZELMÜLLER, G., LOIDL, M., STREICHSBIER, F.: Isolation and characterization of a 2,4-dichlorophenoxyacetic acid-degrading soil bacterium. Appl. Microbiol. Biotechnol. **31** (1989), 93–96.

COR-AAD1-00001294

Critical Reviews in Biochemistry and Molecular Biology, 39:21–68, 2004
Copyright © Taylor & Francis Inc.
ISSN: 1040-9238
DOI: 10.1080/10409230490440541

Taylor & Francis

Fe(II)/α-Ketoglutarate-Dependent Hydroxylases and Related Enzymes

Robert P. Hausinger

Department of Microbiology and Molecular Genetics and Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan, USA

Fe(II)/α-ketoglutarate (αKG)-dependent hydroxylases catalyze an amazing diversity of reactions that result in protein side-chain modifications, repair of alkylated DNA/RNA, biosynthesis of antibiotics and plant products, metabolism related to lipids, and biodegradation of a variety of compounds. These enzymes possess a β-strand "jellyroll" structural fold that contains three metal-binding ligands found in a His^1-X-Asp/Glu-X_n-His^2 motif. The co-substrate, αKG, chelates Fe(II) using its C-2 keto group (binding opposite the Asp/Glu residue) and C-1 carboxylate (coordinating opposite either His^1 or His^2). Oxidative decomposition of αKG forms CO_2 plus succinate and leads to the generation of an Fe(IV)-oxo or other activated oxygen species that hydroxylate the primary substrate. The reactive oxygen species displays alternate reactivity in related enzymes that catalyze desaturations, ring expansions, or ring closures. Other enzymes resemble the Fe(II)/αKG-dependent hydroxylases in terms of protein structure or chemical mechanism but do not utilize αKG as a substrate. This review describes the reactions catalyzed by this superfamily of enzymes, highlights key active site features revealed by structural studies, and summarizes results from spectroscopic and other approaches that provide insights into the chemical mechanisms.

Keywords hydroxylase, dioxygenase, α-ketoglutarate, ferrous ion, metalloenzyme

Editor: James A. Imlay.

Address correspondence to Robert P. Hausinger, 6193 Biomedical Physical Science, Michigan State University, East Lansing, MI 48824-4320, USA. E-mail: hausinge@msu.edu

Abbreviations: **ACC**: 1-aminocyclopropane-1-carboxylate; **αKG**: alpha-ketoglutarate; **ANS**: anthocyanidin synthase; **AtsK**: alkyl sulfatase; **CarC**: carbapenem synthase; **CAS**: clavaminate synthase; **CD**: circular dichroism; **2,4-D**: 2,4-dichlorophenoxyacetic acid; **DACS**: deacetylcephalosporin C synthase; **DAOCS**: deacetoxycephalosporin C synthase; **eIF-5A**: eukaryotic initiation factor 5A; **ENDOR**: electron nuclear double resonance; **EPR**: electron paramagnetic resonance; **ESEEM**: electron spin-echo envelop modulation; **FIH**: factor-inhibiting HIF (HIF-specific asparaginyl hydroxylase); **HIF**: hypoxia-inducible factor; **HPPD**: 4-hydroxyphenylpyruvate dioxygenase; **IPNS**: isopenicillin N synthase; **LMCT**: ligand-to-metal charge transfer; **MCD**: magnetic circular dichroism; **MLCT**: metal-to-ligand charge-transfer; **NO**: nitric oxide; **PDB**: protein database; **rR**: resonance Raman; **TauD**: taurine/αKG dioxygenase; **TfdA**: 2,4-D/αKG dioxygenase; **UV**: ultraviolet; **VHL**: von Hippel-Lindau tumor suppressor protein; **XAS**: X-ray absorption spectroscopy.

I. INTRODUCTION

This review of Fe(II)/α-ketoglutarate (αKG)-dependent hydroxylases and related enzymes begins (see Section II) by describing the amazing diversity of reactions catalyzed by members of this enzyme superfamily. Most representatives couple the oxidative decomposition of αKG (forming CO_2 and succinate) to the hydroxylation of a cosubstrate, as illustrated in Scheme 1 (the incorporated oxygen atom is shown in bold here and in the following schemes). The various types of primary substrates recognized by these enzymes include proteins, methylated nucleotides, lipids, and a wide range of small molecules that, depending on the enzyme, are used for synthesis or undergo decomposition. Other selected family members catalyze desaturation, ring expansion, ring formation, or other types of oxidative reactions. Also discussed are enzymes that function independently of αKG, yet that are related to the Fe(II)/αKG-dependent hydroxylases by sequence or mechanism. This superfamily of enzymes is united by their requirement for Fe(II), which is weakly bound by three amino acid side chains that typically occur in a His^1-X-Asp/Glu-X_n-His^2 motif (Table 1). When utilized as a substrate, αKG chelates the Fe(II) using its C-1 carboxylate and C-2 ketone, with additional binding stabilization provided by the C-5 carboxylate interacting with other side chains. Section III highlights the key active site features revealed by the rapidly expanding number of structural studies being carried out with this enzyme family (Table 2). Finally, insights into the mechanism of these enzymes from structural, spectroscopic, isotope incorporation, and

COR-AAD1-00001295

TABLE 1
Metal ligands and αKG-binding residues of selected Fe(II)/αKG-dependent hydroxylases and related enzymes

Enzyme	Ligand 1	Ligand 2	Spacing	Ligand 3	Spacing	αKG C-5 stabilization
Human prolyl 4-hydroxylase	His 412	Asp 414	68	His 483	9	Lys 493
Human lysyl hydroxylase	His 638	Asp 640	49	His 690	9	Arg 700
Human HIFα-specific prolyl hydroxylase-1	His 135	Asp 137	58	His 196	8	Arg 205
Human FIH	His 199	Asp 201	77	His 279	NA	Lys 214, Tyr 145, Thr 196
Escherichia coli AlkB	His 131	Asp 133	53	His 187	16	Arg 204
Streptomyces clavuligerus CAS	His 144	Glu 146	132	His 279	13	Arg 293, Thr 172
Streptomyces clavuligerus DAOCS	His 183	Asp 185	57	His 243	14	Arg 258, Ser 260
Erwinia carotovora CarC	His 101	Asp 103	147	His 251	11	Arg 263, Thr 130, Arg 253
Penicillium decumbens EpoA	His 131	Asp 133	40	His 174	10	Arg 185
Dactylosporium sp. proline 4-hydroxylase	His 109	Asp 111	103	His 215	10	Arg 226
Streptomyces sp. strain TH1 proline 3-hydroxylase	His 107	Asp 109	48	His 158	9	Arg 168, Ser 170, His 135
Petunia hydrida flavanone 3β-hydroxylase	His 220	Asp 222	55	His 278	9	Arg 288, Ser 290
Citrus unshiu flavonol synthase	His 221	Asp 223	53	His 277	9	Arg 287, Ser 289
Arabidopsis thaliana ANS	His 232	Asp 234	53	His 288	9	Arg 298, Tyr 217, Asn 215
Cucurbita maxima gibberellin 7-oxidase	His 182	Asp 184	53	His 238	9	Arg 248, Ser 250
Cucurbita maxima gibberellin 20-oxidase	His 243	Asp 245	53	His 299	9	Arg 309, Ser 311
Caranthus roseus desacetoxyvindoline-4-hydroxylase	His 308	Asp 310	55	His 366	9	Arg 376, Ser 378
Hyoscyamus niger hyoscyamine 6β-hydroxylase	His 217	Asp 219	54	His 274	9	Arg 284, Ser 286
Human trimethyllysine hydroxylase	His 201	Glu 203	146	His 350	13	Arg 364
Pseudomonas sp. AK1 β-butyrobetaine hydroxylase	His 209	Asp 211	138	His 350	11	Arg 362
Human β-butyrobetaine hydroxylase	His 202	Asp 204	142	His 347	12	Arg 360
Human phytanoyl-CoA hydroxylase	His 175	Asp 177	86	His 264	10	Arg 275
Salmonella typhimurium LpxO	His 155	Asp 157	44	His 202	9	Arg 212
Escherichia coli TauD	His 99	Asp 101	153	His 255	10	Arg 266, Thr 126
Pseudomonas putida S-313 AtsK	His 108	Asp 110	153	His 264	10	Arg 275, Thr 135
Pseudomonas stutzeri WM88 hypophosphite/αKG dioxygenase	His 116	Asp 118	87	His 206	10	Arg 217
Ralstonia eutropha TfdA	His 114	Asp 116	146	His 263	10	Arg 274
Pseudomonas syringae pv. *Phaseolicola* ethylene-forming enzyme	His 189	Asp 191	76	His 268	8	Arg 277
Escherichia coli Gab protein	His 160	Asp 162	129	His 292	12	Arg 305
Aspergillus (Emericella) nidulans IPNS	His 214	Asp 216	53	His 270	NA	NA
Tomato ACC oxidase	His 177	Asp 179	54	His 234	9	Arg 244

NA, not applicable.

biomimetic studies are described in Section IV. Notably, our current understanding of the hydroxylase mechanism represents a detailed refinement of a mechanistic proposal that was published over 20 years ago (Hanauske-Abel & Günzler, 1982). This review serves to update a series of previous surveys of this field (Abbott & Udenfriend, 1974; De Carolis & De Luca, 1994; Prescott, 1993; Prescott & Lloyd, 2000; Schofield & Zhang, 1999; Udenfriend & Cardinale, 1982) and to highlight directions for future investigations.

COR-AAD1-00001296

TABLE 2
Summary of structurally characterized Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes

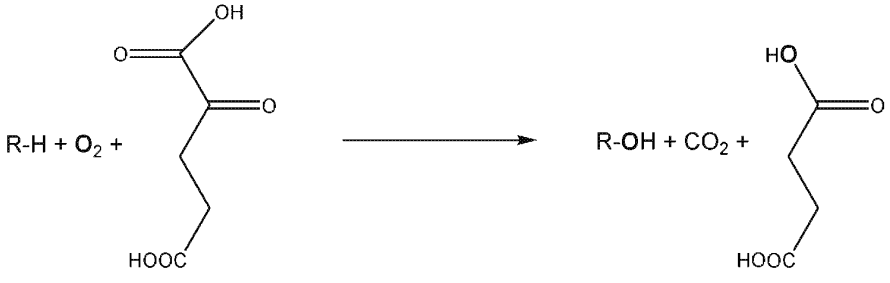
Enzyme	Molecules bound	Resolution (Å)	PDB code	Reference
Human FIH	Apoprotein	2.8	1IZ3	Lee *et al.*, 2003
	Fe(II)	2.2	1MZE	Dann *et al.*, 2002
	Fe(II) + αKG	2.4	1MZF	Dann *et al.*, 2002
	Fe(II) + *N*-oxaloylglycine + HIFα C-terminal peptide	2.15	1H2K	Elkins *et al.*, 2003
	Fe(II) + αKG + HIFα C-terminal peptide	2.25	1H2L	Elkins *et al.*, 2003
	Zn(II) + *N*-oxaloylglycine + HIFα C-terminal peptide	2.5	1H2M	Elkins *et al.*, 2003
	Fe(II) + αKG + HIFα C-terminal peptide	2.82	1H2N	Elkins *et al.*, 2003
Streptomyces clavuligerus CAS	Apoprotein	1.63	1DS0	Zhang *et al.*, 2000
	Fe(II) + αKG	1.08	1DS1	Zhang *et al.*, 2000
	Fe(II) + αKG + N-α-acetyl-L-arginine	1.4	1DRY	Zhang *et al.*, 2000
	Fe(II) + αKG + proclavaminic acid	2.1	1DRT	Zhang *et al.*, 2000
	Fe(II) + αKG + deoxyguanidinoproclavaminic acid + NO	1.54	1GVG	Zhang *et al.*, 2002
Streptomyces clavuligerus DAOCS	Apoprotein	1.3	1DCS	Valegård *et al.*, 1998
	Fe(II)	1.5	1RXF	Valegård *et al.*, 1998
	Fe(II) + αKG	1.5	1RXG	Valegård *et al.*, 1998
	Fe(II) + succinate	1.5	1UO9	Valegård *et al.*, 2004
	Fe(II) + αKG + ampicillin	1.5	1UNB	Valegård *et al.*, 2004
	Fe(II) + penicillin G	1.6	1UOF	Valegård *et al.*, 2004
	Fe(II) + αKG + penicillin G	1.7	1UOB	Valegård *et al.*, 2004
	Fe(II) + deacetoxycephalosporin C	1.7	1UOG	Valegård *et al.*, 2004
DAOCS R258Q	Fe(II) + α-keto-β-methyl-butanoate	1.5	1HJG	Lee *et al.*, 2001b
DAOCS R258Q	Fe(II) + α-keto-β-methyl-pentanoate	1.6	1HJF	Lee *et al.*, 2001b
DAOCS deltaR306	Fe(II) + αKG	2.1	1E5I	Lee *et al.*, 2001a
DAOCS deltaR307A	Fe(II) + succinate + carbon dioxide	1.96	1E5H	Lee *et al.*, 2001a
Erwinia carotovora CarC	Fe(II) + αKG	2.4	1NX4	Clifton *et al.*, 2003
	Fe(II) + αKG + L-*N*-acetylproline	2.3	1NX8	Clifton et al., 2003
Streptomyces sp. strain TH1 proline 3-hydroxylase	Apoprotein	2.3	1E5R	Clifton *et al.*, 2001
	Fe(II)	2.4	1E5S	Clifton *et al.*, 2001
Arabidopsis thaliana ANS	Apoprotein + αKG	2.1	1GP4	Wilmouth *et al.*, 2002
	Fe(II) + αKG + dihydroquercetin	2.2	1GP5	Wilmouth *et al.*, 2002
	Fe(II) + αKG + dihydroquercetin + 30 min oxygen exposure	1.75	1GP6	Wilmouth *et al.*, 2002
Escherichia coli TauD	Apoprotein	1.9	1OTJ	O'Brien *et al.*, 2003
	Fe(II) + αKG + taurine	2.5	1GY9	Elkins *et al.*, 2002
	Fe(II) + αKG + taurine	2.5	1OS7	O'Brien *et al.*, 2003

(Continued on next page)

COR-AAD1-00001297

TABLE 2
Summary of structurally characterized Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes
(Continued)

Enzyme	Molecules bound	Resolution (Å)	PDB code	Reference
Pseudomonas putida S-313 AtsK	Sodium ion	1.9	1OIH	Müller *et al.*, 2004
	Na + αKG	2.07	1OIJ	Müller *et al.*, 2004
	Fe(II) + αKG	2.2	1OII	Müller *et al.*, 2004
	Fe(II) + αKG + 2-ethylhexyl-1-sulfate	2.05	1OIK	Müller *et al.*, 2004
Escherichia coli Gab protein	Apoprotein	2	1JR7	Chance *et al.*, 2002
Aspergillus (Emericella) nidulans IPNS	Mn(II)	2.5	1IPS	Roach *et al.*, 1995
	Fe(II) + δ-(L-α-aminoadipoyl)-L-cysteinyl-D-valine	1.3	1BK0	Roach *et al.*, 1997
	Fe(II) + δ-(L-α-aminoadipoyl)-L-cysteinyl-D-valine + NO	1.45	1BLZ	Roach *et al.*, 1997
	Fe(II) + isopenicillin *N*	1.35	1QJE	Burzlaff *et al.*, 1999
	Fe(II) + δ-(L-α-aminoadipoyl)-L-cysteinyl-L-*S*-cysteine	1.5	1QIQ	Burzlaff *et al.*, 1999
	Fe(II) + δ-(L-α-aminoadipoyl)-L-cysteinyl-L-*S*-cysteine + oxygen	1.4	1QJF	Burzlaff *et al.*, 1999
	Fe(II) + δ-(L-α-aminoadipoyl)-L-cysteinyl-D-α-hydroxyvaleryl ester	1.55	1HB1	Ogle *et al.*, 2001
	Fe(II) + δ-(L-α-aminoadipoyl)-L-cysteinyl-D-α-hydroxyvaleryl ester + oxygen 30 s	1.3	1HB2	Ogle *et al.*, 2001
	Fe(II) + δ-(L-α-aminoadipoyl)-L-cysteinyl-D-α-hydroxyvaleryl ester + oxygen 30 s	1.4	1HB3	Ogle *et al.*, 2001
	Fe(II) + δ-(L-α-aminoadipoyl)-L-cysteinyl-D-α-hydroxyvaleryl ester + oxygen 120 s	1.5	1HB4	Ogle *et al.*, 2001
	Fe(II) + δ-(L-α-aminoadipoyl)-L-cysteinyl-D-α-aminobutyrate	2.2	1OC1	Long *et al.*, 2003
	Fe(II) + δ-(L-α-aminoadipoyl)-L-cysteinyl-D-α-aminobutyrate + NO	1.3	1OBN	Long *et al.*, 2003
Pseudomonas fluorescens HPPD	Fe(II)	2.4	1CJX	Serre *et al.*, 1999



SCHEME 1.

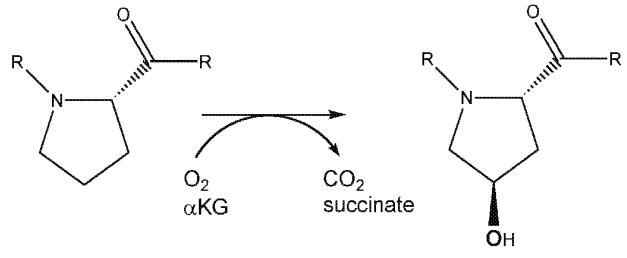
COR-AAD1-00001298

II. REACTIONS OF Fe(II)/α-KETOGLUTARATE-DEPENDENT HYDROXYLASES AND RELATED ENZYMES

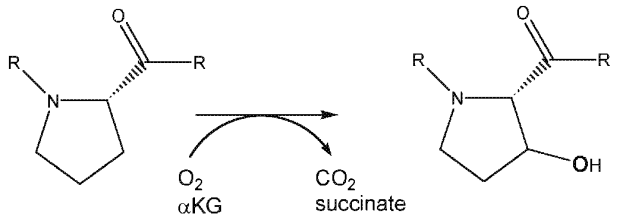
A. Protein Modification

At least four distinct Fe(II)/αKG-dependent hydroxylases utilize the side chains of proteins as substrates, resulting in the modification of proline, lysine, aspartic acid, and asparagine residues. Studies of these enzymes range from the earliest work on Fe(II)/αKG-hydroxylases through very recent investigations related to hypoxia-induced cellular signaling events.

Prolyl 4-hydroxylase, the first Fe(II)/αKG-dependent hydroxylase identified (Hutton *et al.*, 1966), catalyzes the reaction illustrated in Scheme 2 to yield the *trans*-4-hydroxyprolyl group. In mammals, this activity is essential to the formation of collagens, elastins, and several other proteins (Kivirikko & Pihlajaniemi, 1998). For example, type I collagen contains lengthy Gly-X-Y repeats with about 10% of the Y positions occupied by 4-hydroxyproline. The hydroxylated residue forms both inter- and intrachain hydrogen bonds, thus stabilizing the protein's unique triple-helical structure. In plants, prolyl 4-hydroxylase targets precursors of the extensions, repetitive proline-rich proteins, arabinogalactan-proteins, and other key components of the extracellular matrix and reproductive tissues (Kieliszewski & Lamport, 1994; Wu *et al.*, 2001). Genes encoding prolyl 4-hydroxylase have been cloned from human (Helaakoski *et al.*, 1989), plant (Hieta & Myllyharju, 2002), insect (Annunem *et al.*, 1999), nematode (Veijola *et al.*, 1994), and other sources, including *Paramecium bursaria Chlorella* virus-1 (Ericksson *et al.*, 1999), suggesting a role for 4-hydroxyproline in this algal-specific virus. In most cases, the enzyme is comprised of two distinct subunits, the hydroxylase α subunit and a protein disulfide isomerase β subunit which maintains solubility; however, algal and plant prolyl 4-hydroxylases are monomeric proteins (Hieta & Myllyharju, 2002; Kaska *et al.*, 1988). Results from site-directed mutagenesis studies of the human enzyme were interpreted to suggest that His 412, Asp 414, and His 483 coordinate Fe(II), Lys 493 binds the C-5 carboxylate of αKG (with the K493R variant retaining 15% of the wild-type enzyme activity), and His



SCHEME 2.

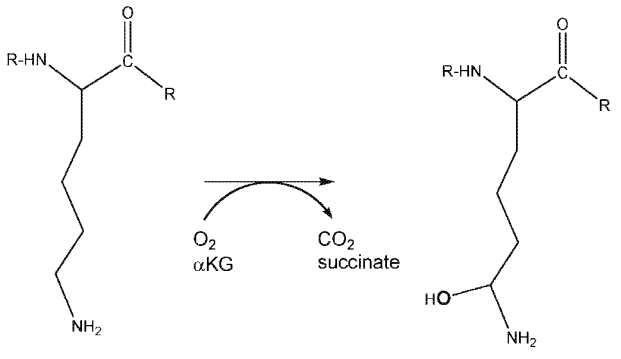


SCHEME 3.

501 having a critical but undefined role (Table 1) (Lamberg *et al.*, 1995; Myllyharju & Kivirikko, 1997). The peptide substrate-binding domain of human collagen prolyl 4-hydroxylase was crystalized (Pekkala *et al.*, 2003), but this domain is separate from the catalytic domain and no structure has been reported.

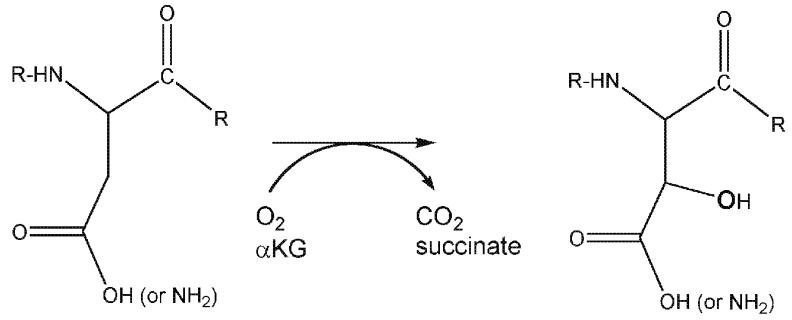
Much less information is available for prolyl 3-hydroxylase, an enzyme catalyzing the reaction shown in Scheme 3. This activity has been resolved from prolyl 4-hydroxylase (Tryggvason *et al.*, 1976) and partially purified from several sources (Risteli *et al.*, 1977; Tryggvason *et al.*, 1979); however, no subunit or sequence information is available. Procollagen is the only known natural substrate of this enzyme, with mature collagen containing 3-hydroxyproline levels up to 10% of the amount of 4-hydroxyproline (Kivirikko & Pihlajaniemi, 1998).

Lysyl hydroxylase is a medically important enzyme that catalyzes the reaction indicated in Scheme 4. The enzyme hydroxylates 0.5% to 7% of the X-Lys-Gly triplets of procollagen, with the modification serving as the attachment site for galactose or glucosylgalactose units (Kivirikko & Pihlajaniemi, 1998). In addition, many other proteins are known to contain this modified amino acid. Defects in lysyl hydroxylase activity are associated with the type VI variant of Ehlers-Danlos syndrome, leading to various connective tissue problems (e.g., joint hypermobility, skin hyperextensibility, and skin fragility). Three related isoforms of lysyl hydroxylase have been identified in humans (Hautala *et al.*, 1992; Passoja *et al.*, 1998; Valtavaara *et al.*, 1997,



SCHEME 4.

COR-AAD1-00001299



SCHEME 5.

1998). Site-directed mutagenesis studies were carried out on one of the human genes, and the results obtained with the variant homodimeric enzymes were interpreted to suggest that His 638, Asp 640, and His 690 (Table 1, residue numbers after processing of the protein) are metal ligands, Arg 700 stabilizes the C-5 carboxylate of αKG (with the K700R variant retaining about 10% activity), and His 688 plays an additional critical role (Passoja *et al.*, 1998; Pirskanen *et al.*, 1996).

A specific lysyl residue of eukaryotic initiation factor 5A (eIF-5A) undergoes a two-step modification, including hydroxylation, to produce hypusine (*N*-ε-(4-amino-2-hydroxybutyl) lysine) (Park *et al.*, 1993). The NAD-dependent enzyme deoxyhypusine synthase first transfers the butylamine group of spermine to the eIF-5A lysyl group in a reversible reaction (Park *et al.*, 2003), then deoxyhypusyl hydroxylase completes the transformation. Initial studies of the hydroxylase activity led to the conclusion that enzyme reactivity was independent of Fe(II) and αKG (Abbruzzese *et al.*, 1986); however, later inhibition studies with metal chelators provided evidence of an essential Fe(II) binding site (Abbruzzese *et al.*, 1991; Clement *et al.*, 2002). More recently, deoxyhypusyl hydroxylase has been purified and shown to be a member of the Fe(II)/αKG-dependent hydroxylases (George McLendon, personal communication). The role of this modified residue is unknown, but the eIF-5A protein is essential for cell proliferation.

Aspartyl (asparaginyl) β-hydroxylase modifies the β-carbon of specific aspartic acid or asparagine residues (Scheme 5) in the epidermal growth factor-like domains of a number of vitamin K-dependent proteins, coagulation factors, and complement factors. The modified residue forms a calcium-binding site, likely to be important for protein–protein interactions. Since its discovery in 1989 (Gronke *et al.*, 1989; Stenflo *et al.*, 1989), the monomeric enzyme was purified and characterized (Wang *et al.*, 1991) and the gene has been cloned and expressed from several sources, including humans (Jia *et al.*, 1992; Korioth *et al.*, 1994). Mutagenesis studies of the bovine gene led to the proposal that His 675 is one of the metal ligands, where substitution of this residue by Asp or Glu retains some activity (Jia *et al.*, 1994; McGinnis *et al.*, 1996). Notably, His 675 is not preceded by a His-X-Asp/Glu sequence, so the identity of additional metal ligands remains unclear. Also of interest, overexpression of human aspartyl (asparaginyl) β-hydroxylase in NIH-3T3 cells is associated with malignant transformation—for unknown reasons. (Ince *et al.*, 2000).

Although protein hydroxylases have been studied for decades, interest in this enzyme group has intensified since the discovery, in 2001, of their involvement in the hypoxic response. Oxygen deprivation, or hypoxia, occurs during tumor formation, diabetes, and other ischemic diseases, leading to hypoxia-inducible factor (HIF)-dependent transcription of a wide variety of genes. The two subunits of this heterodimeric transcription factor (HIFα and HIFβ) are constitutively synthesized, but HIFα is targeted for degradation in cells containing normal oxygen concentrations. Two enzymes catalyzing side-chain hydroxylation are involved in this process (Hewitson *et al.*, 2003).

A prolyl 4-hydroxylase plays a key role in targeting HIFα for degradation on the basis of evidence indicating that Pro 564 of HIFα is hydroxylated in an oxygen-, Fe(II)-, and αKG-dependent reaction (Ivan *et al.*, 2001; Jaakkola *et al.*, 2001). The hydroxylated HIFα forms a complex with the von Hippel-Lindau tumor suppressor protein (VHL), elongin B, and elongin C, resulting in polyubiquitinylation and destruction of the transcription factor subunit. The hypoxia-mimicking effects of cobalt and desferrioxamine are thus understood to result from inhibition of the prolyl 4-hydroxylase activity (Minchenko *et al.*, 2002). Similarly, a lack of VHL-directed degradation of HIFα explains the elevated expression of hypoxia-inducible genes caused by suppression of *Drosophila melanogaster* prolyl 4-hydroxylase by RNA interference (Bruick & McKnight, 2001). The structural basis for the hydroxylation-specific interaction was elucidated by two crystallography studies involving a complex comprised of a Pro 564-containing peptide of HIFα, VHL, and the two elongin proteins (Hon *et al.*, 2002; Min *et al.*, 2002). The hydroxyproline residue nicely fits into a gap of the

COR-AAD1-00001300

VHL core with optimized hydrogen bonding interactions. The VHL protein was shown to ubiquitinylate two independently-acting sites within HIFα (Masson *et al.*, 2001). EGL-9 protein in *C. elegans* was predicted to be an Fe(II)/αKG-dependent hydroxylase (Aravind & Koonin, 2001) and later shown to be a HIFα-specific prolyl 4-hydroxylase (Epstein *et al.*, 2001). A family of isoforms exists in humans—each containing the signature sequence of Fe(II)/αKG-dependent hydroxylases (Table 1) (Bruick & McKnight, 2001; Epstein *et al.*, 2001; Oehme *et al.*, 2002). No crystal structure has yet been reported for a HIFα-specific prolyl hydroxylase; however, two sets of coordinates have been deposited in the Protein Data Bank (PDB; access codes 1IZF and 1IZG), so this information may be forthcoming.

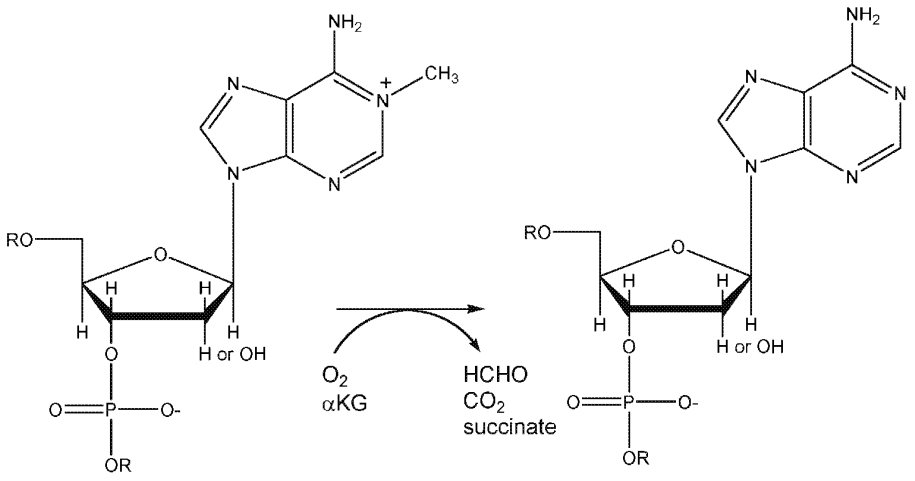
Superimposed on the hypoxia-inducible effects associated with the HIFα-specific prolyl 4-hydroxylase is a process involving hydroxylation of Asn 803. In the absence of hydroxylation the HIFα domain binds to the p300 protein with the target Asn deeply buried at the molecular interface (Dames *et al.*, 2002; Freedman *et al.*, 2002). Hydroxylation of Asn 803 in the carboxyl-terminal transactivation domain of HIFα prevents its interaction with the p300 transcription coactivator (Lando *et al.*, 2002). Interestingly, the stereospecificity of the hydroxylation reaction is distinct from that observed for the well-studied aspartyl/asparaginyl hydroxylases, resulting in the *threo*-isomer rather than the *erythro*-isomer (McNeill *et al.*, 2002a). A yeast two-hybrid approach was used to identify the HIFα-specific asparaginyl hydroxylase that was shown to be identical to the "factor-inhibiting HIF" (FIH), a protein known to interact with HIFα and VHL and to repress HIF transcriptional activity (Hewitson *et al.*, 2002; Mahon *et al.*, 2001). As summarized in Table 2, the structures of the apoprotein (Lee *et al.*, 2003), the Fe(II)- and αKG-bound form of FIH (Dann *et al.*, 2002), and the complex between FIH and a carboxyl-terminal domain of HIFα (Elkins *et al.*, 2003) were elucidated. The metal ligands are His 199, Asp 201, and His 279, whereas Tyr 145, Thr 196, and Lys 214 stabilize binding of the αKG C-5 carboxylate in a unique type of interaction (Table 1). The HIFα Asn 803 is ideally oriented to undergo the observed reaction at the FIH active site, but oxidation during crystallization was prevented by the use of anaerobic conditions.

B. Repair of Alkylated DNA/RNA

The newest function ascribed to the Fe(II)/αKG-dependent hydroxylases is repair of alkylation-damaged DNA and RNA. Cellular polynucleotides are alkylated by endogenous components, such as *S*-adenosylmethionine, and by reaction with two general classes of environmental and laboratory chemicals. S_N1 agents include methylnitrosourea and *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine that react with the N^7 position of guanine, N^3 of adenine, O^6 of guanine, O^2 or O^4 of pyrimidines, and the nonphosphodiester oxygen atoms of the phosphate backbone. In contrast, S_N2 agents such as methylmethanesulfonate and dimethylsulfate react primarily with the N^1 position of adenine and N^3 of cytosine (Sedgwick & Lindahl, 2002). To overcome the mutagenic and toxic effects of these modifications, cells produce a variety of DNA repair enzymes. For example, *Escherichia coli* possesses two distinct 3-methyladenine DNA glycosylases, encoded by *tag* and *alkA*, that remove the offending alkyl groups along with the attached adenines to create abasic sites in the DNA product. Alternatively, *ogt* and *ada* encode O^6-methylguanine methyltransferases, while *ada* also encodes O^4-methylthymine methyltransferase and methylphosphotriester methyltransferase activities. These single-turnover proteins transfer the target methyl groups to their active site cysteine residues, thus inactivating their activities. While *tag* and *ogt* are constitutively expressed, methyl transfer to Cys 69 of the *ada* gene product converts it into an activator that increases expression of *alkA*, *ada*, *aidB* (somehow involved in protection against S_N1 agents), and *alkB*. As described below, *alkB* encodes an Fe(II)/αKG-dependent hydroxylase that directly repairs 1-alkyladenine and 3-alkylcytosine lesions in DNA or RNA (Schemes 6 and 7).

Nearly 20 years were required to identify the function of *alkB* following its initial description as an *E. coli* gene that modestly affects reactivation of methylmethanesulfonate-treated phage lambda (Kataoka *et al.*, 1983). The gene was cloned and sequenced (Kataoka & Sekiguchi, 1985) and the protein purified (Kondo *et al.*, 1986), but without any clarification of its role. Expression of *E. coli alkB* in human cells was shown to confer resistance to high concentrations of S_N2 alkylating agents (Chen *et al.*, 1994), demonstrating that the enzyme functions independently of other cellular components. In contrast to the weak effects observed for *alkB* mutants on reactivation of methylated double-stranded phage (lambda and M13), pronounced deficiencies were noted for reactivation of the methylmethanesulfonate-treated single-stranded DNA viruses (M13, f1, and G4) (Dinglay *et al.*, 2000) and a single-stranded RNA virus (Aas *et al.*, 2003). These results are most compatible with repair of 1-methyladenine and 3-methylcytosine lesions that are preferentially formed by methylation of single-strand polynucleotides. Critical insight into the role of this enzyme was derived from iterative sequence analyses that identified AlkB as a member of the Fe(II)/αKG-dependent dioxygenase superfamily (Aravind & Koonin, 2001). This suspicion was confirmed by studies demonstrating direct repair of methylated DNA by AlkB in an Fe(II)-dependent process that consumes oxygen plus αKG and produces succinate plus formaldehyde (Trewick *et al.*, 2002), thus resulting in release of replication blocks (Falnes *et al.*, 2002). Subsequently, direct

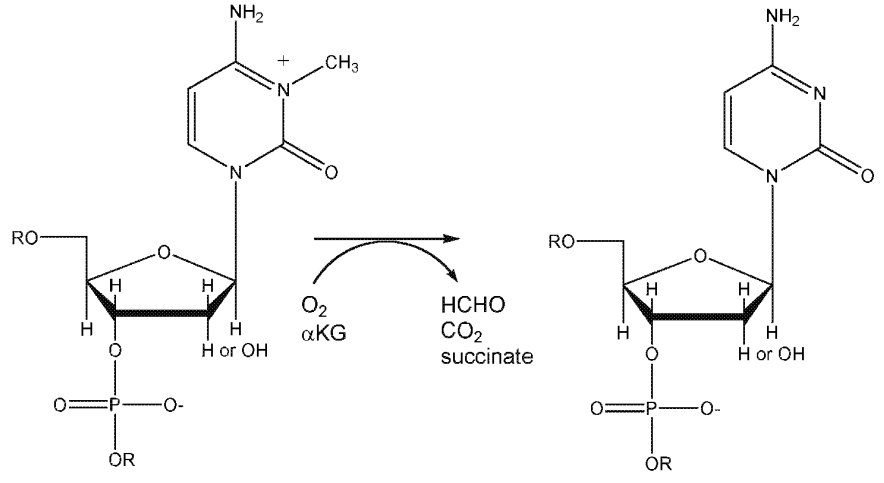
COR-AAD1-00001301



SCHEME 6.

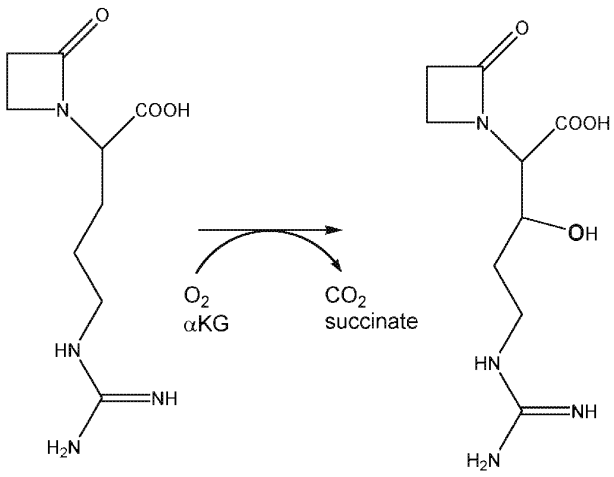
repair of methylated RNA lesions was established (Aas *et al.*, 2003). *In vitro* investigations reveal that the sites repaired by AlkB include 1-methyladenine and 3-methylcytosine in DNA (Trewick *et al.*, 2002) or RNA (Aas *et al.*, 2003), as well as adducts containing ethyl (Duncan *et al.*, 2002), hydroxyethyl, propyl, and hydroxylpropyl groups (Koivisto *et al.*, 2003). Trewick *et al.* (2002) noted that αKG decomposition is partially uncoupled from DNA repair, accounting for the observed ascorbate requirement for AlkB activity (described further in Section IV part I). Only this uncoupled reaction is observed when using the small molecules 1-methyladenosine, 1-methyl-2′-deoxyadenosine, 3-methylcytidine, and 3-methyl-2′-deoxycytidine (Welford *et al.*, 2003). In contrast, demethylation is achieved by the enzyme when using the oligonucleotide trimer d(Tp1meApT) or 5′ phosphorylated 1-me-dAMP (Koivisto *et al.*, 2003). The bacterial enzyme contains the likely metal- and αKG-binding residues His 131, Asp 133, His 187, and Arg 204 (Table 1). Each of the putative metal ligands was individually changed to cysteine by site-directed mutation of the corresponding codons, and the variant proteins were shown to undergo thiol:disulfide exchange reactions with disulfide-containing DNA (synthesized with thiol-substituted cytosine and derivatized with diaminoethane disulfide to create the mixed disulfide) to covalently attach AlkB to the DNA (Mishina & He, 2003). These studies demonstrate that DNA has access to the AlkB active site, but they do not provide any information related to the normal binding mode of this substrate. Although no crystallographic results have been reported yet, a theoretical model of the enzyme structure has been reported (Henshaw *et al.*, 2004).

Eight human homologues of *alkB* have been identified (termed *hABH1* through *hABH8*) (Duncan *et al.*, 2002;



SCHEME 7.

COR-AAD1-00001302



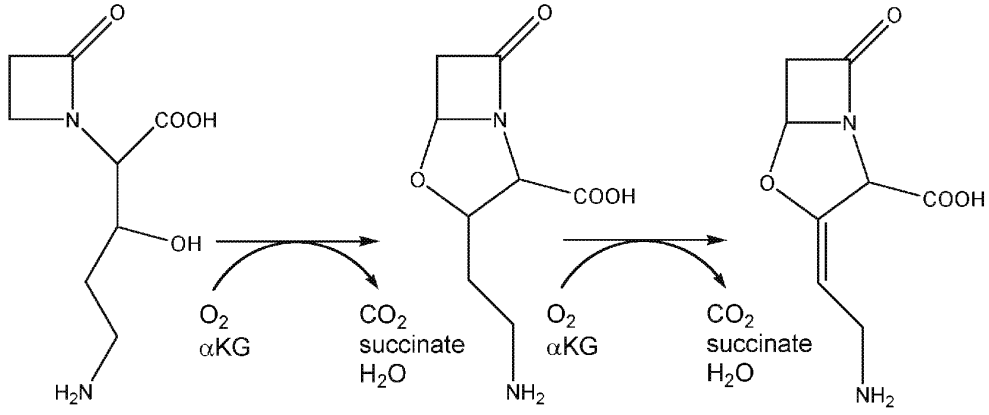
SCHEME 8.

Kurowski *et al.*, 2003). Curiously, a clone containing *hABH1* initially was reported to complement an *E. coli alkB* mutant (Wei *et al.*, 1996), but other researchers could not reproduce this result (Duncan *et al.*, 2002) and the purified gene product was inactive for repair of methylated DNA or RNA (Aas *et al.*, 2003; Duncan *et al.*, 2002). Both *hABH2* and *hABH3* complement an *alkB* mutant and encode enzymes capable of repairing 1-methyladenine, 3-methylcytosine, and 1-ethyladenine lesions in DNA (Duncan *et al.*, 2002). In addition, ABH3 (but not ABH2) repairs alkylated RNA (Aas *et al.*, 2003). In addition, ABH2 and ABH3 (but not ABH1) can form crosslinks to disulfide-containing DNA when these proteins are modified to replace a metal ligand residue with a cysteine; thus confirming their similarities to AlkB (Mishina *et al.*, 2004). No biochemical studies have been reported for the remaining five homologues (Kurowski *et al.*, 2003). A three-dimensional model was constructed for the core of ABH3, but no crystal structures are reported for any of these human proteins.

C. Biosynthesis of Antibiotics

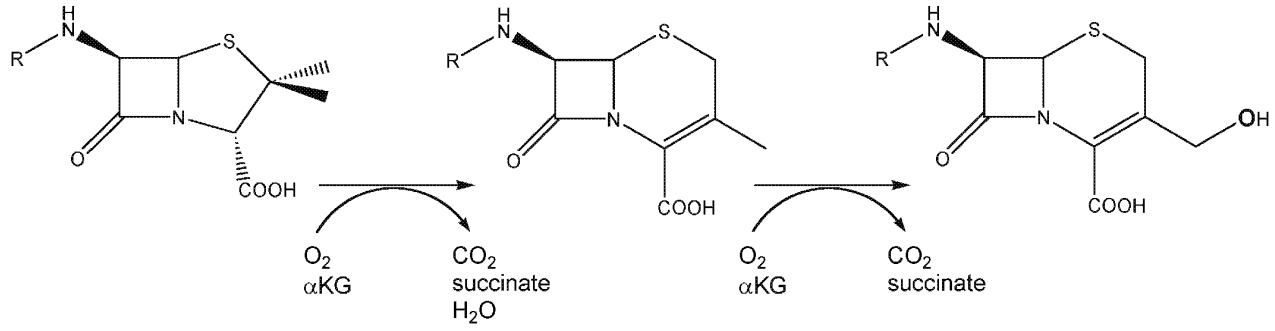
Fe(II)/αKG-dependent dioxygenases function during the biosynthesis of a diverse range of antibiotics. In addition to catalyzing a variety of hydroxylation reactions, some family members catalyze desaturation, ring expansion, ring formation, and other types of chemical transformations. The positions of these enzymes in antibiotic pathways vary from the generation of precursors to the modification of more complete intermediates. The following paragraphs describe biochemical studies on the better-studied representatives of the Fe(II)/αKG-dependent dioxygenases involved in antibiotic synthesis; however, it is important to note that sequence analyses are consistent with similar enzymes participating in a number of other pathways (e.g., those generating mitomycin C, nikkomycin, syringomycin, and viomycin (Chen *et al.*, 2000; Mao *et al.*, 1999; Yin *et al.*, 2003; Zhang *et al.*, 1995)).

Clavaminate synthase (CAS) catalyzes three separate oxidative steps in the synthesis of clavulanic acid and exemplifies the versatility of Fe(II)/αKG-dependent dioxygenases (Lloyd *et al.*, 1999b). The enzyme hydroxylates the β-lactam illustrated in Scheme 8, in a reaction comparable to other Fe(II)/αKG-dependent hydroxylases. Proclavaminate amidinohydrolase then hydrolyzes the guanidino group to yield proclavaminic acid, and clavaminate synthase catalyzes sequential cyclization and desaturation reactions as shown in Scheme 9. Most studies of clavaminate synthase have been carried out using the two isozymes from *Streptomyces clavuligerus* (Busby *et al.*, 1995; Marsh *et al.*, 1992), but the single *Streptomyces antibioticus* CAS isozyme also was purified and shown to exhibit the trifunctional activities (Janc *et al.*, 1995). Results from site-directed mutagenesis studies suggest that His 145 and His 280 are likely to be metal ligands in CAS isozyme 2 from *S. clavuligerus* (Khaleeli *et al.*, 2000). The crystal structure of Fe(II)- and αKG-bound CAS isozyme 1 in the presence of N-α-acetyl-L-arginine or proclavaminic acid (under anaerobic conditions, Table 2) confirms that the



SCHEME 9.

COR-AAD1-00001303



SCHEME 10.

corresponding histidines and a glutamic acid (His 144, Glu 146, and His 279) bind Fe(II) (Zhang *et al.*, 2000). In addition, the structure reveals that the C-5 carboxylate of αKG interacts with Arg 293 and Thr 172 (Table 1). Subtle differences in the binding modes of the two substrates likely relate to their distinct hydroxylase versus ring-forming/desaturase reactivities.

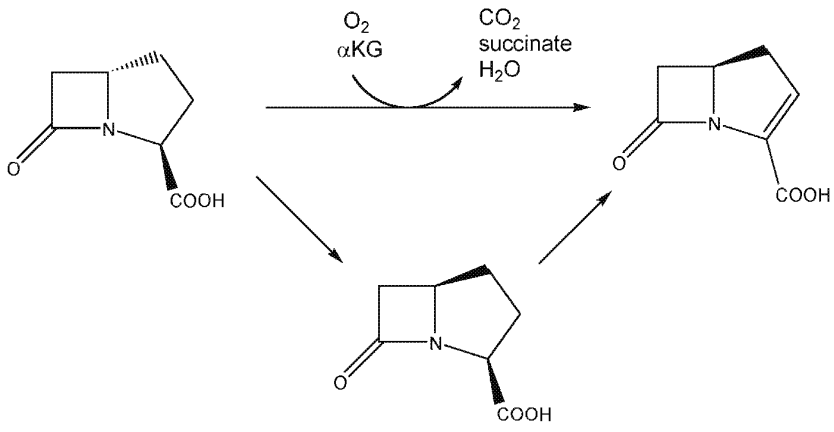
Deacetoxycephalosporin C synthase (DAOCS) and deacetylcephalosporin C synthase (DACS) are Fe(II)/αKG-dependent dioxygenases that catalyze sequential reactions starting with penicillin N in cephalosporin synthesis, as depicted in Scheme 10. A single enzyme carries out the ring expansion and hydroxylase reactions in *Cephalosporium acremonium* according to gene cloning and enzyme isolation studies (Baldwin *et al.*, 1987; Dotzlaf & Yeh, 1987; Samson *et al.*, 1985). In contrast, separate but related enzymes are found in *Streptomyces clavuligerus*. The second enzyme, DACS, was purified from *S. clavuligerus*, found to be a monomer, kinetically characterized, and shown to exhibit slight expandase activity (Baker *et al.*, 1991). The gene was cloned encoding *S. clavuligerus* DAOCS, the first enzyme in the sequence, allowing recombinant enzyme to be produced and extensively characterized (Dotzlaf & Yeh, 1989; Kovacevic *et al.*, 1989). This protein undergoes a monomer-to-trimer equilibrium according to gel filtration chromatography and light-scattering analyses (Lloyd *et al.*, 1999a). Results from extended X-ray absorption fine structure spectroscopy provide evidence for two histidine ligands to the six-coordinate metal site in the absence or presence of αKG (Lloyd *et al.*, 1999a). Crystallographic investigations of anaerobic samples were used to confirm the trimeric nature of the protein, identify the metal ligands as His 183, Asp 185, and His 243, and reveal interactions between the αKG C-5 carboxylate and Arg 258 as well as Ser 260 (Tables 1 and 2) (Valegård *et al.*, 1998). The R258Q variant uses αKG poorly because it lacks the ability to form a salt bridge to the C-5 carboxylate of this substrate; however, the mutant DAOCS exhibits enhanced reactivity with aliphatic α-ketoacids—a process termed "chemical co-substrate rescue" by the authors (Lee *et al.*, 2001b). An array of active site variants has been investigated, providing evidence that Arg 160, Arg 162, and Arg 266 are important for binding penicillin N (Lipscomb *et al.*, 2002) and that changes to several other residues can affect catalysis (Lee *et al.*, 2002). For example, substitution of a single amino acid side chain was shown to control whether the enzyme catalyzes ring expansion, methyl group hydroxylation, or both reactions (Lloyd *et al.*, 2004). Additional DAOCS crystallographic studies have revealed critical features of substrate and product binding, as described in Section III.

Carbapenem synthase (CarC) catalyzes an Fe(II)/αKG-dependent desaturation reaction, similar to CAS and DAOCS, and additionally exhibits an epimerization reaction at the tertiary carbon atom that joins the two rings in the bicyclic structure (Scheme 11). The best-characterized example of this enzyme, derived from *Erwinia carotovora*, is 23% identical to CAS (McGowan *et al.*, 1996). Crystal structures were obtained for the enzyme in the presence of Fe(II) plus αKG and with the substrate analogue L-*N*-acetylproline (Table 2) (Clifton *et al.*, 2003). The structural results show that His 101, Asp 103, and His 251 are metal ligands while Arg 263, Thr 130, and Arg 253 are used to stabilize the C-5 carboxylate of αKG (Table 1).

Another family member appears to catalyze a step in the semisynthetic pathway of the antibiotic fosfomycin (Watanabe *et al.*, 1999). *Penicillium decumbens* was shown to possess an epoxidase activity, encoded by *epoA*, which converts chemically synthesized *cis*-propenylphosphonic acid to the broad-spectrum antibiotic (Scheme 12). The natural function of this enzyme is unclear. The sequence of the gene product contains His 131, Asp 133, His 174, and Arg 185 (the standard Fe(II)/αKG-dependent hydroxylase motif; Table 1) and otherwise resembles the sequence of proline 4-hydroxylase. No activity was detected in cell extracts, but it is unclear whether αKG was included in the assay mixture.

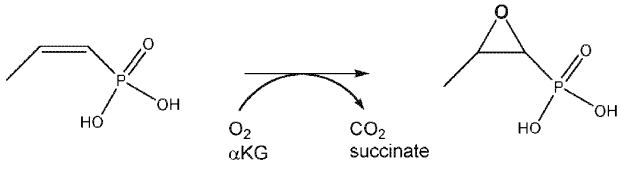
Proline hydroxylases provide a paradigm of Fe(II)/αKG-dependent hydroxylases that participate in antibiotic synthesis by producing hydroxylated precursors. The

COR-AAD1-00001304

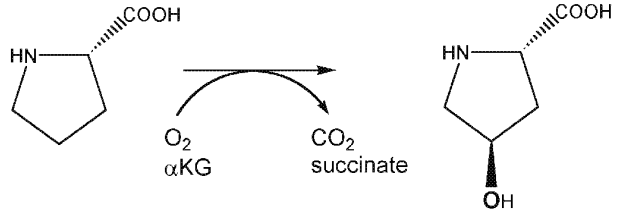


SCHEME 11.

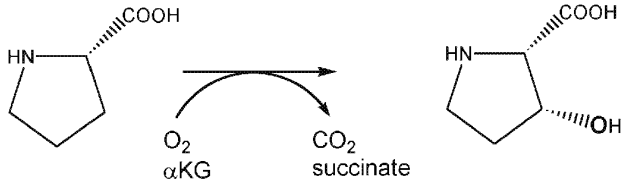
proline 4-hydroxylase from *Streptomyces griseoviridus* uses the free amino acid, as shown in Scheme 13, rather than a protein side chain, as described earlier, to produce *trans*-4-hydroxyproline that is subsequently incorporated into the antibiotic entamycin (Lawrence *et al.*, 1996). Analogous enzymes occur in selected other strains of *Streptomyces*, *Amycolatopsis*, and *Dactylosporangium*, with the gene being cloned (Table 1) and enzyme characterized from one of the latter species (Shibasaki *et al.*, 1999). Similarly, proline 3-hydroxylase activity (Scheme 14) was detected in two *Bacillus* strains and two *Streptomyces* strains, including *S. canus* ATCC 12647—a producer of the peptide antibiotic telomycin (Mori *et al.*, 1996). Of note, the product in this case is the *cis* isomer of 3-hydroxyproline. The gene-encoding proline 3-hydroxylase was cloned and the enzyme characterized from *Streptomyces* sp. strain TH1 (Mori *et al.*, 1997). This enzyme is a monomer like the proline 4-hydroxylase, but distinct from the heteromeric prolyl 4-hydroxylase. The fungus *Glara lozoyensis* possesses two distinct enzymes that produce *trans*-3-hydroxyproline or *trans*-4-hydroxyproline (Petersen *et al.*, 2003), but no sequence information is available for the two isozymes. The structures of *Streptomyces* sp. strain TH1 proline 3-hydroxylase apoprotein and the anaerobic, Fe(II)-bound form reveal His 107, Asp 109, and His 158 as metallocenter ligands, while Arg 168, Ser 170, and His 135 are likely to bind αKG (Tables 1 and 2) (Clifton *et al.*, 2001).



SCHEME 12.



SCHEME 13.



SCHEME 14.

It is reasonable to suggest that Fe(II)/αKG-dependent hydroxylases might provide precursors or otherwise contribute to the synthesis of many additional antibiotics (Chen *et al.*, 2001). For example, a gene in *Streptomyces verticillus* was proposed to encode a histidine hydroxylase, responsible for production of β-hydroxyhistidine in bleomycin, a cancer chemotherapy agent (Calcutt & Schmidt, 1994). Streptothricin synthesis utilizes 3-hydroxyarginine (Martinkus *et al.*, 1983), possibly synthesized by an enzyme analogous to CAS. An alternative arginine hydroxylase might participate in synthesis of the peptide antibiotic K-582, which contains 4-hydroxyarginine (Kawauchi *et al.*, 1983). On the other hand, without studies of the purified enzymes and sequence information one must be cautious in assigning roles to Fe(II)/αKG-dependent hydroxylases during antibiotic synthesis. For

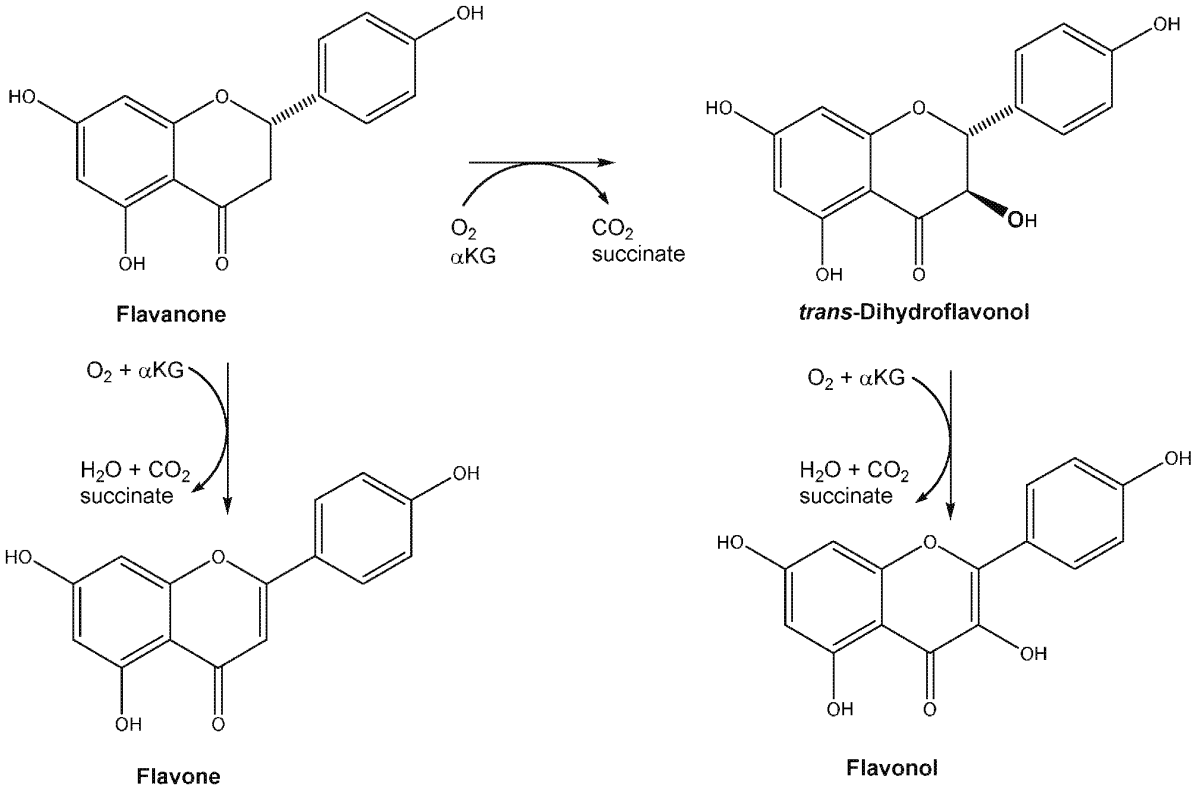
COR-AAD1-00001305

example, investigations carried out with cell extracts of *Streptomyces fradiae* led to the suggestion that two Fe(II)/αKG-dependent hydroxylases participate in the synthesis of the antibiotic tylosin (Omura *et al.*, 1984). Both of the reactions were stated to require αKG and oxygen, and CO_2 production was shown to be stoichiometric with formation of one product. Nevertheless, more recent genetic analyses and DNA sequencing studies (summarized in Cundliffe *et al.*, 2001)) reveal that these reactions are carried out by cytochrome P450-type enzymes.

D. Biosynthesis of Plant Products

Plants synthesize a remarkable variety of compounds by use of Fe(II)/αKG-dependent dioxygenases. The following paragraphs focus on the use of such enzymes during biosynthetic transformations of flavonoids, gibberellins, and alkaloids. The flavonoids include flavones, isoflavones, flavonols, anthocyanins, and other compounds that provide color to plants, protect them from ultraviolet (UV)-induced photodamage, act as signaling molecules in plant-microbe interactions, serve as defense molecules against pathogens, and exhibit a myriad of other plant roles (Bohm, 1998) while also acting as antioxidants, antimalarial medicines, and possible anticancer agents in humans (Bohm, 1998; Le Marchand, 2002). Gibberellins are tetracyclic diterpene phytohormones that regulate growth and development (Hedden, 1999; Hedden & Kamiya, 1997; Hedden & Phillips, 2000). The structurally complex alkaloids include important drugs that have experimental and therapeutic importance. In addition to participating in biosynthesis of these three types of compounds (described below), Fe(II)/αKG-dependent dioxygenases are likely to play additional roles in plants, such as the *Fusarium verticillioides FUM9* gene product that functions in the synthesis of the polyketide mycotoxin fumonisin (Seo *et al.*, 2001).

Flavanone 3β-hydroxylase catalyzes one of the key steps of flavonoid biosynthesis (top reaction in Scheme 15) in which flavanones (such as naringenin depicted here) are converted to *trans*-dihydroflavonols (such as the dihydrokaempferol illustrated). The enzyme from *Petunia hybrida* was purified in 1986 and shown to be an Fe(II)/αKG-dependent hydroxylase (Britsch & Grisebach, 1986). The gene encoding this enzyme has been cloned and sequenced from petunia (Britsch *et al.*, 1992) and several other plants (Britsch *et al.*, 1993; Charrier *et al.*, 1995; Pelletier & Shirley, 1996), and recombinant *P. hybrida* enzyme was purified (Lukacin *et al.*, 2000b). Results derived from mutagenesis studies (Table 1) of the recombinant *P. hybrida*



SCHEME 15.

COR-AAD1-00001306

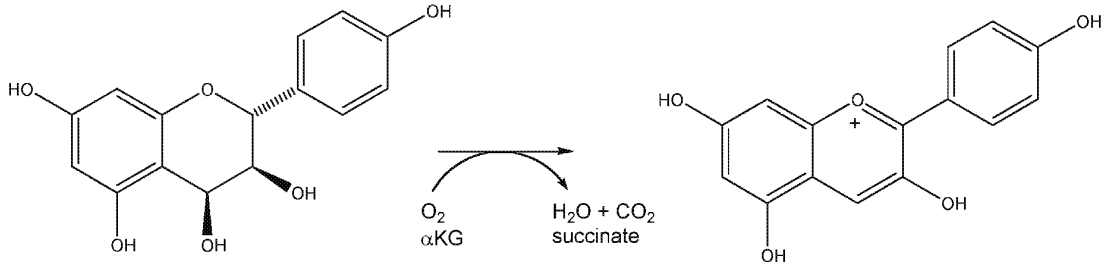
gene are consistent with His 220, Asp 222, and His 278 being metal ligands, while Arg 288 and Ser 290 assist in binding αKG (although some activity is observed with the H220Q, D222N, R288K, R288Q, S290T, S290A, and S290V variants; Lukacin & Britsch, 1997; Lukacin *et al.*, 2000a).

Flavone synthase I also utilizes flavanones as substrates, but this enzyme catalyzes the Fe(II)/αKG-dependent dehydration reaction to introduce a double bond between C-2 and C-3 (producing apigenin for the reaction shown on the left in Scheme 15). The enzyme was purified and the gene cloned from parsley (Britsch, 1990; Martens *et al.*, 2001). No 2-hydroxynaringenin is detected during the reaction, and addition of this compound does not inhibit the reaction or lead to product; thus, 2-hydroxyflavanones are not dissociable intermediates in the reaction (Martens *et al.*, 2003). An alternative P450-type enzyme catalyzes this reaction in many plants and is denoted *flavone synthase II*.

Flavonol synthases catalyze a reaction analogous to that of flavone synthases but use dihydroflavonols as their substrates (e.g., converting dihydrokaempferol to kaempferol, as shown on the right in Scheme 15). The gene-encoding flavonol synthase was first isolated from petunia petals (Holton *et al.*, 1993) but has since been cloned and expressed from several sources including *Arabidopsis thaliana* (Prescott *et al.*, 2002). Antisense expression of the petunia gene in the plant leads to reduced flavonol levels (Holton *et al.*, 1993). The Fe(II)/αKG-dependent enzyme was purified from *Citrus unshiu* and extensively characterized by kinetic and mutagenesis methods (Wellman *et al.*, 2002). The flavonol synthase metal ligands are proposed to include His 221, Asp 223, and His 277, while Arg 287 is suggested to participate in αKG binding (Table 1). In addition to possessing flavonol synthase activity, the recombinant *C. unshiu* protein also exhibits flavanone 3β-hydroxylase activity (Lukacin *et al.*, 2003). A chimeric enzyme comprised of the *P. Hybrida* flavanone 3β-hydroxylase with its C-terminal 52 residences replaced by the corresponding region of *C. unshiv* flavonol synthase exhibited exclusively flavanone 3β-hydroxylase activity, indicating that this region does not define the selectivity of enzyme (Wellman *et al.*, 2004).

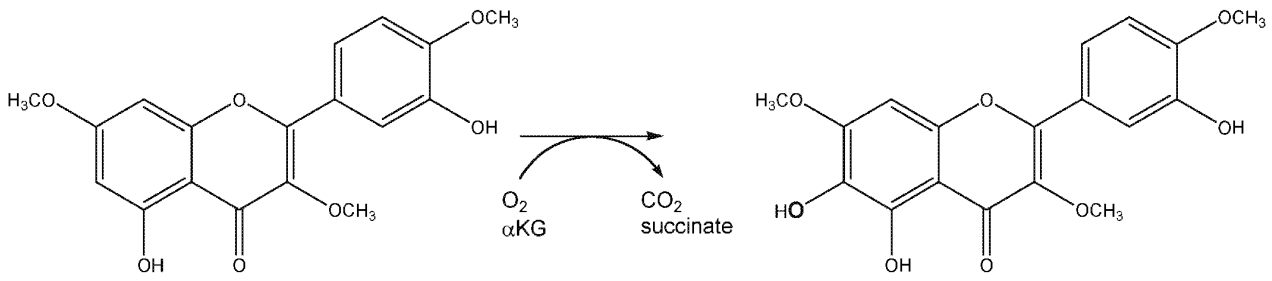
The synthesis of anthocyanins, involving the Fe(II)/αKG-dependent anthocyanidin synthase (ANS), offers an alternative route of processing of *trans*-dihydroflavonols. This pathway includes ketone reduction of the *trans*-dihydroflavonol by dihydroflavonol reductase, conversion of the resulting leucoanthocyanidin to anthocyanidin (Scheme 16), and hydroxyl group glycosylation to form anthocyanin (Nakajima *et al.*, 2001; Springob *et al.*, 2003). Surprisingly, *in vitro* studies with purified ANS reveal quercetin and dihydroquercetin as the major products, a finding that is consistent with formation of flav-2-en-3,4-diol as the *in vivo* product (Turnbull *et al.*, 2003). This intermediate transforms to anthocyanidin under acidic conditions. Genes encoding ANS proteins have been cloned from several sources, but direct evidence for a gene encoding this Fe(II)/αKG-dependent activity was first obtained using recombinant enzyme from mint, *Perilla fructescens* (Saito *et al.*, 1999). The enzyme from *Arabidopsis thaliana* was structurally characterized under anaerobic conditions in the presence of Fe(II), αKG, and substrate (Tables 1 and 2) (Wilmouth *et al.*, 2002). These studies reveal His 232, Asp 234, and His 288 as the metal ligands, Arg 298, Asn 215, and Tyr 217 as being involved in αKG binding, and the presence of two substrate binding sites (containing dihydroquercetin). In addition, a structure was determined for crystals that were exposed to oxygen for 30 min; the results are consistent with one substrate molecule being converted to quercetin and the αKG decomposing to succinate.

Another Fe(II)/αKG-dependent dioxygenase involved in flavonoid biosynthesis is an enzyme catalyzing 6-hydroxylation of methylated flavonols (Scheme 17). This enzyme, isolated from *Chrysosplenium americanum* was purified, kinetically characterized, and shown to utilize quercetin derivatives in the order trimethylated > dimethylated > tetramethylated > methylated > nonmethylated (but still retaining 10% of the activity observed with trimethylated substrate) (Anzellotti & Ibrahim, 2000). Significantly, this is one of only two Fe(II)/αKG-dependent hydroxylases known to insert an oxygen atom into a C–H bond of a carbon atom located in an aromatic ring. The second such reaction (involving the flavonoid-like compound

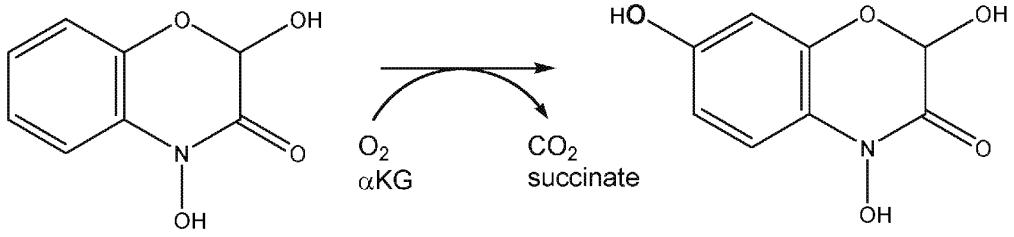


SCHEME 16.

COR-AAD1-00001307



SCHEME 17.

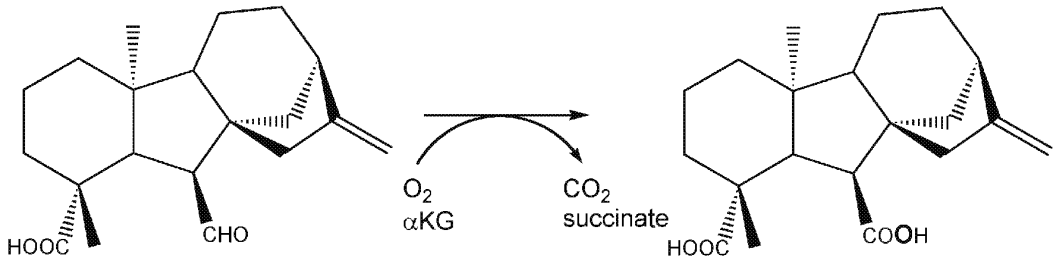


SCHEME 18.

shown in Scheme 18) converts 2,4-dihydroxy-2*H*-1,4-benzoxazin-3(4*H*)-one to 2,4,7-trihydroxyl-2*H*-1,4-benzoxazin-3(4*H*)-one, a natural pesticide produced in grasses (Frey *et al.*, 2003).

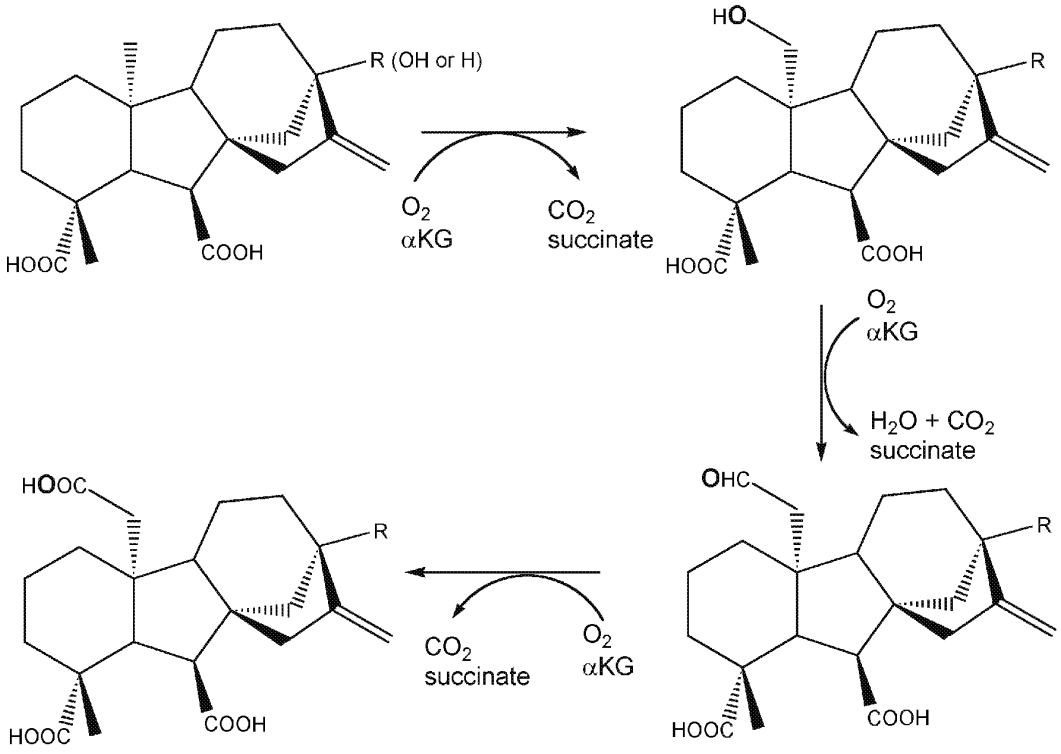
Within the complex network of gibberellin biosynthesis, the first Fe(II)/αKG-dependent step utilizes gibberellin 7-oxidase to catalyze aldehyde oxidation at the C-7 position, as illustrated in Scheme 19. This enzyme was shown to be Fe(II)/αKG-dependent and separated from two other Fe(II)/αKG-dependent dioxygenases in studies using cell-free extracts of developing pumpkin, *Cucurbita maxima* (Lange *et al.*, 1994b), but the activity appears to be cytochrome P450-dependent in other plant species (Hedden & Phillips, 2000). Subsequent cloning of the pumpkin gene allowed production of the recombinant enzyme, which was shown to catalyze the depicted reaction as well as several other less well-characterized activities (Lange, 1997). The protein sequence contains His 182, Asp 184, and His 238 as likely metal ligands along with Arg 248 and Ser 250 as potential αKG-binding residues (Table 1).

Gibberellin 20-oxidase catalyzes the sequential oxidation of the C-20 methyl group to form the alcohol, aldehyde, and (in some species) the carboxylate (Hedden, 1997, 1999), as shown in Scheme 20. Alternatively, the enzyme oxidatively transforms the aldehyde-containing species to a γ-lactone lacking C-20 (Scheme 21). Gibberellin 20-oxidase was first partially purified from seeds of pea (*Pisum sativum*) and shown to require αKG (Lange & Graebe, 1989). Studies of the isolated enzyme from *C. maxima* confirmed the capability of the enzyme to carry out all reactions shown in Schemes 20 and 21 (Lange, 1994). This same source was used for cloning the first gibberellin 20-oxidase gene (Lange *et al.*, 1994a), but many examples of the gene have since been cloned from other sources (including multiple copies for some plants) (Hedden & Phillips, 2000). Analysis of chimeric oxidases generated from fusions of portions of the gene from pumpkin (which makes primarily the carboxylate shown in Scheme 20) with that of the closely related species *Marah macrocarpus* (which favors γ-lactone production, shown



SCHEME 19.

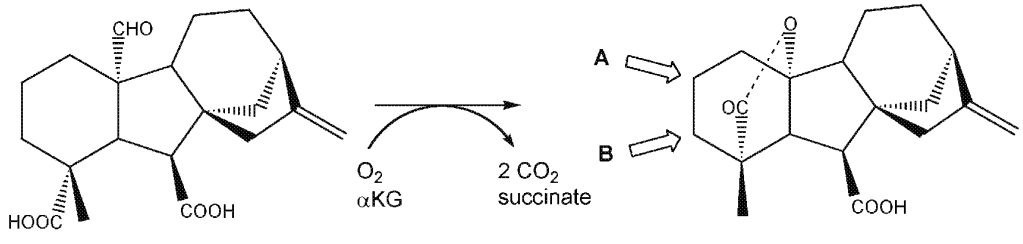
COR-AAD1-00001308



SCHEME 20.

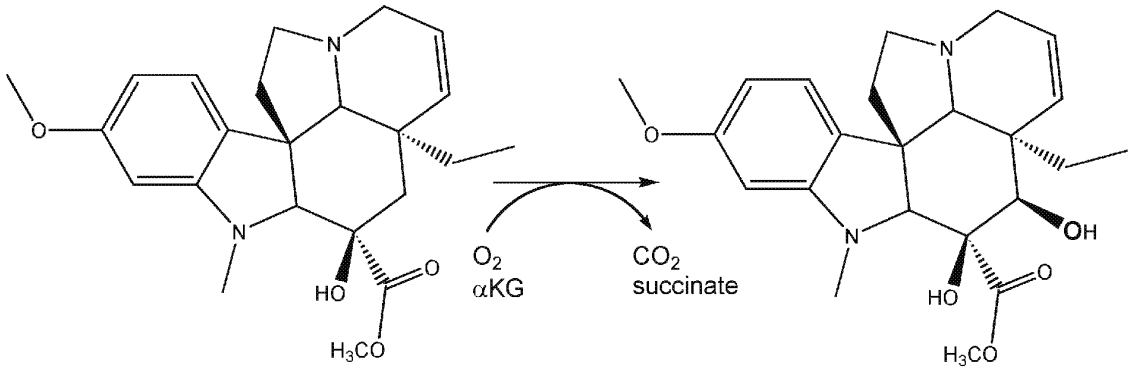
in Scheme 21) highlights the importance of the carboxyl terminus of the protein as the major region controlling product specificity (Lange *et al.*, 1997a). The pumpkin protein sequence includes the likely metal liganding residues His243, Asp 245, and His 299 as well as the possible αKG-stabilizing residues Arg 309 and Ser 311 (Table 1).

Several other Fe(II)/αKG-dependent enzymes utilize the gibberellin γ-lactone as a substrate and catalyze reactions at C-2 and C-3 (indicated by arrows A and B in Scheme 21). Gibberellin 2-oxidase catalyzes initial hydroxylation and then oxidation to form the ketone at C-2, as shown by studies of recombinant enzymes from the runner bean (*Phaseolus coccineus*) (Thomas *et al.*, 1999) and pea (Lester *et al.*, 1999). Another related 2-hydroxylation activity acting on the non-γ–lactone substrate was identified and cloned from *Arabidopsis* (Schomburg *et al.*, 2003). Gibberellin 3β-hydroxylase activity was first demonstrated to be Fe(II)/αKG-dependent after enrichment from pumpkin (Lange *et al.*, 1994b). The gene encoding this activity was isolated initially from *Arabidopsis thaliana* (Chiang *et al.*, 1995), and has since been demonstrated to be associated with the pea stem-length gene (*Le*) studied by Gregor Mendel during his pioneering genetic investigations (Lester *et al.*, 1997; Martin *et al.*, 1997). Rice contains two genes encoding gibberellin 3-hydroxylase, one of which also exhibits 2,3-desaturation activity (Itoh *et al.*, 2001). A gene from pumpkin was shown to encode an enzyme with both 3β-hydroxylase and 2β-hydroxylase activities (Lange *et al.*, 1997b). Of biological significance, modification at C-2 leads to a less active gibberellin.



SCHEME 21.

COR-AAD1-00001309



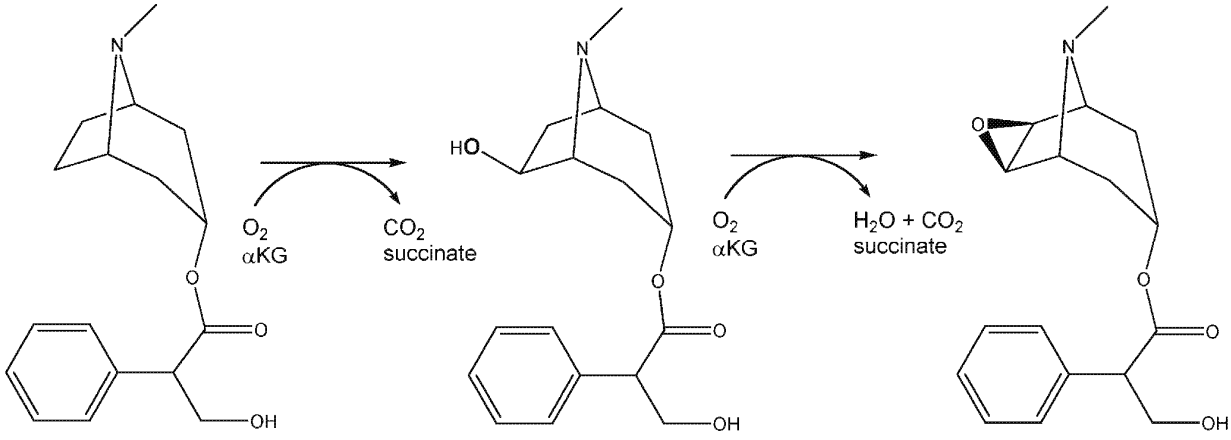
SCHEME 22.

At least two alkaloids require Fe(II)/αKG-dependent enzymes for their biosynthesis. The periwinkle plant, *Caranthus roseus*, synthesizes the vinca alkaloids vinblastine and vincristine that are used to treat Hodgkin's disease and leukemia, respectively. A critical component of these bis-indole alkaloids is vindoline, generated from tryptophan by a complex series of transformations. The second to last step in the vindoline biosynthetic pathway (Scheme 22) is catalyzed by an Fe(II)/αKG-dependent hydroxylase (De Carolis *et al.*, 1990). Similarly, an Fe(II)/αKG-dependent dioxygenase plays a key role in the synthesis of scopolamine, a hallucinogenic tropane alkaloid of henbane, *Hyoscyamus niger*. A single enzyme carries out both hydroxylation and epoxidation reactions (Scheme 23) as shown by using recombinant protein (Hashimoto *et al.*, 1993). Both of the alkaloid biosynthesis-related enzymes were purified (De Carolis & De Luca, 1993; Hashimoto & Yamada, 1987) and their genes cloned (Matsuda *et al.*, 1991; Vazquez-Flota *et al.*, 1997). In each case, metal-binding and αKG-interaction motifs are observed (His 308, Asp 310, His 366, Arg 376, and Ser 378 for deacetoxyvindoline 4-hydroxylase, and His 217, Asp 219, His 274, Arg 284, and Ser 286 of hyoscyamine 6β-hydroxylase; Table 1).

Another plant Fe(II)/αKG-dependent dioxygenase has been characterized from the seeds of fenugreek *Trigonella foenum-graecum* (Haefelè *et al.*, 1997). This enzyme forms 4-hydroxyisoleucine (Scheme 24), an insulinotropic compound that accounts for the use of this plant in traditional medicine for its antidiabetic properties (Sauvaire *et al.*, 1998). The role of this compound in the plant has not been reported.

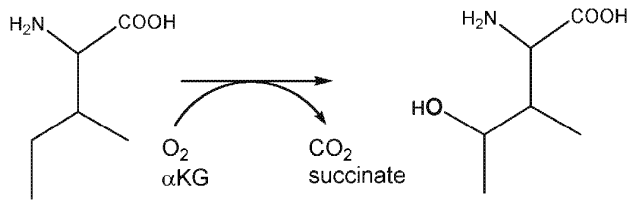
E. Lipid Metabolism

Fe(II)/αKG-dependent hydroxylases play several important roles in lipid metabolism. Two of these enzymes are needed to synthesize carnitine, which transports activated fatty acids across the inner mitochondrial membrane to allow their degradation in the matrix by the β-oxidation



SCHEME 23.

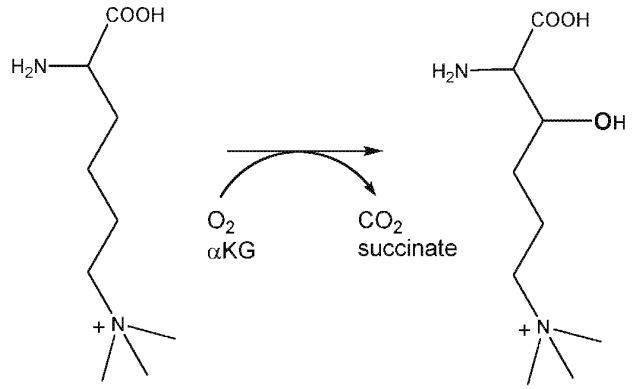
COR-AAD1-00001310



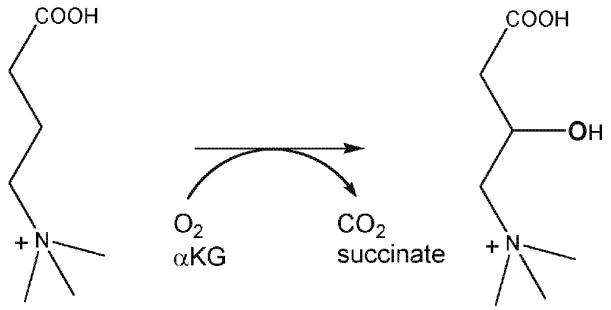
SCHEME 24.

pathway (Vaz & Wanders, 2002). A separate representative is required for degrading the plant lipid phytanic acid, and human deficiencies of this enzyme result in Refsum disease and other disorders (Wierzbicki *et al.*, 2002). Finally, on the basis of sequence results and physiological studies a family member appears to function in the synthesis of lipid A in several pathogenic bacteria (Gibbons *et al.*, 2000).

The synthesis of carnitine initiates with trimethyllysine, derived by hydrolysis of proteins that are modified through the action of an *S*-adenosylmethionine-dependent methyl transferase. As first illustrated for a rat liver mitochondrial sample (Hulse *et al.*, 1978), an Fe(II)/αKG-dependent hydroxylase converts ε-*N*-trimethyl-L-lysine to β-hydroxy-ε-*N*-trimethyl-L-lysine (Scheme 25). The dimeric rat enzyme has been purified and characterized, and the *Neurospora crassa* (Swiegers *et al.*, 2002), rat, mouse, and human (Vaz *et al.*, 2001) cDNAs were cloned and expressed in functional forms. The product of these enzymes is cleaved by an aldolase to produce glycine and γ-trimethylaminobutyraldehyde (γ-butyrobetaine aldehyde), with the latter compound being oxidized by a nicotinamide-dependent dehydrogenase to form γ-butyrobetaine. This product undergoes hydroxylation by a second Fe(II)/αKG-dependent hydroxylase to generate carnitine (Scheme 26). γ-Butyrobetaine hydroxylase was first purified from *Pseudomonas* sp. AK1 (Lindstedt *et al.*, 1977)



SCHEME 25.

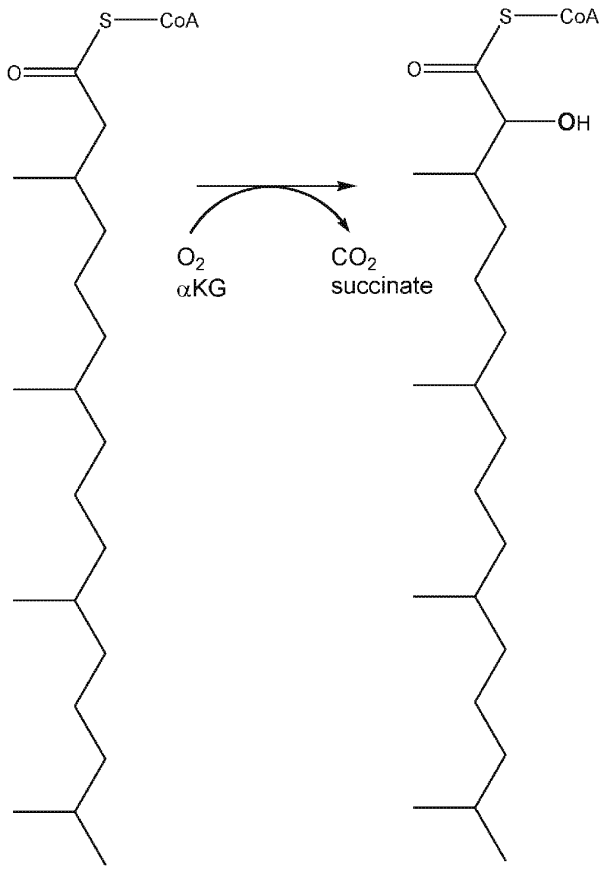


SCHEME 26.

and soon thereafter from calf liver (Kondo *et al.*, 1981). The bacterial protein sequence was determined by Edman degradation of the isolated peptides (Rüetschi *et al.*, 1993), whereas the sequences of animal enzymes were determined by translation of the cDNA sequences (Galland *et al.*, 1999; Vaz *et al.*, 1998). These γ-butyrobetaine hydroxylases are related to each other (Table 1), with the bacterial enzyme possessing His 209, Asp 211, and His 350 as likely metal-binding ligands and Arg 362 positioned for αKG binding. Furthermore, these enzymes are homologous to trimethyllysine hydroxylase, which appears to use glutamic acid rather than aspartic acid to bind the metal ion (Table 1).

Phytanoyl-CoA hydroxylase (Scheme 27) is required for the metabolism of phytanic acid (3,7,11,15-tetramethylhexadecanoic acid), a lipid derived from the phytol group of chlorophyll and found in dairy products, meat, and fish. The presence of the 3-methyl group precludes operation of the β-oxidation pathway; however, this oxidative degradation process can be utilized after removal of the C-1 carbon unit by action of a series of four enzymes (Mukherji *et al.*, 2003). A ligase is used to convert phytanic acid to phytanoyl-CoA (Watkins *et al.*, 1994). The CoA thioester is hydroxylated at the C-2 position by an αKG- and Fe(II)-dependent reaction (Scheme 27), as first demonstrated in rat liver peroxisomes in 1995 (Mihalik *et al.*, 1995) and later for human liver homogenates (Jansen *et al.*, 1996). Interestingly, GTP and ATP stimulate this reaction, although the role of the nucleotides is unclear (Croes *et al.*, 2000), and the participation of sterol carrier protein-2 in this reaction was suggested (Mukherji *et al.*, 2002). A thiamin pyrophosphate-dependent enzyme (Foulon *et al.*, 1999) catalyzes a lyase reaction of the hydroxylated intermediate to release formyl-CoA (Croes *et al.*, 1997b) and generate methylpentadecanal (pristanal) (Croes *et al.*, 1997a; Verhoeven *et al.*, 1997). After undergoing a nicotinamide-dependent oxidation to form pristanic acid, the ligase is used again to produce pristanoyl-CoA that is able to undergo β-oxidative degradation. Elevated levels of phytanic acid are observed in patients

COR-AAD1-00001311



SCHEME 27.

suffering from Zellweger syndrome, rhizomelic chondrodysplasia punctata, and Refsum disease. These disorders are associated with a generalized defect in peroxisome assembly, a disturbance of a type of peroxisomal targeting signal, or a specific defect in phytanoyl-CoA hydroxylase activity, respectively.

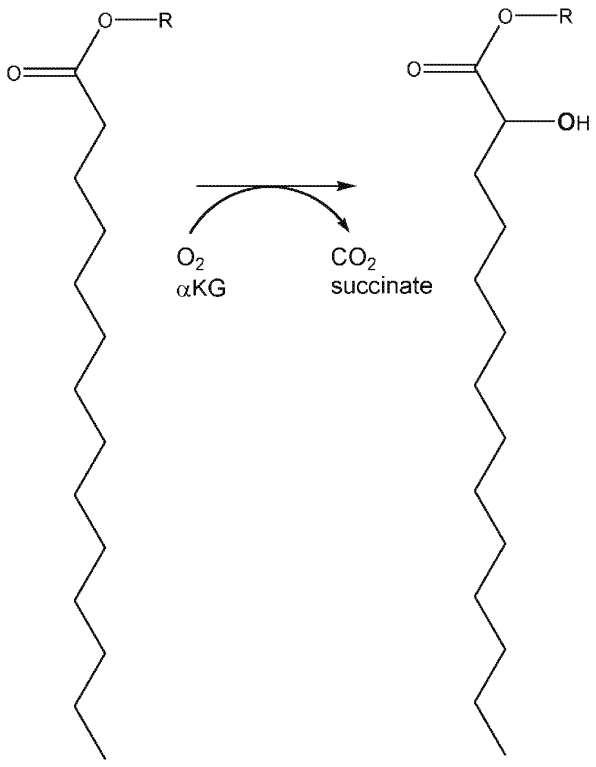
Using the amino terminal sequence and an internal sequence of the rat liver enzyme, expressed sequence tag (EST) cDNAs were used to directly identify mouse and human clones of the phytanoyl-CoA hydroxylase gene, initially termed *PHYH* (Jansen *et al.*, 1997). The gene, renamed *PAHX*, was identified simultaneously by another group that searched for EST cDNAs encoding the required peroxisomal targeting signal, cloned the targeted sequence in an *E. coli* expression vector, and showed that affinity-purified sample exhibited Fe(II)/αKG-dependent phytanoyl-CoA hydroxylase activity (Mihalik *et al.*, 1997). Subsequently, the rat gene was cloned from its cDNA (Jansen *et al.*, 1999) and the complete ~21 kb human gene sequence was determined (Mukherji *et al.*, 2001a), revealing the presence of nine exons and eight introns. Refsum disease patients possess large (111 bp) or small (1 bp) deletions, insertions (3 bp), and a variety of single base changes (Jansen *et al.*, 1997; Mihalik *et al.*, 1997; Mukherji *et al.*, 2001a). Transfection of cDNA encoding the active form of the enzyme restored phytanoyl-CoA hydroxylase activity to fibroblasts isolated from two Refsum disease patients (Chahal *et al.*, 1998). On the basis of sequence comparisons (i.e., that of active human enzyme versus those of defective enzymes from Refsum disease patients), site-directed mutagenesis studies of the recombinant gene in various expression systems, and kinetic analysis of purified enzymes with αKG and alternative α-keto acids, the metal ligands of the human enzyme are likely to be His 175, Asp 177, and His 264, while Arg 275 assists in αKG binding (Table 1) (Mukherji *et al.*, 2001b).

Certain Gram-negative bacteria, including *Salmonella typhimurium* and *Pseudomonas aeruginosa*, appear to utilize an Fe(II)/αKG-dependent hydroxylase during the synthesis of lipid A (Gibbons *et al.*, 2000). When grown under aerobic conditions these pathogenic bacteria possess *S*-2 hydroxyacyl chains that are further decorated by acylation, whereas these modifications do not occur for anaerobically grown cells. Mass spectrometric studies confirm that the added oxygen atom is derived from O_2. Sequence analyses of the *S. typhimurium* genome were used to identify a possible homologue, denoted *lpxO*, to the mammalian gene for aspartyl/asparaginyl β-hydroxylase. This sequence includes His 155, Asp 157, His 202, and Arg 212 (Table 1), resembling the standard Fe(II)/αKG-dioxygenase motif. Curiously, however, these residues do not coincide with the likely metal- and αKG-binding residues of the mammalian hydroxylase used for sequence comparison (it does not contain this motif). Homologues to *lpxO* were noted in many other pathogenic bacteria that are known to make 2-hydroxyacyl-containing lipids. Expression of the *S. typhimurium* gene in *E. coli* leads to synthesis of 2-hydroxymyristate in this organism, thus confirming the function of the gene. Still unclear is whether the enzyme hydroxylates the acyl group after incorporation into the nascent lipid A (i.e., R in Scheme 28 is a glucosamine disaccharide) or whether hydroxylation of the acyl group occurs while bound to acyl carrier protein or CoA (e.g., R in Scheme 28 is a thiol). Notably, the latter situation would be highly reminiscent of the reaction carried out by phytanoyl-CoA hydroxylase (Scheme 27).

F. Other Fe(II)/αKG-Dependent Hydroxylases

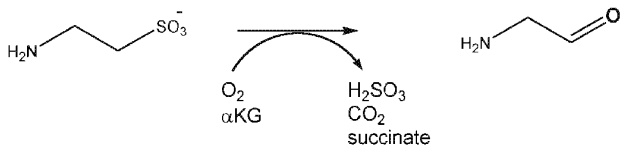
This final grouping of Fe(II)/αKG-dependent hydroxylases describes an eclectic assembly of enzymes with diverse roles ranging from biodegradation (with examples that supply sulfur, phosphorus, or carbon to the cell) to nucleotide recycling to microbial biosynthesis of ethylene to still-unknown functions.

COR-AAD1-00001312

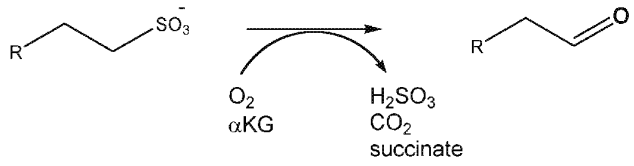


SCHEME 28.

Selected microorganisms decompose alkyl sulfonates or alkyl sulfates by using Fe(II)/αKG-dependent hydroxylases, thus providing sulfur for biosynthetic needs (Kertesz, 1999). For example, *E. coli* synthesizes taurine/αKG dioxygenase (TauD) during times of sulfur starvation (Eichhorn *et al.*, 1997). This enzyme hydroxylates taurine (2-aminoethanesulfonate) to create an unstable intermediate that decomposes to aminoacetaldehyde and sulfite (Scheme 29), which is subsequently utilized as a sulfur source. Structural studies (Elkins *et al.*, 2002; O'Brien *et al.*, 2003) demonstrate that His 99, Asp 101, and His 255 bind Fe(II), while Arg 266 and Thr 126 stabilize the C-5 carboxylate of αKG (Tables 1 and 2). An open reading frame in *Saccharomyces cerevisiae* (YLL057c) encodes a protein with 31.5% identity to TauD. The recombinant yeast protein was purified from *E. coli* and shown to possess weak taurine/αKG dioxygenase activity, but it



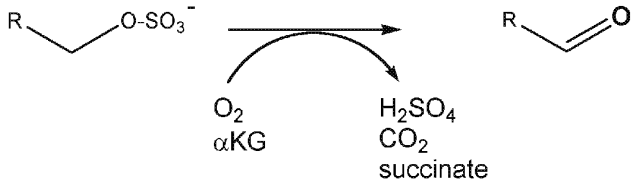
SCHEME 29.



SCHEME 30.

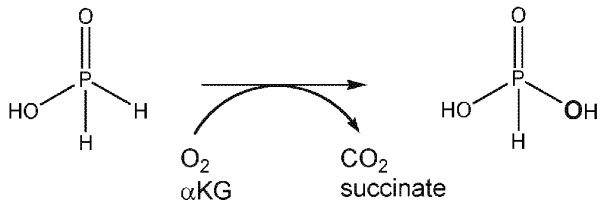
exhibited 50- to 150-fold greater catalytic efficiency when using the bile salt taurocholate, hydroxyethanesulfonate (isethionate), *N*-phenyltaurine, and MOPS or MOPSO buffers as alternative substrates (Hogan *et al.*, 1999). Thus, the yeast enzyme appears to be a broad specificity sulfonate/αKG dioxygenase (Scheme 30). Alkyl sulfates, rather than sulfonates, are the substrates of an enzyme encoded by the *atsK* gene in *Pseudomonas putida* S-313 (Kahnert & Kertesz, 2000). The isolated alkylsulfatase hydroxylates alkyl sulfate esters (C_4 to C_{12} in length) at the C-1 position to create unstable intermediates that release sulfate and form the corresponding aldehydes (Scheme 31). AtsK is closely related to TauD (38% identity) and, as shown by structural studies (Table 2) (Müller *et al.*, 2004), contains the expected key residues (His 108, Asp 110, His 264, and Arg 275; Table 1).

A particularly interesting Fe(II)/αKG-dependent hydroxylase participates in a pathway that provides phosphorus for cellular needs. *Pseudomonas stutzeri* WM88 exemplifies an expanding group of microorganisms that are capable of growth using reduced forms of phosphorus, including hypophosphite (H_3PO_2) and phosphite (H_3PO_3). Molecular genetic analyses of this bacterium identified the *htxA* and *ptxD* genes that encode an Fe(II)/αKG-dependent hydroxylase and a NAD-dependent enzyme, respectively (Metcalf & Wolfe, 1998). Overexpression of *htxA* in *E. coli* allowed the isolation of an enzyme shown to hydroxylate hypophosphite (White & Metcalf, 2002), as illustrated in Scheme 32. PtxD then oxidizes phosphite, with concomitant reduction of NAD, to generate phosphate (Garcia Costas *et al.*, 2001). The sequence of HtxA resembles those of proline 4-hydroxylase, phytanoyl-CoA hydroxylase, and selected other family members, and includes the metal-binding residues His 116, Asp 118, and His 206, along with Arg 217, which likely facilitates αKG



SCHEME 31.

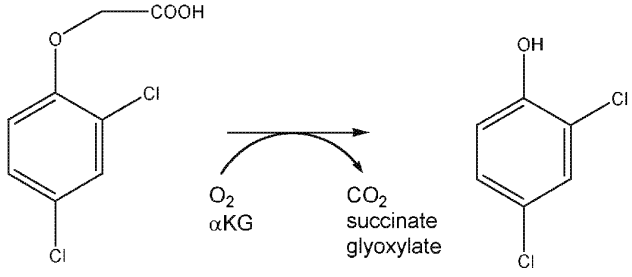
COR-AAD1-00001313



SCHEME 32.

binding (Table 1). Although it seems reasonable to suggest that Fe(II)/αKG-dependent hydroxylases might participate in the degradation of alkyl phosphonates and alkyl phosphates, in parallel to the alkyl sulfonate- and alkyl sulfate-degrading systems, these activities have not been reported.

A herbicide-degrading Fe(II)/αKG-dependent hydroxylase is found in bacteria that metabolize 2,4-dichlorophenoxyacetic acid (2,4-D), including microbes that are capable of using this molecule as their sole carbon source (Hausinger *et al.*, 1997). Catabolism of 2,4-D has long been known to initiate with ether bond cleavage (Tiedje & Alexander, 1969), but the gene encoding the corresponding enzyme (*tfdA*) is unlike other known ether cleavage systems (Streber *et al.*, 1987). By examining the overproduced *Ralstonia eutropha* enzyme in *E. coli*, 2,4-D/αKG dioxygenase (TfdA) was shown to catalyze the reaction shown in Scheme 33 (Fukumori & Hausinger, 1993b); i.e., the oxidative cleavage of αKG to CO_2 plus succinate is coupled to the hydroxylation of 2,4-D with subsequent decomposition of the hydroxy-intermediate to yield glyoxylate and 2,4-dichlorophenol. The purified enzyme decomposes a variety of other phenoxyacetic acid derivatives (Fukumori & Hausinger, 1993a), thiophenoxyacetic acids (Saari & Hausinger, 1998), naphthoxyacetic acids, benzofuran-2-carboxylate, and various cinnamic acids (Dunning Hotopp & Hausinger, 2001), with release of the corresponding substituted phenols, thiophenols, and, for the benzofuran-2-carboxylate and cinnamic acids, the epoxide products. Whereas some ethene-containing cinnamic acids are substrates, the acetylenic compound phenylpropiolic acid is a



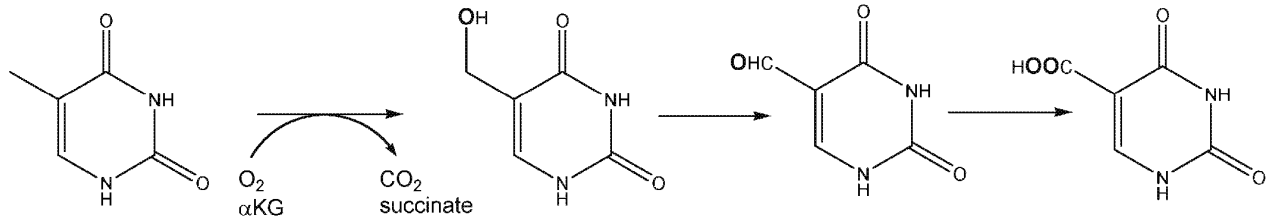
SCHEME 33.

mechanism-based inactivator that forms a covalent bond to the protein (Dunning Hotopp & Hausinger, 2002). In addition, TfdA transforms the *S* enantiomer of the 2-phenoxypropionate, dichloroprop, to the substituted phenol plus pyruvate (Saari *et al.*, 1999). Interestingly, *Alcaligenes denitrificans* possesses an enzyme that utilizes the opposite enantiomer of a related phenoxypropionate (mecoprop) (Tett *et al.*, 1997) and both *Sphingomonas herbicidovorans* MG and *Delftia acidovorans* MC1 have two Fe(II)/αKG-dependent phenoxypropionate-degrading hydroxylases—one specific to each enantiomer (Nickel *et al.*, 1997; Westendorf *et al.*, 2003). Analyses of Fe(II)-bound TfdA and Cu(II)-substituted protein by X-ray absorption spectroscopy (XAS) suggested the presence of two imidazole ligands (Cosper *et al.*, 1999). Electron paramagnetic resonance (EPR) and electron spin-echo envelop modulation (ESEEM) spectroscopic analyses of the Cu-substituted protein similarly provided evidence for His ligation, and showed that binding of αKG and 2,4-D perturbs the metal site (Whiting *et al.*, 1997). The results of site-directed mutagenesis studies of *tfdA* (Dunning Hotopp & Hausinger, 2002; Hogan *et al.*, 2000) are consistent with His 114, Asp 116, and His 263 serving to bind the metal and Arg 274 assisting in binding αKG (Table 1). The modeled structure of the enzyme, generated by comparison to structures of other known family members, is consistent with these interactions (Elkins *et al.*, 2002).

Thymine hydroxylase catalyzes three sequential Fe(II)/αKG-dependent hydroxylations of the methyl group of thymine (Scheme 34). The activity was first studied in *Neurospora crassa* (Abbott *et al.*, 1967; Holme, 1975; Holme *et al.*, 1971) and has been noted in *Aspergillus nidulans* (Shaffer *et al.*, 1984), but recent investigations have focused on the enzyme from *Rhodotorula glutinis*. For example, additional reactivities were uncovered, including epoxidation of 5-vinyluracil, sulfur oxidation of 5-(methylthio)uracil, and conversion of an amine to an aldehyde using 1-methylthymine (Thornburg *et al.*, 1993). Mechanism-based inactivation was observed using the acetylenic analogue of thymine, allowing characterization of the covalently modified peptides (Thornburg & Stubbe, 1989, 1993; Thornburg *et al.*, 1993; Warn-Cramer *et al.*, 1983). On the other hand, the complete sequence is unknown and no structure has been reported for this enzyme.

At least two Fe(II)/αKG-dependent hydroxylases utilize nucleosides as their substrates and catalyze transformations involving the sugar moiety. *N. crassa*, *A. nidulans*, and *R. glutinis* possess pyrimidine deoxyribonucleoside 2′-hydroxylases that convert thymidine or uridine to the ribosides, as shown in Scheme 35 (where X is a methyl group or hydrogen atom, respectively) (Bankel *et al.*, 1972; Shaffer *et al.*, 1968, 1984; Warn-Cramer *et al.*, 1983). In addition, *R. glutinis* contains a deoxyuridine (uridine)

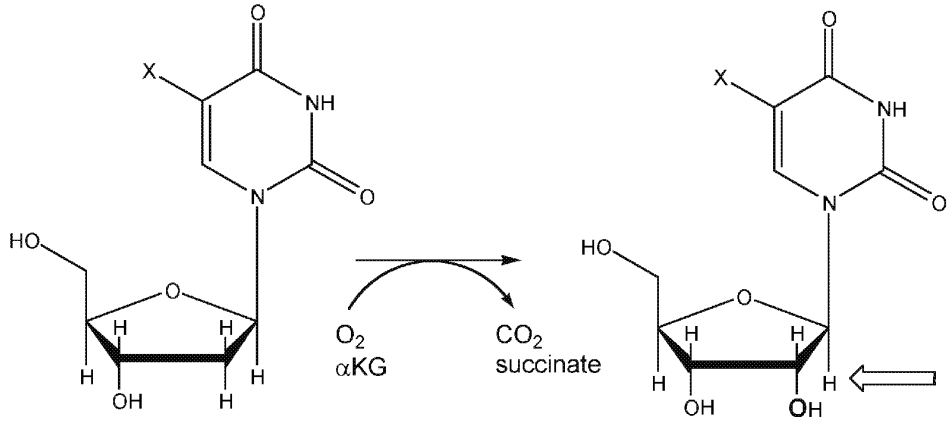
COR-AAD1-00001314



SCHEME 34.

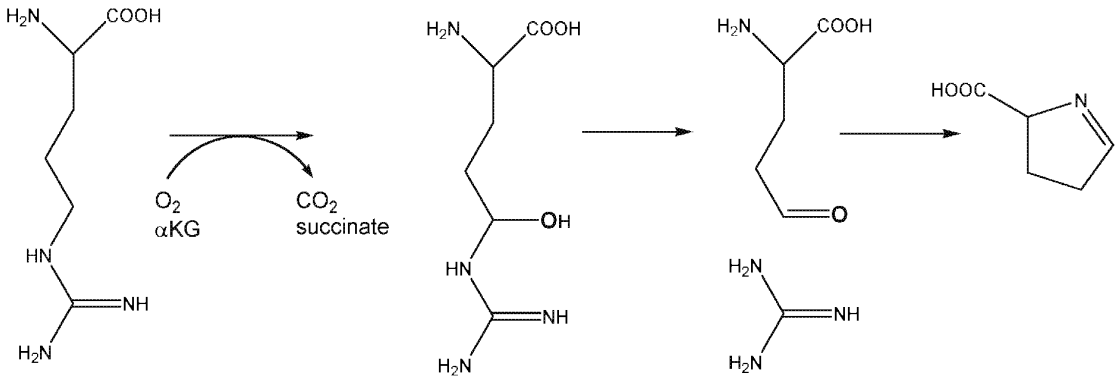
1′-hydroxylase that replaces the hydrogen atom indicated by the arrow with a hydroxyl group for the case where X = H (Stubbe, 1985). No genes have been sequenced or proteins structurally characterized for these enzymes.

Penicillium digitalum and *Pseudomonas syringae* pv. *phaseolicola* PK2 synthesize the plant hormone ethylene by action of an enzyme that is distinct from the ethylene-forming system found in plants (described further in Section II part G). The fungal activity was described in 1986 and shown to require αKG, Fe(II), and L-arginine (Fukuda *et al.*, 1986). Although this enzyme has been purified (Fukuda *et al.*, 1989), little is known about its properties compared to the corresponding bacterial enzyme. A cell-free extract of *P. syringae* initially was suggested to generate ethylene from αKG with no stimulation by arginine (Goto & Hyodo, 1987); however subsequent studies demonstrated that the amino acid and oxygen are essential components of catalysis (Nagahama *et al.*, 1991a). The monomeric enzyme was purified (Nagahama *et al.*, 1991b) and the corresponding gene cloned and expressed in *E. coli* (Fukuda *et al.*, 1992a). The sequence displays the likely metal-binding residues His 189 and Asp 191, along with histidines at positions 268, 284, 305, 309, or 335 and arginines at positions 275, 277, 316, 321, 332, and 346. Site-directed mutagenesis studies revealed that the H189Q variant is inactive, and the H268Q, H305Q, and H335Q variants possess 1.8%, 40%, and 60% of wild-type activity, respectively (Nagahama *et al.*, 1998). On the basis of these results and comparison to sequences of other family members, His 268 and Arg 277 are likely to participate in metal- and αKG-binding (Table 1). The bacterial enzyme appears to catalyze two reactions in a 2:1 stoichiometry (Fukuda *et al.*, 1992b). The minor reaction is more straightforward and involves the Fe(II)/αKG-dependent hydroxylation of arginine with subsequent decomposition of this intermediate and formation of guanidine plus L-Δ^1-pyrroline-5-carboxylate (Scheme 36). The major reaction involves the remarkable oxidative transformation of αKG to three molecules of CO_2 plus one molecule each of H_2O and ethylene (Scheme 37). An intriguing "dual-circuit mechanism" has been proposed in order to account for the two reactions (Fukuda *et al.*, 1992b). In this proposal, αKG forms a Schiff's base with arginine, the coligand binds Fe(II) in a tridentate complex (with only two enzyme side chains binding the metal), and an intermediate formed after reaction with oxygen partitions in such a manner as to give rise to the observed products. Other than the appropriate product stoichiometries, no experimental evidence has been reported to support this biochemically suspect mechanism.



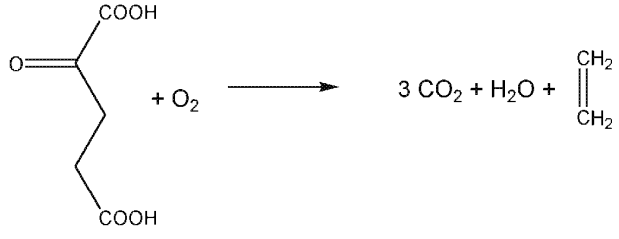
SCHEME 35.

COR-AAD1-00001315



SCHEME 36.

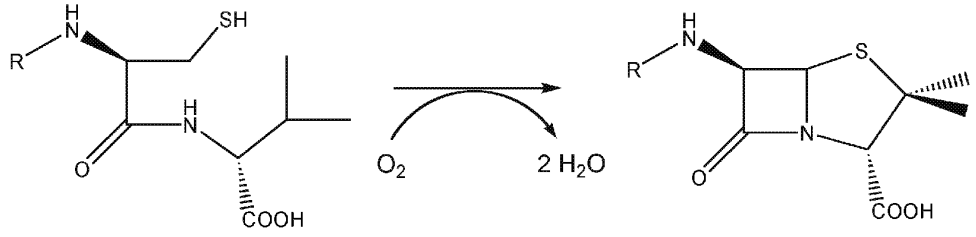
The final putative Fe(II)/αKG-dependent hydroxylase to be discussed is the *E. coli* Gab protein, whose function is still unknown. The structure of this protein was solved as part of a structural genomics effort (Chance *et al.*, 2002). It revealed a fold that, despite the lack of sequence similarity, resembles CAS, DAOCS, and other members of the Fe(II)/αKG-dioxygenase superfamily. The crystallographically observed metal ion, possibly Fe(III) since the protein was crystallized aerobically, is bound to His 160, Asp 162, and His 292. Nearby and appropriately positioned to interact with αKG is Arg 305; however, no electron density for αKG was observed when this molecule was added to the protein. Additional electron density was observed at the position likely to be occupied by substrate, but the identity of this species could not be discerned. The authors speculate that the protein plays a role related to γ-aminobutyrate metabolism since the gene is located in the γ-aminobutyrate operon.



SCHEME 37.

G. Related Enzymes

Isopenicillin *N* synthase (IPNS) is structurally related to the Fe(II)/αKG-dependent dioxygenases, but it does not utilize αKG as a cosubstrate. Rather, this enzyme catalyzes the remarkable reaction illustrated in Scheme 38 in which a linear tripeptide (δ-(L-α-aminoadipoyl)-L-cysteinyl-D-valine) is transformed to a bicyclic structure while carrying out the four-electron reduction of oxygen (Kreisberg-Zakarin *et al.*, 1999; Schofield *et al.*, 1997). The gene encoding this enzyme was first cloned and sequenced from *Cephalosporium acremonium* (Samson *et al.*, 1985), and spectroscopic studies have focused on this recombinant enzyme. The Mössbauer spectrum associated with the high-spin mononuclear Fe(II) site is perturbed by substrate binding, consistent with direct substrate interaction with the metallocenter (Chen *et al.*, 1989). The same study examined the nitric oxide (NO) complex of the enzyme and showed optical spectral changes indicative of substrate thiol binding to the metal. ESEEM studies of Cu(II)-substituted protein (Jiang *et al.*, 1991) were interpreted in terms of two His ligands, whereas ^{1}H nuclear magnetic resonance spectra of Fe(II) and Co(II) derivatives indicated the likely presence of three His ligands (Ming *et al.*,



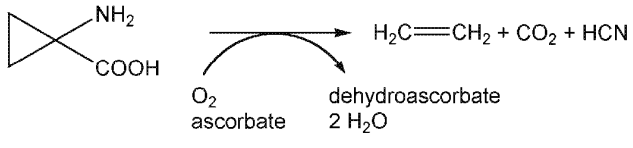
SCHEME 38.

COR-AAD1-00001316

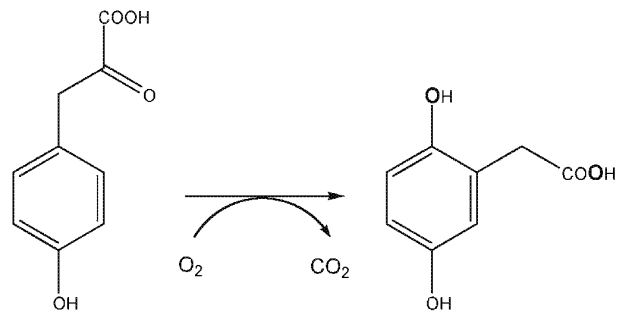
1990, 1991). Results from XAS studies were consistent with 2–3 imidazole ligands along with substrate thiol interaction with the metal (Randall *et al.*, 1993; Scott *et al.*, 1992). Site-directed mutagenesis studies of the genes from *C. acremonium* and *Streptomyces jumonjinensis* identified the likely His^1-X-Asp-X_n-His^2 metal-binding ligands (Borovok *et al.*, 1996; Kreisberg-Zakarin *et al.*, 2000; Tan & Sim, 1996; Tiow-Suan & Tan, 1994) (Table 1), and these assignments were confirmed by crystal structure determination of the *Aspergillus (Emericella) nidulans* enzyme (Table 2) (Roach *et al.*, 1995). An intriguing feature revealed by the structure is that a fourth amino acid side chain, the carboxyl terminal Gln residue, coordinates the metal in the absence of substrate. Additional mutagenesis studies have shown this residue is not important to catalysis (Landman *et al.*, 1997; Sami *et al.*, 1997), a result in accord with additional structural studies showing that substrate displaces the Gln residue (Roach *et al.*, 1997). In addition, crystallographic investigations confirm that substrate thiol binds the metal.

Like IPNS, 1-aminocyclopropane-1-carboxylate (ACC) oxidase (also known as ethylene-forming enzyme) is related by sequence to the Fe(II)/αKG-dependent hydroxylases, yet this enzyme does not utilize αKG as a cosubstrate. ACC is an intermediate in formation of the plant-ripening hormone ethylene from methionine (Adams and Yang, 1979), and ACC oxidase catalyzes the reaction shown in Scheme 39 (Dong *et al.*, 1992). Since the initial cloning and sequencing of the gene encoding tomato ACC oxidase (Hamilton *et al.*, 1991) homologues have been characterized from many plants, and all reveal the typical Fe(II)/αKG dioxygenase superfamily sequence motif (Table 1). Numerous site-directed mutagenesis studies have been carried out (e.g., Lay *et al.*, 1996; Shaw *et al.*, 1996) to support the metal ligand assignments; however, no crystal structure has yet been reported. On the basis of EPR and electron nuclear double resonance (ENDOR) spectroscopic investigations of NO-bound enzyme, alanine (a structural analogue of ACC) was shown to bind the metal ion through both its α-amino and α-carboxylate groups (Rocklin *et al.*, 1999). Such a binding mode for ACC would be reminiscent of the chelate binding observed for αKG in the Fe(II)/αKG-dependent hydroxylases, as described in Sections III and IV. Significantly, the spectroscopic studies of ACC oxidase were interpreted to suggest that ACC and oxygen simultaneously bind to the metal to promote catalysis, whereas the data are incompatible with previously proposed mechanisms involving simultaneous metal binding of ascorbate and oxygen. Kinetic studies of the avocado (Brunhuber *et al.*, 2000) and tomato (Thrower *et al.*, 2001) ACC oxidases are inconsistent in terms of the order of substrate binding to the enzyme, but both suggest the need to bind ascorbate, ACC, and oxygen for multiple rounds of catalysis. An especially intriguing aspect of ACC oxidase is its activation by carbon dioxide (Dong *et al.*, 1992; Pirrung *et al.*, 1993; Poneleit & Dilley, 1993; Smith & John, 1993). On the basis of near-infrared circular dichroism (CD) and magnetic circular dichroism (MCD) spectroscopic studies, the role of CO_2 is now understood in terms of preventing an aberrant reaction that results in enzyme inactivation (Zhou *et al.*, 2002). Thus, binding of ACC to the CO_2-free resting enzyme converts the six-coordinate metal site into a five-coordinate site that reacts with oxygen in an uncoupled reaction, leading to loss of enzyme activity. In contrast, CO_2 stabilizes a six-coordinate ACC-bound state (perhaps by reacting with a metal-bound water molecule to create a metal-bound bicarbonate ligand) until ascorbate is present, at which time the sixth ligand is lost to create a productive oxygen-binding site. Of interest, inactivation of ACC oxidase is accompanied by protein fragmentation due to metal-catalyzed oxidative reactions that cleave the peptide backbone (Barlow *et al.*, 1997; Zhang *et al.*, 1997). It is possible that this aberrant chemistry is catalyzed by an activated oxygen species related to one that normally participates in catalysis.

4-Hydroxyphenylpyruvate dioxygenase (HPPD) is not related in sequence to the Fe(II)/αKG-dependent hydroxylases, but it exhibits close chemical parallels to this enzyme family. As shown in Scheme 40, HPPD catalyzes the oxidative cleavage of an α-keto acid while hydroxylating a substrate, but in this case the α-keto acid group is part of the substrate and the hydroxylation is associated with an "NIH shift" involving acetyl group migration to a different position on the aromatic ring. The enzyme functions in tyrosine catabolism, and inhibitors of HPPD are



SCHEME 39.

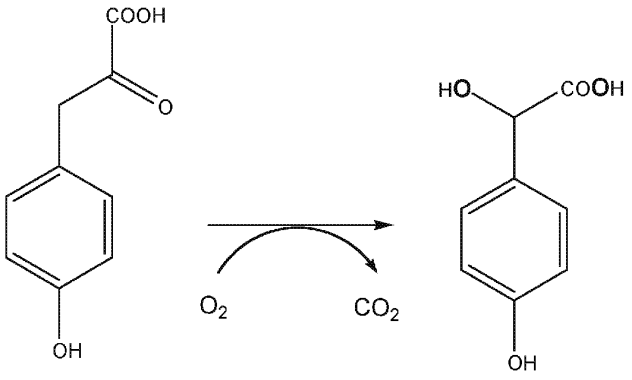


SCHEME 40.

COR-AAD1-00001317

useful as herbicides (preventing the production of plastoquinone) and for human patients with type 1 tyrosinemia (lacking fumarylacetoacetase, the enzyme catalyzing the final step in tyrosine degradation, so that a toxic intermediate accumulates). HPPD was purified from human liver, shown to be a dimer of M_r 87,000, and found to require ferrous ions (Lindblad *et al.*, 1977). Aerobically purified enzyme from *Pseudomonas* sp. strain P.J.874 is inactive and possesses a blue color ($\lambda_{max} \sim 595$ nm, $\varepsilon_{595} = 2.6\ mM^{-1}\ cm^{-1}$) associated with its metal site (0.95 Fe per subunit) (Lindstedt & Rundgren, 1982). Reduction of this enzyme led to the loss of color and gain of activity. On the basis of resonance Raman (rR) spectroscopic investigations, the blue color of oxidized bacterial HPPD was suggested to arise from tyrosinate coordination to Fe(III) (Bradley *et al.*, 1986). Surprisingly, however, the crystal structure of the closely related HPPD from *Pseudomonas fluorescens* reveals a mononuclear iron site with the metal coordinated by His 161, His 240, and Glu 322, and lacking any Tyr residue in its vicinity (Serre *et al.*, 1999). This finding prompted speculation that the enzyme might catalyze a self-hydroxylation reaction to create one or more hydroxy-phenylalanine residues that bind to the oxidized metal ion (Ryle & Hausinger, 2002). Kinetic and spectroscopic studies with recombinant *Streptomyces avermitilis* HPPD apoprotein show an ordered binding of Fe(II), then substrate and oxygen (with a 3600-fold increase in oxygen reactivity if substrate binds first) (Johnson-Winters *et al.*, 2003). These investigations further indicate that substrate binds to the metal in a chelating mode (i.e., via the side chain carboxylate and keto groups) and reveal the presence of an intermediate during its reaction with oxygen (Johnson-Winters *et al.*, 2003). The identity and properties of this intermediate are currently unknown, but are likely to be related to intermediates in the Fe(II)/αKG-dependent hydroxylases (see Section IV).

A reaction closely analogous to that catalyzed by HPPD is carried out by 4-hydroxymandelate synthase (Scheme 41) (Choroba *et al.*, 2000; Hubbard *et al.*, 2000). This enzyme is related in sequence to HPPD and unrelated to the Fe(II)/αKG-dependent hydroxylases. Like HPPD, 4-hydroxymandelate synthase uses 4-hydroxyphenylpyruvate as substrate and decomposes the substrate-associated α-keto acid; however, it hydroxylates the alkyl side chain rather than the aromatic ring. The gene encoding this enzyme is part of the chloroeremomycin gene cluster that also encodes 4-hydroxymandelate oxidase and 4-hydroxyphenylglycine transaminase that provide the 4-hydroxyphenylglycine used for synthesis of the antibiotic. Like HPPD, 4-hydroxymandelate synthase requires ferrous ions for activity. A mutagenesis approach was used to obtain clues about specific protein features responsible for the distinct chemistries of 4-hydroxymandelate synthase and HPPD. Specifically, the *S. avermitilis* HPPD gene was



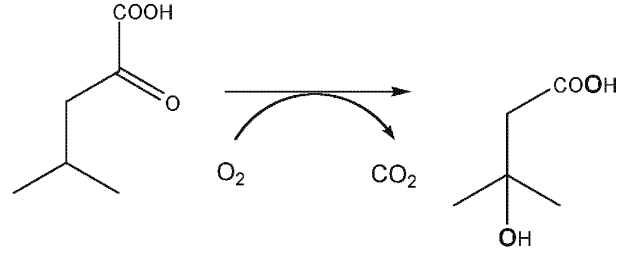
SCHEME 41.

engineered to create a variant enzyme (F337I) that produced 4-hydroxymandelate along with its normal product (Gunsior *et al.*, 2004).

Until recently α-ketoisocaproate dioxygenase also was thought to catalyze a reaction involving hydroxylation of a substrate that provides its own α-keto acid (Scheme 42). Proteins possessing this activity have been purified and characterized from several sources (e.g., Han & Pascal, 1990; Sabourin & Bieber, 1982); however, the rat liver enzyme was shown to be identical to rat liver HPPD (Crouch *et al.*, 1997). Thus, *a separate α-ketoisocaproate dioxygenase may not exist.*

III. STRUCTURAL STUDIES

Structurally characterized Fe(II)/αKG-dependent hydroxylases and related proteins are listed in Table 2 along with their bound metal and substrates, crystallographic resolution, and PDB accession numbers. The structures of the Fe(II)/αKG-dependent proteins are unified by possession of a "jellyroll" structural fold comprised of eight β-strands forming two, four-stranded sides (Figure 1, left). Surrounding this protein core, and likely contributing to its stability, are additional α-helices, β-strands, and protein loops as illustrated for the first representative of this protein family, the non-αKG-dependent enzyme IPNS (Figure 1,



SCHEME 42.

COR-AAD1-00001318

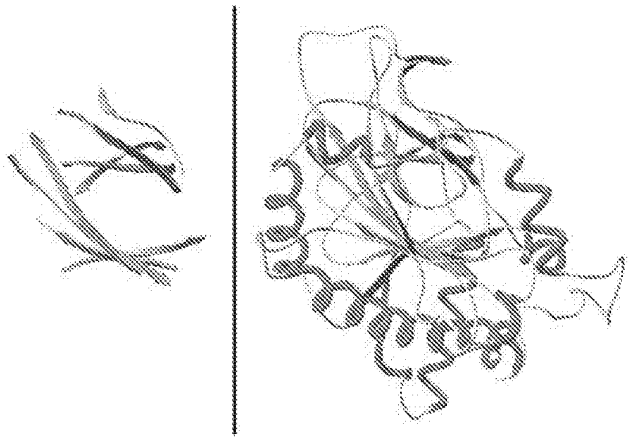


FIG. 1. Jellyroll fold found in all Fe(II)/αKG-dependent dioxygenases. On the left are the eight β-strands that combine to form the core of all Fe(II)/αKG dioxygenases, as first identified in the related enzyme IPNS (Roach *et al.*, 1995). Additional β-strands and α-helices are present at the amino terminus, and inserted into the jellyroll fold and at the carboxyl terminus, as illustrated for the complete IPNS structure shown on the right. These additional features stabilize the core structure, participate in interprotein contacts, and modify the substrate-binding site. The IPNS structure contains a total of 10 α-helices and 16 β-strands, with the jellyroll motif incorporating the β^5 and β^{8-14} secondary structure features.

right). The insertions and extensions to this jellyroll fold differ for each protein and help to define the substrate-binding sites and establish protein–protein contact surfaces. IPNS (Roach *et al.*, 1995), CAS (Zhang *et al.*, 2000), and ANS (Wilmouth *et al.*, 2002) are monomeric proteins, and DAOCS equilibrates between monomeric and trimeric species in solution, but crystallizes as a trimeric protein (Lloyd *et al.*, 1999a). In contrast, FIH (Dann *et al.*, 2002; Elkins *et al.*, 2003; Lee *et al.*, 2003), proline 3-hydroxylase (Clifton *et al.*, 2001), and TauD (Elkins *et al.*, 2002; O'Brien *et al.*, 2003) are dimeric proteins, AtsK (Müller *et al.*, 2004) and the Gab protein (Chance *et al.*, 2002) are tetrameric, and CarC is hexameric (a dimer of trimers) (Clifton *et al.*, 2003).

The active site structures of the Fe(II)/αKG dioxygenases exhibit nearly identical arrangements of the three amino acid side chains (associated with the His^1-X-Asp/Glu-X_n-His^2 motif) that bind one face of the metal. On the basis of the αKG binding mode, the active site structures are resolved into two distinct categories (Figures 2 and 3). In both classes, the C-1 carboxylate and C-2 keto group chelate the metal and the C-5 carboxylate is stabilized by a salt bridge to an Arg residue or by ionic interaction with a Lys side chain. Additional hydrogen-bond interactions are found to participate in C-5 carboxylate binding, and critically positioned Arg residues often help orient the C-1 carboxylate (not shown). One category of enzymes is represented by CAS (Zhang *et al.*, 2000), TauD (Elkins *et al.*, 2002), and FIH (Dann *et al.*, 2002; Elkins *et al.*, 2003; Figure 2), along with AtsK (Müller *et al.*, 2004). In these enzymes, His^1 and the carboxylate ligand are located approximately in plane with αKG (C-1 carboxylate opposite His^1 and C-2 ketone opposite the acidic residue). The primary substrate does not bind to the metal, but rather binds nearby and above this plane. An open metal coordination site exists near the primary substrate and opposite His^2 in this "in line" binding mode. In the second category of enzymes, represented by CarC (Clifton *et al.*, 2003) and ANS (Wilmouth *et al.*, 2002; Figure 3), the C-1 carboxylate of αKG binds opposite His^2, and the keto group again is located opposite the acidic residue. Thus, the open coordination site is located opposite His^1 and oriented away from the substrate in this "off line" binding mode. The two αKG binding modes may simply represent conformational flexibility in binding this cosubstrate, with different protein structures tending to stabilize one mode over the other. For example, the αKG binding mode reverses for CAS (with a different substrate bound) in the presence of the oxygen analogue NO (Figure 4). Furthermore, the C-1 carboxylate of αKG flips from the "in line" to the "off line" mode when Fe(II) is replaced by sodium ion in AtsK (Müller *et al.*, 2004). Thus, it is reasonable to posit that αKG conformational flexibility in ANS and/or CarC may allow oxygen to bind opposite His^2 and near the substrate for productive catalysis. As described below, these chelate binding modes cannot account for the chemistry observed with DAOCS.

Multiple structures of DAOCS and its variants have revealed a series of intriguing features related to this active site. Figure 5 (left) depicts the αKG-bound holoprotein in which the cosubstrate C-1 carboxylate binds opposite His^2 and the keto group coordinates opposite the Asp (Valegård *et al.*, 1998), analogous to the situation in the "off line" enzymes ANS and CarC shown in Figure 3. In contrast to the latter enzymes, however, the substrate-binding site of DAOCS overlaps the binding site of the cosubstrate; thus, simultaneous binding cannot occur (Valegård *et al.*, 2004). To illustrate, in Fe(II)DAOCS/penicillin G the substrate overlaps both the keto group and C-5 carboxylate positions of the αKG (structure not shown). For DAOCS apoprotein crystals soaked with Fe(II), αKG, and penicillin G (or ampicillin), the structures reveal a mixture of two species containing either the cosubstrate or the substrate. Notably, the substrate-binding mode differs for these samples compared to the case where crystals are soaked with only Fe(II) and substrate; i.e., the presence of αKG somehow causes the penicillin G (or ampicillin) to orient so that its ring sulfur atom binds Fe(II) (Valegård

COR-AAD1-00001319

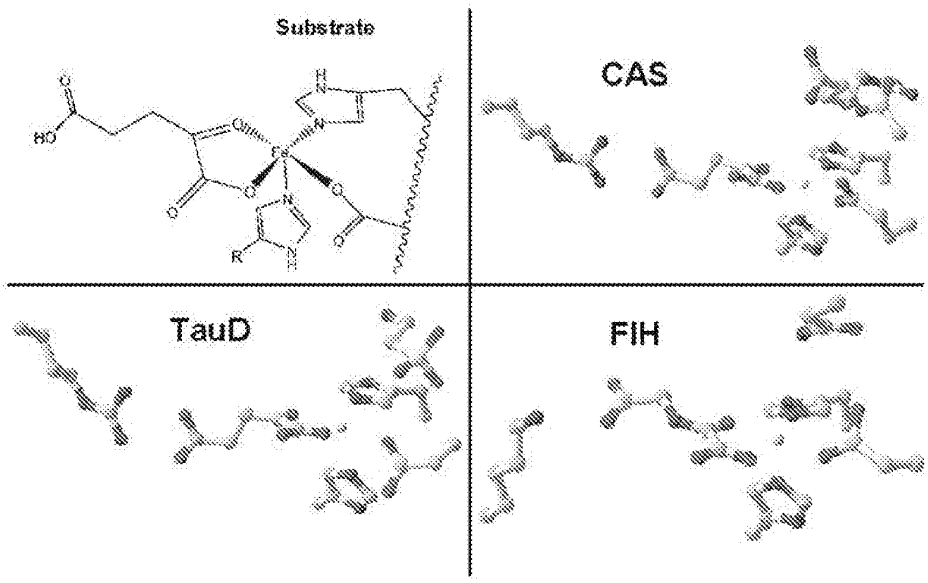


FIG. 2. "In-line" active site structure. Some Fe(II)/αKG-dependent dioxygenases bind their substrates as depicted in this model. The αKG chelates the metal approximately in the same plane as His^1 and the carboxylate ligand of the His^1-X-Asp/Glu-X_n-His^2 motif (with the C-1 carboxylate opposite His^1 and the keto group opposite the acidic residue). The His^2 ligand lies below this plane, leaving an open coordination site above the plane and directed toward the primary substrate. Examples of structures illustrating this substrate-binding mode include CAS (with bound *N*-acetyl-L-arginine) (Zhang *et al.*, 2000), TauD (with bound taurine) (Elkins *et al.*, 2002; O'Brien *et al.*, 2003), FIH (with a bound peptide that includes the Asn residue shown) (Elkins *et al.*, 2003), and AtsK (with bound 2-ethylhexyl-1-sulfate, not shown) (Müller *et al.*, 2004). The αKG C-5 carboxylate forms a salt bridge to Arg residues in CAS, TauD, and AtsK, whereas a Lys residue is used to stabilize this group in FIH. Other residues form hydrogen-bonds to the C-5 carboxylate, and additional Arg residues orient the C-1 carboxylate of αKG (not shown). The FIH structure was obtained by using *N*-oxaloylglycine as a substitute for αKG.

et al., 2004). This binding mode is remarkably similar to that observed for product (isopenicillin N) binding to IPNS (Figure 5, right) (Burzlaff *et al.*, 1999). A product-bound structure of DAOCS reveals similar binding interactions (Valegård *et al.*, 2004). The enzyme was also structurally characterized with another product of the reaction, succinate (Valegård *et al.*, 2004). One carboxylate of this molecule binds Fe(II) opposite the Asp ligand, and the other forms a salt bridge to Arg, a pattern expected to arise upon decarboxylation of αKG. A similar structure was observed for the succinate-bound form of Δ307A DAOCS (with the last four residues replaced by Ala), but the salt bridge is absent and a linear, unhydrated CO_2 molecule (derived from bicarbonate added to buffer) binds opposite His^2 (Lee *et al.*, 2001a).

IPNS is not an Fe(II)/αKG dioxygenase, but it is related in sequence and structure to this enzyme family. The initial IPNS structure, obtained for the Mn-substituted form of this enzyme in the absence of substrate, indicated that the metal is bound via His 214, Asp 216, Gln 330, and His 270 within an eight β strand jellyroll structural motif (Roach *et al.*, 1995). Followup studies showed the Gln residue is displaced by substrate (δ-(L-α-aminoadipoyl)-L-cysteinyl-D-valine) that binds via its thiol directly to the mononuclear Fe(II) active site (Roach *et al.*, 1997). Exposure of these crystals to the oxygen analogue NO identified a likely site of oxygen binding, illustrated in Figure 6. Alternatively, direct exposure of these crystals to high pressures of oxygen afforded the bicyclic product-bound structure, i.e., the IPNS complex with bound isopenicillin N (Burzlaff *et al.*, 1999). Similar treatment with high-pressure oxygen for short time periods was used with crystals containing either of two substrate analogues, δ-(L-α-aminoadipoyl)-L-cysteinyl-L-*S*-methylcysteine and δ-(L-α-aminoadipoyl)-L-cysteinyl-D-α-hydroxyvaleryl ester, leading to structures containing bound monocyclic compound (Burzlaff *et al.*, 1999) or bound thiocarboxylate (Ogle *et al.*, 2001), respectively. The in-crystal formation of the monocyclic product is consistent with the generation of a postulated reaction pathway intermediate, while formation of the thiocarboxylate suggests attack of a thioaldehyde by a hydroperoxide intermediate. These and

COR-AAD1-00001320

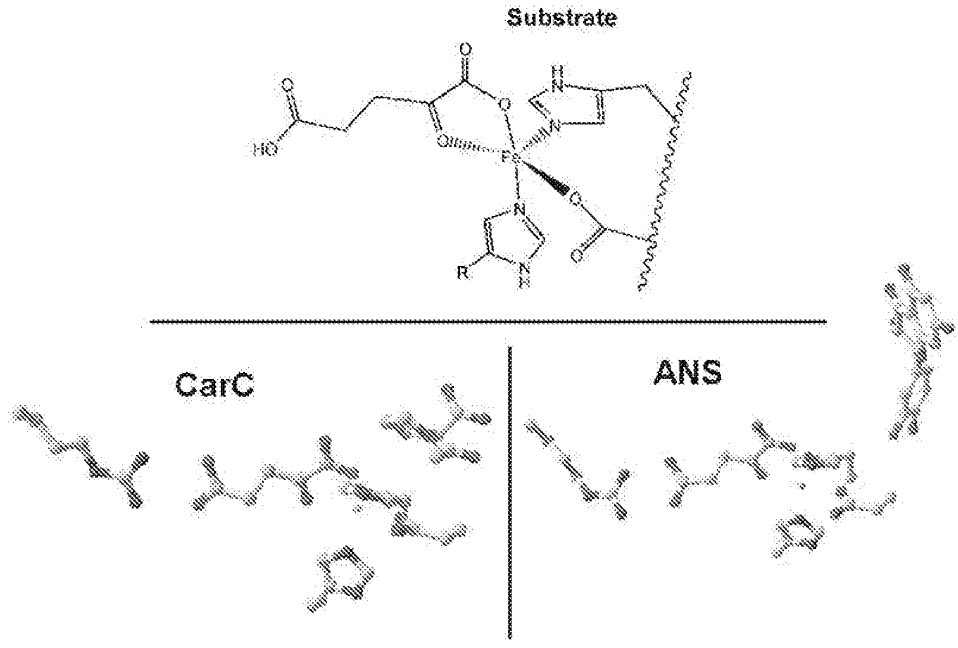


FIG. 3. "Off-line" active site structure. Some Fe(II)/αKG-dependent dioxygenases bind their substrates as depicted in this alternative model. While the αKG keto group is similarly positioned opposite the acidic ligand, the αKG C-1 carboxylate is shifted so as to bind opposite His^2. The open coordination site is thus located opposite His^1, a site that is not directed toward the substrate. Examples of structures illustrating this arrangement include ANS (with bound dihydroquercetin) (Wilmouth *et al.*, 2002) and CarC (with bound L-*N*-acetylproline) (Clifton *et al.*, 2003).

additional structural studies (Long *et al.*, 2003) are providing important insights that will help to define how IPNS carries out its remarkable chemistry.

HPPD is thought to catalyze chemistry similar to the Fe(II)/αKG-dependent hydroxylases, but this protein possesses a clearly distinct structure (Serre *et al.*, 1999) related to the extradiol cleavage enzymes 2,3-dihydroxybiphenyl dioxygenase (Han *et al.*, 1995; Senda *et al.*, 1996) and catechol 2,3-dioxygenase (Kita *et al.*, 1999). Rather than possessing a jellyroll fold, two barrel-like domains are found, with Fe(II) bound to His 161, His 240, and Glu 322 in the second domain. Although the 2 His-1 carboxylate ligand set is similar to the Fe(II)/αKG dioxygenases, the architecture of this site is clearly different. No structure is available for substrate-bound enzyme, but the α-keto acid-containing substrate is likely to chelate the metal ion on the basis of spectroscopic studies (Johnson-Winters *et al.*, 2003).

IV. MECHANISTIC STUDIES

This section examines several proposals for the enzymatic mechanisms of Fe(II)/αKG-dependent hydroxylases and related enzymes while reviewing the associated experimental evidence derived from crystallographic, spectroscopic, and isotope incorporation studies.

A. In-Line Hydroxylation Mechanisms

As a starting point for discussion, I describe the intermediates of the representative hydroxylation mechanism illustrated in Scheme 43. This mechanism has evolved from a model first postulated for prolyl hydroxylase over 20 years ago (Hanauske-Abel & Günzler, 1982).

Intermediate **A** depicts resting enzyme with Fe(II) bound to the three protein side chains found in the His^1-X-Asp/Glu-X_n-His^2 sequence. All Fe(II)/αKG-dependent hydroxylases, with the exception of aspartyl (asparaginyl) β-hydroxylase, contain this signature motif (Table 1). Furthermore, all protein structures available for this class of enzymes disclose this metal-binding mode, with the side chains found in the same orientation on one face of the metal (Table 2). The crystal structure of Fe(II)-bound but substrate- and αKG-free DAOCS (not a hydroxylase, but found in the same superfamily of enzymes) clearly reveals three additional water molecules in the metal coordination sphere (Valegård *et al.*, 1998). Crystal structures also are reported for Fe(II)-bound proline 3-hydroxylase (Clifton *et al.*, 2001) and FIH (Dann *et al.*, 2002). Metal-bound

COR-AAD1-00001321

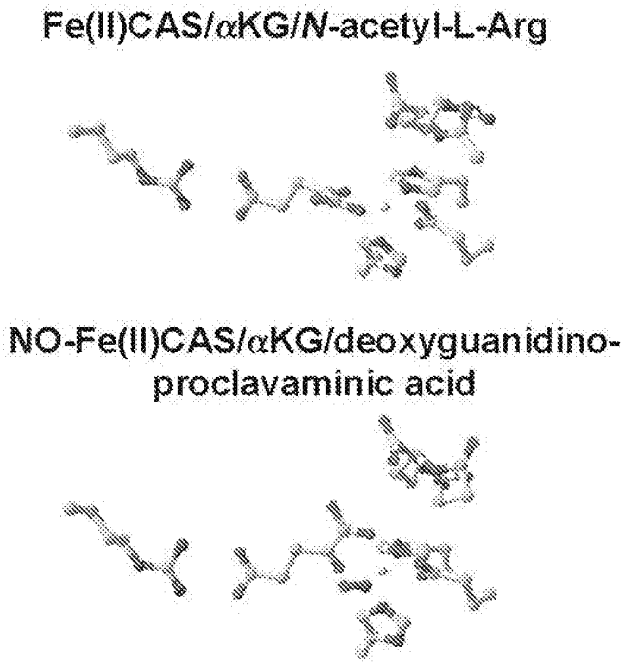


FIG. 4. Conformational flexibility of cosubstrate binding. CAS active site structures are compared for Fe(II)- and αKG-bound forms of the enzyme containing *N*-acetyl-arginine versus deoxyguanidinoproclavaminic acid (Zhang *et al.*, 2000). Significantly, the oxygen analogue NO is bound to the metal ion in the latter structure and the αKG binding mode has shifted.

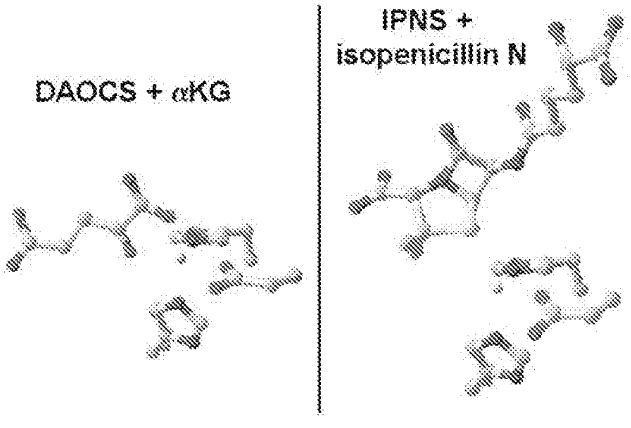


FIG. 5. Active site structures of αKG-bound DAOCS and product-bound IPNS. The cosubstrate- and substrate- (αKG and penicillin) binding sites of DAOCS overlap, precluding simultaneous binding of both molecules (Valegård *et al.*, 2004). The αKG/Fe(II)DAOCS active site structure (Valegård *et al.*, 1998) is illustrated on the left, whereas the substrate-bound DAOCS active site is represented on the right by the very similar structure reported for product (isopenicillin N)-bound IPNS (Burzlaff *et al.*, 1999).

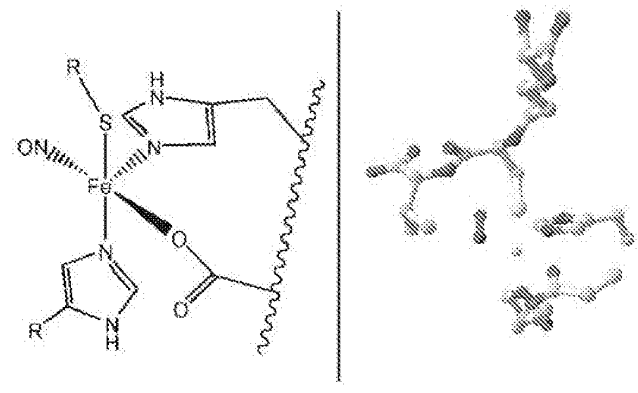
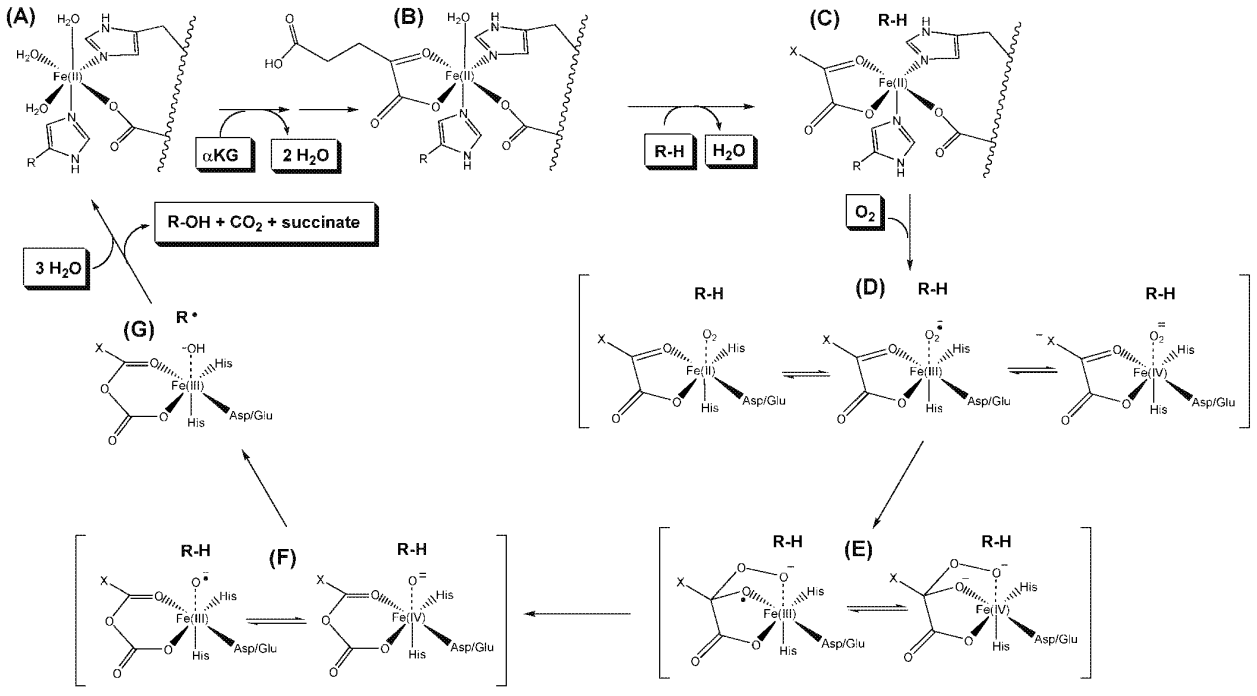


FIG. 6. Substrate-bound IPNS in the presence of the oxygen analogue NO. The active site of IPNS is depicted for enzyme-containing bound δ-(L-α-aminoadipoyl)-L-cysteinyl-D-valine and NO (Roach *et al.*, 1997).

solvent molecules were not observed in these lower resolution structures; although a tartrate molecule (from the crystallization buffer) was found to chelate Fe(II) in the case of FIH. The metal site displays no UV/visible spectral features; however, other types of spectroscopy have been used to characterize the metallocenter. Near-infrared magnetic circular dichroism (MCD) spectroscopy of CAS identified the two d $\rightarrow$ d transitions ($\sim$8,500 cm^{-1} and $\sim$10,000 cm^{-1} at 278 K) expected of a six-coordinate distorted octahedral center (Pavel *et al.*, 1998; Zhou *et al.*, 1998; Zhou *et al.*, 2001). On the basis of Fe(II)-dependent quenching of the intrinsic tryptophan fluorescence of TfdA, a metal ion K_d of 7.45 $\pm$ 0.61 μM was estimated (Dunning Hotopp *et al.*, 2003). Kinetic studies show similar μM Fe(II) affinities of related enzymes. By exploiting the weak metal binding affinity of this site, ^{57}Fe was substituted into TauD, allowing analysis by Mössbauer spectroscopy (Price *et al.*, 2003b). The observed isomer shift and quadrupole splitting (δ and ΔE_Q, 1.27 $\pm$ 0.05 mm/s and 3.06 $\pm$ 0.05 mm/s, respectively) are consistent with high-spin Fe(II) in this holoprotein.

Intermediate **B** corresponds to enzyme with bound Fe(II) and chelated αKG. As summarized in Table 2, structures of this species are available for FIH, CAS, DAOCS, CarC, and AtsK. The five-membered ring formed by αKG chelation of the metal is associated with metal-to-ligand charge-transfer (MLCT) transitions (17,820, 20,820, and 24,070 cm^{-1}), as revealed by UV/visible, CD, and MCD spectroscopy of CAS (Pavel *et al.*, 1998). These CAS studies also reveal the d $\rightarrow$ d transitions expected of a six-coordinate complex (Zhou *et al.*, 2001). The lilac-colored chromophore exhibits a λ_{max} at 530 nm for both αKG/Fe(II)TauD (Ryle *et al.*, 1999) and αKG/Fe(II)TfdA (Hegg *et al.*, 1999b) (ε_{530} of 140 M^{-1} cm^{-1}

COR-AAD1-00001322



SCHEME 43.

and 180 M^{-1} cm^{-1}, respectively), whereas the feature is observed at 500 nm in αKG/Fe(II)AlkB (Trewick *et al.*, 2002). Stopped-flow UV/visible spectroscopic methods were used to monitor the rate of chromophore formation in the case of TauD (Ryle *et al.*, 1999), and the absorption changes were found to be independent of αKG concentration. These findings were interpreted to indicate that αKG binds in two steps, with initial binding of cosubstrate to the metal ion (e.g., by the αKG carboxylate) followed by a rate-determining conformational change (accompanied by metal chelation) to create the chromophore. Excitation into this TauD absorption band gives rise to two resonance Raman (rR) features at 460 cm^{-1} and 1686 cm^{-1}, both of which are sensitive to ^{18}O labeling of the carbonyl group (i.e., analysis in $H_2{}^{18}O$ leads to shifts of these vibrations to 451 cm^{-1} and 1648 cm^{-1}) (Ho *et al.*, 2001).

Intermediate **C** represents an enzyme with cosubstrate coordinated directly to the metal and the primary substrate bound nearby in the active site. Protein structures are available for FIH, CAS, CarC, ANS, TauD, and AtsK with both the cosubstrate and substrate bound (see Table 2); however, Scheme 43 focuses on the "in line" enzymes (described in Section III and shown in Figure 2), where the open coordination site is directed towards the substrate. This situation is found in FIH, CAS, TauD, and AtsK. The crystal structures of FIH (Elkins *et al.*, 2003), TauD (Elkins *et al.*, 2002; O'Brien *et al.*, 2003), and AtsK (Müller *et al.*, 2004) reveal a five-coordinate metal site indicating the loss of a water molecule as substrate binds. The CAS metallocenter, by contrast, remains 6-coordinate in the crystal, but the metal-to-water bond distance increases from 2.2 to 2.35 Å (Zhang *et al.*, 2000). Substrate-induced changes to the metallocenters of Fe(II)/αKG-dependent hydroxylases also are apparent when using various spectroscopic methods. For example, perturbations to the UV/visible spectra of αKG/Fe(II)TauD and αKG/Fe(II)TfdA are observed when the appropriate substrates bind to these enzymes (Hegg *et al.*, 1999b; Ryle *et al.*, 1999). In both cases, the MLCT features increase slightly in intensity, sharpen so as to reveal three overlapping transitions, and undergo a blue shift (resulting in 520 nm (TauD) and 515 nm (TfdA) as λ_{max} of the major transitions). Stopped-flow UV/visible spectroscopic methods demonstrate that substrate binds to αKG/Fe(II)TauD within the dead time of the instrument (7 ms) (Ryle *et al.*, 1999). The binding of taurine to αKG/Fe(II)TauD also leads to upshifts in the rR vibrations (to 470 cm^{-1} and 1688 cm^{-1} in H_2O, or 460 cm^{-1} and 1653 cm^{-1} in $H_2{}^{18}O$) associated with the chelated metal site (Ho *et al.*, 2001). On the basis of comparisons to similar rR changes in model compounds, this result is attributed to the conversion of a six-coordinate site to a five-coordinate metal center. The Mössbauer parameters of this site also shift ($\delta = 1.16 \pm 0.05$ mm/s and $\Delta E_Q = 2.76 \pm 0.05$ mm/s) in accord with a reduction in coordination number (Price *et al.*, 2003b). Convincing

COR-AAD1-00001323

evidence for a substrate-induced conversion was obtained for αKG/Fe(II)CAS by using UV/visible, near-infrared CD/MCD, and variable-temperature variable-field MCD approaches (Zhou *et al.*, 1998, 2001). The d → d transitions in the near infrared at ~5,200 cm^{-1} and ~8,700 cm^{-1} indicate a five-coordinate site, and the observed MLCT transitions (350 and 500 nm) confirm metal ion chelation by αKG. Substrate-induced creation of an open coordination site on the metal center to allow for oxygen binding is an appealing mechanism for minimizing oxygen reactivity with Fe(II) until substrate is present.

The reaction of oxygen with intermediate **C** and subsequent chemistry leading to substrate hydroxylation are the least well characterized aspects of the mechanism of Fe(II)/αKG-dependent hydroxylases. Oxygen addition initially yields the isoelectronic Fe(II)-O_2, Fe(III)-superoxo, or Fe(IV)-peroxo species (intermediates **D**). Nucleophilic attack on the keto group of αKG results in an Fe(III) or Fe(IV) peroxyhemiketal bicyclic complex (intermediates **E**). By inserting an oxygen atom into the C1-C2 carbon–carbon bond of αKG, the cosubstrate is transformed to a carbonate-succinate mixed anhydride bound to an Fe(III)-hydroxyl radical or Fe(IV)-oxo species (intermediates **F**). Finally, the activated oxygen species positioned near the substrate abstracts a hydrogen atom and forms a substrate radical plus an Fe(III)-hydroxide species (intermediate **G**). Subsequent recombination (joining the hydroxyl group and the substrate radical) results in hydroxylated substrate and restores the Fe(II) form of the enzyme. These last steps in the reaction are comparable to the radical rebound reaction observed in heme-containing oxygenases (Sono *et al.*, 1996). Experimental evidence for an Fe(IV)-oxo intermediate has been obtained for TauD and will be described in Section IV part B.

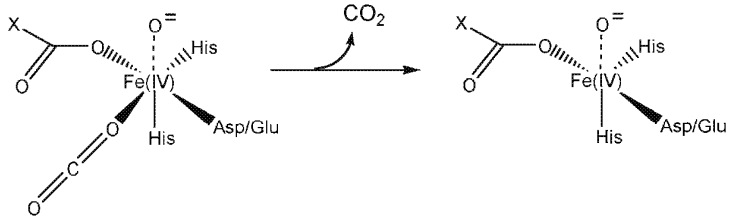
It is important to note that the carbonate-succinate mixed-anhydride species shown in Scheme 43 (and later) could reasonably transform to yield the separately bound succinate and CO_2 products. Alternatively, αKG may decompose directly to these products without the intermediacy of the anhydride. In either case, the resulting intermediates would be expected to exhibit facile loss of CO_2 to expose an extra metal coordination site as illustrated in Scheme 44.

B. Evidence for an Fe(IV)-oxo Intermediate in TauD

As stated earlier, the least well-understood steps of the hydroxylase mechanism occur after oxygen addition. TauD is the only Fe(II)/αKG-dependent hydroxylase for which an intermediate has been directly detected at this stage of the reaction. UV/visible stopped-flow spectroscopy was used to reveal the existence of a species exhibiting an absorption maximum near 318 nm that forms 20–25 ms after mixing taurine/αKG/Fe(II)TauD with oxygen at 5°C (Price *et al.*, 2003b). This intermediate was trapped by a rapid freeze-quench approach, and the novel species was characterized by Mössbauer and EPR spectroscopies. On the basis of its well-resolved quadrupole doublet ($\delta = 0.31 \pm 0.03$ mm/s and $\Delta E_Q = 0.88 \pm 0.03$ mm/s, assigned to an integer spin species with $S \geq 2$) and the observation of an Fe(III) EPR signal after cryoreduction, the intermediate was identified as some type of Fe(IV)-containing species. Significantly, the rate of decay of the intermediate exhibits a 37-fold decrease when using deuterated taurine. This substantial kinetic isotope effect indicates that the Fe(IV) intermediate is directly responsible for hydrogen atom abstraction from the substrate (Price *et al.*, 2003a). Continuous-flow rR difference spectra (i.e., the rR spectrum of sample mixed with $^{16}O_2$ minus that of sample mixed with $^{18}O_2$) acquired at −38°C after a 0.22 s delay reveals isotope-sensitive vibrations assigned to an Fe(IV)-oxo species (787 cm^{-1} for ^{18}O and 821 cm^{-1} for ^{16}O) as well as additional vibrations (555 cm^{-1} for ^{18}O and 583 cm^{-1} for ^{16}O) that are not yet assigned but may be associated with an Fe(III)-superoxo species (Proshlyakov *et al.*, 2004). Analysis of TauD active site variants provides evidence that Fe(IV)-oxo intermediate is stabilized by a more hydrophobic environment (Ryle *et al.*, unpublished data). Because clear evidence for an Fe(IV)-oxo intermediate is available for TauD, putative mechanistic pathways for other enzymes also will emphasize this species.

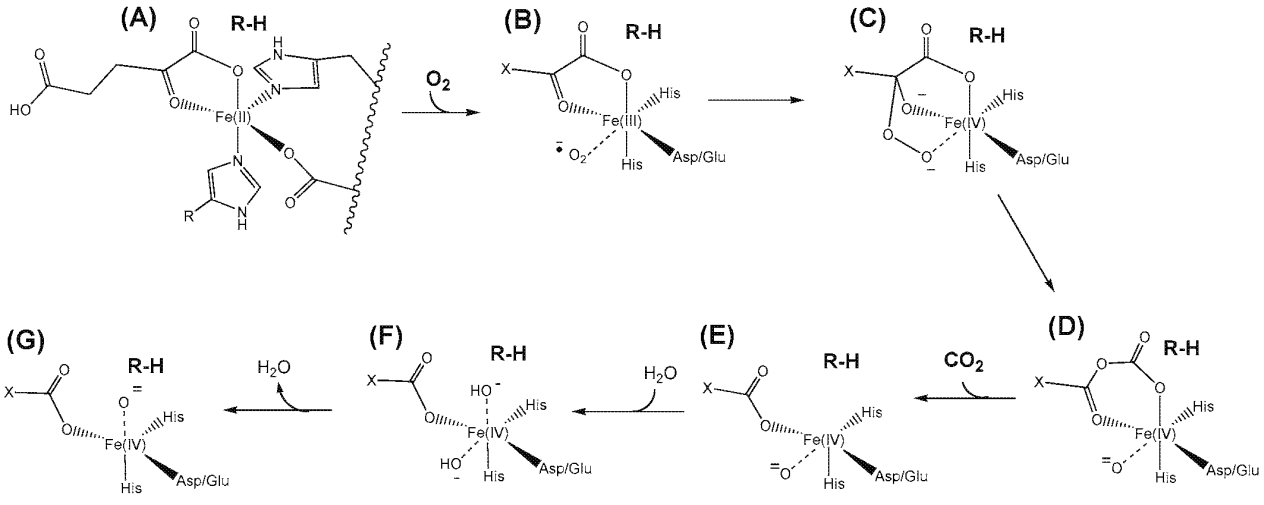
C. Off-Line Hydroxylation Mechanisms

For αKG-dependent dioxygenases classified in the "off line" αKG binding mode (Figure 3), such as CarC (Clifton *et al.*, 2003) and ANS (Wilmouth *et al.*, 2002), two



SCHEME 44.

COR-AAD1-00001324

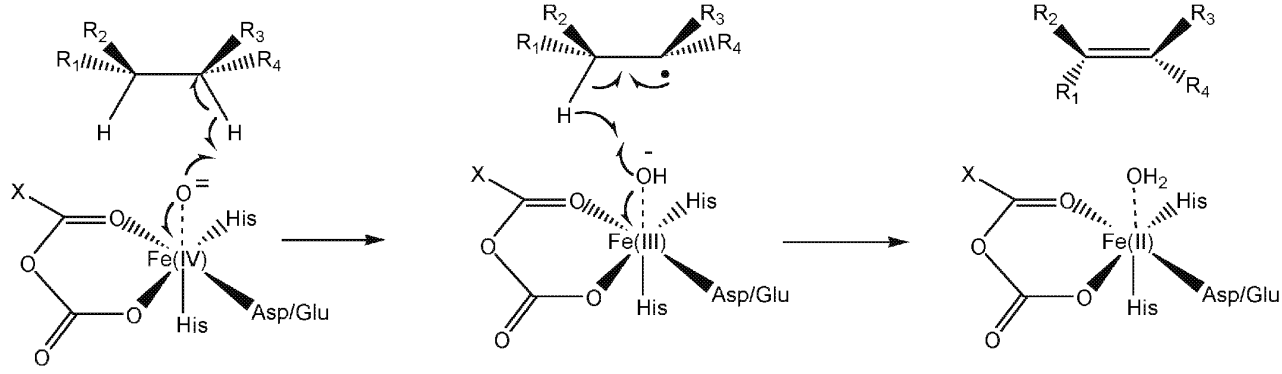


SCHEME 45.

reasonable options can explain how an activated oxygen atom becomes appropriately positioned near the substrate. One possibility is that conformational flexibility of αKG allows it to reorient to adopt an "in line" geometry, so that subsequent steps are equivalent to those in Scheme 43. Alternatively, the "off line" enzymes may undergo a series of steps analogous to those of Scheme 43, but resulting in an Fe(IV)-oxo (or Fe(III)-hydroxyl radical) species positioned away from the substrate (Scheme 45, species **A** to **D**; note that only one possible species is shown here for each intermediate). Loss of CO_2 provides an open coordination site to which water can bind (**D** to **F**). Subsequent loss of water from the dihydroxylated metal site (**F** to **G**) results in a "ferryl flip" (Zhang *et al.*, 2002) to position the activated oxygen near the substrate. Hydrogen atom abstraction and radical recombination can then proceed as described earlier to achieve substrate hydroxylation.

D. Desaturation Mechanisms

The mechanism of αKG-dependent desaturases is likely to be closely related to that of the αKG-dependent hydroxylases. As described in Section II, these enzymes include CarC, flavone synthase, flavonol synthase, and ANS, as well as one reaction catalyzed by CAS. One reasonable mechanism to explain this chemistry invokes an Fe(IV)-oxo abstraction of a hydrogen atom, radical recombination to give the hydroxylated product, followed by a dehydration reaction to form the double bond. Arguing against this mechanism, the hydroxylated intermediate generally is not observed as a side product of these enzymes, nor can it be used as a substrate for dehydration reactions. An alternative mechanism also begins with formation of the Fe(IV)-oxo species, but it is followed by two hydrogen atom transfers from the substrate to directly produce the product as illustrated in Scheme 46.



SCHEME 46.

COR-AAD1-00001325

E. Ring-Expansion Mechanisms

DAOCS most probably utilizes an Fe(IV)-oxo (or Fe(III)-hydroxyl radical) intermediate to catalyze an αKG-dependent ring expansion reaction. Structural and kinetic studies suggest that αKG and the substrate penicillin N cannot simultaneously bind to the active site (Valegård *et al.*, 2004), hence this enzyme does not fall into the "in line" or "off line" categories. Rather, it appears that oxygen reacts with αKG/Fe(II)DAOCS (depicted in Figure 5, left panel) to produce an activated oxygen intermediate in the absence of substrate; CO_2 and succinate then dissociate and penicillin N later binds to allow further chemistry. Assuming that the DAOCS protein environment is capable of adequately stabilizing the standard activated oxygen species, the first portion of the DAOCS reaction is able to be accommodated by the chemistry shown in steps **A** through **E** of Scheme 45. Alternatively, the Fe(IV)-oxo species has been proposed to derive from a mechanism involving a persuccinic acid intermediate (Scheme 47) (Valegård *et al.*, 2004). *Ab initio* quantum molecular dynamics and simulated annealing calculations indicate that the planar peracid is less reactive than a ferryl species and should be a more stable intermediate. Arguing against a more general mechanism involving a peracid for this family of enzymes is the lack of observed activity when synthetic persuccinic acid (in place of αKG plus oxygen) is added to prolyl hydroxylase or pyrimidine deoxyribonucleoside 2′-hydroxylase (Abbott & Udenfriend, 1974). It is possible that the peracid behaves differently when added to the enzyme than when generated *in situ*, or it may be that different Fe(II)/αKG-dependent dioxygenases use distinct pathways to generate the terminal oxygenating intermediate.

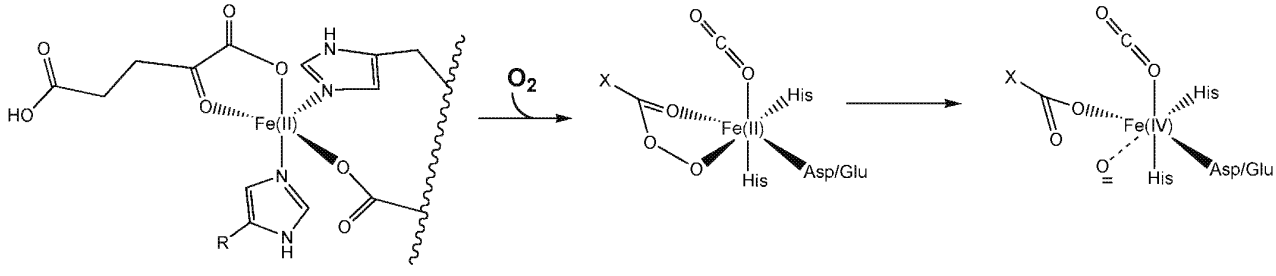
Once DAOCS has generated a reactive oxygen species, the second half of the reaction takes place. One possible mechanism to account for the ring-expansion reaction is illustrated in Scheme 48 (Valegård *et al.*, 2004). This sequence suggests that CO_2 and succinate dissociate (**A** to **B**), penicillin N binds (**B** to **C**, yielding a structure analogous to that shown in Figure 5, right panel), hydrogen atom abstraction yields a substrate radical and Fe(III)-OH species (**C** to **D**), the radical rearranges (**D** to **F**), and a second hydrogen atom transfer (**F** to **G**) yields product plus Fe(II)-OH_2. Note that the latter step is analogous to that proposed for desaturases and shown in Scheme 46.

F. IPNS Mechanism

IPNS also is likely to use an Fe(IV)-oxo intermediate during its remarkable bicyclization reaction to form the antibiotic β-lactam unit. In this case, the critical intermediate is generated without use of αKG. Extensive studies over the years, primarily carried out by the Oxford group (Baldwin & Abraham, 1988; Schofield *et al.*, 1997), have led to the proposed mechanism shown in Scheme 49. An end-on Fe(IV)-peroxo species is suggested to deprotonate the methylene position of the cysteinyl residue (**C**). The O—O bond of the Fe(II)-hydroperoxide is heterolytically cleaved as the amide proton is removed and the four-membered ring forms (**D**). The resulting Fe(IV)-oxo species carries out hydrogen atom abstraction, leading to formation of the five-membered ring (**E** to **G**) (Burzlaff *et al.*, 1999; Ogle *et al.*, 2001; Roach *et al.*, 1995, 1997). As noted in Section III, structures are available depicting several "snapshots" of this reaction including bound substrate, bound substrate in the presence of NO as an O_2 analogue (Figure 6), a structure representing the monocyclic intermediate, and bound product (Figure 5, right).

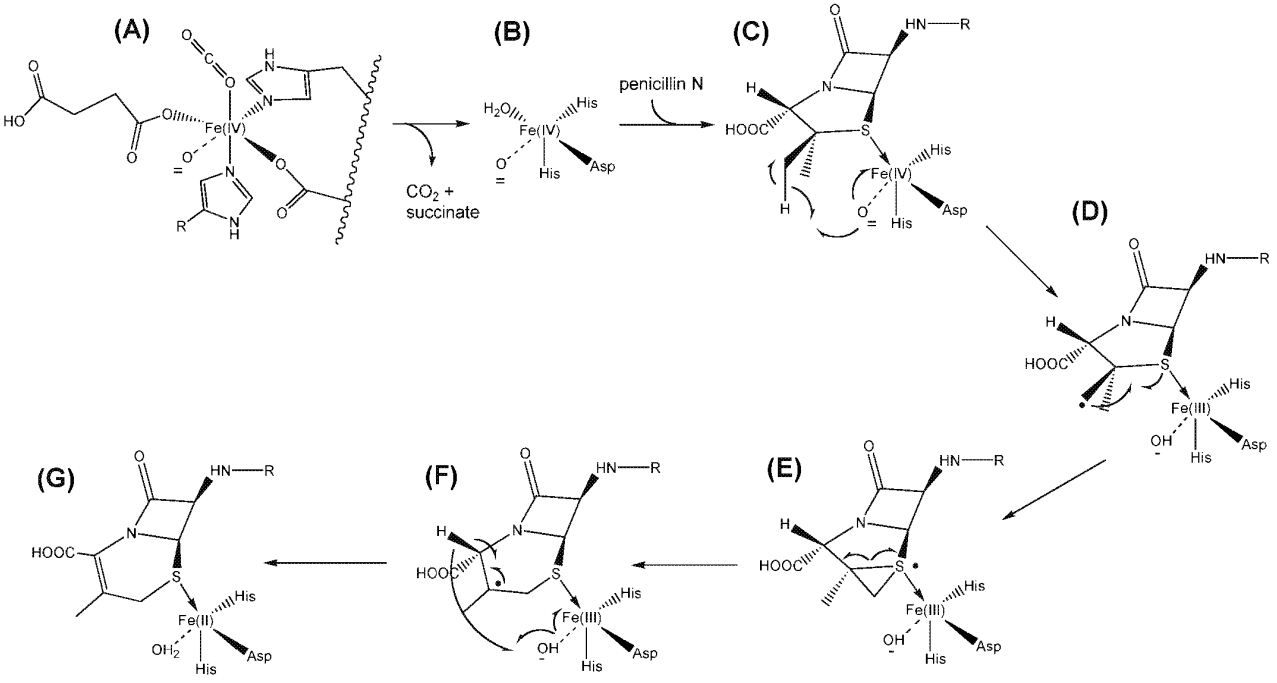
G. ACC Oxidase Mechanism

The ethylene-forming enzyme of plants, ACC oxidase, utilizes ascorbic acid rather than αKG as the two-electron reductant. Several potential mechanisms have been proposed for this extraordinary reaction (e.g., Barlow *et al.*, 1997; Pirrung *et al.*, 1998; Rocklin *et al.*, 1999), but a single hypothesis that most closely parallels the other reactions of this enzyme family is discussed here (Scheme 50). CO_2 plays a critical role in maximizing activity of the enzyme by minimizing the extent of autoinactivation,

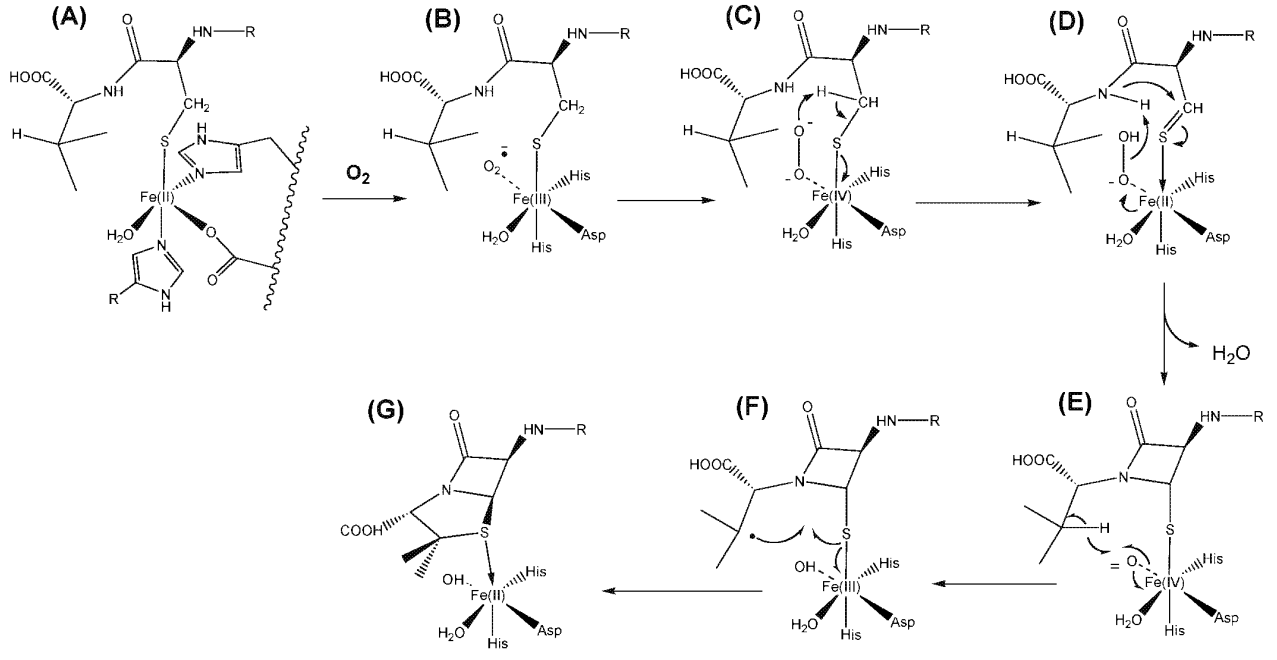


SCHEME 47.

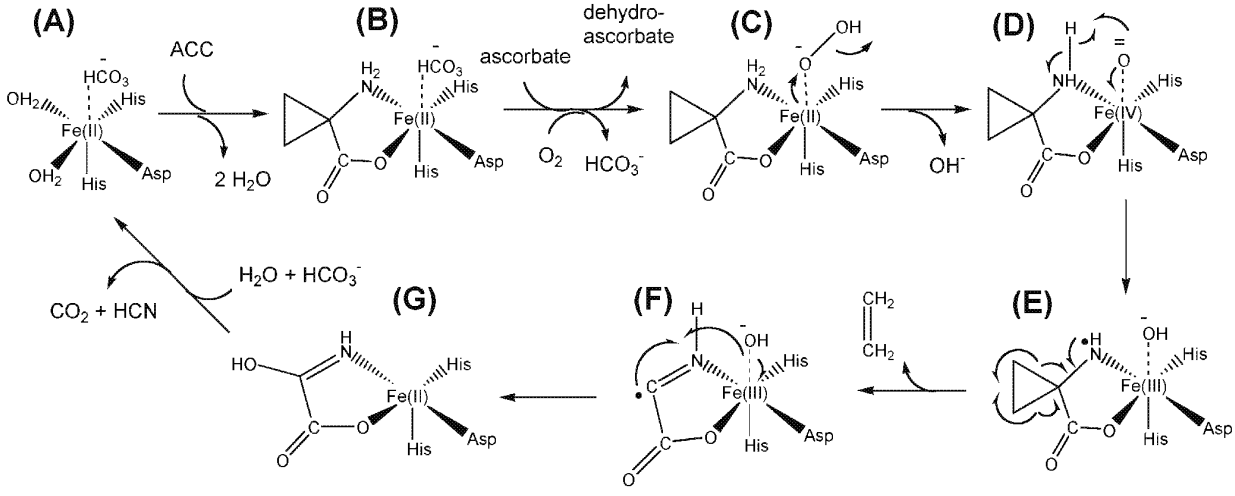
COR-AAD1-00001326



SCHEME 48.



SCHEME 49.



SCHEME 50.

presumably by reacting with a metal-bound solvent molecule to generate the bicarbonate-bound species (**A**) (Zhou *et al.*, 2002). ACC is suggested to chelate the metal ion (**B**) on the basis of ENDOR studies of NO-bound enzyme in the presence of isotopically labeled substrate analogues (Rocklin *et al.*, 1999). Binding of ascorbic acid stimulates loss of the bound bicarbonate and enhances reactivity with oxygen. As oxygen binds, ascorbate donates two electrons by an outer sphere mechanism to generate an Fe(II)-hydroperoxo species (**C**). Heterolytic O-O cleavage (analogous to that just described for IPNS) results in loss of hydroxide and formation of the Fe(IV)-oxo species (**D**). Hydrogen atom abstraction, radical rearrangement, and hydroxyl radical transfer yields a metal-bound intermediate (**G**) that decomposes to the observed products.

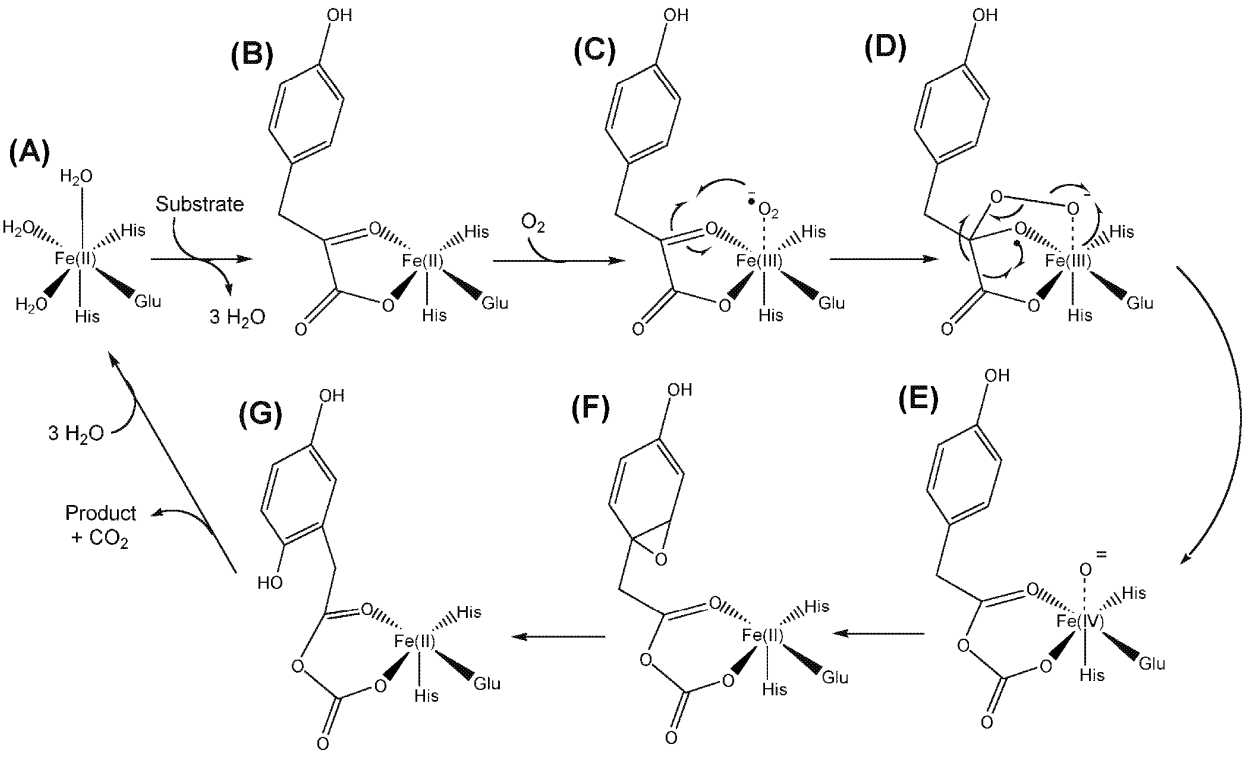
H. HPPD Mechanism

Although unrelated to the αKG-dependent dioxygenases in sequence or structure, HPPD catalyzes a chemically similar mechanism and is likely to utilize analogous intermediates. A reasonable mechanism of HPPD is illustrated in Scheme 51. Resting enzyme contains Fe(II) bound to a 2 His-1 carboxylate motif (**A**). The α-keto acid is proposed to chelate the metal (**B**), thus greatly enhancing oxygen reactivity (Johnson-Winters *et al.*, 2003). Oxygen binds to form an Fe(III)-superoxo or analogous species (**C**) followed by nucleophilic attack to form a bicyclic species (**D**). Still analogous to the αKG-dependent enzymes, O—O bond cleavage provides an Fe(IV)-oxo intermediate and the 4-hydroxyphenylacetate-bicarbonate anhydride (**E**, or its decarboxylation products CO_2 and 4-hydroxyphenylacetate). The activated oxygen inserts into the aromatic ring to form the arene oxide (**F**). This species can rearrange to form oxepinone (Gunsior *et al.*, 2004), providing support for the epoxide, or undergoes an "NIH shift" and rearranges further to yield the known products.

I. Nonproductive Reactions with Oxygen

Further complicating analyses of the Fe(II)/αKG-dependent dioxygenase mechanisms are nonproductive reactions that take place with oxygen. For example, the metal sites of substrate-free forms of these enzymes are oxidized by oxygen, resulting in inactive Fe(III)-containing enzymes. Alternatively, αKG-bound enzymes may exhibit "uncoupled turnover" in which αKG decomposes to CO_2 and succinate (with no corresponding substrate transformation) as the enzyme becomes inactivated. This situation has long been known to occur in prolyl hydroxylase (e.g., Counts *et al.*, 1978; Rao & Adams, 1978; Tuderman *et al.*, 1977), where the uncoupling reaction was shown to result in the metal becoming oxidized (Jong & Kemp, 1984) in an ascorbic acid-reversible manner (Myllylä *et al.*, 1978, 1984). Similar ascorbate-reversible uncoupling reactions occur in several other representatives of this enzyme family, especially if inhibitors or poor substrates are included (e.g., lysyl hydroxylase (Puistola *et al.*, 1980), thymine hydroxylase (Hsu *et al.*, 1981), γ-butyrobetaine hydroxylase (Holme *et al.*, 1984), TfdA (Saari & Hausinger, 1998), and AlkB (Trewick *et al.*, 2002; Welford *et al.*, 2003)). The ability of ascorbic acid to reverse metal oxidation accounts for its widespread use in buffers for assaying these enzymes. Whereas the rates of product formation often decrease rapidly over time for these enzymes, inclusion of ascorbate yields more linear progress curves (e.g., Fukumori & Hausinger, 1993). Significantly, the rates of

COR-AAD1-00001328

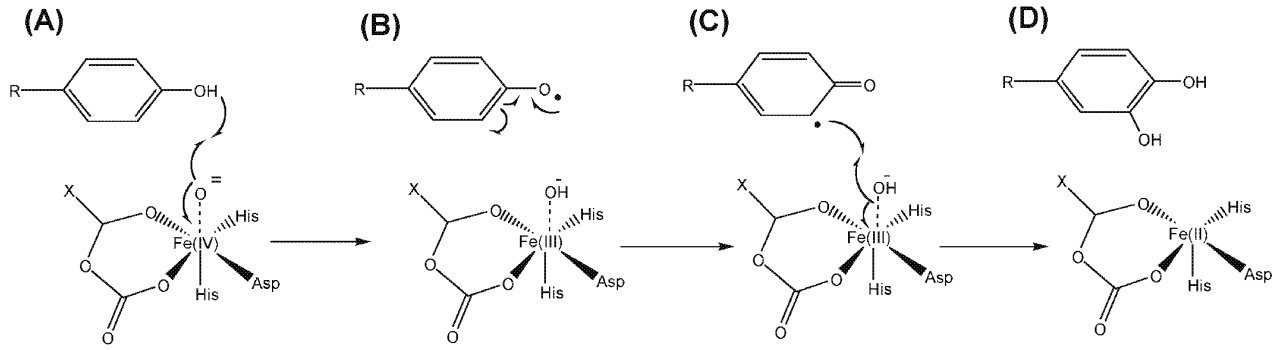


SCHEME 51.

metal center oxidation by direct or uncoupled reaction pathways are generally much slower than the rates of substrate transformation.

In addition to the ascorbate-reversible inactivating reactions associated with uncoupled turnover, uncoupling reactions leading to irreversible protein modifications have been noted in TfdA, AlkB, and TauD. Exposure to oxygen of αKG/Fe(II)TfdA and αKG/Fe(II)AlkB converts the samples with pink chromophores (due to MLCT transitions of the Fe(II)/αKG chelates) to blue species (λ_{max} of 580 nm and 595 nm with $\varepsilon_{580-590}$ of ~1000 M^{-1} cm^{-1} and 960 M^{-1} cm^{-1}, respectively) (Henshaw *et al.*, 2004; Liu *et al.*, 2001). Several lines of evidence were used to assign the new chromophores to ligand-to-metal charge-transfer (LMCT) transitions arising from hydroxy-tryptophan (OH-Trp) coordination to Fe(III). The metal-oxidation state of TfdA was identified as Fe(III) by EPR identification. In addition, specific rR vibrations (corresponding to hydroxyindole-Fe(III) features observed in model compounds) were observed for TfdA. Finally, OH-Trp residues were identified in each protein on the basis of mass spectrometric evidence. The modified residue in TfdA (Trp 113) lies adjacent to the His^1 metal ligand (His 114) (Liu *et al.*, 2001), whereas the target for self hydroxylation of AlkB (Trp 178) is located nine residues before His^2 in the sequence (His 187). Despite these differences, a predicted structure of AlkB positions Trp 178 in the active site environment and only 4.7 Å from the metal (Henshaw *et al.*, 2004). TauD exhibits a distinct form of protein modification in which a tyrosine side chain is transformed into a catechol (Ryle *et al.*, 2003a, 2003b). Stopped-flow UV/visible spectroscopy revealed that αKG/Fe(II)TauD reacts with oxygen to generate an intermediate species (λ_{max} at 408 nm) followed by a final protein product (λ_{max} of 550 nm). The 408 nm species was assigned to a tyrosyl radical species on the basis of EPR characterization, and the 550 nm species was identified as a catecholate-Fe(III) LMCT transition by using rR spectroscopy (Ryle *et al.*, 2003b). Both the tyrosyl radical and the modified tyrosine were shown to derive from Tyr 73, a residue located 6.5 Å from the metal ion. A distinct type of catecholate-Fe(III) species (λ_{max} at ~700 nm) also involving Tyr 73 was identified in hydrogen peroxide-treated enzyme sample (Ryle *et al.*, 2003a). The spectroscopically distinct samples were interconverted depending on whether bicarbonate was bound to the metal, leading to the suggestion that αKG-derived CO_2/bicarbonate remains associated with the oxidized metallocenter. As illustrated in Scheme 52, formation of the tyrosyl radical and catechol species are readily accommodated in mechanisms that

COR-AAD1-00001329



SCHEME 52.

utilize Fe(IV)-oxo intermediates. An analogous reaction is plausible for generating the OH-Trp residues in TfdA and AlkB. Notably, the uncoupling reactions leading to aberrant self-hydroxylation chemistry occur at rates that are much slower than the catalytic reactions involving substrate.

J. Insights from ^{18}O Isotope Labeling Studies

Experimental evidence from $^{18}O_2$ and $H_2{}^{18}O$ isotope studies has been used to identify the sources of the oxygen atoms incorporated into succinate and the hydroxylated substrate of various representatives of this enzyme class and to provide additional insight into the possible intermediacy of an Fe(IV)-oxo species. For all cases examined, O_2 supplies the majority of the oxygen incorporated into succinate (e.g., studies of γ-butyrobetaine hydroxylase (Lindblad *et al.*, 1969), DAOCS (Baldwin *et al.*, 1989), thymine hydroxylase (Holme *et al.*, 1971; Thornburg *et al.*, 1993), HPPD (Lindblad *et al.*, 1970), and α-ketoisocaproate dioxygenase (Sabourin & Bieber, 1982); note that the latter two activities were later shown to be associated with the same enzyme (Crouch *et al.*, 1997)). Oxygen gas also is the predominant source of oxygen incorporated into the products (or, in some cases, the side products) of several αKG-dependent dioxygenases; e.g., collagen-specific prolyl hydroxylase (Kikuchi *et al.*, 1983; Min *et al.*, 2000), HIF-specific prolyl hydroxylase (McNeill *et al.*, 2002b), FIH (Hewitson *et al.*, 2002), CAS (Lloyd *et al.*, 1999b), DAOCS (Baldwin *et al.*, 1993), thymine hydroxylase (Holme *et al.*, 1971), anthocyanidin synthase, flavonol synthase, and flavanone 3β-hydroxylase (Turnbull *et al.*, 2004). Significantly, many of these reactions were shown to be less than 100% efficient, with the remainder of the oxygen presumed to derive from water. Furthermore, only a small percentage of the incorporated oxygen is supplied from O_2 for lysyl hydroxylase (Kikuchi *et al.*, 1983), HPPD (Lindblad *et al.*, 1970), and α-ketocaproate dioxygenase (where this activity is associated with HPPD (Crouch *et al.*, 1997)) (Sabourin & Bieber, 1982). Moreover, the slow self-hydroxylation reactions noted earlier for TfdA and TauD (Liu *et al.*, 2001; Ryle *et al.*, 2003a) were shown by rR spectroscopy of the resulting chromophores to incorporate oxygen primarily from water. (The source of oxygen incorporated into the normal products of these enzymes has not yet been elucidated). Clearly, the latter hydroxylation reactions must utilize an intermediate that exchanges with solvent.

A reasonable hypothesis to explain the divergence in solvent exchange rates of αKG-dependent dioxygenases is that a critical intermediate differs in lifetime and solvent accessibility among these proteins. For example, the postulated succinate-bicarbonate anhydride intermediates may decompose with different kinetics, allowing a range of rates associated with CO_2 loss and solvent access. Alternatively, one must consider that different enzymes utilize distinct activated oxygen species to hydroxylate their substrates. For example, some enzymes may utilize Fe(IV)-oxo (capable of solvent exchange) as their activated oxygen species, others may exploit the Fe(III)-hydroxyl radical species (likely to be less capable of exchanging with solvent), or equilibration may occur between these species in selected enzymes. Potential support for a nonexchangeable intermediate is offered by ANS, an "off line" enzyme according to structural studies (Wilmouth *et al.*, 2002), that shows no exchange with solvent despite a possible requirement for a ferryl flip to orient the reactive oxygen species towards the substrate. It is also possible, however, that αKG reorients in the ANS active site to allow proper positioning of the reactive oxygen, which rapidly reacts with substrate before significant levels of solvent exchange can occur.

K. Biomimetic Studies

Additional evidence for the proposed Fe(IV)-oxo intermediate in the enzymatic reaction has been obtained by studies of high-valent Fe compounds containing nonporphyrin

COR-AAD1-00001330

ligands. Two such Fe(IV)-oxo compounds have been described, including one which was sufficiently stable to allow crystallographic characterization (Lim *et al.*, 2003; Rohde *et al.*, 2003). Other chemical models containing α-keto acids exhibit chemical reactivity similar to the Fe(II)/αKG-dependent hydroxylases, including catalysis of ligand hydroxylation (much like the enzyme self-hydroxylation reactions (Henshaw *et al.*, 2004; Liu *et al.*, 2001; Ryle *et al.*, 2003a)) and are thought to generate Fe(IV)-oxo intermediates (Hegg *et al.*, 1999a; Jensen *et al.*, 2003; Mehn *et al.*, 2003). Activated oxygen intermediates other than the Fe(IV)-oxo species also must still be considered, and Fe(III)-oxo (MacBeth *et al.*, 2000), Fe(III)-peroxide (e.g., Chen & Que, 1999; Hazell *et al.*, 2002; Ho *et al.*, 1999a, 1999b; Jensen *et al.*, 1999; Roelfes *et al.*, 1999, 2000; Simaan *et al.*, 1999; Wada *et al.*, 1999), and other species have been characterized as summarized in an excellent review on this topic (Solomon *et al.*, 2000).

V. CONCLUDING REMARKS

The Fe(II)/αKG-dependent hydroxylases and related enzymes participate in a vast array of biologically important reactions. Ongoing genomic and proteomic studies, followed up by appropriate biochemical analyses, are certain to uncover additional reactions catalyzed by this superfamily of enzymes. Structural studies have advanced very rapidly for this group of proteins, but further structural investigations (buttressed by site-directed mutagenesis analyses) are needed to better define the features important for substrate recognition and choice of reaction pathway, and to identify critical intermediates of these enzymes. An Fe(IV)-oxo species is likely to be the key reactive intermediate in at least some of these enzymes and has been characterized by pre-steady–state techniques, but further spectroscopic and other efforts are needed to elucidate additional or alternative intermediates in the various family members.

ACKNOWLEDGMENTS

I thank Michael Kertesz and Inger Andersson for communication of unpublished information. Studies related to Fe(II)/αKG-dependent hydroxylases carried out in the Hausinger laboratory were supported by National Institutes of Health Grant GM063584.

REFERENCES

Aas, P.A., Otterlei, M., Falnes, P.O., Vagbe, C.B., Skorpen, F., Akbari, M., Sundheim, O., Bjoras, M., Slupphaug, G., Seeberg, E., and Krokan, H.E. 2003. Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature* **421:**859–863.

Abbott, M.T., Schandl, E.K., Lee, R.F., Parker, T.S., and Midgett, R.J. 1967. Cofactor requirements of thymine hydroxylase. *Biochim Biophys Acta* **132:**525–528.

Abbott, M.T. and Udenfriend, S. 1974. α-Ketoglutarate-coupled dioxygenases, In: *Molecular Mechanisms of Oxygen Activation.* pp. 167–214, O. Hayaishi, Ed., Academic Press, New York.

Abbruzzese, A., Hanauske-Able, H.M., Park, M.H., Henke, S., and Folk, J.E. 1991. The active site of deoxyhypusyl hydroxylase: use of catecholpeptides and their component chelator and peptide moieties as molecular probes. *Biochim Biophys Acta* **1077:**159–166.

Abbruzzese, A., Park, M.H., and Folk, J.E. 1986. Deoxyhypusine hydroxylase from rat testis. Partial purification and characterization. *J Biol Chem* **261:**3085–3089.

Adams, D.O. and Yang, S.F. 1979. Ethylene biosynthesis: identification of 1-aminopropane-1-carboxylic acid as an intermediate in the conversion of methionine to ethylene. *Proc Natl Acad Sci USA* **76:**170–174.

Annunem, P., Koivunen, P., and Kivirikko, K.I. 1999. Cloning of the α-subunit of prolyl 4-hydroxylase from *Drosophila* and expression and characterization of the corresponding enzyme tetramer with some unique properties. *J Biol Chem* **274:**6790–6796.

Anzellotti, D. and Ibrahim, R.K. 2000. Novel flavonol 2-oxoglutarate dependent dioxygenase: affinity purification, characterization, and kinetic properties. *Arch Biochem Biophys* **382:**161–172.

Aravind, L. and Koonin, E.V. 2001. The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate-dependent- and iron-dependent dioxygenases. *Genomebiology.com* **2:**7.1–7.8.

Baker, B.J., Dotzlaf, J.E., and Yeh, W.-K. 1991. Deacetoxycephalosporin C hydroxylase of *Streptomyces clavuligerus*. Purification, characterization, bifunctionality, and evolutionary implication. *J Biol Chem* **266:**5087–5093.

Baldwin, J.E. and Abraham, E. 1988. The biosynthesis of penicillins and cephalosporins. *Nat Prod Rep* **5:**129–145.

Baldwin, J.E., Adlington, R.M., Coates, J.B., Crabbe, M.J.C., Crouch, N.P., Keeping, J.W., Knight, G.C., Schofield, C.J., Ting, H.-H., Vallejo, C.A., Thorniley, M., and Abraham, E.P. 1987. Purification and initial characterization of an enzyme with deacetoxycephalosporin C synthetase and hydroxylase activities. *Biochem J* **245:**831–841.

Baldwin, J.E., Adlington, R.M., Crouch, N.P., and Pereira, I.A.C. 1993. Incorporation of ^{18}O-labelled water into oxygenated products produced by the enzyme deacetoxy/deacetylcephalosporin C synthase. *Tetrahedron* **49:**7499–7518.

Baldwin, J.E., Adlington, R.M., Schofield, C.J., Sobey, W.J., and Wood, M.E. 1989. The role of α-ketoglutarate in cephalosporin biosynthesis. *Chem Commun* 1012–1015.

Bankel, L., Lindstedt, G., and Lindstedt, S. 1972. Thymidine 2′-hydroxylation in *Neurospora crassa*. *J Biol Chem* **247:**6128–6134.

Barlow, J.N., Zhang, Z., John, P., Baldwin, J.E., and Schofield, C.J. 1997. Inactivation of 1-aminocyclopropane-1-carboxylate oxidase involves oxidative modification. *Biochemistry* **36:**3563–3569.

Bohm, B.A. 1998. *Introduction to Flavonoids*. Harwood Academic Publishers, Amsterdam, The Netherlands.

Borovok, I., Landman, O., Kreisberg-Zakarin, R., Aharonowitz, Y., and Cohen, G. 1996. Ferrous active site of isopenicillin N synthase: genetic and sequence analysis of the endogenous ligands. *Biochemistry* **35:**1981–1987.

COR-AAD1-00001331

Bradley, F.C., Lindstedt, S., Lipscomb, J.D., Que, L., Jr., Roe, A.L., and Rundgren, M. 1986. 4-Hydroxyphenylpyruvate dioxygenase is an iron-tyrosinate protein. *J Biol Chem* **261:**11693–11696.

Britsch, L. 1990. Purification and characterization of flavone synthase I, a 2-oxoglutarate-dependent desaturase. *Arch Biochem Biophys* **282:**152–160.

Britsch, L., Dedio, J., Saedler, H., and Forkmann, G. 1993. Molecular characterization of flavanone 3ß-hydroxylases. Consensus sequence, comparison with related enzymes and the role of conserved histidine residues. *Eur J Biochem* **217:**745–754.

Britsch, L. and Grisebach, H. 1986. Purification and characterization of (2*S*)-flavanone 3-hydroxylase from *Petunia hybrida*. *Eur J Biochem* **156:**569–577.

Britsch, L., Ruhnau-Brich, B., and Forkmann, G. 1992. Molecular cloning, sequence analysis, and in vitro expression of flavanone 3β-hydroxylase from *Petunia hybrida*. *J Biol Chem* **267:**5380–5387.

Bruick, R.K. and McKnight, S.L. 2001. A conserved family of prolyl-4-hydroxylases that modify HIF. *Science* **294:**1337–1340.

Brunhuber, N.M.W., Mort, J.L., Christoffersen, R.E., and Reich, N.O. 2000. Steady-state kinetic mechanism of recombinant avocado ACC oxidase: initial velocity and inhibitor studies. *Biochemistry* **39:**10730–10738.

Burzlaff, N.I., Rutledge, P.J., Clifton, I.J., Hensgens, C.M.H., Pickford, M., Adlington, R.M., Roach, P.L., and Baldwin, J.E. 1999. The reaction cycle of isopenicillin *N* synthase observed by X-ray diffraction. *Nature* **401:**721–724.

Busby, R.W., Chang, M.D.-T., Busby, R.C., Wimp, J., and Townsend, C.A. 1995. Expression and purification of two isozymes of clavaminate synthase and initial characterization of the iron binding site. General error analysis in polymerase chain reaction amplification. *J Biol Chem* **270:**4262–4269.

Calcutt, M.J. and Schmidt, F.J. 1994. Gene organization in the bleomycin-resistance region of the producer organism. *Streptomycin verticillus*. *Gene* **151:**17–21.

Chahal, A., Khan, M., Pai, S.G., Barbosa, E., and Singh, I. 1998. Restoration of phytanic acid oxidation in Refsum disease fibroblasts from patients with mutations in the phytanoyl-CoA hydroxylase gene. *FEBS Lett* **429:**119–122.

Chance, M.R., Bresnick, A.R., Burley, S.K., Jiang, J.-S., Lima, C.D., Sali, A., Almo, S.C., Bonanno, J.B., Buglino, J.A., Boulton, S., Chen, H., Eswar, N., He, G., Huang, R., Iiyin, V., McMahan, L., Pieper, U., Ray, S., Vidal, M., and Wang, L.K. 2002. Structural genomics: a pipeline for providing structures for the biochemist. *Prot Science* **11:**723–738.

Charrier, B., Coronado, C., Kondorosi, A., and Ratet, P. 1995. Molecular characterization and expression of alfalfa (*Medicago sativa* L.) flavanone 3-hydroxylase and dihydroflavonol-4-reductase encoding genes. *Plant Molec Biol* **29:**773–786.

Chen, B.J., Carroll, P., and Samson, L. 1994. The *Escherichia coli* AlkB protein protects human cells against alkylation-induced toxicity. *J Bacteriol* **176:**6255–6261.

Chen, H., Thomas, M.G., O'Conner, S.E., Hubbard, B.K., Burkart, M.D., and Walsh, C.T. 2001. Aminoacyl-*S*-enzyme intermediates in β-hydroxylations and α,β-desaturations of amino acids in peptide antibiotics. *Biochemistry* **40:**1651–11659.

Chen, K. and Que, L., Jr. 1999. Evidence for the participation of a high-valent iron-oxo species in stereospecific alkane hydroxylation by a non-heme iron catalyst. *Chem Commun* 1375–1376.

Chen, V.J., Orville, A.M., Harpel, M.R., Frolik, C.A., Surerus, K.K., Münck, E., and Lipscomb, J.D. 1989. Spectroscopic studies of isopenicillin N synthase. A mononuclear nonheme Fe^{2+} oxidase with metal coordination sites for small molecules and substrate. *J Biol Chem* **264:**21677–21681.

Chen, W., Zeng, H., and Tan, H. 2000. Cloning, sequencing, and function of *sanF*: a gene involved in nikkomycin biosythesis of *Streptomyces ansochromogenes*. *Curr Microbiol* **41:**312–316.

Chiang, H.H., Hwang, I., and Goodman, H.M. 1995. Isolation of the *Arabidopsis GA4* locus. *Plant Cell* **7:**195–201.

Choroba, O.W., Williams, D.H., and Spencer, J.B. 2000. Biosynthesis of the vancomycin group of antibiotics: involvement of an unusual dioxygenase in the pathway to (*S*)-4-hydroxyphenylglycine. *J Am Chem Soc* **122:**5389–5390.

Clement, P.M., Hanauske-Abel, H.M., Wolff, E.C., Kleinman, H.K., and Park, M.H. 2002. The antifungal drug ciclopirox inhibits deoxyhypusine and proline hydroxylation, endothelial cell growth and angiogenesis in vitro. *Int J Cancer* **100:**491–498.

Clifton, I.J., Doan, L.X., Sleeman, M.C., Topf, M., Suzuki, H., Wilmouth, R.C., and Schofield, C.J. 2003. Crystal structure of carbapenem synthase (CarC). *J Biol Chem* **278:**20843–20850.

Clifton, I.J., Hsueh, L.-C., Baldwin, J.E., Harlos, K., and Schofield, C.J. 2001. Structure of proline 3-hydroxylase. Evolution of the family of 2-oxoglutarate dependent dioxygenases. *Eur J Biochem* **268:**6625–6636.

Cosper, N.J., Stälhandske, C.M.V., Saari, R.E., Hausinger, R.P., and Scott, R.A. 1999. X-ray absorption spectroscopic analysis of Fe(II) and Cu(II) forms of an herbicide-degrading α-ketoglutarate dioxygenase. *J Biol Inorg Chem* **4:**122–129.

Counts, D.F., Cardinale, G.J., and Udenfriend, S. 1978. Prolyl hydroxylase half reaction: peptidyl prolyl-independent decarboxylation of α-ketoglutarate. *Proc Natl Acad Sci USA* **75:**2145–2149.

Croes, K., Casteels, M., Asselberghs, S., Herdewijn, P., Mannaerts, G.P., and Van Veldhoven, P.P. 1997a. Formation of 2-methyl-branched fatty aldehyde during peroxisomal α-oxidation. *FEBS Lett* **412:**643–645.

Croes, K., Van Veldhoven, P.P., Mannaerts, G.P., and Casteels, M. 1997b. Production of formyl-CoA during peroxisomal α-oxidation of 3-methyl-branched fatty acids. *FEBS Lett* **407:**197–200.

Croes, K., Foulon, V., Casteels, M., Van Veldhoven, P.P., and Mannaerts, G.P. 2000. Phytanoyl-CoA hydroxylase: recognition of 3-methyl-branched acyl-CoAs and requirement for GTP or ATP and Mg^{2+} in addition to its known hydroxylation cofactors. *J Lipid Res* **41:**629–636.

Crouch, N.P., Adlington, R.M., Baldwin, J.E., Lee, M.-H., and MacKinnon, C.H. 1997. A mechanistic rationalization for the substrate specificity of recombinant mammalian 4-hydroxyphenylpyruvate dioxygenase (HPPD). *Tetrahedron* **53:**6993–7010.

Cundliffe, E., Bate, N., Butler, A., Fish, S., Gandecha, A., and Merson-Davies, L. 2001. The tylosin-biosynthetic genes of *Streptomyces fradiae*. *Antonie van Leeuwenhoek* **79:**229–234.

Dames, S.A., Martinez-Yamout, M., De Guzman, R.N., Dyson, H.J., and Wright, P.E. 2002. Structural basis for HIF-1α/CBP recognition in the cellular hypoxic response. *Proc Natl Acad Sci USA* **99:**5271–5276.

Dann, C.E., III, Bruick, R.K., and Deisenhofer, J. 2002. Structure of a factor-inhibiting hypoxia-inducible factor 1: an essential asparaginyl hydroxylase involved in the hypoxic response pathway. *Proc Natl Acad Sci USA* **99:**15351–15356.

COR-AAD1-00001332

De Carolis, E., Chan, F., Balsevich, J., and De Luca, V. 1990. Isolation and characterization of a 2-oxoglutarate dependent dioxygenase involved in the second-to-last step in vindoline biosynthesis. *Plant Physiol* **94:**1323–1329.

De Carolis, E. and De Luca, V. 1993. Purification, characterization, and kinetic analysis of a 2-oxoglutarate-dependent dioxygenase involved in vindoline biosynthesis from *Catharanthus roseus*. *J Biol Chem* **268:**5504–5511.

De Carolis, E. and De Luca, V. 1994. 2-Oxoglutarate-dependent dioxygenases and related enzymes: biochemical characterization. *Phytochem* **36:**1093–1107.

Dinglay, S., Trewick, S.C., Lindahl, T., and Sedgwick, B. 2000. Defective processing of methylated single-stranded DNA by *E. coli alkB* mutants. *Genes Develop* **14:**2097–2105.

Dong, J.G., Fernàndez-Maculet, J.C., and Yang, S.F. 1992. Purification and characterization of 1-aminoacyclopropane-1-carboxylate oxidase from apple fruit. *Proc Natl Acad Sci USA* **89:**9789–9793.

Dotzlaf, J.E. and Yeh, W.-K. 1987. Copurification and characterization of deacetoxycephalosporin C synthetase/hydroxylase from *Cephalosporium acremonium*. *J Bacteriol* **169:**1611–1618.

Dotzlaf, J.E. and Yeh, W.-K. 1989. Purification and properties of deacetoxycephalosporin C synthase from recombinant *Escherichia coli* and its comparison to the native enzyme purified from *Streptomyces clavuligerus*. *J Biol Chem* **264:**10219–10227.

Duncan, T., Trewick, S.C., Koivisto, P., Bates, P.A., Lindahl, T., and Sedgwick, B. 2002. Reversal of DNA alkylation damage by two human dioxygenases. *Proc Natl Acad Sci USA* **99:**16660–16665.

Dunning Hotopp, J.C., Auchtung, T.A., Hogan, D.A., and Hausinger, R.P. 2003. Intrinsic tryptophan fluorescence as a probe of metal and α-ketoglutarate binding to TfdA, a mononuclear non-heme iron dioxygenase. *J Inorg Biochem* **93:**66–70.

Dunning Hotopp, J.C. and Hausinger, R.P. 2001. Alternative substrates for 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. *J Molec Catalysis B* **15:**155–162.

Dunning Hotopp, J.C. and Hausinger, R.P. 2002. Probing the 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase substrate binding site by site-directed mutagenesis and mechanism-based inactivation. *Biochemistry* **41:**9787–9794.

Eichhorn, E., van der Ploeg, J.R., Kertesz, M.A., and Leisinger, T. 1997. Characterization of α-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. *J Biol Chem* **272:**23031–23036.

Elkins, J.M., Hewitson, K.S., McNeill, L.A., Seibel, J.F., Schlemminger, I., Pugh, C.W., Ratcliffe, P.J., and Schofield, C.J. 2003. Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1α. *J Biol Chem* **278:**1802–1806.

Elkins, J.M., Ryle, M.J., Clifton, I.J., Dunning Hotopp, J.C., Lloyd, J.S., Burzlaff, N.I., Baldwin, J.E., Hausinger, R.P., and Roach, P.L. 2002. X-ray crystal structure of *Escherichia coli* taurine/α-ketoglutarate dioxygenase complexed to ferrous iron and substrates. *Biochemistry* **41:**5185–5192.

Epstein, A.C.R., Gleadle, J.M., McNeill, L.A., Hewitson, K.S., O'Rourke, J., Mole, D.R., Mukherji, M., Metzen, E., Wilson, M.I., Dhanda, A., Tian, Y.-M., Masson, N., Hamilton, D.L., Jaakkola, P., Barstead, R., Hodgkin, J., Maxwell, P.H., Pugh, C.W., Schofield, C.J., and Ratcliffe, P.J. 2001. *C. elegans* EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107:**43–54.

Ericksson, M., Myllyharju, J., Tu, H., Hellman, M., and Kivirikko, K.I. 1999. Evidence for 4-hydroxyproline in viral proteins. Characterization of a viral prolyl 4-hydroxylase and its peptide substrates. *J Biol Chem* **274:**22131–22134.

Falnes, P.O., Johansen, R.F., and Seeberg, E. 2002. AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature* **419:**178–182.

Foulon, V., Antonenkov, V.D., Croes, K., Waelkens, E., Mannaerts, G.P., Van Veldhoven, P.P., and Casteels, M. 1999. Purification, molecular cloning, and expression of 2-hydroxyphytanoyl-CoA lyase, a peroxisomal thiamin pyrophosphate-dependent enzyme that catalyzes the carbon-carbon cleavage during α-oxidation of 3-methyl-branched fatty acids. *Proc Natl Acad Sci USA* **96:**10039–10044.

Freedman, S.J., Sun, Z.-Y.J., Poy, F., Kung, A.L., Livingston, D.M., Wagner, G., and Eck, M.J. 2002. Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1α. *Proc Natl Acad Sci USA* **99:**5367–5372.

Frey, M., Huber, K., Park, W.J., Sicker, D., Lindberg, P., Meeley, R.B., Simmons, C.R., Yalpani, N., and Gierl, A. 2003. A 2-oxoglutarate-dependent dioxygenase is integrated in the DIMBOA-biosynthesis. *Phytochem* **62:**371–376.

Fukuda, H., Fujii, T., and Ogawa, T. 1986. Preparation of a cell-free ethylene-forming system from *Penicillium digitatum*. *Agric Biol Chem* **50:**977–981.

Fukuda, H., Kitajima, H., Fujii, T., Tazaki, M., and Ogawa, T. 1989. Purification and some properties of a novel ethylene-forming enzyme produced by *Penicillium digitatum*. *FEMS Microbiol Lett* **59:**1–6.

Fukuda, H., Ogawa, T., Ishihara, K., Fujii, T., Nagahama, K., Omata, T., Inoue, Y., Tanase, S., and Morino, Y. 1992a. Molecular cloning in *Escherichia coli*, expression, and nucleotide sequence of the gene for the ethylene-forming enzyme of *Pseudomonas syringae* pv. *phaseolicola* PK2. *Biochem Biophys Res Commun* **188:**826–832.

Fukuda, H., Ogawa, T., Tazaki, M., Nagahama, K., Fujii, T., Tanase, S., and Morino, Y. 1992b. Two reactions are simultaneously catalyzed by a single enzyme: the arginine-dependent simultaneous formation of two products, ethylene and succinate, from 2-oxoglutarate by an enzyme from *Pseudomonas syringae*. *Biochem Biophys Res Commun* **188:**483–489.

Fukumori, F. and Hausinger, R.P. 1993a. Purification and characterization of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. *J Biol Chem* **268:**24311–24317.

Fukumori, F. and Hausinger, R.P. 1993b. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetate monooxygenase" is an α-ketoglutarate-dependent dioxygenase. *J Bacteriol* **175:**2083–2086.

Galland, S., Le Borgne, F., Bouchard, F., Georges, B., Clouet, P., Grand-Jean, F., and Demarquoy, J. 1999. Molecular cloning and characterization of the cDNA encoding rat liver γ-butyrobetaine hydroxylase. *Biochim Biophys Acta* **1441:**85–92.

Garcia Costas, A.M., White, A.K., and Metcalf, W.W. 2001. Purification and characterization of a novel phosphorus-oxidizing enzyme from *Pseudomonas stutzeri* WM88. *J Biol Chem* **276:**17429–17436.

Gibbons, H.S., Lin, S., Cotter, R.J., and Raetz, C.R.H. 2000. Oxygen requirement for the biosynthesis of the *S*-2-hydroxymyristate moiety in *Salmonella typhimurium* lipid A. Function of LpxO, a new Fe^{2+}/α-ketoglutarate-dependent dioxygenase homologue. *J Biol Chem* **275:**32940–32949.

Goto, M. and Hyodo, H. 1987. Ethylene production by cell-free extract of the kudzu strain of *Pseudomonas syringae* pv. *phaseolicola*. *Plant Cell Physiol* **28:**405–414.

COR-AAD1-00001333

Gronke, R.S., VanDusen, W.J., Garsky, V.M., Jacobs, J.W., Sardana, M.K., Stern, A.M., and Friedman, P.A. 1989. Aspartyl β-hydroxylase: *in vitro* hydroxylation of a synthetic peptide based on the structure of the first growth factor-like domain of human factor IX. *Proc Natl Acad Sci USA* **86:**3609–3613.

Gunsior, M., Ravel, J., Challis, G.L., and Townsend, C.A. 2004. Engineering *p*-hydroxyphenylpyruvate dioxygenase to a *p*-hydroxymandelate synthase and evidence for the proposed benzene oxide intermediate in homogentisate formation. *Biochemistry* **43:**663–674.

Haefelè, C., Bonfils, C., and Sauvaire, Y. 1997. Characterization of a dioxygenase from *Trigonella foenum-graecum* involved in 4-hydroxyisoleucine biosythesis. *Phytochem* **44:**563–566.

Hamilton, A.J., Bouzayen, M., and Grierson, D. 1991. Identification of a tomato gene for the ethylene-forming enzyme by expression in yeast. *Proc Natl Acad Sci USA* **88:**7434–7437.

Han, H. and Pascal, R.A., Jr. 1990. α-Ketoisocaproate dioxygenase: the stereochemical course of the hydroxylation reaction. *J Org Chem* **55:**5173–5176.

Han, S., Eltis, L.D., Timmis, K.N., Muchmore, S.W., and Bolin, J.T. 1995. Crystal structure of the biphenyl-cleaving extradiol dioxygenase from a PCB-degrading pseudomonad. *Science* **270:**976–980.

Hanauske-Abel, H.M. and Günzler, V. 1982. A stereochemical concept for the catalytic mechanism of prolylhydroxylase. Applicability to classification and design of inhibitors. *J Theor Biol* **94:**421–455.

Hashimoto, T., Matsuda, J., and Yamada, Y. 1993. Two-step epoxidation of hyoscyamine to scopolamine is catalyzed by bifunctional hyoscyamine 6ß-hydroxylase. *FEBS Lett* **329:**35–39.

Hashimoto, T. and Yamada, Y. 1987. Purification and characterization of hyoscyamine 6β-hydroxylase from root cultures of *Hyoscyamus niger* L. Hydroxylase and epoxidase activities of enzyme preparations. *Eur J Biochem* **164:**277–285.

Hausinger, R.P., Fukumori, F., Hogan, D.A., Sassanella, T.M., Kamagata, Y., Takami, H., and Saari, R.E. 1997. Biochemistry of 2,4-D degradation: evolutionary implications, In *Microbial Diversity and Genetics of Biodegradation*, pp. 35–51, K. Horikoshi, M. Fukuda, and T. Kudo, Eds., Japan Scientific Press, Tokyo.

Hautala, T., Byers, M.G., Eddy, R.L., Shows, T.B., Kivirikko, K.I., and Myllylä, R. 1992. Cloning of human lysyl hydroxylase: complete cDNA-derived amino acid sequence and assignment of the gene (*PLOD*) to chromosome 1p36.1–p36.2. *Genomics* **13:**62–69.

Hazell, A., McKensie, C.J., Nielsen, L.P., Schindler, S., and Weitzer, M. 2002. Mononuclear non-heme iron(III) peroxide complexes: syntheses, characterization, mass spectrometric and kinetic studies. *J Chem Soc Dalton Trans* 310–317.

Hedden, P. 1997. The oxidases of gibberellin biosynthesis: their function and mechanism. *Physiol Plant* **101:**709–719.

Hedden, P. 1999. Recent advances in gibberellin biosynthesis. *J Exper Bot* **50:**553–563.

Hedden, P. and Kamiya, Y. 1997. Gibberellin biosynthesis: enzymes, genes and their regulation. *Annu Rev Plant Physiol Plant Molec Biol* **48:**431–460.

Hedden, P. and Phillips, A.L. 2000. Gibberellin metabolism: new insights revealed by the genes. *Trend Plant Sci* **5:**523–528.

Hegg, E.L., Ho, R.Y.N., and Que, L., Jr. 1999a. Oxygen activation and arene hydroxylation by functional mimics of α-keto acid-dependent iron(II) dioxygenases. *J Am Chem Soc* **121:**1972–1973.

Hegg, E.L., Whiting, A.K., Saari, R.E., McCracken, J., Hausinger, R.P., and Que, L., Jr. 1999b. Herbicide-degrading α-keto acid-dependent enzyme TfdA: metal coordination environment and mechanistic insights. *Biochemistry* **38:**16714–16726.

Helaakoski, T., Vuori, K., Myllylä, R., Kivirikko, K.I., and Pihlajaniemi, T. 1989. Molecular cloning of the α-subunit of human prolyl 4-hydroxylase: the complete cDNA-derived amino acid sequence and evidence for alternative splicing of RNA transcripts. *Proc Natl Acad Sci USA* **86:**4392–4396.

Henshaw, T.F., Feig, M., and Hausinger, R.P. 2004. Aberrant activity of the DNA repair enzyme AlkB. *J Inorg Biochem* in press:

Hewitson, K.S., McNeill, L.A., Elkins, J.M., and Schofield, C.J. 2003. The role of iron and 2-oxoglutarate oxygenases in signalling. *Biochem Soc Trans* **31:**510–515.

Hewitson, K.S., McNeill, L.A., Riordan, M.V., Tian, Y.-M., Bullock, A.N., Welford, R.W., Elkins, J.M., Oldham, N.J., Bhattacharya, S., Gleadle, J.M., Ratcliffe, P.J., Pugh, C.W., and Schofield, C.J. 2002. Hypoxia-inducible factor (HIF) asparagine hydroxylase is identical to factor inhibiting HIF (FIH) and is related to the cupin structural familly. *J Biol Chem* **277:**26351–26355.

Hieta, R. and Myllyharju, J. 2002. Cloning and characterization of a low molecular weight prolyl 4-hydroxylase from *Arabidopsis thaliana*. Effective hydroxylation of proline-rich, collagen-like, and hypoxia-inducible transcription factor α-like peptides. *J Biol Chem* **277:**23965–23971.

Ho, R.Y.H., Roelfe, G., Hermant, R., Hage, R., Feringa, B.L., and Que, L., Jr. 1999a. Resonance Raman evidence for the interconversion between an $[Fe^{III}\text{-}n^{I}\text{-}O_2]^{+}$ species and mechanistic implications thereof. *Chem Commun* 2161–2162.

Ho, R.Y.N., Mehn, M.P., Hegg, E.L., Liu, A., Ryle, M.A., Hausinger, R.P., and Que, L., Jr. 2001. Resonance Raman studies of the iron(II)-α-keto acid chromophore in model and enzyme complexes. *J Am Chem Soc* **123:**5022–5029.

Ho, R.Y.N., Roelfes, G., Feringa, B.L., and Que, L., Jr. 1999b. Raman evidence for a weakened O—O bond in mononuclear low-spin iron(III)-hydroperoxides. *J Am Chem Soc* **121:**264–265.

Hogan, D.A., Auchtung, T.A., and Hausinger, R.P. 1999. Cloning and characterization of a sulfonate/α-ketoglutarate dioxygenase from *Saccharomyces cerevisiae*. *J Bacteriol* **181:**5876–5879.

Hogan, D.A., Smith, S.R., Saari, E.A., McCracken, J., and Hausinger, R.P. 2000. Site-directed mutagenesis of 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase. Identification of residues involved in metallocenter formation and substrate binding. *J Biol Chem* **275:**12400–12409.

Holme, E. 1975. A kinetic study of thymine 7-hydroxylase from *Neurospora crassa*. *Biochemistry* **14:**4999–5003.

Holme, E., Lindstedt, G., Lindstedt, S., and Tofft, M. 1971. ^{18}O studies of the 2-ketoglutarate-dependent sequential oxygenation of thymine to 5-carboxyuracil. *J Biol Chem* **246:**3314–3319.

Holme, E., Lindstedt, S., and Nordin, I. 1984. Uncoupling and isotope effects in γ-butyrobetaine hydroxylation. *Biosci Rep* **4:**433–440.

Holton, T.A., Brugliera, F., and Tanaka, Y. 1993. Cloning and expression of flavonol synthase from *Petunia hybrida*. *Plant J* **4:**1003–1010.

Hon, W.-C., Wilson, M.I., Harlos, K., Claridge, T.D.W., Schofield, C.J., Pugh, C.W., Maxwell, P.H., Ratcliffe, P.J., Stuart, D.I., and Jones, E.Y. 2002. Structural basis for the recognition of hydroxyproline in HIF-1α by pVHL. *Nature* **417:**975–978.

Hsu, C.-A., Saewert, M.D., Polsinelli, L.F., Jr., and Abbott, M.T. 1981. Uracil's uncoupling of the decarboxylation of α-ketoglutarate in the

COR-AAD1-00001334

thymine 7-hydroxylase reaction of *Neurospora crassa*. *J Biol Chem* **256:**6098–6101.

Hubbard, B.K., Thomas, M.G., and Walsh, C.T. 2000. Biosynthesis of L-*p*-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics. *Chem Biol* **7:**931–942.

Hulse, J.D., Ellis, S.R., and Henderson, L.M. 1978. Carnitine biosynthesis. β-Hydroxylation of trimethyllysine by an α-ketoglutarate-dependent mitochondrial dioxygenase. *J Biol Chem* **253:**1654–1659.

Hutton, J.J., Jr., Trappel, A.L., and Udenfriend, S. 1966. Requirement for α-ketoglutarate, ferrous ion and ascorbate for collagen proline hydroxylase. *Biochem Biophys Res Commun* **24:**179–184.

Ince, N., de la Monte, S.M., and Wands, J.R. 2000. Overexpression of human aspartyl (asparaginyl) β-hydroxylase is associated with malignant transformation. *Cancer Res* **60:**1261–1266.

Itoh, H., Ueguchi-Tanaka, M., Sentoku, N., Kitano, H., Matsuoka, M., and Kobayashi, M. 2001. Cloning and functional analysis of two gibberellin 3β-hydroxylase genes that are differently expressed during growth of rice. *Proc Natl Acad Sci USA* **98:**8909–8914.

Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J.M., Lane, W.S., and Kaelin, W.G., Jr. 2001. HIFα targeted for VHL-mediated destruction by proline hydroxylation: implications for O_2 sensing. *Science* **292:**464–467.

Jaakkola, P., Mole, D.R., Tian, Y.-M., Wilson, M.I., Gielbert, J., Gaskell, S.J., von Kriegsheim, A., Hebestreit, H.F., Mukherji, M., Schofield, C.J., Maxwell, P.H., Pugh, C.W., and Ratcliffe, P.J. 2001. Targeting of HIF-α to the von Hippel-Lindau ubiquitylation complex by O_2-regulated prolyl hydroxylation. *Science* **292:**468–472.

Janc, J.W., Egan, L.A., and Townsend, C.A. 1995. Purification and characterization of clavaminate synthase from *Streptomyces antibioticus*. A multifunctional enzyme of clavam biosynthesis. *J Biol Chem* **270:**5399–5404.

Jansen, G.A., Mihalik, S.J., Watkins, P.A., Moser, H.W., Jakobs, C., Denis, S., and Wanders, R.J.A. 1996. Phytanoyl-CoA hydroxylase is present in human liver, located in peroxisomes, and deficient in Zellweger syndrome: direct, unequivocal evidence for the new, revised pathway of phytanic acid α-oxidation in humans. *Biochem Biophys Res Commun* **229:**205–210.

Jansen, G.A., Ofman, R., Denis, S., Ferdinandusse, S., Hogenhout, E.M., Jakobs, C., and Wanders, R.J.A. 1999. Phytanoyl-CoA hydroxylase from rat liver: protein purification and cDNA cloning with implications for the subcellular localization of phytanic acid α-oxidation. *J Lipid Res* **40:**2244–2254.

Jansen, G.A., Ofman, R., Ferdinandusse, S., Ijlst, L., Muijsers, A.O., Skjeldal, O.H., Stokke, O., Jakobs, C., Besley, G.T., Wraith, J.E., and Wanders, R.J.A. 1997. Refsum disease is caused by mutations in the phytanoyl-CoA hydroxylase gene. *Nature Genet* **17:**190–193.

Jensen, K.B., McKensie, C.J., Nielsen, L.P., Pedersen, J.Z., and Svendsen, H.M. 1999. Deprotonation of low-spin mononuclear iron(III)-hydroperoxide complexes give transient blue species assigned to high-spin iron(III)-peroxide complexes. *Chem Commun* 1313–1314.

Jensen, M.P., Lange, S.J., Mehn, M.P., Que, E.L., and Que, L., Jr. 2003. Biomimetic aryl hydroxylation derived from alkyl hydroperoxide at a nonheme iron center. Evidence for an $Fe^{IV} = O$ oxidant. *J Am Chem Soc* **125:**2113–2128.

Jia, S., McGinnis, K., VanDusen, W.J., Burke, C.J., Kuo, A., Griffin, P.R., Sardana, M.K., Elliston, K.O., Stern, A.M., and Friedman, P.A. 1994. A fully active catalytic domain of bovine aspartyl (asparginyl) ß-hydroxylase expressed in *Escherichia coli*: characterization and evidence for the identification of an active-site region in vertebrate α-ketoglutarate-dependent dioxygenases. *Proc Natl Acad Sci USA* **91:**7227–7231.

Jia, S., VanDusen, W.J., Diehl, R.E., Kohl, N.E., Dixon, R.A.F., Elliston, K.O., Stern, A.M., and Friedman, P.A. 1992. cDNA cloning and expression of bovine aspartyl (asparginyl) β-hydroxylase. *J Biol Chem* **267:**14322–14327.

Jiang, F., Peisach, J., Ming, L.-J., Que, L., Jr., and Chen, V.J. 1991. Electron spin echo envelope modulation studies of the Cu(I)-substituted derivative of isopenicillin N synthase: a structural and spectroscopic model. *Biochemistry* **30:**11437–11445.

Johnson-Winters, K., Purpero, V.M., Kavana, M., Nelson, T., and Moran, G.R. 2003. (4-Hydroxyphenyl)pyruvate dioxygenase from *Streptomyces avermitilis*: the basis for ordered substrate addition. *Biochemistry* **42:**2072–2080.

Jong, E. and Kemp, A. 1984. Stoichiometry and kinetics of the prolyl 4-hydroxylase partial reaction. *Biochim Biophys Acta* **787:** 105–111.

Kahnert, A. and Kertesz, M.A. 2000. Characterization of a sulfur-regulated oxygenative alkylsulfatase from *Pseudomonas putida* S-313. *J Biol Chem* **275:**31661–31667.

Kaska, D.D., Myllylä, R., Günzler, V., Gibor, A., and Kivirikko, K.I. 1988. Prolyl 4-hydroxylase from *Volvox carteri*. A low-M_r enzyme antigenically related to the α subunit of the invertebrate enzyme. *Biochem J* **256:**257–263.

Kataoka, H. and Sekiguchi, M. 1985. Molecular cloning and characterization of the *alkB* gene of *Escherichia coli*. *Molec Gen Genet* **198:**263–269.

Kataoka, H., Yamamoto, Y., and Sekiguchi, M. 1983. A new gene (*alkB*) of *Escherichia coli* that controls sensitivity to methyl methane sulfonate. *J Bacteriol* **153:**1301–1307.

Kawauchi, H., Tohno, M., Tsuchiya, Y., Hayashida, M., Adachi, Y., Mukai, T., Hayashi, I., Kimura, S., and Kondo, S. 1983. Studies on the structure of a novel peptide antibiotic, K-582. *Int J Pept Prot Res* **21:**546–554.

Kertesz, M.A. 1999. Riding the sulfur cycle-metabolism of sulfonates and sulfate esters in Gram-negative bacteria. *FEMS Microbiol Rev* **24:**135–175.

Khaleeli, N., Busby, R.W., and Townsend, C.A. 2000. Site-directed mutagenesis and biochemical analysis of the endogenous ligands in the ferrous active site of clavaminate synthase. The His-3 variant of the 2-His-1-carboxylase model. *Biochemistry* **39:**8666–8673.

Kieliszewski, M.J. and Lamport, D.T. 1994. Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. *Plant J* **5:**157–172.

Kikuchi, Y., Suzuki, Y., and Tamiya, N. 1983. The source of oxygen in the reaction catalyzed by collagen lysyl hydroxylase. *Biochem J* **213:**507–512.

Kita, A., Kita, S.-I., Fujisawa, I., Inaka, K., Ishida, T., Horiike, K., Nozaki, M., and Miki, K. 1999. An archetypical extradiol-cleaving catecholic dioxygenase: the crystal structure of catechol 2,3-dioxygenase (metapyrocatechase) from *Pseudomonas putida* mt-2. *Structure* **7:**25–34.

Kivirikko, K.I. and Pihlajaniemi, T. 1998. Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylase. *Adv Enzymol Rel Areas Mol Biol* **72:**325–398.

Koivisto, P., Duncan, T., Lindahl, T., and Sedgwick, B. 2003. Minimal methylated substrate and extended substrate range of *Escherichia*

COR-AAD1-00001335

coli AlkB protein, a 1-methyladenine-DNA dioxygenase. *J Biol Chem* **278:**44348–44354.

Kondo, A., Blanchard, J.S., and Englard, S. 1981. Purification and properties of calf liver γ-butyrobetaine hydroxylase. *Arch Biochem Biophys* **212:**338–346.

Kondo, H., Nakabeppu, Y., Kataoka, H., Kuhara, S., Kawabata, S., and Sekiguchi, M. 1986. Structure and expression of the *alkB* gene of *Escherichia coli* related to the repair of alkylated DNA. *J Biol Chem* **261:**15772–15777.

Korioth, F., Gieffers, C., and Frey, J. 1994. Cloning and characterization of the human gene encoding aspartyl β-hydroxylase. *Gene* **150:**395–399.

Kovacevic, S., Weigel, B.J., Tobin, M.B., Ingolia, T.D., and Miller, J.R. 1989. Cloning, characterization, and expression in *Escherichia coli* of the *Streptomyces clavuligerus* gene encoding deacetoxycephalosporin C synthetase. *J Bacteriol* **171:**754–760.

Kreisberg-Zakarin, R., Borovok, I., Yanko, M., Aharonowitz, Y., and Cohen, G. 1999. Recent advances in the structure and function of isopenicillin N synthase. *Antonie van Leeuwenhoek* **75:** 33–39.

Kreisberg-Zakarin, R., Borovok, I., Yanko, M., Frolow, F., Aharonowitz, Y., and Cohen, G. 2000. Structure-function studies of non-heme iron active site of isopenicillin N synthase: some implications for catalysis. *Biophys Chem* **86:**109–118.

Kurowski, M.A., Bhagwat, A.S., Papaj, G., and Bujnicki, J.M. 2003. Phylogenomic identification of five new human homologues of the DNA repair enzyme AlkB. *BMC Genomics* **4:**48.

Lamberg, A., Pihlajaniemi, T., and Kivirikko, K.I. 1995. Site-directed mutagenesis of the α subunit of human prolyl 4-hydroxylase. Identification of three histidine residues critical for catalytic activity. *J Biol Chem* **270:**9926–9931.

Landman, O., Borovok, I., Aharonowitz, Y., and Cohen, G. 1997. The glutamine ligand in the ferrous iron active site of isopenicillin N synthase of *Streptomyces jumonjinensis* is not essential for catalysis. *FEBS Lett* **405:**172–174.

Lando, D., Peet, D.J., Whelan, D.A., Gorman, J.J., and Whitelaw, M.L. 2002. Asparagine hydroxylation of the HIF transactivation domain: a hypoxic switch. *Science* **295:**858–861.

Lange, T. 1994. Purification and partial amino-acid sequence of gibberellin 20-oxidase from *Cucurbita maxima* L. endosperm. *Planta* **195:**108–115.

Lange, T. 1997. Cloning gibberellin dioxygenase genes from pumpkin endosperm by heterologous expression of enzyme activities in *Escherichia coli*. *Proc Natl Acad Sci USA* **94:**6553–6558.

Lange, T. and Graebe, J. E. 1989. The partial purification and characterization of a gibberellin C-20 hydroxylase from immature *Pisum sativum* L. seeds. *Planta* **179:**211–221.

Lange, T., Hedden, P., and Graebe, J.E. 1994a. Expression cloning of a gibberellin 20-oxidase, a multifunctional enzyme involved in gibberellin synthesis. *Proc Natl Acad Sci USA* **91:**8552–8556.

Lange, T., Kegler, C., Hedden, P., Phillips, A.L., and Graebe, J.E. 1997a. Molecular characterization of gibberellin 20-oxidases. Structure-function studies on recombinant enzymes and chimaeric proteins. *Physiol Plant* **100:**543–549.

Lange, T., Robatzek, S., and Frisse, A. 1997b. Cloning and expression of a gibberellin 2β,3β-hydroxylase cDNA from pumpkin endosperm. *The Plant Cell* **9:**1459–1467.

Lange, T., Schweimer, A., Ward, D.A., Hedden, P., and Graebe, J.E. 1994b. Separation and characterization of three 2-oxoglutarate-dependent dioxygenases from *Cucurbita maxima* L. endosperm involved in gibberellin biosynthesis. *Planta* **195:**98–107.

Lawrence, C.C., Sobey, W.J., Field, R.A., Baldwin, J.E., and Schofield, C.J. 1996. Purification and initial characterization of proline 4-hydroxylase from *Streptomyces griseoviridus* P8648: a 2-oxoacid, ferrous-dependent dioxygenase involved in etamycin biosynthesis. *Biochem J* **313:**185–191.

Lay, V.J., Prescott, A.G., Thomas, P.G., and John, P. 1996. Heterologous expression and site-directed mutagenesis of the 1-aminocyclopropane-1-carboxylate oxidase from kiwi fruit. *Eur J Biochem* **242:**228–234.

Le Marchand, L. 2002. Cancer prevention effects of flavonoids–a review. *Biomed Pharmacother* **56:**296–301.

Lee, C., Kim, S.J., Jeong, D.G., Lee, S.M., and Ryu, S.E. 2003. Structure of human FIH-1 reveals a unique active site pocket and interaction sites for HIF-1 and von Hipple-Lindau. *J Biol Chem* **278:**7558–7563.

Lee, H.J., Lloyde, M.D., Clifton, I.J., Baldwin, J.E., and Schofield, C.J. 2001a. Kinetic and crystallographic studies on deacetoxycephalosporin C synthase (DAOCS). *J Mol Biol* **308:**937–948.

Lee, H.-J., Lloyd, M.D., Clifton, I.J., Harlos, K., Dubus, A., Baldwin, J.E., Frere, J.-M., and Schofield, C.J. 2001b. Alteration of the co-substrate selectivity of deacetoxycephalosporin C synthase. The role of arginine 258. *J Biol Chem* **276:**18290–18295.

Lee, H.J., Schofield, C.J., and Lloyde, M.D. 2002. Active site mutations of recombinant deacetoxycephalosporin C synthase. *Biochem Biophys Res Commun* **292:**66–70.

Lester, D.R., Ross, J.J., Davies, P.J., and Reid, J.B. 1997. Mendel's stem length gene (*Le*) encodes a gibberellin 3β-hydroxylase. *The Plant Cell* **9:**1435–1443.

Lester, D.R., Ross, J.J., Smith, J.J., Elliott, R.C., and Reid, J.B. 1999. Gibberellin 2-oxidation and the *SLN* gene of *Pisum sativum*. *Plant J* **19:**65–73.

Lim, M.H., Rohde, J.-U., Stubna, A., Bukowski, M.R., Costas, M., Ho, R.Y.N., Münck, E., Nam, W., and Que, L., Jr. 2003. An $Fe^{IV} = O$ complex of a tetradentate tripodal nonheme ligand. *Proc Natl Acad Sci USA* **100:**3665–3670.

Lindblad, B., Lindstedt, G., and Lindstedt, S. 1970. The mechanism of enzymic formation of homogentisate from *p*-hydroxyphenylpyruvate. *J Am Chem Soc* **92:**7446–7449.

Lindblad, B., Lindstedt, G., Lindstedt, S., and Rundgren, M. 1977. Purification and some properties of human 4-hydroxyphenylpyruvate dioxygenase. *J Biol Chem* **252:**5073–5084.

Lindblad, B.G.L., Tofft, M., and Lindstedt, S. 1969. The mechanism of α-ketoglutarate oxidation in coupled enzymatic oxygenations. *J Am Chem Soc* **91:**4604–4606.

Lindstedt, G., Lindstedt, S., and Nordin, I. 1977. Purification and properties of γ-butyrobetaine hydroxylase from *Pseudomonas* sp. AK1. *Biochemistry* **16:**2181–2188.

Lindstedt, S. and Rundgren, M. 1982. Blue color, metal content, and substrate binding in 4-hydroxyphenylpyruvate dioxygenase from *Pseudomonas* sp. strain P. J. 874. *J Biol Chem* **257:**11922–11931.

Lipscomb, S.J., Lee, H.-J., Mukherji, M., Baldwin, J.E., Schofield, C.J., and Lloyd, M.D. 2002. The role of arginine residues in substrate binding and catalysis by deacetoxycephalosporin C synthase. *Eur J Biochem* **269:**2735–2739.

Liu, A., Ho, R.Y.N., Que, L., Jr., Ryle, M.J., Phinney, B.S., and Hausinger, R.P. 2001. Alternative reactivity of an α-ketoglutarate-dependent iron(II) oxygenase: enzyme self-hydroxylation. *J Am Chem Soc* **123:**5126–5127.

COR-AAD1-00001336

Lloyd, M.D., Lee, H.-J., Harlos, K., Zhang, Z.-H., Baldwin, J.E., Schofield, C.J., Charnock, J.M., Garner, C.D., Hara, T., Terwisscha Van Scheltinga, A.C., Valegård, K., Viklund, J.A.C., Hajdu, J., Andersson, I., Danielsson, Å., and Bhikhabhai, R. 1999a. Studies on the active site of deacetoxycephalosporin C synthase. *J Mol Biol* **287:**943–960.

Lloyd, M.D., Merrit, K.D., Lee, V., Sewell, T.J., Wha-Son, B., Baldwin, J.E., Schofield, C.J., Elson, S.W., Baggaley, K.H., and Nicholson, N.H. 1999b. Product-substrate engineering by bacteria: studies on clavaminate synthase, a trifunctional dioxygenase. *Tetrahedron* **55:**10201–10220.

Lloyd, M.D., Lipscomb, S.J., Hewitson, K.S., Hensgenes, C.M., Baldwin, J.E., and Schofield, C.J. 2004. Controlling the substrate selectivity of deacetoxy/deacetylcephalosporin C synthase. *J Biol Chem*, In press.

Long, A.J., Clifton, I.J., Roach, P.L., Baldwin, J.E., Schofield, C.J., and Rutledge, P.J. 2003. Structural studies on the reaction of isopenicillin N synthase with the substrate analogue δ-(L-α-aminoadipoyl)-L-cysteinyl-D-α-aminobutyrate. *Biochem J* **372:**687–693.

Lukacin, R. and Britsch, L. 1997. Identification of strictly conserved histidine and arginine residues as part of the active site in *Petunia hybrida* flavanone 3ß-hydroxylase. *Eur J Biochem* **249:**748–757.

Lukacin, R., Groning, I., Pieper, U., and Matern, U. 2000a. Site-directed mutagenesis of the active site serine290 in flavanone 3β-hydroxylase from *Petunia hybrida*. *Eur J Biochem* **267:**853–860.

Lukacin, R., Gröning, I., Schiltz, E., Britsch, L., and Matern, U. 2000b. Purification of recombinant flavanone 3β-hydroxylase from Petunia hybrida and assignment of the primary site of proteolytic degradation. *Arch Biochem Biophys* **375:**364–370.

Lukacin, R., Wellman, F., Britsch, L., Martens, S., and Matern, U. 2003. Flavonol synthase from *Citrus unshiu* is a bifunctional dioxygenase. *Phytochem* **62:**287–292.

MacBeth, C.E., Golombek, A.P., Young, V.G., Jr., Yang, C., Kuczera, K., Hendrich, M.P., and Borovik, A.S. 2000. O_2 activation by non-heme iron complexes: a monomeric Fe(III)-oxo complex derived from O_2. *Science* **289:**938–941.

Mahon, P.C., Hirota, K., and Semenza, G.L. 2001. FIH-1: a novel protein that interacts with HIF-1α and VHL to mediate repression of HIF-1 transcriptional activity. *Genes Dev* **15:**2675–2686.

Mao, Y., Varoglu, M., and Sherman, D.H. 1999. Molecular characterization and analysis of the biosynthetic gene cluster for the antitumor antibiotic mitomycin C from *Streptomyces lavendulae* 2564. *Chem Biol* **6:**251–263.

Marsh, E.N., Chang, M.D.-T., and Townsend, C.A. 1992. Two isozymes of clavaminate synthase central to clavulanic acid formation: cloning and sequencing of both genes from *Streptomyces clavuligerus*. *Biochemistry* **31:**12648–12657.

Martens, S., Forkmann, G., Britsch, L., Wellmann, F., Matern, U., and Lukacin, R. 2003. Divergent evolution of flavonoid 2-oxoglutarate-dependent dioxygenases in parsley. *FEBS Lett* **544:**93–98.

Martens, S., Forkmann, G., Matern, U., and Lukacin, R. 2001. Cloning of parsley flavone synthase I. *Phytochem* **58:**43–46.

Martin, D.N., Proebsting, W.M., and Hedden, P. 1997. Mendel's dwarfing gene: cDNAs from the *Le* alleles and function of the expressed proteins. *Proc Natl Acad Sci USA* **94:**8907–8911.

Martinkus, K.J., Tann, C.-H., and Gould, S.J. 1983. The biosynthesis of the streptolidine moiety in streptothricin F. *Tetrahedron* **39:**3493–3505.

Masson, N., William, C., Maxwell, P.H., Pugh, C.W., and Ratcliffe, P.J. 2001. Independent function of two destructive domains in hypoxia-inducible factor-α chains activated by prolyl hydroxylation. *EMBO J* **20:**5197–5206.

Matsuda, J., Okabe, S., Hashimoto, T., and Yamada, Y. 1991. Molecular cloning of hyoscyamine 6β-hydroxylase, a 2-oxoglutarate-dependent dioxygenase, from cultured roots of *Hyoscyamus niger*. *J Biol Chem* **266:**9460–9464.

McGinnis, K., Ku, G.M., VanDusen, W.J., Fu, J., Garsky, V., Stern, A.M., and Friedman, P.A. 1996. Site-directed mutagenesis of residues in a conserved region of bovine aspartyl (asparaginyl) ß-hydroxylase: evidence that histidine 675 has a role in binding Fe^{2+}. *Biochemistry* **35:**3957–3962.

McGowan, S.J., Sebaihia, M., Porter, L.E., Stewart, G.S., Williams, P., Bycroft, B.W., and Salmond, G.P. 1996. Analysis of bacterial carbapenam antibiotic production genes reveals a novel β-lactam biosynthesis pathway. *Molec Microbiol* **22:**415–426.

McNeill, L.A., Hewitson, K.S., Claridge, T.D., Seibel, J.F., Horsfall, L.E., and Schofield, C.J. 2002a. Hypoxia-inducible factor asparginyl hydroxylase (FIH-1) catalyses hydroxylation at the β-carbon of asparagine-803. *Biochem J* **367:**571–575.

McNeill, L.A., Hewitson, K.S., Gleadle, J.M., Horsfall, L.E., Oldham, N.J., Maxwell, P.H., Pugh, C.W., Ratcliffe, P.J., and Schofield, C.J. 2002b. The use of dioxygen by HIF prolyl hydroxylase (PHD1). *Bioorg Med Chem Lett* **12:**1547–1550.

Mehn, M.P., Fujisawa, K., Hegg, E.L., and Que, L., Jr. 2003. Oxygen activation by nonheme iron(II) complexes: α-keto carboxylate versus carboxylate. *J Am Chem Soc* **124:**7828–7842.

Metcalf, W.W. and Wolfe, R.S. 1998. Molecular genetic analysis of phosphite and hypophosphite oxidation by *Pseudomonas stutzeri* WM88. *J Bacteriol* **180:**5547–5558.

Mihalik, S.J., Morrell, J.C., Kim, D., Sacksteder, K.A., Watkins, P.A., and Gould, S.J. 1997. Identification of *PAHX*, a Refsum disease gene. *Nature Genet* **17:**185–189.

Mihalik, S.J., Rainville, A.M., and Watkins, P.A. 1995. Phytanic acid α−oxidation in rat liver peroxisomes. Production of α-hydroxyphytanoyl-CoA and formate is enhanced by dioxygenase cofactors. *Eur J Biochem* **232:**545–551.

Min, J.-H., Yang, H., Ivan, M., Gertler, F., Kaelin, W.G., Jr., and Pavletich, N.P. 2002. Structure of an HIF-1α-pVHL complex: hydroxyproline recognition in signaling. *Science* **296:**1886–1889.

Min, W.-C., Begley, T.J., Myllyharju, J., and Kivirikko, K.I. 2000. Mechanistic studies on prolyl-4-hydroxylase: demonstration that the ferryl intermediate does not exchange with water. *Bioorgan Chem* **28:**261–265.

Minchenko, A., Leshchinsky, I., Opentanova, I., Sang, N., Srinivas, V., Armstead, V., and Caro, J. 2002. Hypoxia-inducible factor-1-mediated expression of the 6-phosphofructo-2-kinase/fructose-2,6-bisphosphate (*PFKFB3*) gene. Its possible role in the Warburg effect. *J Biol Chem* **277:**6183–6187.

Ming, L.-J., Que, L., Jr., Kriauciunas, A., Frolik, C.A., and Chen, V.J. 1990. Coordination chemistry of the metal binding site of isopenicillin N synthase. *Inorg Chem* **29:**1111–1112.

Ming, L.-J., Que, L., Jr., Kriauciunas, A., Frolik, C.A., and Chen, V.J. 1991. NMR studies of the active site of isopenicillin N

COR-AAD1-00001337

synthase, a non-heme iron(II) enzyme. *Biochemistry* **30:**11653–11659.

Mishina, Y. and He, C. 2003. Probing the structure and function of the *Escherichia coli* DNA alkylation repair AlkB protein through chemical cross-linking. *J Am Chem Soc* **125:**8730–8731.

Mishina, Y., Lee, C.-H.J., and He, C. 2004. Interaction of human and bacterial AlkB proteins with DNA as probed through chemical cross-linking studies. *Nucleic Acids Res* **32:**1548–1554.

Mori, H., Shibasaki, T., Uozaki, Y., Ochiai, K., and Ozaki, A. 1996. Detection of novel proline 3-hydroxylase activities in *Streptomyces* and *Bacillus* spp. by regio- and stereospecific hydroxylation of L-proline. *Appl Environ Microbiol* **62:**1903–1907.

Mori, H., Shibasaki, T., Yano, K., and Ozaki, A. 1997. Purification and cloning of a proline 3-hydroxylase, a novel enzyme which hydroxylates free L-proline to cis-3-hydroxy-L-proline. *J Bacteriol* **179:**5677–5683.

Mukherji, M., Kershaw, N.J., Schofield, C.J., Wierzbicki, A.S., and Lloyd, M.D. 2002. Utilization of sterol carrier protein-2 by phytanoyl-CoA 2-hydroxylase in the peroxisomal α oxidation of phytanic acid. *Chem Biol* **9:**597–605.

Mukherji, M., Schofield, C.J., Wierzbicki, A.S., Jansen, G.A., Wanders, R.J.A., and Lloyd, M.D. 2003. The chemical biology of branched-chain lipid metabolism. *Prog Lipid Res* **42:**359–376.

Mukherji, M., Chien, W., Kershaw, N.J., Clifton, I.J., Schofield, C.J., Wierzbicki, A.S., and Lloyd, M.D. 2001a. Structure-function analysis of phytanoyl-CoA 2-hydroxylase mutations causing Refsum's disease. *Human Molec Genet* **10:**1971–1982.

Mukherji, M., Kershaw, N.J., MacKinnon, C.H., Clifton, I.J., Wierzbicki, A.S., Schofield, C.J., and Lloyd, M.D. 2001b. "Chemical co-substrate rescue" of phytanoyl-CoA hydroxylase mutants causing Refsum's disease. *Chem Commun* 972–973.

Müller, I., Kahnert, A., Pape, T., Sheldrick, G.M., Meyer-Klaucke, W., Dierks, T., Kertesz, M., and Usòn, I. 2004. Crystal structure of the alkylsulfatase AtsK: insights into the catalytic mechanism of the Fe(II), α-ketoglutarate dependent dioxygenase superfamily. *Biochemistry* **43:**3075–3088.

Myllyharju, J. and Kivirikko, K.I. 1997. Characterization of the iron- and 2-oxoglutarate-binding sites of human prolyl 4-hydroxylase. *EMBO J* **16:**1173–1180.

Myllylä, R., Kuutti-Saolainen, E.-R., and Kivirikko, K.I. 1978. The role of ascorbate in the prolyl hydroxylase reaction. *Biochem Biophys Res Commun* **83:**441–448.

Myllylä, R., Majamaa, K., Günzler, V., Hanauske-Abel, H.N., and Kivirikko, K.I. 1984. Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. *J Biol Chem* **259:**5403–5405.

Nagahama, K., Ogawa, T., Fujii, T., Tazaki, M., Goto, M., and Fukuda, H. 1991a. *L*-Arginine is essential for the formation *in vitro* of ethylene by an extract of *Pseudomonas syringae*. *J Gen Microbiol* **137:**1641–1646.

Nagahama, K., Ogawa, T., Fujii, T., Tazaki, M., Tanase, S., Morino, Y., and Fukuda, H. 1991b. Purification and properties of an ethylene-forming enzyme from *Pseudomonas syringae* pv. *phaseolica* PK2. *J Gen Microbiol* **137:**2281–2286.

Nagahama, K., Yoshini, K., Matsuoka, M., Tanase, S., Ogawa, T., and Fukuda, H. 1998. Site-directed mutagenesis of histidine residues in the ethylene-forming enzyme from *Pseudomonas syringae*. *J Ferm Bioeng* **85:**255–258.

Nakajima, J.-I., Yoshikaru, Y., Yamazaki, M., and Saito, K. 2001. Reaction mechanism from leucoanthocyanidin to anthocyanidin 3-glucoside, a key reaction for coloring in anthocyanin biosynthesis. *J Biol Chem* **276:**25797–25803.

Nickel, K., Suter, M.J.-F., and Kohler, H.-P.E. 1997. Involvement of two α-ketoglutarate-dependent dioxygenases in enantioselective degradation of (*R*)- and (*S*)-mecoprop by *Sphingomonas herbicidovorans* MH. *J Bacteriol* **179:**6674–6679.

O'Brien, J.R., Schuller, D.J., Yang, V.S., Dillard, B.D., and Lanzilotta, W.N. 2003. Substrate-induced conformational changes in *Escherichia coli* taurine/α-ketoglutarate dioxygenase and insight into the oligomeric structure. *Biochemistry* **42:**5547–5554.

Oehme, F., Ellinghaus, P., Kolkhof, P., Smith, T.J., Ramakrishnan, S., Hütter, J., Schramm, M., and Flamme, I. 2002. Overexpression of PH-4, a novel putative proline 4-hydroxylase, modulates activity of hypoxia-inducible transcription factors. *Biochem Biophys Res Commun* **296:**343–349.

Ogle, J.M., Clifton, I.J., Rutledge, P.J., Elkins, J.M., Burzlaff, N.I., Adlington, R.M., Roach, P.L., and Baldwin, J.E. 2001. Alternative oxidation by isopenicillin *N* synthase observed by X-ray diffraction. *Chem Biol* **8:**1231-1237.

Omura, S., Tomoda, H., Yamamoto, S., Tsukui, M., and Tanaka, H. 1984. Studies on two dioxygenases involved in the synthesis of tylosin in *Streptomyces fradiae*. *Biochim Biophys Acta* **802:**141–147.

Park, M.H., Wolff, E.C., and Folk, J.E. 1993. Hypusine: its post-translational formation in eukaryotic initiation factor 5A and its potential role in cellular regulation. *Biofactors* **4:**95–104.

Park, J.H., Wolff, E.C., Folk, J.E., and Park, M.H. 2003. Reversal of the deoxyhypusine synthesis reaction. Generation of spermidine or homospermidine from deoxyhypusine by deoxyhypusine synthase. *J Biol Chem* **278:**32683–32691.

Passoja, K., Myllyharju, J., Pirskanen, A., and Kivirikko, K.I. 1998a. Identification of arginine-700 as the residue that binds the C-5 carboxyl group of 2-oxoglutarate in human lysyl hydroxylase 1. *FEBS Lett* **434:**145–148.

Passoja, K., Rautavuoma, K., Ala-Kokko, L., Kosonen, T., and Kivirikko, K.I. 1998b. Cloning and characterization of a third human lysyl hydroxylase isoform. *Proc Natl Acad Sci USA* **95:**10482–10486.

Pavel, E.G., Zhou, J., Busby, R.W., Gunsior, M., Townsend, C.A., and Solomon, E.I. 1998. Circular dichroism and magnetic circular dichroism spectroscopic studies of the non-heme ferrous active site in clavaminate synthase and its interaction with α-ketoglutarate cosubstrate. *J Am Chem Soc* **120:**743–753.

Pekkala, M., Hieta, R., Kursula, P., Kivirikko, K.I. Wierenga, R.K., and Myllyharju, J. 2003. Crystallization of the proline-rich-peptide binding domain of human type I collagen prolyl 4-hydroxylase. *Acta Cryst* **D59:**940–942.

Pelletier, M.K. and Shirley, B.W. 1996. Analysis of flavanone 3-hydroxylase in *Arabidopsis* seedlings. *Plant Physiol* **111:**339–345.

Petersen, L., Olewinski, R., Salmon, P., and Conners, N. 2003. Novel proline hydroxylase activities in the pneumocandin-producing fungus *Glarea lozoyensis* responsible for the formation of *trans* 3- and *trans* 4-hydroxyproline. *Appl Microbiol Biotechnol* **62:**263–267.

Pirrung, M.C., Cao, J., and Chen, J.1998. Ethylene biosynthesis: processing of a substrate analog supports a radical mechanism for the ethylene-forming enzyme. *Chem Biol* **5:**49–57.

COR-AAD1-00001338

Pirrung, M.C., Kaiser, L.M., and Chen, J. 1993. Purification and properties of the apple fruit ethylene-forming enzyme. *Biochemistry* **32:**7445–7450.

Pirskanen, A., Kaimio, A.-M., Myllylä, R., and Kivirikko, K.I. 1996. Site-directed mutagenesis of human lysyl hydroxylase expressed in insect cells. Identification of histidine residues and an aspartic acid residue critical for catalytic activity. *J Biol Chem* **271:**9398–9402.

Poneleit, L.S. and Dilley, D.R. 1993. Carbon dioxide activation of 1-aminocyclopropane-1-carboxylate (ACC) oxidase in ethylene biosynthesis. *Postharvest Biol Technol* **3:**191–199.

Prescott, A.G. 1993. A dilemma of dioxygenases (or where biochemistry and molecular biology fail to meet). *J Exper Bot* **44:**849–861.

Prescott, A.G. and Lloyd, M.D. 2000. The iron(II) and 2-oxoacid-dependent dioxygenases and their role in metabolism. *Nat Prod Rep* **17:**367–383.

Prescott, A.G., Stamford, N.P.J., Wheeler, G., and Firmin, J.L. 2002. In vitro properties of a recombinant flavonol synthase from *Arabidopsis thaliana*. *Phytochem* **60:**589–593.

Price, J.C., Barr, E.W., Glass, T.E., Krebs, C., and Bollinger, J.M., Jr. 2003a. Evidence for hydrogen abstraction from C1 of taurine by the high-spin Fe(IV) intermediate detected during oxygen activation by taurine:α-ketoglutarate dioxygenase (TauD). *J Am Chem Soc* **125:**13008–13009.

Price, J.C., Barr, E.W., Tirupati, B., Bollinger, J.M., Jr., and Krebs, C. 2003b. The first direct characterization of a high-valent iron intermediate in the reaction of an α-ketoglutarate-dependent dioxygenase: a high-spin Fe(IV) complex in taurine/α-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. *Biochemistry* **42:**7497–7508.

Proshlyakov, D.A., Henshaw, T.F., Monterosso, G.R., Ryle, M.J., and Hausinger, R.P. 2004. Direct detection of oxygen intermediates in the non-heme Fe enzyme taurine/α-ketoglutarate dioxygenase. *J Am Chem Soc* **126:**1022–1023.

Puistola, U., Turpeenniemi-Hujanen, T.M., Myllylä, R., and Kivirikko, K. 1980. Studies on the lysyl hydroxylase reaction. 1. Initial velocity kinetics and related aspects. *Biochim Biophys Acta* **611:**40–50.

Randall, C.R., Zang, Y., True, A.E., Que, L., Jr., Charnock, J.M., Garner, C.D., Fujishima, Y., Schofield, C.J., and Baldwin, J.E. 1993. X-ray absorption studies of the ferrous active site of isopenicillin N synthase and related model complexes. *Biochemistry* **32:**6664–6673.

Rao, N.V. and Adams, E. 1978. Partial reaction of prolyl hydroxylase. $(Gly\text{-}Gly\text{-}Pro)_n$ stimulates α-ketoglutarate decarboxylation without prolyl hydroxylase. *J Biol Chem* **253:**6327–6330.

Risteli, J., Tryggvason, K., and Kivirikko, K.I. 1977. Prolyl 3-hydroxylase: partial purification of the enzyme from rat kidney cortex. *Eur J Biochem* **73:**485–492.

Roach, P.L., Clifton, I.J., Fülöp, V., Harlos, K., Barton, G.J., Hajdu, J., Andersson, K., Schofield, C.J., and Baldwin, J.E. 1995. Crystal structure of isopenicillin *N* synthase is the first from a new structural family of enzymes. *Nature (London)* **375:**700–704.

Roach, P.L., Clifton, I.J., Hensgens, C.M.H., Shibata, N., Schofield, C.J., Hajdu, J., and Baldwin, J.E. 1997. Structure of isopenicillin *N* synthase complexed with substrate and the mechanism of penicillin formation. *Nature (London)* **387:**827–830.

Rocklin, A.M., Tierney, D.L., Kofman, V., Brunhuber, N.M.W., Hoffman, B.M., Christoffersen, R.E., Reich, N.O., Lipscomb, J.D., and Que, L., Jr. 1999. Role of the nonheme Fe(II) center in the biosynthesis of the plant hormone ethylene. *Proc Natl Acad Sci USA* **96:**7905–7090.

Roelfes, G., Lubben, M., Chen, K., Ho, R.Y.N., Meetsma, A., Genseberger, S., Hermant, R.M., Hage, R., Mandal, S.K., Young, V.G., Jr., Zang, Y., Kooijman, H., Spek, A.L., Que, L., Jr., and Feringa, B.L. 1999. Iron chemistry of a pentadentate ligand that generates a metastable Fe^{III}-OOH intermediate. *Inorg Chem* **38:**1929–1936.

Roelfes, G., Lubben, M., Hage, R., Que, L., Jr., and Feringa, B.L. 2000. Catalytic oxidation with a non-heme iron complex that generates a low-spin $Fe^{III}OOH$ intermediate. *Chem Eur J* **6:**2152–2159.

Rohde, J.-U., In, J.-H., Lim, M.H., Brennessel, W.W., Bukowski, M.R., Stubna, A., Münck, E., and Que, L., Jr. 2003. Crystallographic and spectroscopic characterization of a nonheme Fe(IV)=O complex. *Science* **299:**1037–1039.

Rüetschi, U., Nordin, I., Odelhög, B., Jörnvall, H., and Lindstedt, S. 1993. γ-Butyrobetaine hydroxylase. Structural characterization of the *Pseudomonas* enzyme. *Eur J Biochem* **213:**1075–1080.

Ryle, M.J. and Hausinger, R.P. 2002. Non-heme iron oxygenases. *Curr Opinion Chem Biol* **6:**193–201.

Ryle, M.J., Koehntop, K.D., Liu, A., Que, L., Jr., and Hausinger, R.P. 2003a. Interconversion of two oxidized forms of TauD, a non-heme iron hydroxylase: evidence for bicarbonate binding. *Proc Natl Acad Sci USA* **100:**3790–3795.

Ryle, M.J., Liu, A., Muthukumaran, R.B., Ho, R.Y.N., Koehntop, K.D., McCracken, J., Que, L., Jr., and Hausinger, R.P. 2003b. O_2- and α-ketoglutarate-dependent tyrosyl radical formation in TauD, an α-keto acid-dependent non-heme iron dioxygenase. *Biochemistry* **42:**1854–1862.

Ryle, M.J., Padmakumar, R., and Hausinger, R.P. 1999. Stopped-flow kinetic analysis of *Escherichia coli* taurine/α-ketoglutarate dioxygenase: interactions with α-ketoglutarate, taurine, and oxygen. *Biochemistry* **38:**15278–15286.

Saari, R.E. and Hausinger, R.P. 1998. Ascorbic acid-dependent turnover and reactivation of 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase using thiophenoxyacetic acid. *Biochemistry* **37:**3035–3042.

Saari, R.E., Hogan, D.A., and Hausinger, R.P. 1999. Stereospecific degradation of the phenoxypropionate herbicide dichlorprop. *J Molec Catalysis B* **6:**421–428.

Sabourin, P.J. and Bieber, L.L. 1982. The mechanism of α-ketoisocaproate oxygenase. Formation of β-hydroxyisovalerate from α-ketoisocaproate. *J Biol Chem* **257:**7468–7471.

Saito, K., Kobayashi, M., Gong, Z., Tanaka, Y., and Yamazaki, M. 1999. Direct evidence for anthocyanidin synthase as a 2-oxoglutarate-dependent oxygenase: molecular cloning and functional expression of cDNA from a red forma of *Perilla fructescens*. *Plant J* **17:**181–189.

Sami, M., Brown, T.J.N., Roach, P.L., Schofield, C.J., and Baldwin, J.E. 1997. Glutamine-330 is not essential for activity in isopenicillin N synthase from *Aspergillus nidulans*. *FEBS Lett* **405:**191–194.

Samson, S.M., Belagaje, R., Blankenship, D.T., Chapman, J.L., Perry, D., Skatrud, P.L., VanFrank, R., Abraham, E.P., Baldwin, J.E., Queener, S.W., and Ingolia, T.D. 1985. Isolation, sequence determination and expression in *Escherichia coli* of the isopenicillin N synthetase gene from *Cephalosporium acremonium*. *Nature (London)* **318:**191–194.

Sauvaire, Y., Petit, P., Broca, C., Manteghetti, M., Baissac, Y., Fernandez-Alverez, J., Gross, R., Roye, M., Leconte, A., Gomis, R., and Ribes, G. 1998. 4-Hydroxyisoleucine: a novel amino acid potentiator of insulin secretion. *Diabetes* **47:**206–210.

COR-AAD1-00001339

Schofield, C.J., Baldwin, J.E., Byford, M.F., Clifton, I.J., Hensgens, C., and Roach, P.L. 1997. Proteins of the penicillin biosynthesis pathway. *Curr Opinion Struct Biol* **7:**857–864.

Schofield, C.J. and Zhang, Z. 1999. Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Curr Opinion Struct Biol* **9:**722–731.

Schomburg, F.M., Bizzell, C.M., Lee, D.J., Zeevaart, J.A.D., and Amasino, R.M. 2003. Overexpression of a novel class of gibberellin 2-oxidases decreases gibberellin levels and creates dwarf plants. *The Plant Cell* **15:**151–163.

Scott, R.A., Wang, S., Eidsness, M.K., Kriauciunas, A., Frolik, C.A., and Chen, V.J. 1992. X-ray absorption spectroscopic studies of the high-spin iron(II) active site of isopenicillin N synthase: evidence for Fe-S interaction in the enzyme-substrate complex. *Biochemistry* **31:**4596–4601.

Sedgwick, B. and Lindahl, T. 2002. Recent progress on the Ada response for inducible repair of DNA alkylation damage. *Oncogene* **21:**8886–8894.

Senda, T., Sugiyama, K., Narita, H., Yamamoto, T., Kimbara, K., Fukuda, M., Sato, M., Yano, K., and Mitsui, Y. 1996. Three-dimensional structures of free form and two substrate complexes of an extradiol ring-cleavage type dioxygenase, the BphC enzyme from *Pseudomonas* sp. strain KKS102. *J Mol Biol* **255:**735–752.

Seo, J.-A., Proctor, R.H., and Plattner, R.D. 2001. Characterization of four clustered and coregulated genes associated with fumonisin biosynthesis in *Fusarium verticillioides*. *Fungal Genet Biol* **34:**155–165.

Serre, L., Sailland, A., Sy, D., Boudec, P., Rolland, A., Pebay-Peyroula, E., and Cohen-Addad, C. 1999. Crystal structure of *Pseudomonas fluorescens* 4-hydroxyphenylpyruvate dioxygenase: an enzyme involved in the tyrosine degradation pathway. *Structure* **7:**977–988.

Shaffer, P.M., Cairns, J.R., Dorman, D.C., and Lott, A.M. 1984. Separation and characterization of pyrimidine deoxyribonucleoside 2′-hydroxylase and thymine 7-hydroxylase from *Aspergillus nidulans*. *Int J Biochem* **16:**429–434.

Shaffer, P.M., McCroskey, R.P., Palmatier, R.D., Midgett, R.J., and Abbott, M.T. 1968. The cell-free conversion of a deoxyribonucleoside to a ribonucleoside without detachment of the deoxyribose. *Biochem Biophys Res Commun* **33:**806–811.

Shaw, J.-F., Chou, Y.-S., Chang, R.-C., and Yang, S.F. 1996. Characterization of the ferrous ion binding sites of apple 1-aminocyclopropane-1-carboxylate oxidase by site-directed mutagenesis. *Biochem Biophys Res Commun* **225:**697–700.

Shibasaki, T., Mori, H., Chiba, S., and Ozaki, A. 1999. Microbial proline 4-hydroxylase screening and gene cloning. *Appl Environ Microbiol* **65:**4028–4031.

Simaan, A.J., Banse, F., Mialane, P., Boussac, A., Un, S., Kargar-Grisel, T., Bouchoux, G., and Girerd, J.J. 1999. Characterization of a non-heme mononuclear peroxoiron(III) intermediate by UV vis and EPR spectroscopy and mass spectrometry. *Eur J Inorg Chem* **6:**993–999.

Smith, J.J. and John, P. 1993. Activation of 1-aminocyclopropane-1carboxylate oxidase by bicarbonate/carbon dioxide. *Phytochem* **32:**1381–1386.

Solomon, E.I., Brunold, T.C., Davis, M.I., Kemsley, J.N., Lee, S.-K., Lehnert, N., Neese, F., Skulan, A.J., Yang, Y.-S., and Zhou, J. 2000. Geometric and electronic structure/function correlations in non-heme iron enzymes. *Chem Rev* **100:**235–349.

Sono, M., Roach, M.P., Coulter, E.D., and Dawson, J.H. 1996. Heme-containing oxygenases. *Chem Rev* **96:**2841–2887.

Springob, K., Nakajima, J.I., Yamazaki, M., and Saito, K. 2003. Recent advances in the biosynthesis and accumulation of anthocyanins. *Nat Prod Rep* **20:**288–303.

Stenflo, J., Holme, E., Lindstedt, S., Chandramouli, N., Huang, L.H.T., Tam, J.P., and Merrifield, R.B. 1989. Hydroxylation of aspartic acid in domains homologous to the epidermal growth factor precursor is catalyzed by a 2-oxoglutarate-dependent dioxygenase. *Proc Natl Acad Sci USA* **86:**444–447.

Streber, W.R., Timmis, K.N., and Zenk, M.H. 1987. Analysis, cloning, and high level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JPM134. *J Bacteriol* **169:**2950–2955.

Stubbe, J. 1985. Identification of two α-ketoglutarate-dependent dioxygenases in extracts of *Rhodotorula glutinis* catalyzing deoxyuridine hydroxylation. *J Biol Chem* **260:**9972–9975.

Swiegers, J.H., Vaz, F.M., Pretorius, I.S., Wanders, R.J.A., and Bauer, F.F. 2002. Carnitine biosynthesis in *Neurospora crassa*: identification of a cDNA coding for ε-*N*-trimethyllysine hydroxylase and its functional expression in *Saccharomyces cerevisiae*. *FEMS Microbiol Lett* **210:**19–23.

Tan, D.S.H. and Sim, T.-S. 1996. Functional analysis of conserved histidine residues in *Cephalosporium acremonium* isopenicillin N synthase by site-directed mutagenesis. *J Biol Chem* **271:**889–894.

Tett, V.A., Willetts, A.J., and Lappin-Scott, H.M. 1997. Biodegradation of the chlorophenoxy herbicide (*R*)-(+)-mecoprop by *Alcaligenes denitrificans*. *Biodegradation* **8:**43–52.

Thomas, S.G., Phillips, A.L., and Hedden, P. 1999. Molecular cloning and functional expression of gibberellin 2-oxidases, multifunctional enzymes involved in gibberellin deactivation. *Proc Natl Acad Sci USA* **96:**4698–4703.

Thornburg, L.D., Lai, M.-T., Wishnok, J.S., and Stubbe, J. 1993. A non-heme iron protein with heme tendencies: an investigation of the substrate specificity of thymine hydroxylase. *Biochemistry* **32:** 14023–14033.

Thornburg, L.D. and Stubbe, J. 1989. Mechanism-based inhibition of thymine hydroxylase. *J Am Chem Soc* **111:**7632–7633.

Thornburg, L.D. and Stubbe, J. 1993. Mechanism-based inactivation of thymine hydroxylase, an α-ketoglutarate-dependent dioxygenase, by 5-ethynyluracil. *Biochemistry* **32:**14034–14042.

Thrower, J.S., Blalock, R. III, and Klinman, J.P. 2001. Steady-state kinetics of substrate binding and iron release in tomato ACC oxidase. *Biochemistry* **40:**9717–9724.

Tiedje, J.M. and Alexander, M. 1969. Enzymatic cleavage of the ether bond of 2,4-dichlorophenoxyacetate. *J Agric Food Chem* **17:**1080–1084.

Tiow-Suan, S. and Tan, D.S.H. 1994. Histidine-272 of isopenicillin N synthase of *Cephalosporium acremonium*, which is possibly involved in iron binding, is essential for its catalytic activity. *FEMS Microbiol Lett* **120:**241–248.

Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T., and Sedgwick, B. 2002. Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature* **419:**174–178.

Tryggvason, K., Majamaa, K., Risteli, J., and Kivirikko, K.I. 1979. Partial purification and characterization of chick-embryo prolyl 3-hydroxylase. *Biochem J* **183:**303–307.

COR-AAD1-00001340

Tryggvason, K., Risteli, J., and Kivirikko, K.I. 1976. Separation of prolyl 3-hydroxylase and 4-hydroxylase activities and the 4-hydroxyproline requirement for synthesis of 3-hydroxyproline. *Biochem Biophys Res Commun* **76:**275–281.

Tuderman, L., Myllylä, R., and Kivirikko, K.I. 1977. Mechanism of prolyl hydroxylase reaction. 1. Role of co-substrates. *Eur J Biochem* **80:**314–348.

Turnbull, J.J., Nagle, M.J., Seibel, J.F., Welford, R.W.D., Grant, G.H., and Schofield, C.J. 2003. The C-4 stereochemistry of leucocyanidine substrates for anthocyanidin synthase affects product selectivity. *Bioorg Med Chem Lett* **13:**3853–3857.

Turnbull, J.J., Nakajima, J.I., Welford, R.W., Yamazaki, M., Saito, K., and Schofield, C.J. 2004. Mechanistic studies on three 2-oxoglutarate dependent oxygenases of flavonoid biosynthesis: anthocyanidin synthase, flavonol synthase, and flavaonone 3b hydroxylase. *J Biol Chem* **279:**1206–1216.

Udenfriend, S. and Cardinale, G. 1982. α-Ketoglutarate-coupled dioxygenases. In *Oxygenases and Oxygen Metabolism*, pp. 99–110, M. Nozaki, S. Yamamoto, Y. Ishimura, M.J. Coon, L. Ernster, and R.W. Estabrook, Eds., Academic Press, New York.

Valegård, K., Terwisscha van Scheltinga, A.C., Dubus, A., Ranghino, G., Öster, L.M., Hajdu, J., and Andersson, I. 2004. The structural basis of cephalosporin formation in a mononuclear ferrous enzyme. *Nat Struct Biol* in press.

Valegård, K., Terwisscha van Scheltinga, A.C., Lloyd, M.D., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Lee, W.-J., Baldwin, J.E., Schofield, C.J., Hajdu, J., and Andersson, I. 1998. Structure of a cephalosporin synthase. *Nature (London)* **394:**805–809.

Valtavaara, M., Papponen, H., Pirttila, A.M., Hiltunen, K., Helander, H., and Myllylä, R. 1997. Cloning and characterization of a novel human lysyl hydroxylase isoform highly expressed in pancreas and muscle. *J Biol Chem* **272:**6831–6834.

Valtavaara, M., Szpirer, C., Szpirer, J., and Myllylä, R. 1998. Primary structure, tissue distribution, and chromosomal localization of a novel isoform of lysyl hydroxylase (lysyl hydroxylase 3). *J Biol Chem* **273:**12881–12886.

Vaz, F.M., Ofman, R., Westinga, K., Back, J.W., and Wanders, R.J.A. 2001. Molecular and biochemical characterization of rat *N*-trimethyllysine hydroxylase, the first enzyme of carnitine biosynthesis. *J Biol Chem* **276:**33512–33517.

Vaz, F.M., van Gool, S., Ofman, R., Ijlst, L., and Wanders, R.J.A. 1998. Carnitine biosynthesis: identification of the cDNA encoding human γ-butyrobetaine hydroxylase. *Biochem Biophys Res Commun* **250:**506–510.

Vaz, F.M. and Wanders, R.J.A. 2002. Carnitine biosynthesis in mammals. *Biochem J* **361:**417–429.

Vazquez-Flota, F., De Carolis, E., Alarco, A.-M., and De Luca, V. 1997. Molecular cloning and characterization of desacetoxyvindoline-4-hydroxylase, a 2-oxoglutarate dependent-dioxygenase involved in the biosythesis of vindoline in *Catharanthus roseus* (L.) G. Don. *Plant Molec Biol* **34:**935–948.

Veijola, J., Koivunen, P., Annunem, P., Pihlajaniemi, T., and Kivirikko, K.I. 1994. Cloning, Baculovirus expression, and characterization of the α subunit of prolyl 4-hydroxylase from the nematode *Caenorhaditis elegans*. This α subunit forms an active αβ dimer with the human protein disulfide isomerase/β subunit. *J Biol Chem* **269:**26746–26753.

Verhoeven, N.M., Schor, D.S.M., ten Brink, H.J., Wanders, R.J.A., and Jakobs, C. 1997. Resolution of the phytanic acid α-oxidation pathway: identification of pristanal as product of the decarboxylation of 2-hydroxyphytanoyl-CoA. *Biochem Biophys Res Commun* **237:**33–36.

Wada, A., Ogo, S., Watanabe, Y., Mukai, M., Kitagawa, T., Jitsukawa, K., Masuda, H., and Einaga, H. 1999. Synthesis and characterization of novel alkylperoxo mononuclear iron(III) complexes with a tripodal pyridylamine ligand: a model for peroxo intermediates in reactions catalyzed by non-heme iron enzymes. *Inorg Chem* **38:**3592–3593.

Wang, Q., VanDusen, W.J., Petroski, C.J., Garsky, V.M., Stern, A.M., and Friedman, P.A. 1991. Bovine liver aspartyl β-hydroxylase. Purification and characterization. *J Biol Chem* **266:**14004–14010.

Warn-Cramer, B.J., Macrander, L.A., and Abbott, M.T. 1983. Markedly different ascorbate dependencies of the sequential α-ketoglutarate dioxygenase reactions catalyzed by an essentially homogeneous thymine 7-hydroxylase from *Rhodotorula glutinis*. *J Biol Chem* **258:**10551–10557.

Watanabe, M., Sumida, N., Murakami, S., Anzai, H., Thompson, C.J., Tateno, Y., and Murakami, T. 1999. A phosphonate-induced gene which promotes *Penicillium*-mediated bioconversion of *cis*-propenylphosphonic acid to fosfomycin. *Appl Environ Microbiol* **65:**1036–1044.

Watkins, P.A., Howard, A.E., and Mihalik, S.J. 1994. Phytanic acid must be activated to phytanoyl-CoA prior to its α-oxidation in rat liver peroxisomes. *Biochim Biophys Acta* **1214:**288–294.

Wei, Y.-F., Carter, K.C., Wang, R.-P., and Shell, B.K. 1996. Molecular cloning and functional analysis of a human cDNA encoding an *Escherichia coli* AlkB homolog, a protein involved in DNA alkylation damage repair. *Nucl Acids Res* **24:**931–937.

Welford, R.W.D., Schlemminger, I., McNeill, L.A., Hewitson, K.S., and Schofield, C.J. 2003. The selectivity and inhibition of AlkB. *J Biol Chem* **278:**10157–10161.

Wellman, F., Lukacin, R., Moriguchi, T., Britsch, L., Schiltz, E., and Matern, U. 2002. Functional expression and mutational analysis of flavonol synthase from *Citrus unshiu*. *Eur J Biochem.* **269:**4134–4142.

Wellman, F., Matern, U., and Lukacin, R. 2004. Significance of C-terminal sequence elements for *Petunia* flavanone 3β-hydroxylase activity. *FEBS Lett* **561:**149–154.

Westendorf, A., Müller, A., and Babel, W. 2003. Purification and characterization of the enantiospecific dioxygenases from *Delftia acidovorans* MC1 initiating the degradation of phenoxypropionate and phenoxyacetate herbicides. *Acta Biotechnol* **23:**3–17.

White, A.K. and Metcalf, W.W. 2002. Isolation and biochemical characterization of hypophosphite/2-oxoglutarate dioxygenase. A novel phosphorus-oxidizing enzyme from *Pseudomonas stutzeri* WM88. *J Biol Chem* **277:**38262–38271.

Whiting, A.K., Que, L., Jr., Saari, R.E., Hausinger, R.P., Fredrick, M.A., and McCracken, J. 1997. Metal coordination environment of a Cu(II)-substituted α-keto acid-dependent dioxygenase that degrades the herbicide 2,4-D. *J Am Chem Soc* **119:**3413–3414.

Wierzbicki, A.S., Lloyd, M.D., Schofield, C.J., Feher, M.D., and Gibberd, F.B. 2002. Refsum's disease: a peroxisomal disorder affecting phytanic acid α-oxidation. *J Neurochem* **80:**727–735.

Wilmouth, R.C., Turnbull, J.J., Welford, R.W.D., Clifton, I.J., Prescott, A.G., and Schofield, C.J. 2002. Structure and mechanism of

COR-AAD1-00001341

anthocyanidin synthase from *Arabidopsis thaliana*. *Structure* **10:**93–103.

Wu, H., de Graaf, B., Mariani, C., and Cheung, A.Y. 2001. Hydroxyproline-rich glycoproteins in plant reproductive tissues: structure, functions and regulation. *Cell Molec Life Sci* **58:**1418–1429.

Yin, X., O'Hare, T., Gould, S.J., and Zabriskie, T.M. 2003. Identification and cloning of genes encoding viomycin biosynthesis from *Streptomyces vinaceus* and evidence for involvement of a rare oxygenase. *Gene* **312:**215–224.

Zhang, J.H., Quigley, N.B., and Gross, D.C. 1995. Analysis of *syrB* and *syrC* genes of *Pseudomonas syringae* pv. *syringae* indicates that syringomycin is synthesized by a template mechanism. *J Bacteriol* **177:**4009–4020.

Zhang, Z., Barlow, J.N., Baldwin, J.E., and Schofield, C.J. 1997. Metal-catalyzed oxidation and mutagenesis studies on the iron(II) binding site of 1-aminocyclopropane-1-carboxylate oxidase. *Biochemistry* **36:**15999–16007.

Zhang, Z., Ren, J., Stammers, D.K., Baldwin, J.E., Harlos, K., and Schofield, C.J. 2000. Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase. *Nat Struct Biol* **7:**127–133.

Zhang, Z., Ren, J.-S., Harlos, K., McKinnon, C.H., Clifton, I.J., and Schofield, C.J. 2002. Crystal structure of a clavaminate synthase-Fe(II)-2-oxoglutarate-substrate-NO complex: evidence for metal centred rearrangements. *FEBS Lett* **517:**7–12.

Zhou, J., Gunsior, M., Bachmann, B.O., Townsend, C.A., and Solomon, E.I. 1998. Substrate binding to the α-ketoglutarate-dependent non-heme iron enzyme clavaminate synthase 2: coupling mechanism of oxidative decarboxylation and hydroxylation. *J Am Chem Soc* **120:**13539–13540.

Zhou, J., Rocklin, A.M., Lipscomb, J.D., Que, L., Jr., and Solomon, E.I. 2002. Spectroscopic studies of 1-aminocyclopropane-1-carboxylic acid oxidase: molecular mechanism and CO_2 activation in the biosynthesis of ethylene. *J Am Chem Soc* **124:**4602–4609.

Zhou, J., Kelly, W.L., Bachmann, B.O., Gunsior, M., Townsend, C.A., and Solomon, E.I. 2001. Spectroscopic studies of substrate interactions with clavaminate synthase 2, a multifunctional α-KG-dependent non-heme iron enzyme: correlation with mechanisms and reactivities. *J Am Chem Soc* **123:**7388–7398.

COR-AAD1-00001342

Journal of Industrial Microbiology & Biotechnology (1999) 23, 336–340
© 1999 Society for Industrial Microbiology 1367-5435/99/$15.00
http://www.stockton-press.co.uk/jim

Sphingomonas herbicidovorans MH: a versatile phenoxyalkanoic acid herbicide degrader

HPE Kohler

EAWAG, Swiss Federal Institute for Environmental Sciences and Technology, Ueberlandstrasse 133, CH-8600 Dübendorf, Switzerland

Sphingomonas herbicidovorans **MH was isolated from a dichlorprop-degrading soil column. It is able to grow on phenoxyalkanoic acid herbicides, such as mecoprop, dichlorprop, 2,4-D, MCPA, and 2,4-DB. Strain MH utilizes both enantiomers of the chiral herbicides mecoprop and dichlorprop as sole carbon and energy sources. Enantiomer-specific uptake systems are responsible for transporting the acidic substrates across the cell membrane. Catabolism is initiated by two enantiomer-specific α-ketoglutarate-dependent dioxygenases that catalyze the cleavage of the ether bond of the respective enantiomer to yield the corresponding phenol and pyruvate. Therefore selective degradation of the enantiomers of mecoprop and dichlorprop by strain MH is not only due to enantioselective catabolism but also to enantioselective transport.**

Keywords: *Sphingomonas herbicidovorans* MH; phenoxyalkanoic acid herbicides; enantioselectivity; transport; α-ketoglutarate-dependent dioxygenase

Introduction

Sphingomonas herbicidovorans MH has been recently isolated and described [12,28]. Our studies of strain MH, which I summarize in this short review, are mainly driven by two reasons. First, strain MH degrades phenoxyalkanoic acid herbicides such as mecoprop, dichlorprop, 2,4-D, MCPA, and 2,4-DB (Figure 1, Table 1) [12,27,28]. The broad catabolic capabilities of strain MH with respect to the degradation of these herbicides, which are often found in leachates of landfills [5,10,13,29] and as groundwater contaminants [7], make this strain a potential candidate for bioremediation. Second, strain MH metabolizes the chiral phenoxypropanoic acid herbicides mecoprop and dichlorprop in an enantioselective manner. The effect of chirality on the metabolism of chiral pollutants, the importance of treating enantiomers separately, and the necessity of considering stereochemistry when studying degradation and environmental fate of chiral compounds have been emphasized in recent reviews [2,3,14,16]. Strain MH is an ideal system for studying the effect of chirality on the bacterial metabolism of chiral herbicides.

The chlorine-substituted phenoxyacetic acids 2,4-D, 2,4,5-T, and MCPA (Figure 1) were introduced as selective weed killers in the 1940s [18,26]. They are active against many broad-leaved weeds and are commonly used for the control of weeds in cereal crops, in grass pastures, and in lawns. The chiral 2-phenoxypropanoic acids mecoprop, dichlorprop, and fenoprop (Figure 1) were introduced in the 1950s for controlling certain weed species that were unaffected by the phenoxyacetic acid herbicides [18]. The herbicidal action of the phenoxyalkanoic acid herbicides is based on their ability to mimic natural plant growth regulators or auxins [18]. These synthetic auxins induce several abnormalities in growth and plant structure; they cause dedifferentiation and initiation of cell division in certain mature cells and inhibit cell division in primary meristems, processes that ultimately lead to cell death. As early as 1953, studies on the stereospecificity of plant growth regulators revealed that the herbicidal activity of mecoprop and dichlorprop is associated only with the *R* enantiomer [19]. The phenoxyalkanoic acid herbicides are among the most widely used herbicides to control weeds in cereal crops throughout the world [26]. In the European Union (EU), the estimated total usage of 2,4-D, MCPA, mecoprop, and dichlorprop exceeded 11 000 tons in 1991 [7].

Strain MH belongs to the genus *Sphingomonas* [23–25], which appears to be ubiquitous in soil, water, and sediments. *Sphingomonas* strains isolated from these environments have broad catabolic capabilities that make them potentially useful for bioremediation and waste treatment [4]. The unique ability of *S. herbicidovorans* MH to cleave the ether bond of various phenoxyalkanoic acid herbicides makes it an interesting and fitting addition to this genus.

Taxonomic characteristics of strain MH

Sphingomonas herbicidovorans MH was isolated by Horvath *et al* [12]. They repeatedly cultivated a sample from a dichlorprop-degrading soil column (nine transfers) and used the sample to inoculate a chemostat. After 3 weeks of incubation in continuous culture, they isolated a dichlorprop-degrading strain from the effluent of the chemostat. Strain MH consists of Gram-negative, catalase- and oxidase- positive, yellow-pigmented, non-motile, and non-spore-forming rods [12]. It does not ferment carbohydrates, does not accumulate poly-β-hydroxybutyric acid under nitrogen-deficient conditions, does not use nitrate as an electron acceptor, and does not utilize complex carbohydrates [12]. Horvath *et al* [12] also noticed that strain

Correspondence: HPE Kohler, Department of Microbiology, EAWAG, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland
Received 7 May 1999; accepted 11 August 1999

COR-AAD1-00001343

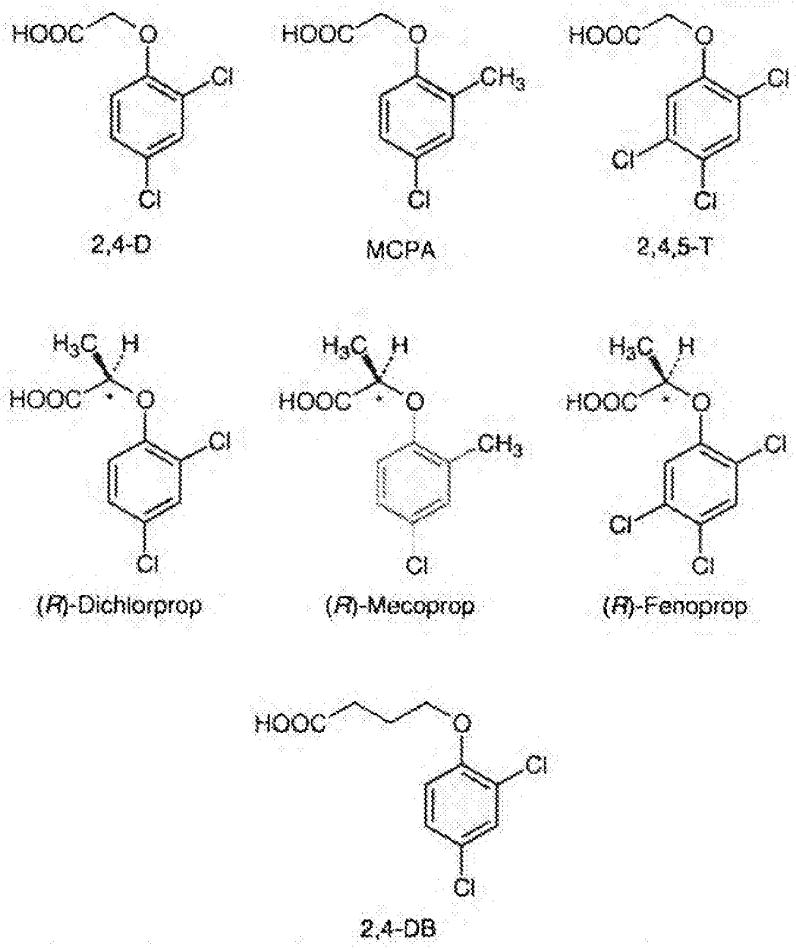


Figure 1 Structural formulas of the achiral phenoxyacetic acid herbicides 2,4-dichlorophenoxyacetic acid (2,4-D), 4-chloro-2-methylphenoxyacetic acid (MCPA), 2,4,5-trichlorophenoxyacetic acid (2,4,5-T), of the achiral phenoxybutanoic acid herbicide 4-(2,4-dichlorophenoxy)butanoic acid (2,4-DB), and of the chiral phenoxypropanoic acid herbicides (*R*)-2-(2,4-dichlorophenoxy)propanoic acid [(*R*)-dichlorprop], (*R*)-2-(4-chloro-2-methylphenoxy)propanoic acid [(*R*)-mecoprop], and (*R*)-2-(2,4,5-trichlorophenoxy)propanoic acid [(*R*)-fenoprop]. Asterisks mark the stereogenic centers.

MH releases a brownish water-soluble pigment into complex nutrient media. They tentatively identified strain MH as a *Flavobacterium* sp. Additional work on the taxonomic characteristics of strain MH by Zipper *et al* [28] and the Deutsche Stammsammlung von Mikroorganismen und Zellkulturen (DSM) [28] revealed that in addition to the characteristics mentioned above, strain MH is aminopeptidase-positive and alcohol dehydrogenase- and urease-negative, lyses only esculin but not gelatin or DNA, is not able to grow anaerobically, and does not form acids from sugars. Analysis of the cellular fatty acids showed that strain MH belongs to the *Sphingomonas* group within the α-subgroup of the proteobacteria [28]. The partial sequence of the gene for the 16S rRNA has the highest degree of similarity to the one for the 16S rRNA of *Rhizomonas suberifaciens* and *S. yanoikuyae*. DNA-DNA hybridization experiments between strain MH and the type strains of these two species showed that strain MH belongs to a new species within the group [28]. Although the highest identity was observed with *R. suberifaciens* (37.4%), the strain was named *S. herbicidovorans* because the genus name *Rhizomonas* is a *nomen rejiciendum* (opinion 14) and, therefore, usage of this name is in violation of rule 28a of the *International Code of Nomenclature of Bacteria* [17]. The name *Sphingomonas herbicidovorans* MH was validated by inclusion on validation list No. 61 of the *International Journal of Systematic Bacteriology* [1].

Growth characteristics

Growth at 30°C on nutrient agar plates produces circular, bright-yellow colonies after 3 days; when the plates are incubated longer, the colonies change to a flower-like shape. Cells secret a brownish pigment when growing on solid nutrient medium. In mineral salts medium with mecoprop as the sole carbon and energy source, strain MH requires supplements of vitamins and peptone for reproducible growth [28].

Table 1 lists growth substrates of *S. herbicidovorans* MH. Where available, growth rates, yields, and substrate consumption rates are included in the table. Strain MH is able to grow on the most important phenoxyalkanoic acid

COR-AAD1-00001344

Table 1 Growth characteristic of *S. herbicidovorans* MH

Growth substrate	Growth rate (h^{-1})	Substrate consumption rate ($\mu M\ h^{-1}$)	Yield (g dry weight g^{-1} substrate)	Reference
(*R*)-Dichlorprop	0.082 ± 0.006	28	0.37 ± 0.06	[27]
(*S*)-Dichlorprop	0.036 ± 0.001	12	0.35 ± 0.02	[27]
Racemic mixture of dichlorprop[a]	NA[b]	NA	0.29 ± 0.03	[27]
-(*R*)-Dichlorprop[a]	0.043 ± 0.003	22	NA	[27]
-(*S*)-Dichlorprop[a]	0.029 ± 0.003	13	NA	[27]
Racemic mixture of dichlorprop	0.2[c]	35	ND[d]	[12]
(*R*)-Mecoprop	0.032	14	0.30 ± 0.15	[28]
(*S*)-Mecoprop	0.039	13	0.35 ± 0.05	[28]
Racemic mixture of mecoprop[e]	0.016[f]	NA	0.30 ± 0.04	[28]
-(*R*)-Mecoprop[e]	NA	6	NA	[28]
-(*S*)-Mecoprop[e]	NA	5	NA	[28]
2,4-D	0.133 ± 0.003	54	0.19 ± 0.05	[27]
MCPA	+[g]	ND	ND	[28]
2,4-DB	+	ND	ND	[12]
2,4,5-T	–[h]	NA	NA	[12]

[a]Diauxic growth; a growth rate was calculated for each enantiomer.
[b]NA, not applicable.
[c]This growth rate is probably not correct. Recalculation of the data given in the reference showed that the growth rate must be between 0.065 and 0.115 per h.
[d]ND, not determined.
[e]No indication of diauxic growth; one growth rate was calculated.
[f]In this case the growth rate was affected by the concentration, which for each enantiomer was only half the value of what it was in the experiments with pure mecoprop enantiomers. Additional experiments showed increased growth rates in experiments with higher (*RS*)-mecoprop concentrations [28].
[g]+, Strain MH grew on the compound.
[h]–, Strain MH did not grow on the compound.

herbicides, such as mecoprop, dichlorprop, 2,4-D, MCPA, and 2,4-DB (Table 1, Figure 1).

S. herbicidovorans MH utilizes both enantiomers of mecoprop and dichlorprop as sole carbon and energy sources for growth [12,27,28]. However, growth experiments inoculated with cells that were grown on complex media and with the single pure enantiomers or the racemic mixtures as substrates revealed enantioselectivity; strain MH always degrades the *S* enantiomers before the *R* enantiomers. Growth on (*S*)-mecoprop is complete after about 3.5 days whether the *R* enantiomer is present or not. With racemic mecoprop as the growth substrate, degradation of the enantiomers is sequential, with (*S*)-mecoprop being degraded first [28]. However, the growth curve does not show indications of diauxic growth and, therefore, only one growth rate is reported in Table 1 for racemic mecoprop. (*R*)-Mecoprop is only slowly degraded during the first 50 h of such experiments, but its degradation rate markedly increases after the complete consumption of (*S*)-mecoprop. With pure (*R*)-mecoprop as the sole carbon and energy source, cells start to grow after a lag phase of about 7 days. Once growth commences, the cells grow as fast with (*S*)-mecoprop as the growth substrate (Table 1) [28]. Basically the same results were obtained with dichlorprop as the sole carbon and energy source [27]. As the growth substrate, (*S*)-dichlorprop completely disappears from the culture medium after 43 h of incubation, while in experiments with (*R*)-dichlorprop as the growth substrate, a lag period of about 60 h is evident. When strain MH is offered racemic dichlorprop as the growth substrate, it degrades both enantiomers; the *S* enantiomer is preferentially degraded but at a slower rate (Table 1) [27]. In contrast to growth on racemic mecoprop, growth on racemic dichlorprop is diauxic and, therefore, two growth rates are reported in Table 1. Generally, chloride release is stoichiometric and strain MH does not excrete any detectable metabolites [12,27,28]. The levels of dissolved organic carbon at the beginning and at the end of a growth experiment usually correspond well to a complete conversion of dichlorprop and mecoprop [28].

Additional growth experiments showed that strain MH grows best on 2,4-D [27] (Table 1). Typically, growth on 2,4-D starts without a lag period and is complete after about 15 h of incubation.

Enantioselective metabolism of mecoprop and dichlorprop

Biochemical investigations focused on enzyme activities that turn over the mecoprop enantiomers [20]. *S. herbicidovorans* MH was grown on pure enantiomers and the harvested cells were disrupted. The crude cell extracts were tested for the existence of enzyme activities that were able to convert either (*R*)- or (*S*)-mecoprop. Cell extracts of *Sphingomonas herbicidovorans* MH grown on (*R*)-mecoprop contain an enzyme activity that selectively converts (*R*)-mecoprop to 4-chloro-2-methylphenol, whereas the enzyme activity in cell extracts of (*S*)-mecoprop-grown cells is selective for the *S* enantiomer. Both reactions depend on α-ketoglutarate and ferrous ions. Besides 4-chloro-2-methylphenol, pyruvate and succinate are formed as products of the reactions [20]. Labeling experiments with $^{18}O_2$ revealed that both enzyme activities catalyze a dioxygenation reaction. The appearance of ^{18}O label in pyruvate and not in phenol suggests that the oxidative cleavage of the ether-bond proceeds through hydroxylation at the C-2 carbon of the propanoic acid side chain. The resulting unstable intermediate should then easily decompose to 4-chloro-2-methylphenol and pyruvate [20]. This suggestion was also in agreement with a general reaction mechanism for α-ketoglutarate-dependent dioxygenases. The reaction is thought to start with the formation of a ferryl oxidant ($Fe_{IV}=O$) by the binding of oxygen to the iron cofactor, a process that is linked to oxidative decarboxylation of α-ketoglutarate to succinate [6,11,21,22]. Therefore, Nickel *et al* [20] proposed a degradation scheme based on two distinct α-ketoglutarate-dependent dioxygenases for the initial transformation of (*R*)- and (*S*)-mecoprop to 4-chloro-2-methylphenol and pyruvate by *S. herbicidovorans* MH (Figure 2) [20]. Both enzyme activities seem to share the overall chemistry with 2,4-D dioxygenase (TfdA), the enzyme catalyzing the first step of the degradation of 2,4-D by *R. eutropha* JMP134 [8,9], but not the substrate specificity. Whereas TfdA preferentially transforms 2,4-D to 2,4-dichlorophenol and glyoxylate (Figure 2) [8,9], the enzyme activities present in cell extracts of *S. herbicidovorans* MH specifically turn over either the *R* or the *S* enantiomer of mecoprop and show

COR-AAD1-00001345

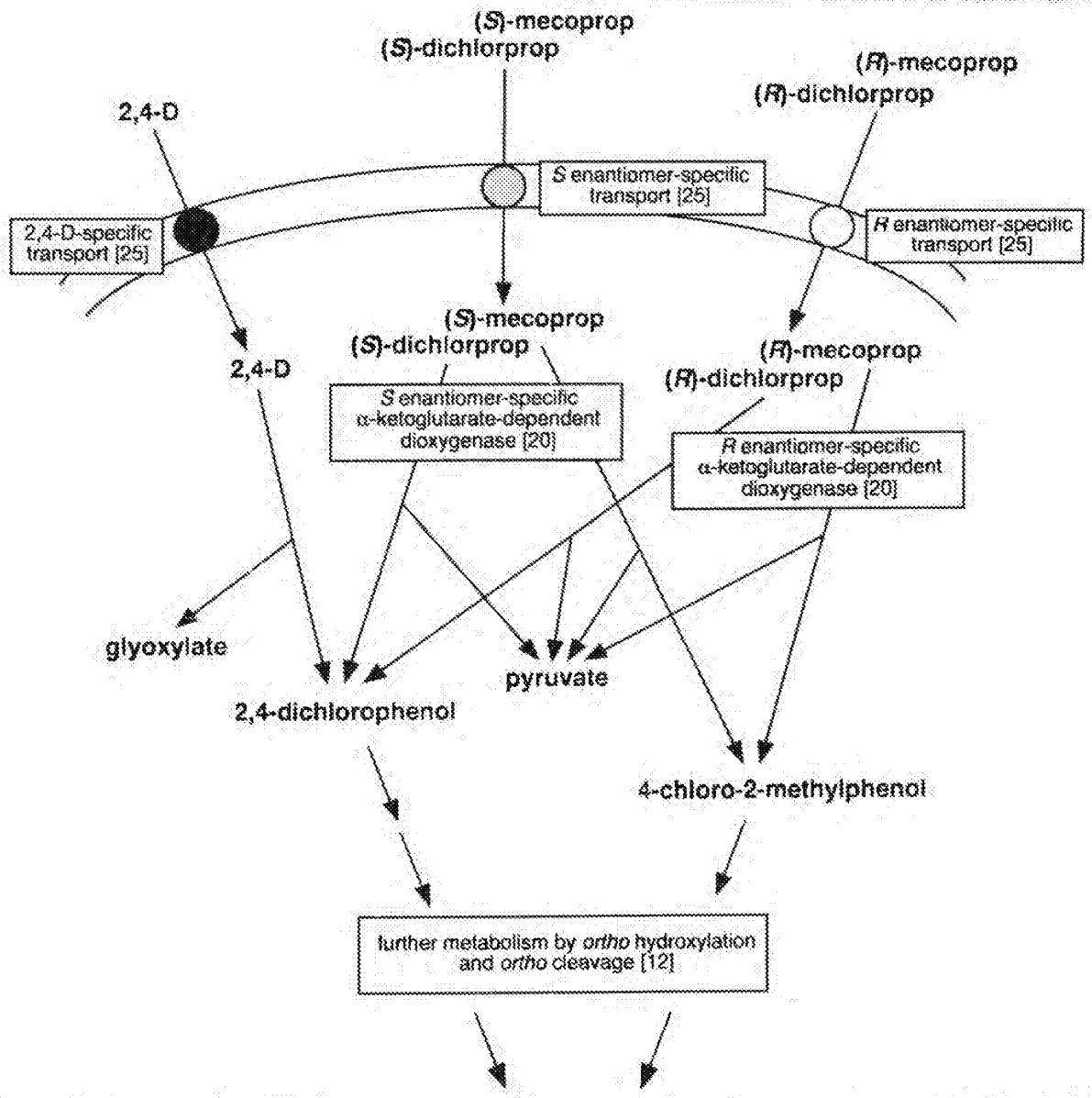


Figure 2 Model proposed for the uptake and metabolism of (*R*)- and (*S*)-dichlorprop, (*R*)- and (*S*)-mecoprop, and 2,4-D by *S. herbicidovorans* MH. Strain MH has three inducible proton gradient-driven uptake systems; one for (*R*)-dichlorprop and (*R*)-mecoprop, one for (*S*)-dichlorprop and (*S*)-mecoprop, and a third one for 2,4-D [27]. Once the compounds enter the cell, specific α-ketoglutarate-dependent dioxygenases enantioselectively catalyze cleavage of the ether bonds [20].

less activity with 2,4-D as the substrate [20]. (*R*)- and (*S*)-dichlorprop are also substrates of the (*R*)-enantiomer-specific enzyme and the (*S*)-enantiomer-specific enzyme, respectively; in these cases, 2,4-dichlorophenol and pyruvate were the common products (Figure 2) [15,20].

Enantioselective uptake of mecoprop and dichlorprop

Transport studies with ^{14}C-labeled dichlorprop enantiomers and with ^{14}C-labeled 2,4-D revealed that the uptake of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D by *S. herbicidovorans* MH is inducible [27]. The results from these transport studies clearly demonstrated that the uptake of dichlorprop is enantioselective and that there must be a separate uptake system for each enantiomer [27]. Additionally, Zipper *et al* [27] showed that the rates of uptake of 2,4-D by 2,4-D-grown cells in response to increasing 2,4-D concentrations are in accordance with saturation kinetics. The same is true for the transport of (*R*)-dichlorprop and the transport of (*S*)-dichlorprop into (*R*)-dichlorprop-grown and (*S*)-dichlorprop-grown cells, respectively. Experiments on the substrate specificity of the putative transport systems revealed that (*R*)-dichlorprop uptake is inhibited by (*R*)-

COR-AAD1-00001346

mecoprop but not by (*S*)-mecoprop, (*S*)-dichlorprop, or 2,4-D. On the other hand, (*S*)-dichlorprop transport is inhibited by (*S*)-mecoprop but not by (*R*)-dichlorprop, (*R*)-mecoprop, or 2,4-D [27]. From these results, Zipper *et al* [27] concluded that the (*R*)-dichlorprop carrier also transports (*R*)-mecoprop and that the (*S*)-dichlorprop carrier also transports (*S*)-mecoprop. Apparently, the transport systems are specific with respect to the substitution and the configuration at the stereogenic center (C-2 atom of the propanoic acid side chain) and rather unspecific with respect to the aromatic moiety of the phenoxypropanoic acids [27]. Moreover, transport of (*R*)-dichlorprop, (*S*)-dichlorprop, and 2,4-D is completely inhibited by various uncouplers and by the ionophore nigercin but is only marginally inhibited by the ionophore valinomycin and the ATPase inhibitor *N,N'*-dicyclohexylcarbidiimine [27]; these results are strong indications for the involvement of active transport. Zipper *et al* [27] concluded and summarized that the first step in the degradation of dichlorprop, mecoprop, and 2,4-D in *S. herbicidovorans* MH is active transport and that three inducible, proton-gradient driven uptake systems exist; one for (*R*)-dichlorprop and (*R*)-mecoprop, another for (*S*)-dichlorprop and (*S*)-mecoprop, and a third for 2,4-D (Figure 2).

Concluding remarks

The reaction sequence illustrated in Figure 2 demonstrates that selective degradation of the enantiomers of mecoprop and dichlorprop by the soil isolate *S. herbicidovorans* MH [28] is not only due to enantioselective catabolism [20] but also to enantioselective transport [27].

Taking into consideration the fact that the toxicity and the effect on biological targets might strongly differ for the enantiomers of a chiral pesticides or drug [16], study of enantioselective metabolism on the molecular level well deserves further emphasis.

References

1 Anonymous. 1997. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. List No. 61. Int J Syst Bacteriol 47: 601–602.

2 Ariëns EJ. 1989. Racemates—an impediment in the use of drugs and agrochemicals. In: Chiral Separations by HPLC (Krstulovic AM, ed), pp 31–80, Ellis Horwood, Chichester.

3 Armstrong DW, GL Reid, ML Hilton and C-D Chang. 1993. Relevance of enantiomeric separations in environmental science. Environ Pollut 79: 51–58.

4 Balkwill DL, GR Drake, RH Reeves, JK Fredrickson, DC White, DB Ringelberg, DP Chandler, MF Romine, DW Kennedy and CM Spadoni. 1997. Taxonomic study of aromatic-degrading bacteria from deep-terrestrial-subsurface sediments and description of *Sphingomonas aromaticivorans* sp nov, *Sphingomonas subterranea* sp nov, and *Sphingomonas stygia* sp nov. Int J Syst Bacteriol 47: 191–201.

5 Christensen TH, P Kjeldsen, HJ Albrechtsen, G Heron, PH Nielsen, PL Bjerg and PE Holm. 1994. Attenuation of landfill leachate pollutants in aquifers. Crit Rev Environ Sci Technol 24: 119–202.

6 Feig AL and SJ Lippard. 1994. Reactions of non-heme iron(II) centers with dioxygen in biology and chemistry. Chem Rev 94: 759–805.

7 Fielding M, D Barcelo, A Helweg, S Galassi, L Torstensson, P Van Zoonen, R Wolter and G Angeletti. 1991. Pesticides in ground and drinking water. Review 27. Commission of the European Communities, Brussels.

8 Fukumori F and RP Hausinger. 1993. *Alcaligenes eutrophus* JMP134 '2,4-dichlorophenoxyacetate monooxygenase' is an α-ketoglutarate-dependent dioxygenase. J Bacteriol 175: 2083–2086.

9 Fukumori F and RP Hausinger. 1993. Purification and characterization of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. J Biol Chem 268: 24311–24317.

10 Gintautas PA, SR Daniel and DL Macalady. 1992. Phenoxyalkanoic acid herbicides in municipal landfill leachates. Environ Sci Technol 26: 517–521.

11 Hanauske-Abel HM and V Günzler. 1982. A stereochemical concept for the catalytic mechanism of prolylhydroxylase. Applicability to classification and design of inhibitors. J Theor Biol 94: 421–455.

12 Horvath M, G Dinzelmüller, M Loidl and F Streichsbier. 1990. Isolation and characterization of a 2-(2,4-dichlorophenoxy)propionic acid-degrading bacterium. Appl Microbiol Biotechnol 33: 213–216.

13 Kjeldsen P. 1993. Groundwater pollution source characterization of an old landfill. J Hydrol 142: 349–371.

14 Kohler HPE, W Angst, W Giger, C Kanz, S Müller and MJF Suter. 1997. Environmental fate of chiral pollutants—the necessity of considering stereochemistry. Chimia 51: 947–951.

15 Kohler HPE, K Nickel, M Bunk and C Zipper. 1999. Microbial transformation of the chiral pollutants mecoprop and dichlorprop—the necessity of considering stereochemistry. In: Novel Approaches for Bioremediation of Organic Pollution (Fass R, Y Flashner and S Reuveny, eds), pp 13–20, Kluwer Academic/Plenum Press, New York.

16 Kohler HPE, K Nickel and C Zipper. 1999. Effect of chirality on the microbial degradation and environmental fate of chiral pollutants. In: Advances in Microbial Ecology. Vol 16 (Schink B, ed), Plenum Press, New York, in press.

17 Lapage SP, PHA Sneath, EF Lessel, VBD Skerman, HPR Seeliger and WA Clark (eds). 1992. International Code of Nomenclature of Bacteria (1990 Revision). American Society for Microbiology, Washington, DC.

18 Loos MA. 1975. Phenoxyalkanoic acids. In: Herbicides. Chemistry, Degradation and Mode of Action, 2nd edn (Kearney PC and DD Kaufman, eds), pp 1–128, Marcel Dekker, New York.

19 Matell M. 1953. Stereochemical studies on plant growth regulators. VII. Optically active α-(2-methyl-4-chlorophenoxy)propionic acid and α-(2,4-dichlorophenoxy)-n-butyric acid and their steric relations. Arkiv foer Kemi 6: 365–373.

20 Nickel K, MJF Suter and HPE Kohler. 1997. Involvement of two α-ketoglutarate-dependent dioxygenases in the enantioselective degradation of (*R*)- and (*S*)-mecoprop by *Sphingomonas herbicidovorans* MH. J Bacteriol 179: 6674–6679.

21 Prescott AG. 1993. A dilemma of dioxygenases (or where biochemistry and molecular biology fail to meet). J Exp Bot 44: 849–861.

22 Roach PL, IJ Clifton, V Fülöp, K Harlos, G Barton, J Hajdu, I Andersson, CJ Schofield and JE Baldwin. 1995. Crystal structure of isopenicillin *N* synthase is the first from a new structural family of enzymes. Nature 375: 700–704.

23 Takeuchi M, H Sawada, H Oyaizu and A Yokota. 1994. Phylogenetic evidence for *Sphingomonas* and *Rhizomonas* as nonphotosynthetic members of the alpha-4 subclass of the *Proteobacteria*. Int J Syst Bacteriol 44: 308–314.

24 van Bruggen AHC, KN Jochimsen, EM Steinberger, P Segers and M Gillis. 1993. Classification of *Rhizomonas suberifaciens*, unnamed *Rhizomonas* species, and *Sphingomonas* spp in rRNA superfamily IV. Int J Syst Bacteriol 43: 1–17.

25 White DC, SD Sutton and DR Rigelberg. 1996. The genus *Sphingomonas*: physiology and ecology. Curr Opin Biotech 7: 301–306.

26 Worthing CR and RJ Hance. 1991. The Pesticide Manual—A World Compendium, 9th edn. The British Crop Council, Farnham, UK.

27 Zipper C, M Bunk, AJB Zehnder and HPE Kohler. 1998. Enantioselective uptake and degradation of the chiral herbicide dichlorprop [(*RS*)-2-(2,4-dichlorophenoxy)propanoic acid] by *Sphingomonas herbicidovorans* MH. J Bacteriol 180: 3368–3374.

28 Zipper C, K Nickel, W Angst and HPE Kohler. 1996. Complete microbial degradation of both enantiomers of the chiral herbicide mecoprop [(*RS*)-2-(4-chloro-2-methylphenoxy)propionic acid] in an enantioselective manner by *Sphingomonas herbicidovorans* sp nov. Appl Environ Microbiol 62: 4318–4322.

29 Zipper C, MJF Suter, SB Haderlein, M Grubl and HPE Kohler. 1998. Changes in the enantiomeric ratio of (*R*)- to (*S*)-mecoprop indicate in situ biodegradation of this chiral herbicide in a polluted aquifer. Environ Sci Technol 32: 2070–2076.

COR-AAD1-00001347

Characterization of the First Enzyme in 2,4-Dichlorophenoxyacetic Acid Metabolism

Robert P. Hausinger[1,2] and Fumiyasu Fukumori[2,3]

[1]Departments of Microbiology and Biochemistry, and [2]Center for Microbial Ecology, Michigan State University, East Lansing, Michigan; [3]Research Development Corporation of Japan, Chiyoda-ku, Tokyo, Japan

This paper reviews the properties of the *Alcaligenes eutrophus* JMP134 *tfdA* gene product, the enzyme responsible for the first step in 2,4-dichlorophenoxyacetic acid (2,4-D) biodegradation. The gene was overexpressed in *Escherichia coli* and several of its enzymatic properties were characterized. Although this enzyme catalyzes a hydroxylation reaction, it is not a monooxygenase. Rather, TfdA is an Fe(II) and α-ketoglutarate-dependent dioxygenase that metabolizes the latter cosubstrate to succinate and carbon dioxide. A variety of other phenoxyacetates and α-ketoacids can be used by the enzyme, but the greatest catalytic efficiencies were found using 2,4-D and α-ketoglutarate. The enzyme possesses multiple essential histidine residues, whereas catalytically essential cysteine and lysine groups do not appear to be present. — Environ Health Perspect 103(Suppl 5):37–39 (1995)

Key words: 2,4-dichlorophenoxyacetic acid; 2,4-D; biodegradation; chloroaromatic compounds; α-ketoglutarate; dioxygenase

The biodegradation of 2,4-dichlorophenoxyacetic acid (2,4-D), a broadleaf herbicide, has been shown to involve side chain removal, hydroxylation of the resulting 2,4-dichlorophenol (2,4-DCP), *ortho* cleavage of 3,5-dichlorocatechol, conversion of 2,4-dichloro-*cis,cis*-muconate to succinate, and subsequent metabolism of this intermediate by the cells (Figure 1) as reviewed by Häggblom (*1*). The genes encoding the enzymes involved in these processes in *Alcaligenes eutrophus* JMP134 have been localized to the pJP4 plasmid, cloned, and sequenced (*2*). Whereas the sequences of the *tfdB* gene (encoding 2,4-DCP hydroxylase) and the *tfdC, tfdD,* and *tfdE* genes (responsible for dichlorocatechol degradation) exhibit similarities to sequences of genes involved in mineralization of nonchlorinated or monochlorinated analogues (*3*), the sequence of the *tfdA* gene (*4*) exhibits no obvious evolutionary relationship to other known sequences. Furthermore, until recently no biochemical characterization of the *tfdA* gene product (TfdA) had been reported. Here, we describe our efforts to characterize the first enzyme in the 2,4-D pathway. We have shown that the enzyme is not a 2,4-D monooxygenase, as commonly stated in the literature but is rather a ferrous and α-ketoglutarate (α-KG)-dependent dioxygenase.

A 3.1-kilobase pair (kbp) *Sac*I fragment of pJP4 containing the *tfdA* gene was subcloned into the *Sac*I site of pUC19, followed by elimination of a 1.5-kbp *Xba*I fragment to yield plasmid pUS311. This plasmid was transformed into *Escherichia coli* JM109 and the recombinant cells were shown to synthesize high levels of a peptide with relative molecular mass (M_r) 32,000. Despite the abundance of TfdA, the 2,4-D-degrading activity in cell extracts was very low (0.003 μmoles of 2,4-D converted to 2,4-DCP/min/mg protein) compared to the published rate of degradation in whole cells of *A. eutrophus* JMP134(pJP4) (0.105 μmole/min/mg) (*5*). The trace level of activity was abolished upon addition of chelators and restored upon addition of ferrous ion, consistent with a Fe(II) requirement for the enzyme. The presence or absence of reducing agents had no effect on activity, which was inconsistent with the behavior of a monooxygenase. Rather, we found (*6*) that the enzyme is a dioxygenase that requires α-KG as a cosubstrate and converts this compound to carbon dioxide and succinate, as illustrated in Figure 2.

The thermolabile enzyme (stable only up to 30° C) was purified to apparent homogeneity (specific activity of 16.9 μmoles substrate converted per minute per milligram of protein) by a simple two-step procedure (Table 1) and extensively characterized (*7*). The presence of protease inhibitors during early stages of TfdA purification enhances the stability of the enzyme by preventing conversion of the subunit to an inactive TfdA fragment of apparent M_r 27,000. Whereas N-terminal sequence analysis of the nondegraded subunit revealed the residues (S–V–V–A–N–) expected from DNA sequence analysis, the amino-terminal sequence of the proteolytic fragment of TfdA (F–K–Y–A–E–L–) was consistent with hydrolytic cleavage after arginine (residue #77). By using anti-TfdA IgG in Western blot analysis of various samples, the conversion was found to occur after cell disruption rather than during the cell cultivation period. The same methods were used to demonstrate that the M_r of 32,000 form of the protein was present in cell extracts of *A. eutrophus* JMP134. The native protein has an apparent M_r of 50,000 ± 2500, which is consistent with a homodimeric structure. The enzyme exhibits maximum activity at pH 6.5 to 7; however, it is stable over a pH range of 6.5 to 11.

Ferrous ion is absolutely required for activity and cannot be replaced by Co(II), Cu(II), Li(II), Mg(II), Mn(II), Ni(II), or Zn(II). As shown by a time-dependent decrease in enzyme activity, however, ferrous ion alone is unable to sustain enzyme catalysis over long time periods. The rate of activity loss was greatly reduced (although not completely eliminated) by inclusion of ascorbic acid in the assay.

This paper was presented at the Conference on Biodegradation: Its Role in Reducing Toxicity and Exposure to Environmental Contaminants held 26–28 April 1993 in Research Triangle Park, North Carolina. Manuscript updated: fall 1994; manuscript received: January 23, 1995; manuscript accepted: February 13, 1995.

This work was supported by the NSF-Science and Technology Center grant BIR 9120006 and the Michigan State University Agricultural Experiment Station.

Address correspondence to Dr. Robert P. Hausinger, Department of Microbiology, Michigan State University, East Lansing, MI 48824. Telephone (517) 353-9675. Fax (517) 353-8957.

COR-AAD1-00001348

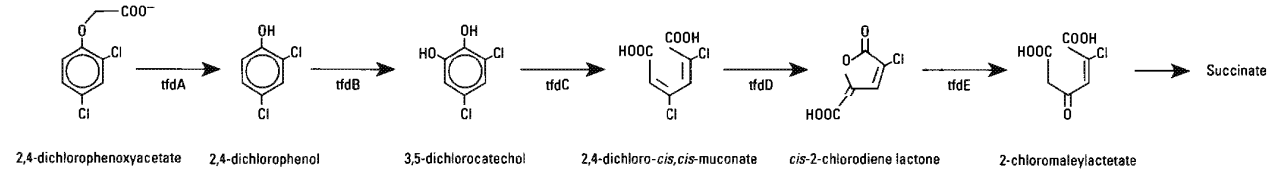


Figure 1. Pathway for degradation of 2,4-D in *Alcaligenes eutrophus* JMP134. Modified from Fukumori and Hausinger (*7*).

Catalytic turnover of the enzyme is not required for inactivation, as shown by the loss of activity during enzyme incubation with ferrous ion before addition of substrate. Although not completely characterized, the inactivation clearly results from a metal ion-mediated event as demonstrated by the retention of activity when the enzyme is stored in the absence of metal ions and the presence of EDTA. To minimize enzyme inactivation in kinetic studies, reactions were initiated by addition of enzyme to assay mixtures, and short assay periods were used to calculate initial rates.

Although capable of hydroxylating a wide range of phenoxyacetates and related compounds (Table 2), the enzyme exhibits the greatest affinity and highest catalytic efficiency for 2,4-D. Nonhalogenated phenoxyacetate possesses a larger K_m value than the halogenated substrates. Similarly, the K_m value for 2-phenoxypropionate is substantially larger than that of 2-(2,4-dichlorophenoxy)propionate. The additional methyl group in the side chain of these two compounds, however, greatly decreases their k_{cat} values for hydroxylation, perhaps due to the change from a secondary to a tertiary carbon atom. Although 3-phenoxypropionate is a very poor substrate, it is hydroxylated by the enzyme, demonstrating that the substrate binding site can accommodate one extra methylene carbon in the side chain. In contrast, TfdA exhibits no activity toward 4-(2,4-dichlorophenoxy)butyrate, 2-phenoxybenzoate, 2-phenoxyethanol, hydrocinnamic acid, indolylacetic acid, and methyl esters of phenoxyacetate and 4-chlorophenoxyacetate.

Although α-KG is the preferred cosubstrate for the enzyme, TfdA can use a range of other α-ketoacids with lower efficiency (Table 3). The non-α-ketoacid carboxyl group is not required for recognition by the enzyme; the K_m values for the two substrates possessing a second acidic group are significantly less, and the catalytic rates are generally higher compared to those for the other substrates. Addition of an extra methylene group between the α-ketoacid group and the free carboxyl group, as in α-ketoadipate, leads to small changes in the kinetic constants. In contrast, removal of one of the methylene groups, as in oxalacetate, led to an ineffective substrate. Furthermore, β-ketoglutarate, malonate, succinate, and glutarate were unable to support hydroxylation. Finally, in the absence of 2,4-D, no decomposition of α-KG was observed.

Chemical modification studies were used to provide evidence consistent with the absence of essential thiol or arginine residues and the presence of multiple essential histidine residues in the enzyme. Whereas iodoacetamide, *N*-ethylmaleimide, and butanedione failed to affect TfdA activity, the addition of diethylpyrocarbonate (DEP), a histidine-selective reagent, led to rapid pseudo-first-order loss of activity. The ability of 2,4-D, α-KG, Fe(II) plus ascorbate and combinations of these substances to protect the enzyme against DEP inactivation was examined. Whereas none of the individual compounds are able to significantly protect the enzyme from inactivation by DEP, the combinations of 2,4-D plus Fe(II) or α-KG with Fe(II) decrease the inactivation rate. Furthermore, the combined presence of 2,4-D and α-KG is very effective in protecting the enzyme from inactivation by DEP.

We interpret results from the above studies in terms of the model illustrated in Figure 3. Consistent with the expected requirements for positive charges at the binding sites of 2,4-D and α-KG, we propose that essential histidine residues are present at the binding sites for each of

Table 1. Purification of 2,4-D/α-ketoglutarate dioxygenase.[a]

Purification step	Total protein, mg	Total activity, U (%)	Activity, U/mg	Purification, -fold
Cell extracts	750	1690 (100)	2.2	1
$(NH_4)_2SO_4$ precipitation	225	1400 (83)	6.2	2.8
Mono Q	58	980 (58)	16.9	7.7

[a]Reproduced from Fukumori and Hausinger (*7*).

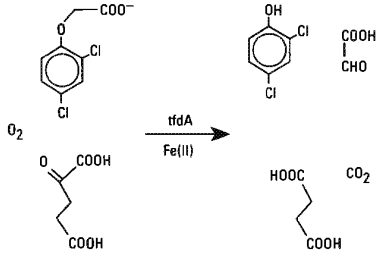


Figure 2. Overall reaction for 2,4-D/α-ketoglutarate dioxygenase.

Table 2. Phenoxyacetate substrate specificity and kinetic parameters for 2,4-D/α-ketoglutarate dioxygenase.[a]

Substrate	K_m, μM	k_{cat}, min^{-1}	k_{cat}/K_m, $min^{-1} \times mM^{-1}$
Phenoxyacetate	460	443	960
2-Chlorophenoxyacetate	110	380	3,450
4-Chlorophenoxyacetate	117	595	5,090
2,3-Dichlorophenoxyacetate	102	288	2,820
2,4-Dichlorophenoxyacetic acid	17.5	529	30,200
3,4-Dichlorophenoxyacetate	219	307	1,400
2,4,5-Trichlorophenoxyacetic acid	59.6	96	1,610
4-Chloro-2-methylphenoxyacetate	89.0	233	2,620
DL-2-Phenoxypropionate	1,170	5.1	4.4
2-(2,4-Dichlorophenoxy)propionate	191	61	320
3-Phenoxypropionate	12,900	3.2	0.25

[a]All experiments were performed at 30° C in 10 mM imidazole buffer (pH 6.75) containing 1 mM α-KG, 50 μM ascorbate, and 50 μM $(NH_4)_2Fe(SO_4)_2$. Modified from Fukumori and Hausinger (*7*).

COR-AAD1-00001349

Table 3. α-Ketoacid substrate specificity and kinetic parameters for 2,4-D/α-ketoglutarate dioxygenase.[a]

Substrate	K_m, μM	k_{cat}, min^{-1}	k_{cat}/K_m, $min^{-1} \times mM^{-1}$
α-KG	3.20	643	2.0×10^5
α-Ketoadipate	20.6	290	14,100
Pyruvate	1020	58	60
α-Ketobutyrate	464	89	190
α-Ketovalerate	607	404	660
α-Ketocaproate	583	158	270
α-Ketoisovalerate	745	16	20

[a] Enzyme reactions were performed at 30°C in 10 mM imidazole buffer (pH 6.75) containing 1 mM 2,4-D, 50 μM ascorbate, and 50 μM $(NH_4)_2Fe(SO_4)_2$. Modified from Fukumori and Hausinger (*7*).

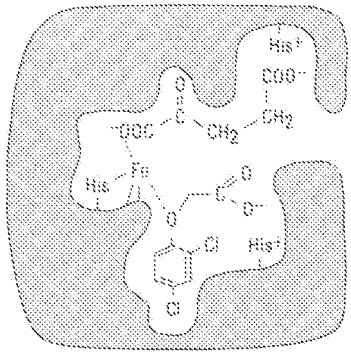


Figure 3. Model of the TfdA active site. Modified from Fukumori and Hausinger (*7*).

these substrates. In addition, we propose that one or more additional histidine residues may be buried in the protein at the Fe(II) binding site. Binding of 2,4-D and α-KG protects the histidine residues at the substrate binding sites and additionally may protect the Fe(II) ligands by steric constraints. Binding of either 2,4-D plus Fe(II) or both α-KG and Fe(II) might lead to the observed reduced rate of inactivation. In contrast, addition of any one compound alone is unable to protect the three or more distinct sites of inactivation postulated in this model.

The specific requirements for α-KG and Fe(II) in 2,4-D degradation and the stimulation of activity by ascorbate are typical characteristics of α-KG-dependent dioxygenases (*8*); however, the TfdA sequence exhibited no significant similarity to any of the known α-KG-dependent dioxygenase sequences (e.g., prolyl hydroxylase, lysyl hydroxylase, aspartyl hydroxylase, hyoscyamine hydroxylase, deacetoxycephalosporin hydroxylase, or the mechanistically related *p*-hydroxyphenylpyruvate hydroxylase) available in GenBank. We speculate that TfdA may have evolved from a gene involved in biodegradation of a plant-derived compound containing an aromatic ring in ether linkage to an acidic sidechain. One intriguing possibility is that a TfdA-like enzyme is involved in lignin degradation; i.e., lignin peroxidases and manganese peroxidases degrade the complex polyaromatic lignin substrate to smaller pieces, and those fragments that retain ether linkages may subsequently be degraded by an α-KG-dependent dioxygenase.

REFERENCES

1. Häggblom MM. Microbial breakdown of halogenated aromatic pesticides and related compounds. FEMS Microbiol Rev 103:29–72 (1992).
2. Perkins EJ, Gordon MP, Caceres O, Lurquin PF. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J Bacteriol 172: 2351–2359 (1990).
3. van der Meer JR, de Vos WM, Harayama S, Zehnder AJB. Molecular mechanisms of genetic adaption to xenobiotic compounds. Microbiol Rev 56:677–694 (1992).
4. Streber WR, Timmis KN, Zenk MH. Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J Bacteriol 169:2950–2955 (1987).
5. Pieper DH, Engesser K-H, Knackmuss H-J. Regulation of catabolic pathways of phenoxyacetic acids and phenol in *Alcaligenes eutrophus* JMP134. Arch Microbiol 151:365–371 (1989).
6. Fukumori F, Hausinger RP. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetate monooxygenase" is an α-ketoglutarate-dependent dioxygenase. J Bacteriol 175:2083–2086 (1993).
7. Fukumori F, Hausinger RP. Purification and characterization of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. J Biol Chem 268:24311–24317 (1993).
8. Abbott MT, Udenfriend S. α-Ketoglutarate-coupled dioxygenases. In: Molecular Mechanisms of Oxygen Activation (O Hayaishi, ed). New York:Academic Press, 1974;167–214.

COR-AAD1-00001350

Journal of Bacteriology

Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in Alcaligenes eutrophus JMP134(pJP4).

R H Don, A J Weightman, H J Knackmuss and K N Timmis

J. Bacteriol. 1985, 161(1):85.

Updated information and services can be found at:
http://jb.asm.org/content/161/1/85

These include:

CONTENT ALERTS

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to to another ASM Journal go to: http://journals.asm.org/site/subscriptions/

Journals.ASM.org

Downloaded from http://jb.asm.org/ on June 3, 2013 by guest

JOURNAL OF BACTERIOLOGY, Jan. 1985, p. 85–90
0021-9193/85/010085-06$02.00/0
Copyright © 1985, American Society for Microbiology
Vol. 161, No. 1

Transposon Mutagenesis and Cloning Analysis of the Pathways for Degradation of 2,4-Dichlorophenoxyacetic Acid and 3-Chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4)

R. H. DON,[1]* A. J. WEIGHTMAN,[1]† H.-J. KNACKMUSS,[2] AND K. N. TIMMIS[1]

Department of Medical Biochemistry, University of Geneva, Geneva, Switzerland,[1] and Lehrstuhl für Chemische Mikrobiologie, Universität-Gesamthochschule Wuppertal, Federal Republic of Germany[2]

Received 7 June 1984/Accepted 8 October 1984

Plasmid pJP4 permits its host bacterium, strain JMP134, to degrade and utilize as sole sources of carbon and energy 3-chlorobenzoate and 2,4-dichlorophenoxyacetic acid (R. H. Don and J. M. Pemberton, J. Bacteriol. 145:681–686, 1981). Mutagenesis of pJP4 by transposons Tn*5* and Tn*1771* enabled localization of five genes for enzymes involved in these catabolic pathways. Four of the genes, *tfdB*, *tfdC*, *tfdD*, and *tfdE*, encoded 2,4-dichlorophenol hydroxylase, dichlorocatechol 1,2-dioxygenase, chloromuconate cycloisomerase, and chlorodienelactone hydrolase, respectively. No function has been assigned to the fifth gene, *tfdF*, although it may encode a *trans*-chlorodiene-lactone isomerase. Inactivation of genes *tfdC*, *tfdD*, and *tfdE*, which encode the transformation of dichlorocatechol to chloromaleylacetic acid, prevented host strain JMP134 from degrading both 3-chlorobenzoate and 2,4-dichlorophenoxyacetic acid, which indicates that the pathways for these two substrates utilize common enzymes for the dissimilation of chlorocatechols. Studies with cloned catabolic genes from pJP4 indicated that whereas all essential steps in the degradation of 2,4-dichlorophenoxyacetic acid are plasmid encoded, the conversion of 3-chlorobenzoate to chlorocatechol is specified by chromosomal genes.

Chlorinated phenoxy herbicides have been used extensively in agriculture for over 40 years (10). Recently, several groups of workers have reported the isolation of bacteria belonging to different genera that are capable of degrading these compounds, especially 2,4-dichlorophenoxyacetic acid (2,4-D) and 2-methyl-4-chlorophenoxyacetic acid (1, 8), a finding which indicates rapid evolution and dissemination of these degradative properties in bacterial populations. In 1978, Fisher et al. (9) reported the involvement of plasmid-borne genes in the degradation of 2,4-D by a strain of *Alcaligenes paradoxus*. Since then, six other strains of soil bacteria have been isolated which contain plasmid-borne genes that encode the degradation of 2-methyl-4-chlorophenoxyacetic acid or 2,4-D or both (4). One such plasmid, pJP4, a 75-kilobase, broad-host-range, P1 incompatibility group plasmid, which was first isolated from *Alcaligenes eutrophus* strain JMP134 (4), carries genes essential for the degradation of 2,4-D, 2-methyl-4-chlorophenoxyacetic acid, and 3-chlorobenzoate (3CB) by strain JMP134. The ease of isolation and manipulation of pJP4 and the fact that it has been characterized physically (R. H. Don and J. M. Pemberton, J. Bacteriol., in press) make it a particularly attractive model for studies on the genetics of degradation of haloaromatics and for the construction in the laboratory of novel strains that are able to degrade recalcitrant haloaromatics (K. N. Timmis, P. R. Lehrbach, S. Harayama, R. H. Don, N. Mermod, S. Bas, R. Leppik, A. J. Weightman, W. Reineike, and H.-J. Knackmuss, *in* S. Kaplan, J. Konisky, and A. Hollaender, ed., *Plasmids in Bacteria*, in press). In this report, we describe the identification, localization, and cloning of genes which encode the catabolism of 2,4-D and 3CB by *A. eutrophus* JMP134.

* Corresponding author.
† Present address: Department of Applied Biology, University of Wales, Institute of Science and Technology, Cardiff, Wales.

MATERIALS AND METHODS

Strains. *A. eutrophus* JMP134 (4) contains plasmid pJP4 and degrades 2,4-D (2,4-D$^+$) and 3CB (3CB$^+$). Strain JMP221 is a pJP4-negative, cured derivative of strain JMP134 and is 2,4-D$^-$ 3CB$^-$ (4). Strain JMP289 (2,4-D$^-$ 3CB$^-$ Rif^r) (4) and strain JMP222 (2,4-D$^-$ 3CB$^-$ Sm^r) (4) are spontaneous mutants of strain JMP221 that are resistant to rifampin and streptomycin, respectively. Strain RD520 (2,4-D$^-$ 3CB$^-$ Sm^r Km^r) and strain RD521 (2,4-D$^-$ 3CB$^-$ SM^r Tc^r) are derivatives of strain JMP222 containing in their chromosomes transposon Tn*5* (specifies resistance to kanamycin) and transposon Tn*1771* (specifies resistance to tetracycline), respectively. Strain RD530 is a pJP4-containing 2,4-D$^+$ 3CB$^+$ derivative of strain RD520; strain RD531 is a pJP4-containing 2,4-D$^+$ 3CB$^+$ derivative of strain RD521. The *Escherichia coli* K-12 derivative used in these studies was strain LE392, *hsdR514 supE44 supF58 lacY1 galK2 galT22 metB1 trpR55* (16). All strains were cultivated at 30°C. Liquid cultures were incubated with aeration on an orbital shaker platform. L broth and antibiotic medium no. 3 (Difco Laboratories, Detroit, Mich.) agar were used as complete media, and minimal salts (11) containing 15 mM pyruvate, 1 mM 2,4-D (E. Merck AG, Darmstadt, Germany), or 2 mM 3CB (Merck) was used as the minimal medium. Antibiotics were incorporated into media at the following final concentrations: tetracycline, 10 μg/ml; kanamycin, 100 μg/ml; streptomycin, 1,000 μg/ml; rifampin, 50 μg/ml.

Transposon mutagenesis of pJP4. Two different transposons, Tn*5* (14) and Tn*1771* (18), were used to minimize potential problems associated with specificity of insertion sites. Plasmid pJP4 was transferred by conjugation from strains that were derived from strain JMP134 and carried either Tn*5* (strain RD520) or Tn*1771* (strain RD521) in their chromosomes to strain JMP298, a plasmid-free rifampin-resistant derivative of strain JMP134. Km^r transconjugants were selected and subsequently screened for the absence of pJP4-encoded degradative functions by replica plating on

Downloaded from http://jb.asm.org/ on June 3, 2013 by guest

COR-AAD1-00001352

minimal media containing either 2,4-D or 3CB as a carbon source.

Accumulation of metabolites by insertion mutant derivative bacteria. Mutant bacteria were grown for 18 h in minimal medium containing 20 mM pyruvate and 0.2 mM 2,4-D and were then removed by centrifugation. Coloration of the medium due to polymerization of accumulated chlorocatechol was minimized by adjusting the pH of the medium to 6.0. Metabolites accumulated by mutant bacteria were identified from spectra taken of undiluted media over the range from 220 to 350 nm at pH 2, 6, and 12. The references used were either spectra of pure 2,4-D, 2,4-dichlorophenol (Merck), and 3-chlorocatechol or literature values of absorption maxima for 3,5-dichlorocatechol, chloromuconate, and chlorodienelactone (8).

Enzyme assays. 2,4-D monooxygenase and 2,4-dichlorophenol hydroxylase were assayed by determining oxygen uptake by whole cells in a Rank oxygen electrode. Chlorocatechol 1,2-dioxygenase and 4-carboxymethylenebut-2-en-4-olide (chlorodienelactone) hydrolase were assayed in cell extracts by using previously described procedures (7, 17).

Induced cells were prepared for enzymes assays by inoculating 500 ml of minimal medium containing 1 mM pyruvate (pH 6.0) with 1 ml of an overnight L-broth culture and subsequent incubation for 6 h. 3CB was then added to a final concentration of 1 mM, and the cultures were incubated for an additional 14 h, followed by twofold dilution with minimal medium containing 15 mM pyruvate plus 1 mM 3CB. After incubation for 4 h, the cells were harvested by centrifugation at 5°C. For oxygen uptake assays, cells were washed three times in minimal medium and resuspended in the same medium at an appropriate cell density. For preparation of cell extracts, cells were washed in 20 mM Tris-hydrochloride (pH 8.0), resuspended in the same buffer, and disrupted by sonication with a Branson model B30 sonifier. Cell debris was sedimented by centrifugation at 20,000 × *g* for 90 min, after which the supernatant was removed and used for the assays. Similar conditions were used to prepare uninduced cells, except that 3CB was not added to the medium.

Plasmid isolation and restriction endonuclease cleavage analysis of insertion mutants. Plasmid DNA was isolated by the procedure of Hansen and Olsen (12) and was subsequently digested with restriction endonucleases under the conditions recommended by the manufacturers. DNA fragments produced in this way were analyzed by electrophoresis through horizontal 0.7% agarose gels. *Hin*dIII-generated fragments of bacteriophage lambda DNA served as molecular weight standards.

Transposon Tn*5* contains two *Hin*dIII cleavage sites that are 1,150 base pairs from each end and a single *Bam*HI site that is 2,600 base pairs from one end (14). These two enzymes were used to locate Tn*5* elements relative to the *Bam*HI and *Hin*dIII sites in pJP4::Tn*5* mutant derivatives. Transposon Tn*1771* contains two *Eco*RI sites less than 50 base pairs from each end, a central *Eco*RI site, and a single *Hin*dIII site 1,300 base pairs from one end (18). Therefore, *Hin*dIII and *Eco*RI were used to locate the Tn*1771* elements in pJP4::Tn*1771* mutant derivatives.

Cloning. Fragments of pJP4 were cloned into vector pKT231 by codigesting 2 to 5 μg of pJP4 DNA and 0.5 μg of pKT231 DNA, followed by phenol extraction, ethanol precipitation, and ligation in buffer (50 mM Tris-hydrochloride, pH 7.4, 10 mM $MgCl_2$, 1.5 mM dithiothreitol, 0.5 mM ATP) supplemented with 2 U of T4 DNA ligase. The ligation mixture was used to transform competent *E. coli* strain LE392 cells as previously described (15), and hybrid plasmids in transformant clones were analyzed by cleavage analysis of plasmid DNA isolated by the method of Holmes and Quigley (13).

TABLE 1. Phenotypic classes of pJP4 transposon mutant derivative plasmids

Mutant class	Growth on:		Metabolite accumulated
	2,4-D	3CB	
A	—	+	2,4-D
B	—	+	2,4-Dichlorophenol
C	—	—	3,5-Dichlorocatechol
D	—	—	Chloromuconate
E	—	—	Chlorolactone
F	+/—	+	ND[a]

[a] ND, Not determined.

2,4-D transport studies. Cells were induced as described above for enzyme assays, washed four times in minimal medium, and resuspended in minimal medium. To 0.9 ml of prewarmed (30°C) cells we added 0.1 ml of 10 μM 2,4-dichlorophenoxy [2-^{14}C]acetic acid (specific activity, 30 to 60 mCi/mmol; Amersham Corp., Arlington Heights, Ill.) and the suspension was incubated for 15, 30, or 60 s. 2,4-D uptake was terminated by filtering the cultures through membrane filters (0.45 μm; Millipore Corp., Bedford, Mass.) presoaked in minimal medium, and by immediate washing of the collected cells with 5 ml of fresh, nonradioactive minimal medium.

RESULTS

Transposon insertion mutants of pJP4. We isolated 17 kanamycin-resistant and 12 tetracycline-resistant strain JMP289 transconjugants containing Tn*5* or Tn*1771* transposon mutant derivatives of pJP4 that were unable to grow on 2,4-D. On the basis of their growth phenotypes and the metabolites which they accumulated when they were grown in minimal medium supplemented with pyruvate and 2,4-D, these transconjugants were divided into five classes (Table 1). The metabolites accumulated by mutants in classes B, C, D, and E corresponded to those described by Evans et al. (8) for the 2,4-D metabolic pathway in *Psudomonas* (Fig. 1). Mutants in class F did not accumulate any intermediate detectable by UV spectral analysis. Mutants in this class were 3CB$^+$, but growth on 2,4-D was severely retarded. We did not isolate any mutants that were unable to convert 2,4-D to 2,4-dichlorophenol.

Mutants that accumulated 2,4-dichlorophenol were capable of growing on 3CB. However, those mutants that accumulated chlorocatechol or lower metabolites were unable to grow on either substrate, which indicated a convergence of the pathways at chlorocatechol. We isolated no mutant plasmid derivatives that resulted in inactivation of the 3CB pathway prior to convergence with the 2,4-D pathway (i.e., derivatives having the phenotype 3CB$^-$ 2,4-D$^+$). To determine whether the initial steps in 3CB degradation are specified by the bacterial chromosome, strain JMP222, a cured derivative of strain JMP134, was grown in minimal medium containing pyruvate and 3CB. A comparison of the UV spectrum of the culture supernatant with that of a solution of 3-chlorocatechol indicated that plasmid-free strain JMP222 cells metabolized 3CB to 3-chlorocatechol.

Organization of the degradative genes. A physical map of the transposon insertions is shown in Fig. 2. As expected, in every case the mutations of each class were clustered.

Downloaded from http://jb.asm.org/ on June 3, 2013 by guest

COR-AAD1-00001353

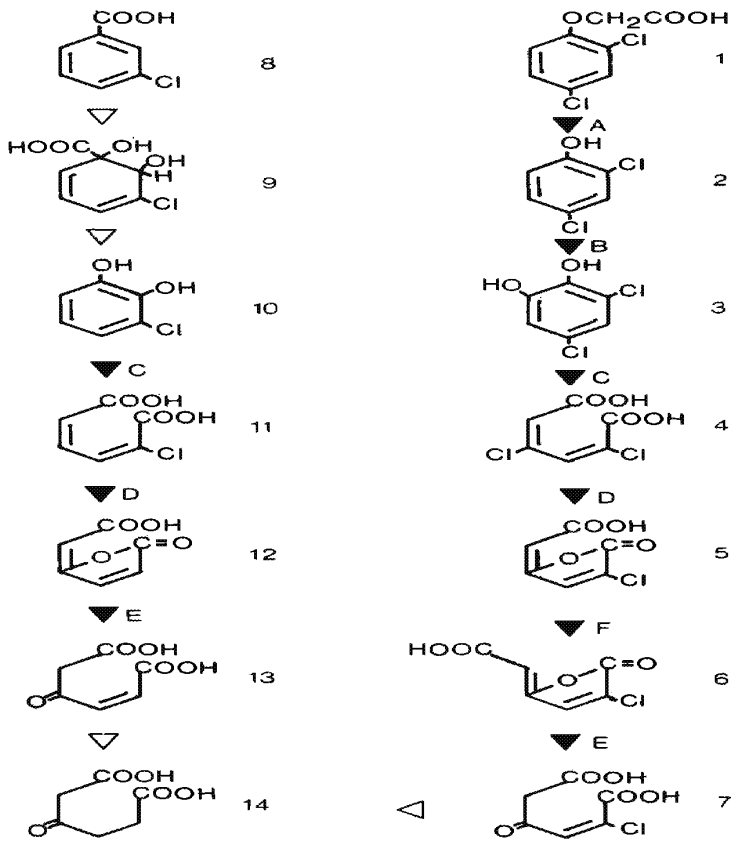


FIG. 1. Proposed pathways for the degradation of 3CB and 2,4-D in strain JMP134. These pathways are taken from references 3, 7, 8, and 17. The open arrowheads show pathway steps specified by chromosomal genes, and the solid arrowheads indicate plasmid-encoded steps. The enzymes involved are as follows: A, 2,4-D monooxygenase; B, 2,4-dichlorophenol hydroxylase; C, chlorocatechol-1,2-dioxygenase; D, chloromuconate cycloisomerase; E, 4-carboxymethylenebut-2-en-4-olide (diene-lactone) hydrolase; F, *trans*-chlorodiene-lactone isomerase. The following metabolites are involved: 1, 2,4-D; 2, 2,4-dichlorophenol; 3, 3,5-dichlorocatechol; 4, 2.4-dichloromuconate; 5, *trans*-2-chlorodiene-lactone; 6, *cis*-2-chlorodiene-lactone; 7, chloromaleylacetic acid; 8, 3CB; 9, 3-chlorodihydrodihydroxybenzoate; 10, 3-chlorocatechol; 11, 2-chloromuconate; 12, dienelactone; 13, maleylacetic acid; 14, 3-oxoadipate. Plasmid pJP4 genes B, C. D, and E code for enzymes B, C, D, and E, respectively. Although a gene function has not been assigned to gene F, it may encode a trans-chlorodiene-lactone isomerase (enzyme F).

In order to confirm the specific enzyme defects suggested by the mutant phenotypes, representative insertion mutants from each class were assayed for their uninduced and induced levels of 2,4-D monooxygenase, 2,4-dichlorophenol hydroxylase, chlorocatechol 1,2-dioxygenase, and chlorodienelactone hydrolase (Table 2). Athough both 3CB and 2,4-D induced expression of all 2,4-D pathway enzymes, 3CB was chosen as an inducer because accumulation of 2,4-dichlorophenol by mutants in class B proved to be lethal for the cells.

With the exception of strains RD324 and RD341, all of the mutants were defective for only one of the enzymes assayed. With all of the mutants except strain RD318, the enzyme inactivated was the enzyme required to metabolize the product which accumulated when the strain was cultured in the presence of 2,4-D; i.e., mutants belonging to class B (strains RD405 and RD406), class C (strains RD312 and RD313), and class E (strains RD330 and RD331) were inactivated in 2,4-dichlorophenol hydroxylase, chlorocatechol 1,2-dioxygenase, and chlorodienelactone hydrolase, respectively.

It was not possible to assay chloromuconate cycloisomerase due to unavailability of the substrate. However, strain RD324, which belongs to class D, accumulated chloromuconate and had wild-type activities for all of the enzymes assayed. Therefore, we assumed that *tfdD* encodes this enzyme. Strain RD318 also accumulated chloromuconate, but in addition no catechol 1,2-dioxygenase activity was found in cell extracts. The Tn*5* insertion in this strain mapped close to *tfdC*, which specifies catechol 1,2-dioxygenase, and it is possible that a transposon-induced rearrangement altered the terminal sequence of this gene without changing the enzyme activity of its product in vivo, but rendered it unstable in cell extracts. Therefore, this mutant was assigned to class D or gene *tfdD*.

Strain RD341 was a class F mutant and grew poorly on 2,4-D but well on 3CB. This mutant showed wild-type activities for all of the enzymes assayed. We have not yet been able to assign a function to the gene defined by type F mutations, but a role in 2,4-D transport has been ruled out; the rates of uptake of ^{14}C-labeled 2,4-D by induced mutants belonging to this class were comparable to those of the induced wild-type strain (Fig. 3).

Cloning and expression of genes of the *ortho* cleavage pathway. DNA fragments that carry genes encoding enzymes of the 2,4-D pathway were identified by digesting Tn*5* mutant plasmids with a variety of restriction endonucleases. We found that all insertions mapped in 14-kilobase *Eco*RI fragment B (Fig. 2, and insertions in genes *tfdC*, *tfdD*, *tfdE*, and *tfdF* were located in an 8-kilobase *Sst*I fragment. Both of these fragments were cloned in *E. coli* LE392 by using broad-host-range cloning vector pKT231. Hybrid plasmids were subsequently transferred to strain JMP298, a cured derivative of strain JMP134, by RP4-mediated mobilization. Transconjugants were selected on minimal medium containing rifampin and either 2,4-D or 3CB. The presence of cloned *Eco*RI fragment B in strain JMP298 enabled it to grow on both 2,4-D and 3CB, whereas the presence of the cloned *Sst*I fragment permitted growth only on 3CB. Both types of hybrid plasmid were unstable in strain JMP298, and the degradation phenotype was lost at a rate of 10^{-2} per generation in nonselective media.

DISCUSSION

By using transposon mutagenesis and DNA cloning techniques, we established the plasmid locations of genes encod-

Downloaded from http://jb.asm.org/ on June 3, 2013 by guest

COR-AAD1-00001354

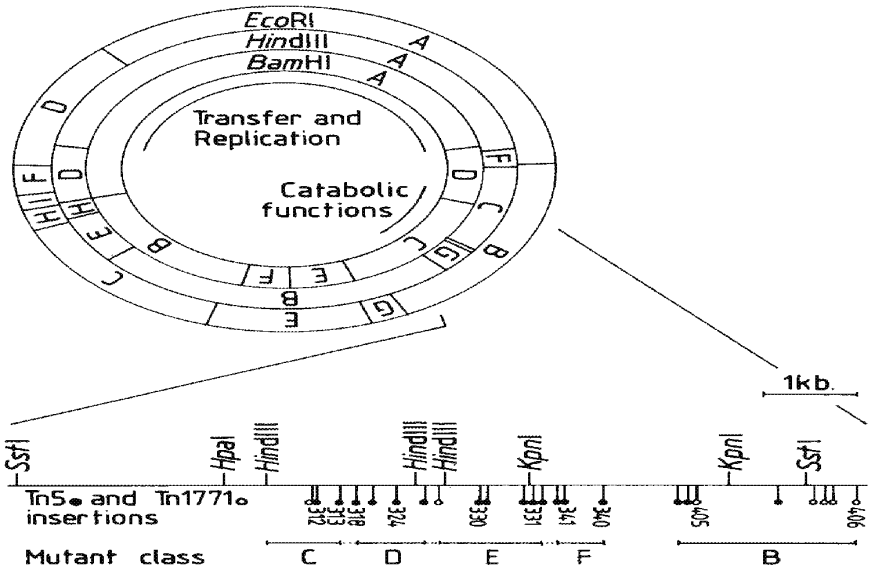


FIG. 2. Functional map of pJP4. The locations of *Eco*RI, *Hin*dIII, and *Bam*HI restriction sites are from a study reported elsewhere (Don and Pemberton, in press). The expanded map shows the insertion sites of Tn*5* and Tn*1771* transposons which inactivated the 2,4-D degradative pathway. The mutants were divided into five classes (classes B through F) on the basis of phenotype and metabolite accumulation (Table 2). The mutants used for enzyme studies (Table 3) are numbered. kb, Kilobase.

TABLE 2. Enzyme activities of *A. eutrophus* cells carrying mutant plasmids

Strain	Mutant class	Inducer[a]	Relative activities[b]			
			Mon	Hlase	Catase	Hase
JMP134	Wild type	−	0	0	0	0
		+	100	100	100	100
RD405	B	−	0	0	0	0
		+	76	0	72	46
RD406	B	−	0	0	0	0
		+	74	0	69	46
RD312	C	−	0	0	0	190
		+	105	95	0	200
RD313	C	−	0	0	0	210
		+	97	98	0	200
RD318	D	−	0	0	0	160
		+	93	67	0	160
RD324	D	−	0	0	0	190
		+	99	75	90	130
RD330	E	−	0	0	0	0
		+	94	90	90	0
RD331	E	−	0	0	0	0
		+	99	93	76	0
RD341	F	−	0	0	0	0
		+	84	81	84	90

[a] The inducer used was 2 mM 3CB.
[b] Expressed as percentages of the activities obtained with strain JMP134. Mon, Monooxygenase; Hlase, hydroxylase; Catase, catechol 1,2-dioxygenase; Hase, hydrolase.

ing five steps in the 2,4-D degradation pathway. Four of these genes were identified as encoding the second, third, fourth, and fifth steps in 2,4-D degradation according to the pathway proposed by Evans and co-workers (8).

The inability to isolate mutants in the first step of 2,4-D degradation, 2,4-D monooxygenase, could be due to either specificity of insertion of the transposons, resulting in lack of insertions in a plasmid-borne gene, or the fact that the gene is chromosomally encoded. The first possibility appears more likely because cured derivatives of strain JMP134 are

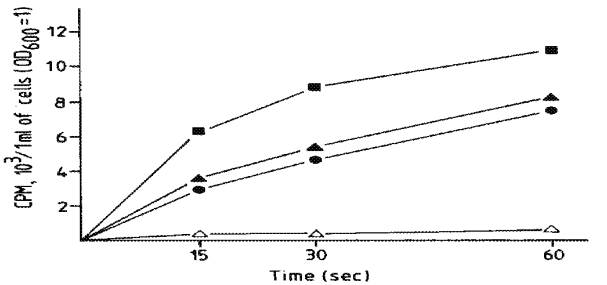


FIG. 3. Uptake of ^{14}C-labeled 2,4-D by induced (▲) and uninduced (△) cells of strain JMP134 and cells of induced mutant strains RD340 (■) and RD341 (●). OD_{600}, Optical density at 600 nm.

Downloaded from http://jb.asm.org/ on June 3, 2013 by guest

COR-AAD1-00001355

not able to effect this step (Don and Pemberton, in press). Also, 2,4-D is the inducer of the 2,4-D monoxygenase (D. Pieper, unpublished data). When strain JMP134 is grown in a chemostat with 2,4-D and 2,4-dichlorophenol, the 2,4-D cannot be completely replaced by 2,4-dichlorophenol or the culture is washed out. However, a constitutive mutant of strain JMP134 gives a stable culture in the absence of 2,4-D. This eliminates the possibility that the 2,4-D monooxygenase is chromosomally encoded and induced by a metabolite produced by a plasmid-encoded enzyme.

Interestingly, all of the insertions in *tfdC* and *tfdD* led to constitutive expression of *tfdE*. The expression of *tfdB*, which is located downstream from *tfdF*, was not influenced by mutations in other genes. These results suggest (i) that, unlike the situation in *E. coli* (2), in *A. eutrophus* Tn*5* insertions are not polar and result in constitutive expression of genes downstream from the site of insertion and (ii) that genes encoding degradation of 2.4-D are organized into two operons, one encoding metabolism of 3-chlorocatechol to chloromaleylacetic acid (genes C through E; i.e., those genes also essential for 3CB degradation) and another encoding 2,4-dichlorophenol hydroxylase (gene B; and perhaps also gene A, which has not been localized).

Transfer to strain JMP222 of a cloned fragment of pJP4 containing genes encoding the first three enzymes of chlorocatechol degradation enabled complete metabolism of 3CB. These enzymes, which catalyze conversion of chlorocatechol to maleylacetic acid, have previously been shown to be critical steps in chloroaromatic degradation (6). An additional enzyme, maleylacetic acid reductase, is necessary for transformation of maleylacetic acid to 3-oxoadipate before the chlorobenzoate pathway can converge with the pathway for the degradation of unsubstituted benzoates. Our failure to identify a fourth plasmid gene necessary for the catabolism of 3CB suggests that this enzyme is specified by a chromosomal gene.

Inactivation of a sixth gene, *tfdF*, adjacent to the gene encoding lactone hydrolase, severely retarded growth of strain JMP134 on 2,4-D, but did not affect its growth on 3CB. A function has not been assigned to this gene yet, although it has been shown not to be involved in 2,4-D uptake. Gene *tfdF* does not have any significant homology with the first three genes of the chlorocatechol pathway (Don, unpublished data), which argues against a possible duplication of a gene encoding an enzyme specific for a metabolite of a monosubstituted chlorocatechol during evolution of an enzyme specific for the metabolites of dichlorocatechol degradation.

Recent observations (U. Schwien, personal communication) have suggested that *tfd*F may be a *trans*-chlorodiene-lactone isomerase. The only product produced by the action of dichloromuconate cycloisomerase on 2.4-dichloromuconate is *trans*-chlorodiene-lactone, and the lactone hydrolase of strain JMP134 displays no activity toward this compound. An isomerase activity which is clearly distinguished from hydrolase activity by chromotographic techniques has been shown to transform *trans*- to *cis*-chlorodiene-lactone, which is a good substrate for the lactone hydrolase. As spontaneous *trans*-to-*cis* isomerization is rather slow, inactivation of an isomerase would severely retard growth on 2,4-D without affecting utilization of 3CB.

The cometabolism of chloroaromatic compounds to chlorocatechols has been observed frequently with a variety of bacteria, but the complete degradation of these compounds can only be accomplished by a much smaller number of organisms (6, 19). A key feature of the mineralization of chloroaromatics by such organisms is their possession of an efficient *ortho* cleavage pathway for chlorocatechols. By characterizing and cloning the genes specifying the critical steps in chloro- and dichlorocatechol degradation, we have laid the groundwork for constructing bacteria capable of degrading a variety of xenobiotic chloroaromatic compounds.

ACKNOWLEDGMENTS

We thank our catabolic colleagues in Geneva, Switzerland, and Göttingen, Federal Republic of Germany, for stimulating discussions.

This work was supported by the Bundesministerium für Forschung und Technologie, Bonn, Federal Republic of Germany. A.J.W. gratefully acknowledges receipt of a postdoctoral fellowship from the Royal Society of Great Britain.

LITERATURE CITED

1. **Alexander, M.** 1969. Microbial degradation and biological effects of pesticides in the soil. Soil Biol. Rev. Res. Nat. Resources Res. UNESCO **9**:209–240.
2. **Berg, D. E., A. Weiss, and L. Crossland.** 1980. Polarity of Tn*5* insertion mutations in *Escherichia coli*. J. Bacteriol. **142**:439–446.
3. **Chapman, P. S.** 1976. Microbial degradation of halogenated compounds. Biochem. Soc. Trans. **4**:463–466.
4. **Don, R. H., and J. M. Pemberton.** 1981. Properties of six pesticide degradation plasmids isolated from *Alcaligenes eutropus* and *Alcaligenes paradoxus*. J. Bacteriol. **145**:681–686.
5. **Dorn, E., M. Hellwig, W. Reineke, and H.-J. Knackmuss.** 1974. Isolation and characterization of a 3-chlorobenzoate degrading *Pseudomonas*. Arch. Microbiol. **99**:61–70.
6. **Dorn, E., and H.-J. Knackmuss.** 1978. Chemical structure and biodegradability of halogenated aromatic compounds. Substituent effects of 1,2-dioxygenation of catechol. Biochem. J. **174**:85–94.
7. **Dorn, E., and H.-J. Knackmuss.** 1978. Chemical structure and biodegradability of halogenated aromatic compounds. Two catechol 1,2-dioxygenase from a 3-chlorobenzoate grown pseudomonad. Biochem. J. **174**:73–84.
8. **Evans, W. C., B. S. W. Smith, H. N. Zernley, and J. I. Davies.** 1971. Bacterial metabolism of 2,4-dichlorophenoxyacetate. Biochem. J. **122**:543–551.
9. **Fisher, P. R., J. Appleton, and J. M. Pemberton.** 1978. Isolation and characterization of the pesticide-degrading plasmid, pJP1, from *Alcaligenes paradoxus*. J. Bacteriol. **135**:798–804.
10. **Fletcher, W. W.** 1978. The pest war. Basil Blackwell, Oxford.
11. **Gilardi, G. L.** 1971. Characterization of nonfermentative, nonfastidious gram-negative bacteria encountered in medical bacteriology. J. Appl. Bacteriol. **34**:623–644.
12. **Hansen, J. B., and R. H. Olsen.** 1978. Isolation of large bacterial plasmids and characterization of P-2 incompatibility group plasmids pMG1 and pMG5. J. Bacteriol. **135**:227–238.
13. **Holmes, D. S., and M. Quigley.** 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114**:193.
14. **Jorgensen, R. A., S. J. Rothotein, and W. S. Reznikoff.** 1979. A restriction enzyme cleavage map of Tn*5* and location of a region encoding neomycin resistance. Mol. Gen. Genet. **177**:65–72.
15. **Mandel, M., and A. Higa.** 1970. Calcium dependent bacteriophage DNA infection. J. Mol. Biol. **53**:159–162.
16. **Murray, N. E., W. J. Brammer, and K. Murray.** 1977. Lamboid phages that simplify the recovery of *in vitro* recombinants. Mol. Gen. Genet. **150**:53–61.
17. **Schmidt, E., and H.-J. Knackmuss.** 1980. Chemical structure and biodegradability of halogenated aromatic compounds. Conversion of chlorianted muconic acids into maleylacetic acid. Biochem. J. **192**:339–347.

Downloaded from http://jb.asm.org/ on June 3, 2013 by guest

COR-AAD1-00001356

18. **Schöffl, F., and A. Pühler.** 1979. The plasmid system of *Escherichia coli* strain UR12644: identification and molecular characteristics of transposons involved in the generation of endonegous R plasmids. Genet. Res. **24**:287–301.

19. **Wigmore, G. J., and D. W. Ribbons.** 1980. *p*-Cymene pathway in *Pseudomonas putida*: selective enrichment of defective mutants by using halogenated substrate analog. J. Bacteriol. **143**:816–824.

Downloaded from http://jb.asm.org/ on June 3, 2013 by guest

COR-AAD1-00001357

Journal of Bacteriology

Genetic and physical map of the 2,4-dichlorophenoxyacetic acid-degradative plasmid pJP4.

R H Don and J M Pemberton
J. Bacteriol. 1985, 161(1):466.

Updated information and services can be found at:
http://jb.asm.org/content/161/1/466

These include:

CONTENT ALERTS

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to to another ASM Journal go to: http://journals.asm.org/site/subscriptions/

Journals.ASM.org

Downloaded from http://jb.asm.org/ on June 5, 2013 by guest

COR-AAD1-00001358

JOURNAL OF BACTERIOLOGY, Jan. 1985, p. 466–468
0021-9193/85/010466-03$02.00/0
Copyright © 1985, American Society for Microbiology
Vol. 161, No. 1

Genetic and Physical Map of the 2,4-Dichlorophenoxyacetic Acid-Degradative Plasmid pJP4

R. H. DON[1] AND J. M. PEMBERTON[2]*

Department de Biochimie Medicale, Centre Medical Universitaire, Universite de Geneve, Switzerland,[1] and the Department of Microbiology, University of Queensland, St. Lucia, Australia 4067[2]

Received 21 July 1983/Accepted 8 February 1984

Plasmid pJP4 is an 80-kilobase, IncP1, broad-host-range conjugative plasmid of *Alcaligenes eutrophus* encoding resistance to mercuric chloride and phenyl mercury acetate and degradation of 2,4-dichlorophenoxyacetic acid, 2-methyl-4-chlorophenoxyacetic acid, and 3-chlorobenzoate. By the use of cloning, transposon mutagenesis, and restriction endonuclease analysis, a biophysical and genetic map of pJP4 was generated.

Synthetically produced chemicals such as insecticides, fungicides, herbicides, and other pesticides have become an integral part of modern agriculture. Although complex in nature, many pesticides, such as the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D), are readily degraded by microorganisms that are present in soil and water (7). The first indication that plasmids were involved in the environmental degradation of pesticides came from the report by Pemberton and Fisher (10) that a strain of *Alcaligenes paradoxus*, a ubiquitous saprophytic soil bacterium, possessed a transmissible IncP1 plasmid that conferred the ability to degrade 2,4-D and 2-methyl-4-chlorophenoxyacetic acid. Seven such plasmids have now been isolated in this laboratory from strains of either *A. paradoxus* or *Alcaligenes eutrophus* (4, 8, 9). In view of the significance of these plasmids in the recycling of synthetic phenoxyherbicides, one plasmid, pJP4, was chosen for more detailed analysis (Table 1). Plasmid pJP4 has a broad host range, is self transmissible, belongs to the IncP1 group, confers resistance to mercuric ions and phenyl mercury acetate (PMA), and encodes the degradation of 2,4-D, 2-methyl-4-chlorophenoxyacetic acid, and 3-chlorobenzoate (3CB). By using transposon mutagenesis, deletion analysis, gene cloning, and restriction analyses, a genetic and biophysical map of pJP4 was constructed (Fig. 1). Details of the compilation of this map have been given elsewhere (R. G. Don, Ph.D. thesis, University of Queensland, Brisbane, Australia, 1983).

Transposon mutagenesis of both chromosomal and plasmid genes has contributed much to the location and cloning of a variety of procaryotic genes (2). Insertion mutations of pJP4 were obtained from matings between *A. eutrophus* JMP222(pJP4) carrying a genomic insertion of either Tn*1771* (tetracycline resistance) or Tn*5* (kanamycin resistance) and a streptomycin-resistant derivative of JMP222; selection for insertion mutants was made on peptone-yeast extract agar containing 100 μg of streptomycin sulfate per ml with either 10 μg of tetracycline per ml (for Tn*1771* inserts) or 5 μg of kanamycin per ml (for Tn*5* inserts); intrastrain transfer of these transposons occurred at frequencies of 10^{-6} to 10^{-7} per donor cell. Among 500 Tn*1771* insertion mutants of pJP4, none were found to have inactivated mercury or to have resistance, five mutants were unable to degrade 2,4-D, and five mutants were unable to degrade 2,4-D or 3CB. The location and phenotypes of these 10 Tn*1771* insertion mutations are given in Fig. 2. Of the five 2,4-D$^-$ 3CB$^+$ mutations, four mapped at 33.4 kilobases (kb) and the other mapped at 34.0 kb on the pJP4 map. Of the five 2,4-D$^-$ 3CB$^-$ mutations, two mapped at 35.8 kb, two mapped at 38.0 kb, and one mapped at 43.9 kb.

Of 400 Tn*5* insertion mutations of pJP4 that were screened, none inactivated the ability to degrade either 2,4-D or 3CB, whereas 10 mutations resulted in the concomitant loss of resistance to mercuric ions and PMA. Analysis of two of these mutations showed that one mapped at 70.5 kb and the other mapped at 71.7 kb. The *Hin*dIII fragment carrying both these sites, when cloned into pBR322, conferred resistance to mercuric ions and PMA on *Escherichia coli*; this result confirms the location of these resistance genes.

Two Tn*1771* insertion mutants of pJP4 were found to segregate Tn*1771*-encoded tetracycline resistance when grown on nonselective media. One tetracycline-sensitive clone from each segregant was screened for degradation of 2,4-D and 3CB, resistance to mercuric chloride and PMA, and conjugative ability. Both strains were resistant to mercuric chloride and PMA and could transfer this ability by conjugation, but they were unable to degrade either 2,4-D or 3CB. Restriction analyses revealed that both plasmids had suffered apparently identical 40-kb deletions covering the region from 25.6 to 65.7 kb.

The two Tn*1771*-generated deletions of pJP4 were identical and resulted from the insertion of Tn*1771* in an identical site on each plasmid. However, as they occurred after the insertion mutants were separately purified, they could not be regarded as siblings. Specific termini for transposon-generated deletions have been identified in other plasmids (2), and data presented here suggest that such termini may exist on pJP4.

The remaining 40-kb segment was shown to carry the conjugative, incompatibility, replication, maintenance, and mercuric ion resistance functions of pJP4 (Fig. 1). On the basis of a random distribution of restriction sites, this region has fewer restriction sites than would be expected. A similar paucity of restriction sites in regions encoding these functions has been observed in other wide-host-range plasmids belonging to the IncP1, IncW, and IncQ groups (1, 11, 12). It has been postulated that as these plasmids can transfer to a wide variety of bacterial genera with different restriction systems, a selection would be present for mutations, resulting in the loss of restriction sites (11). A greater number of restriction sites in regions not encoding maintenance and

* Corresponding author.

Downloaded from http://jb.asm.org/ on June 5, 2013 by guest

COR-AAD1-00001359

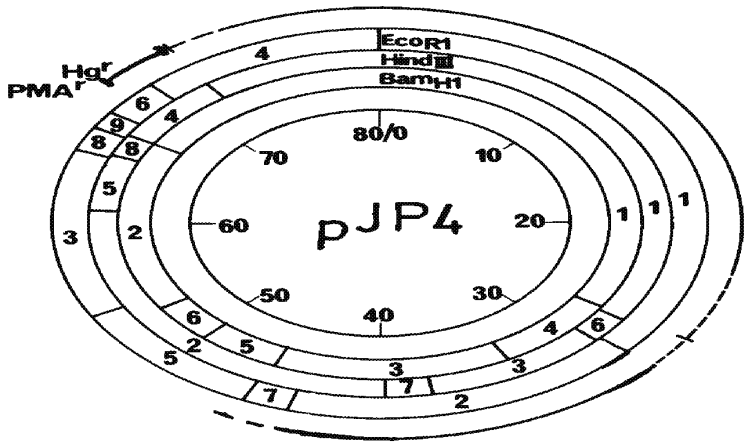


FIG. 1. Genetic and biophysical map of pJP4. The map designations are PMA^r, resistance to PMA; Hg^r, resistance to mercury.

transfer functions suggests the relatively recent acquisition of these functions compared with the evolution of a basic replicon.

Insertions of Tn*1771* in the degradative functions of pJP4 yielded two different phenotypes, either the inability to degrade 2,4-D only or the inability to degrade both 2,4-D and 3CB. Inactivation of both degradative abilities by a single insertion suggests that these pathways are convergent and that the insertion is into a gene encoding a degradative function that is common to both pathways. Chlorocatechol is a catabolic metabolite common to both pathways (*5*, 6), and convergence of the pathways may occur after formation of this compound. Inactivation of only the 2,4-D degradative abilities indicates that the early steps of 2,4-D degradation are plasmid encoded and are adjacent to those encoding enzymatic steps common to both pathways. Mutants incapable of degrading only 3CB were not isolated; therefore, it was not possible to determine whether the initial 3CB degradative steps are chromosomally encoded or, whether due to regional specificity of insertion, transposons did not insert into plasmid-borne genes encoding these functions.

TABLE 1. Sizes of individual restriction fragments of pJP4 generated with endonucleases *Eco*RI, *Hin*dIII, and *Bam*HI[a]

Enzyme	DNA fragment sizes (kb) and designations[b]	Sum of restriction fragments (kb)
*Hin*dIII	HI, 34.7; H2, 22.2; H3, 8.0; H4, 6.4; H5, 4.1; H6, 2.0; H7, 1.6; H8, 1.4	80.4
*Eco*RI	E1, 28.8; E2, 14.5; E3, 10.6; E4, 9.9; E5, 8.3; E6, 3.2; E7, 1.9; E8, 1.7; E9, 1.2	80.1
*Bam*HI	B1, 39.4; B2, 14.4; B3, 11.9; B4, 6.5; B5, 3.6; B6, 3.3	79.10

[a] Fragment sizes (in kb) were calculated by the method of Danna (3) with bacteriophage lambda DNA cleaved with either *Hin*dIII or *Eco*RI as internal standards.

[b] Fragment designations are *Eco*RI (E), *Hin*dIII (H), and *Bam*HI (B). The fragments are numbered sequentially from largest to smallest, e.g., H1 to H8.

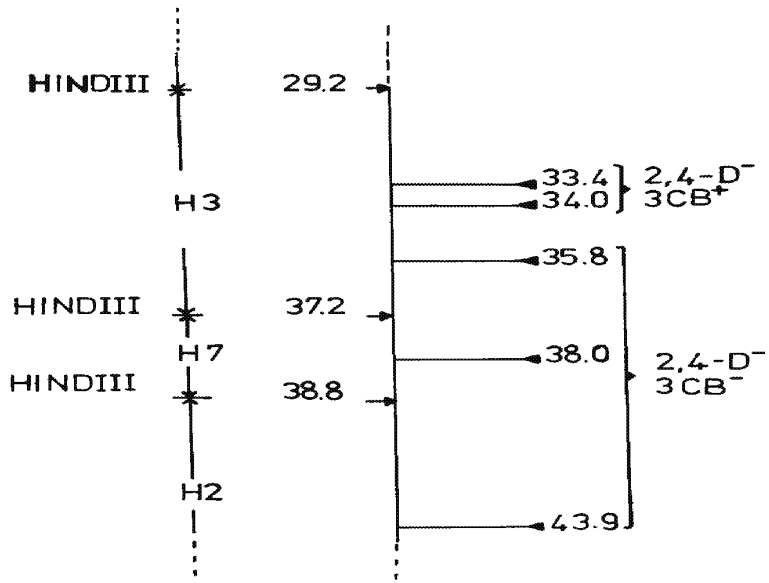


FIG. 2. Location of Tn*1771* insertion mutations which alter the 2,4-D and 3CB degradative genes of pJP4. Insertion mutation sites are designated by —◂, and the *Hin*dIII restriction sites are designated by —▸.

Localization of genes encoding 2,4-D degradation and part of the 3CB degradative functions will enable fine-structure analysis of this region of the plasmid to determine the regulation and physical arrangement of these genes. Once this has been achieved, these DNA sequences may prove to be useful in constructing bacterial strains for wastewater treatment or possibly for constructing herbicide-resistant plants (8).

LITERATURE CITED

1. **Barth, P. T.** 1979. RP4 and R300B as wide host range plasmid cloning vehicles, p. 399–410. *In* K. Timmis and A. Puhler (ed.), Plasmids of medical, environmental and commercial importance. Elsevier/North-Holland Publishing Co., Amsterdam.
2. **Calos, M. P., and J. H. Miller.** 1980. Transposable elements. Cell **20**:579–595.
3. **Danna, K. H.** 1980. Determination of fragment order through partial digests and multiple enzyme digests. Methods Enzymol. **65**:449–467.
4. **Don, R. H., and J. M. Pemberton.** 1981. Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J. Bacteriol. **145**:681–686.
5. **Dorn, E., and H.-J. Knackmuss.** 1978. Chemical structure and biodegradability of halogenated aromatic compounds: two catechol 1,2-dioxygenases from a 3-chlorobenzoate grown Pseudomonas. Biochem. J. **174**:73–84.
6. **Kearney, P. C., and D. D. Kauffman.** 1972. Microbial degradation of some chlorinated pesticides, p. 166–189. *In* Degradation of synthetic organic molecules in the biosphere. National Academy of Science, Washington D.C.
7. **Pemberton, J. M.** 1981. Genetic engineering and biological

COR-AAD1-00001360

Downloaded from http://jb.asm.org/ on June 5, 2013 by guest

detoxification of environmental pollutants. Residue Rev. **78**:1–11.
8. **Pemberton, J. M.** 1983. Degradative plasmids. Int. Rev. Cytol. **84**:155–183.
9. **Pemberton, J. M., and R. H. Don.** 1981. Plasmids of agricultural and environmental importance. Agric. Environ. **6**:23–32.
10. **Pemberton, J. M., and P. R. Fisher.** 1977. 2,4-D plasmids and persistence. Nature (London) **268**:732–733.
11. **Thomas, C. M., R. Meyer, and D. R. Helinski.** 1980. Regions of broad host range plasmid Rk2 which are essential for replication and maintenance. J. Bacteriol. **141**:213–222.
12. **Ward, J. M., and J. Grinsted.** 1978. Mapping of functions in the R plasmid R388 by examination of deletion mutants generated *in vitro*. Gene **3**:87–95.

Downloaded from http://jb.asm.org/ on June 5, 2013 by guest

COR-AAD1-00001361

ELSEVIER

Journal of Molecular Catalysis B: Enzymatic 15 (2001) 155–162

JOURNAL OF MOLECULAR CATALYSIS B: ENZYMATIC

www.elsevier.com/locate/molcatb

Alternative substrates of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase

Julie C. Dunning Hotopp [a], Robert P. Hausinger [a,b,*]

[a] *Department of Microbiology and Molecular Genetics, 160 Giltner Hall, Michigan State University, East Lansing, MI 48824, USA*
[b] *Department of Biochemistry and Molecular Biology, 160 Giltner Hall, Michigan State University, East Lansing, MI 48824, USA*

Received 29 January 2001; received in revised form 6 March 2001; accepted 6 March 2001

Abstract

2,4-Dichlorophenoxyacetate/α-ketoglutarate dioxygenase (TfdA), the first enzyme in the catabolic pathway for the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D), oxidizes α-ketoglutarate (α-kG) to CO_2 and succinate while hydroxylating 2,4-D to yield an unstable hemiacetal that decomposes into 2,4-dichlorophenol and glyoxylate. In an effort to extend the potential biotechnological utility of this enzyme, a variety of non-phenoxyacetate compounds were examined as potential substrates. 2-Naphthoxyacetic acid was the best alternative substrate tested, followed by benzofuran-2-carboxylic acid, 2,4-dichlorocinnamic acid, 2-chlorocinnamic acid, 1-naphthoxyacetic acid, and 4-chlorocinnamic acid. TfdA appeared to oxidize the olefin bond of the cinnamic acids and benzofuran-2-carboxylate to form the corresponding epoxides. Whole cells were observed to also catalyze a TfdA-dependent oxidation of 2,4-dichlorocinnamic acid. Based on the ability of TfdA to metabolize chlorinated cinnamic acids, we speculate that *tfdA*-like sequences present in 2,4-D non-degrading natural isolates may function in metabolism of substituted cinnamic acids. These results support the use of TfdA and related enzymes in the specific oxidation of non-phenoxyacetate substrates. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: 2,4-Dichlorophenoxyacetic acid; Cinnamic acid; Dioxygenase; α-Ketoglutarate; Epoxide

1. Introduction

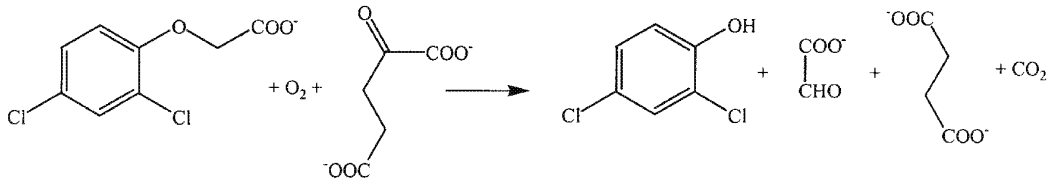
,4-Dichlorophenoxyacetic acid (2,4-D), a broad leaf herbicide, is mineralized by a wide variety of environmental isolates [1] including the best-studied example, *Ralstonia eutropha* JMP134 (pJP4) (formerly *Alcaligenes eutrophus*). The *tfdABCDEF* genes, carried on the pJP4 plasmid in this microorganism [2], encode all of the enzymes necessary for the degradation of 2,4-D to chloromaleylacetate. Chromosomally borne genes encode enzymes that carry out the subsequent metabolism of chloromaleylacetate. Since its discovery and characterization, pJP4 has been used to study the evolution and environmental dispersal of the genes encoding a catabolic pathway for xenobiotic degradation [1].

TfdA catalyzes the first step in this 2,4-D degradation pathway. This ferrous ion and α-ketoglutarate (α-kG)-dependent dioxygenase hydroxylates 2,4-D while converting α-kG to succinate plus CO_2 [3] (Scheme 1). Hydroxylation of 2,4-D at the C-2 position yields an unstable hemiacetal that decomposes into 2,4-dichlorophenol and glyoxylate [3]. TfdA is related in sequence (~30% amino acid identity) to a group of sulfonate degrading α-kG dependent dioxygenases [4,5] and an alkyl sulfate ester degrading α-kG

* Corresponding author. Tel.: +1-517-353-9675; fax: +1-517-353-8957.
E-mail address: hausinge@msu.edu (R.P. Hausinger).

1381-1177/01/$ – see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S1381-1177(01)00018-2

COR-AAD1-00001362



Scheme 1. Reaction catalyzed by TfdA with 2,4-D.

dependent dioxygenase [6]. In addition to these protein homologues, nucleotide sequences closely homologous to *tfdA* are present in approximately 30% of bacterial isolates obtained from agricultural plots. The role of the *tfdA*-like genes containing these sequences is unknown since these soil microorganisms do not degrade 2,4-D [7].

The current study explores alternative substrates for TfdA in order to better define the substrate profile and explore potential biotechnological uses for the enzyme. In particular, we demonstrate that purified enzyme or whole cells can be used to produce synthetically valuable epoxides [8,9] of chlorinated cinnamic acids and benzofuran-2-carboxylic acid. Furthermore, since the 2,4-D-degrading enzyme may retain residual activities for the natural substrate, this work may provide clues about the role of TfdA-like proteins in nature.

2. Experimental

2.1. Purification and assay of TfdA

Escherichia coli DH5α (pUS311) carries *tfdA* on a pUC19-derived plasmid [3]. These cells were grown at 30°C in LB medium for 16–20 h, harvested by centrifugation, and suspended in 30 ml of TE buffer (10 mM Tris, 1 mM EDTA, pH 7.7) containing 1 mM phenylmethylsulfonyl fluoride and 10 μg/ml leupeptin. Crude cell extracts were obtained by passing the cells twice through a pre-cooled French pressure cell (16,000 psi) and clarifying the debris by centrifugation (100,000 × *g* for 45 min) at 4°C. After splitting the sample into two pools, TfdA was enriched by chromatography at 4°C on a column of DEAE-Sepharose (2.5 cm × 19 cm) using TE buffer and a 400 ml linear gradient to 200 mM NaCl. The enzyme eluted at approximately 100 mM NaCl. TfdA-containing fractions were dialyzed in TE buffer and chromatographed on a Mono Q column (HR10/10) (Pharmacia) at room temperature. The enzyme eluted at about 40 mM NaCl when using a 100 ml linear gradient to 200 mM NaCl in TE buffer. When necessary, TfdA was further purified by phenyl-Sepharose chromatography. The sample was adjusted to 1 M in ammonium sulfate, applied to a column (HR10/10) equilibrated with TE buffer containing the same concentration of this salt, and chromatographed with a 100 ml linear gradient from 1 to 0 M ammonium sulfate. The enzyme eluted from this resin at approximately 350 mM salt. TfdA was routinely assayed with 2,4-D as the substrate by using the previously described 4-aminoantipyrene spectrophotometric assay [10]. Protein concentrations were determined by using a commercial protein assay (Bio-Rad) with bovine serum albumin as a standard.

2.2. Kinetics of non-phenoxyacetic acid substrates

The activity of TfdA toward non-phenoxyacetic acid potential substrates was determined at 30°C in 20 mM imidazole buffer (pH 6.8) containing 1 mM α-kG, 50 μM $(NH_4)_2Fe(SO_4)_2$, and 200 μM ascorbic acid by using one or more of four different assays. For all methods, stock solutions of $(NH_4)_2Fe(SO_4)_2$ and ascorbic acid were made fresh prior to each set of experiments.

A YSI model 5300 biological oxygen monitor was used to determine rates of oxygen consumption. The probe was equilibrated using air-saturated MilliQ water, and control runs included samples lacking in ascorbic acid, iron, enzyme, α-kG, or substrate.

Samples analyzed by HPLC (Hewlett-Packard 1050) were resolved by using a Merck Lichrosorb RP-18 column. The peak areas or peak heights for samples absorbing at 230, 254, and/or 280 nm

COR-AAD1-00001363

were determined by using the Hewlett-Packard ChemStation software. Reaction mixtures containing 2,4-dichlorocinnamic acid were quenched with NaOH (0.1 M final concentration), neutralized with HCl (0.1N final concentration), and analyzed by utilizing a mixture of 65:35:0.1 methanol:water:phosphoric acid (effluent A). A standard curve for the reaction product was prepared by complete conversion of selected concentrations of substrate. The area under the product peak was assumed to correspond to the original concentration of substrate. Reaction mixtures that contained naphthoxyacetic acid were quenched with EDTA (5 mM final concentration) and analyzed by using a 50:50:0.1 mixture of methanol:water:phosphoric acid (effluent B) as the mobile phase. Standard curves for the degradation of 1-naphthoxyacetic acid and 2-naphthoxyacetic acid were prepared by using 1-naphthol and 2-naphthol, respectively. All other potential substrates tested were quenched with NaOH (0.1 M final concentration) and neutralized with HCl (0.1N final concentration). Standard curves for the reaction products were created by correlating the loss of substrates with gain of products at various concentrations of substrate. Effluent B was used as the mobile phase, and the kinetics were determined in all cases by analyzing the gain of product, except for the 4-chlorocinnamic acid reaction kinetics, which were estimated from the loss of substrate.

Two spectrophotometric assays also were utilized. Analysis of 1-naphthoxyacetic acid was carried out by using the 4-aminoantipyrene assay with 1-naphthol as the standard. For 2,4-dichlorocinnamic acid, an alternative spectrophotometric assay made use of the 250 nm absorption associated with the conjugated system ($\varepsilon_{250} = 7800\,M^{-1}\,cm^{-1}$). The loss of absorbance at this wavelength was monitored in studies designed to assess the ascorbic acid dependence of the reaction.

As previously documented [10,11], TfdA activity decreased over time under assay conditions, especially when ascorbic acid was absent or at low concentration. To analyze the kinetics of non-phenoxyacetic acid substrates, progress curves were analyzed by fitting the data to the following equation

$$P_t = V_i(1 - e^{-k(\text{inact})t})k(\text{inact})^{-1}$$

where P_t is the accumulated product at time t, V_i the initial velocity, and k(inact) the inactivation rate constant [11].

2.3. Identification of the 2,4-dichlorocinnamic acid and benzofuran-2-carboxylic acid metabolites

Samples for NMR analysis were generated at 30°C in D_2O containing 40 mM phosphate buffer (pH 6.8), 200 μM α-kG, 100 μM substrate, 50 μM $(NH_4)_2Fe(SO_4)_2$, 50 μM ascorbic acid, and 1.6 μM TfdA. The metabolites generated from 2,4-dichlorocinnamic acid and benzofuran-2-carboxylic acid were analyzed by ^{1}H-NMR using a Varian VXR 500 MHz NMR spectrometer.

The 2,4-dichlorocinnamic acid metabolite stability was assessed with a sample prepared by enzymatic conversion of 10 μmol of 2,4-dichlorocinnamic acid. The reaction mixture was applied to a Pharmacia PEP-RPC HR10/10 column, washed with 0.1% trifluoroacetic acid, and eluted by using a 40:60:0.1 mixture of methanol:water:trifluoroacetic acid while monitoring the absorbance at 254 nm. Aliquots of the sample were incubated for varying time periods using specified conditions, and the concentrations of the remaining metabolite and the degradation products were measured by HPLC analysis with effluent A.

2.4. Assessing non-phenoxyacetic acid substrates with whole cells

R. eutropha JMP134 (pJP4) carries the pJP4 plasmid containing *tfdABCDEF* [2]. *R. eutropha* JMP228 (pBH501aE) is a derivative of *R. eutropha* JMP134 (pJP4) where the *tfdA* gene has been interrupted by transposon mutagenesis [12]. *R. eutropha* JMP228 is the strain lacking the plasmid. These cells were grown to late-exponential phase at 30°C in MMO minimal medium [13] amended with the indicated carbon source. Cells were centrifuged and resuspended ($A_{260} = 1.5$) in fresh MMO with no carbon source. The cell suspensions were aerated on a stir plate. Carbon sources were added and samples were removed at various times, centrifuged, diluted, and transferred to HPLC vials. Isocratic HPLC with effluent A was used to analyze the samples for loss of the substrate peak and gain of the product peak.

2.5. Computing curve fits

KaleidaGraph for Windows by Abelbeck Software was used for computing all curve fits.

COR-AAD1-00001364

3. Results

3.1. Identification of non-phenoxyacetic acid substrates

Prior studies have evaluated the ability of purified TfdA to degrade various phenoxyacetic acids, thiophenoxyacetic acids, and phenoxypropionic acids [10,11,14]. The substrate associated with the highest catalytic efficiency was identified as the xenobiotic compound, 2,4-D. Here, we investigated a range of naturally occurring, non-phenoxyacid, aromatic compounds as substrates of the enzyme, along with some related synthetic derivatives of these compounds. In particular, cinnamic acids, auxin-like compounds, and naphthoxyacetic acids were tested by using oxygen electrode, HPLC, and spectrophotometric methods.

The oxygen electrode assay showed clearly enhanced levels of oxygen consumption over background for TfdA (1.25 μM) assay mixtures containing 2-chloro, 4-chloro, and 2,4-dichlorocinnamic acids, 1- and 2-naphthoxyacetic acids, and benzofuran-2-carboxylic acid (each at 300 μM). When each enzyme-containing sample was adjusted to 200 μM 2,4-D, oxygen consumption immediately increased to match that observed for samples containing only 2,4-D. Thus, none of these compounds exhibited significant inhibition or inactivation of 2,4-D hydroxylation activity by TfdA. For other compounds, oxygen consumption rates were close to background levels. This assay did not directly measure substrate conversion (e.g. some substrates may have uncoupled oxygen consumption from substrate oxidation), and spurious results were observed for at least one sample: 3,4-hydroxycinnamic acid. This compound formed a blue color when mixed with ferrous ion (consistent with metal chelation by the catechol group) and consumed oxygen in the absence of enzyme when ascorbic acid was present, presumably similar to the non-enzymatic oxygen consumption observed for EDTA (not shown) [15]. Given these concerns, the potential substrates were quantitatively studied by using HPLC and spectrophotometric methods, which provided similar results for conversion of the potential substrates.

Under conditions equivalent to those used with the oxygen electrode, 2-chloro, 4-chloro, and 2,4-dichlorocinnamic acids, 1- and 2-naphthoxyacetic acids, and benzofuran-2-carboxylic acid were converted to products. In addition, small and variable amounts of 3,4-dihydroxycinnamic acid (caffeic acid), 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid), 4-hydroxy-3-methoxycinnamic acid (ferulic acid), 2,4-dimethoxycinnamic acid, 3,5-dimethoxycinnamic acid, unsubstituted cinnamic acid, 4-methoxycinnamic acid, and 3-methoxycinnamic acid were converted to products. No indication of substrate loss or product formation was observed for 2-hydroxycinnamic acid, 3-hydroxycinnamic acid, 4-hydroxycinnamic acid (coumaric acid), or 2-methoxycinnamic acid. Similarly, chromone-2-carboxylic acid, indole-2-carboxylic acid, indole-3-carboxylic acid, indole-3-acetic acid (auxin), indole-3-acrylic acid, hippuric acid, and phenylpropionic acid were not substrates of TfdA.

3.2. Kinetic analysis of non-phenoxyacetic acid substrates

Detailed studies of the enzyme kinetics were carried out with the test compounds that were convincingly shown to be substrates. Characterization of the kinetic parameters was complicated due to an irreversible inactivation of the ferrous containing enzyme that occurs during exposure to oxygen [11]. An average $k(\text{inact}) = 0.49 \pm 0.26\,\text{min}^{-1}$ was determined for all substrates, regardless of the substrate concentration. Initial rates were obtained by fitting the progress curves for substrate conversion, as described in the Materials and Methods. The enzyme kinetics were further complicated by the dependence of the reaction on ascorbic acid. As illustrated in Fig. 1, the concentration of this reductant had a large effect on the initial rate of 2,4-dichlorocinnamic acid utilization. Similar results were previously reported for conversion of the poor substrate, thiophenoxyacetic acid [11]. To overcome this requirement, 200 μM ascorbic acid was used for all kinetics studies.

The results of the kinetic investigations are provided in Table 1. For comparison, data also were obtained for TfdA metabolism of 2,4-D. The values of K_m, k_{cat}, and k_{cat}/K_m (the catalytic efficiency) associated with 2,4-D were similar to, and more accurate than, previous reports ($17.5 \pm 1.0\,\mu\text{M}$, $529 \pm 16\,\text{min}^{-1}$, and $30{,}200\,\text{min}^{-1}\,\mu\text{M}^{-1}$, respectively) [10]. Each of the non-phenoxyacetic acid substrates was associated with a higher K_m, a lower k_{cat}, and a lower catalytic efficiency than was observed for 2,4-D.

COR-AAD1-00001365

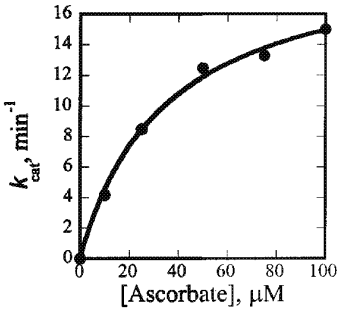


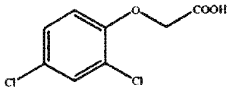
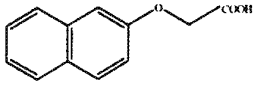
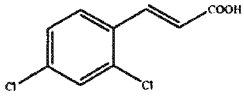
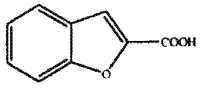
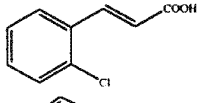
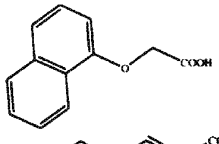
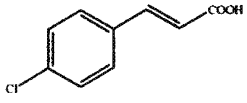
Fig. 1. Ascorbic acid dependence of 2,4-dichlorocinnamic acid utilization by TfdA. The effect of ascorbic acid concentration on the initial rate of 2,4-dichlorocinnamic acid loss was monitored by using a continuous spectrophotometric assay. The standard assay buffer contained 391 nM TfdA dimer and 100 μM 2,4-dichlorocinnamic acid.

3.3. Metabolite analysis

Similar to the conversion of substituted phenoxyacetic acids to their corresponding phenols, 1-naphthoxyacetic acid and 2-naphthoxyacetic acid were shown by co-elution from an HPLC column to be converted to 1- and 2-naphthol (respectively).

^{1}H-NMR analysis of the sample mixture after complete transformation of 2,4-dichlorocinnamic acid by TfdA revealed (a) the disappearance of two doublets at δ 6.4 ppm and δ 7.6 ppm associated with protons bound to carbon atoms participating in the olefin bond and (b) the appearance of two new doublets at δ 3.5 ppm and δ 4.2 ppm (data not shown). The epoxide protons of phenylglycidic acid (the epoxide of cinnamic acid) exhibit identical resonances [16]; thus, we conclude that TfdA oxidized the side chain double bond of 2,4-dichlorocinnamic acid to produce 2,4-dichlorophenylglycidic acid. In the case of benzofuran-2-carboxylic acid, NMR evidence revealed an analogous change in chemical shift of the proton on carbon-3 from δ 7.2 ppm to δ 3.6 ppm (data not shown). This result was consistent with similar formation of an epoxide during substrate conversion.

Table 1
Kinetics of alternative substrates of TfdA

Structure	Substrate	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m (min^{-1} mM^{-1})
	2,4-D	20.0 ± 5.4	1020 ± 90	51000
	2-Naphthoxyacetic acid	134 ± 32	263 ± 5	1960
	2,4-Dichlorocinnamic acid	190 ± 56	52.5 ± 5.1	276
	Benzofuran-2-carboxylic acid	254 ± 120	105 ± 16	413
	2-Chlorocinnamic acid	264 ± 52	22.1 ± 1.3	83.9
	1-Naphthoxyacetic acid	622 ± 310	15.7 ± 2.9	25.2
	4-Chlorocinnamic acid	>900	>38	

COR-AAD1-00001366

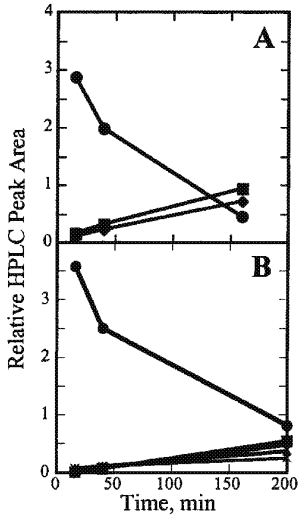


Fig. 2. pH stability and nucleophilic reactivity of the TfdA reaction product derived from 2,4-dichlorocinnamic acid. The peak areas at 230 nm were determined for the 2,4-dichlorocinnamic acid reaction product (●) and degradation products associated with retention times of 2.5 (◆), 4 (■), 6 (▲), and 7 min (X) of samples incubated at pH 3 (A) and pH 4 with 2 mM hydroxylamine (B).

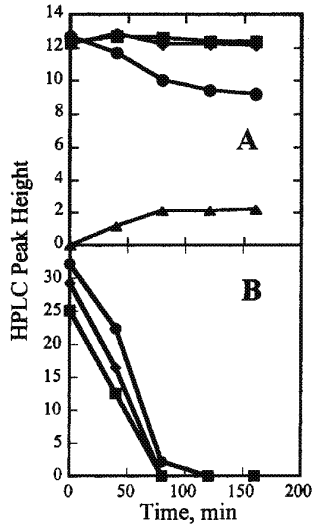


Fig. 3. Degradation of cinnamic acids by intact cells. The time dependence was examined for decomposition of 2,4-dichlorocinnamic acid (panel A) and cinnamic acid (panel B) by *R. eutropha* JMP134 (pJP4) (●), *R. eutropha* JMP228 (■), and *R. eutropha* JMP228 (pBH501aE) (◆) by using HPLC methods and analysis of the resulting absorbances at 230 nm. In addition, the time dependence was determined for production of the metabolite derived from 2,4-dichlorocinnamic acid by *R. eutropha* JMP134 (pJP4) (▲) in panel A.

To further test whether the product of the 2,4-dichlorocinnamic acid is 2,4-dichlorophenylglycidic acid, we examined its pH stability and nucleophilic reactivity. Acidic conditions (pH 3) led to the rapid decomposition of the sample as two degradation products were formed (Fig. 2, panel A). In contrast, the reaction product was stable for at least 8 days at pH 13 (data not shown). Inclusion of the strong nucleophile hydroxylamine at low pH led to the formation of two new degradation products (Fig. 2, panel B). The low pH reactivity and distinct product profile generated in the presence of hydroxylamine are consistent with an epoxide being the initial product of 2,4-dichlorocinnamic acid metabolism by TfdA.

3.4. Whole cell studies

The potential of using intact bacterial cells to metabolize 2,4-dichlorocinnamic acid was examined using *R. eutropha* JMP134 (pJP4), *R. eutropha* JMP228 (pBH501aE), and *R. eutropha* JMP228 (Fig. 3). The first strain encodes the entire 2,4-D pathway, whereas the second and third cultures lack *tfdA* and the pJP4 plasmid, respectively. Degradation of 2,4-dichlorocinnamic acid was observed only in the first isolate, consistent with a requirement for *tfdA* (panel A). Notably, product accumulation mirrored substrate disappearance in agreement with the inability of this strain to grow on 2,4-dichlorocinnamic acid as a sole carbon source. In contrast to the 2,4-dichlorocinnamic acid results, all three strains decomposed unsubstituted cinnamic acid (panel B). These results demonstrate that a cinnamic acid degrading system exists in the cells and does not involve *tfdA* or other loci on the pJP4 plasmid.

4. Discussion

4.1. Expansion of the substrate range of TfdA

We demonstrated that several substituted cinnamic acids are utilized (albeit poorly) as substrates by TfdA. The best of this group of substrates is 2,4-dichlorocinnamic acid, whereas the 2- or 4-monochloro

COR-AAD1-00001367

derivatives are used with less efficiency and the unsubstituted cinnamic acid is metabolized even more poorly. This pattern of substrate preference parallels the situation for phenoxyacetic acids, where the 2,4-dichloro species is preferred over the monochlorinated derivatives, which are more readily metabolized than the unsubstituted species [10]. Based on NMR evidence and stability/reactivity studies of the 2,4-dichlorocinnamic acid metabolite, we propose that the TfdA-generated products arising from the cinnamic acids are the side chain epoxides. Thus, TfdA is not limited to inserting oxygen into unactivated C–H bonds — it also can catalyze the easier oxidation of a C–C double bond.

Because 2,4-D chemically mimics auxin in its action as an herbicide [17], we tested whether the converse situation may occur; i.e. can auxin or auxin-like compounds chemically mimic 2,4-D and serve as a substrate for TfdA? We showed that the natural plant hormone, indole-3-acetic acid, is not a substrate of the enzyme. Furthermore, we demonstrated that indole-3-carboxylate, indole-3-acrylate, and indole-2-carboxylate also are not transformed by the enzyme. Interestingly, however, benzofuran-2-carboxylate (the analogue of indole-2-carboxylate in which the indole N is replaced by O) does serve as a substrate of TfdA. Perhaps related to this reaction, benzofuran-2-carboxylate mimics phenoxyacetic acid with the side chain linked to the aromatic ring via a methenyl carbon. NMR evidence suggests that the double bond of this substrate is converted to an epoxide by TfdA. The ability of TfdA to selectively catalyze the transformation of benzofuran-2-carboxylate while not oxidizing indole-2-carboxylate may relate to the differences in resonance energies between the pyrrole (22 kcal/mol) and furan (16 kcal/mol) rings [18]. Thus, oxygen insertion into the double bond of benzofuran-2-carboxylate is more favorable than for the indole-2-carboxylate. Another compound related in structure to both benzofuran-2-carboxylate and phenoxyacetic acid is chromone-2-carboxylate, which is not a substrate of TfdA.

1- and 2-naphthoxyacetic acids, a third class of test compounds, were shown to be substrates of the enzyme. These compounds closely resemble phenoxyacetic acid, but contain a larger aromatic ring. The fact that TfdA is capable of transforming these compounds demonstrates that the active site is sufficiently large to allow entry of these species. Of interest, 2-naphthoxyacetic acid was the herbicidal predecessor to 2,4-D and was replaced because smaller amounts of 2,4-D could be applied to achieve the same benefits [19].

4.2. The ancestral role of TfdA?

Soon after 2,4-D was introduced into the environment in the early 1940s, reports suggested that 2,4-D was rapidly decomposed in soil (e.g. [20]). To account for decomposition of this xenobiotic compound, we earlier proposed the existence of an ancestral gene encoding a degradative enzyme with greatest specificity toward a natural product, but capable of utilizing 2,4-D with low efficiency [1]. According to this scenario, mutations arose over time to enhance the specificity of the ancestral enzyme toward 2,4-D and the new gene recombined with genes for 2,4-dichlorophenol degradation to create a 2,4-D degradation pathway. In support of this hypothesis, approximately a third of soil bacterial isolates possess a *tfdA*-like gene, but do not measurably degrade 2,4-D [7]. Thus, environmental isolates possess *tfdA*-like genes that appear to play another, yet unidentified, role besides 2,4-D catabolism. Based on the studies reported above, we propose that the ancestral role of TfdA may have involved degradation of cinnamic acids. Ring-substituted cinnamic acids are widely distributed in nature and are present in sizable amounts. For example, 4-hydroxycinnamic acid is found at a concentration of 8.7 mg carbon per gram carbon in the top meter of peat soil [21]. Plants synthesize large amounts of cinnamic acids by phenylalanine deamination, and they catalyze a variety of ring substitution, condensation, degradation, reduction, or conjugation reactions to produce a host of cinnamic acid-related compounds including flavonoids, hydroxybenzoates, and lignin [22]. Degradation of substituted cinnamic acids occurs by several pathways (for a review, see [23]), but it would not be surprising to learn that another pathway remains to be described. Future studies will test the cinnamic acid degrading abilities of the enzyme products from *tfdA*-like genes in environmental isolates.

4.3. Biocatalysis

Because of the wide range of enantiospecific chemical reactions in which epoxides can participate, these

COR-AAD1-00001368

compounds serve as intermediates in a variety of medically important chemical syntheses [24]. For example, phenylglycidic acid, the epoxide of unsubstituted cinnamic acid, has been used as an intermediate in the production of taxol [8] and (2*S*,3*S*)-diltiazem [9]. In order to further develop TfdA for biocatalytic production of epoxides from substituted cinnamic acids, three issues must be addressed: chemical yield, cost, and enantiospecificity. Because of the low k_{cat} values observed for *R. eutropha* enzyme, it may be necessary to isolate mutants or obtain isozymes from other environmental sources to increase the chemical yield. With regard to cost, we note that pyruvic acid can replace α-kG for in vitro assays [10]; hence, this less expensive α-ketoacid potentially could be used in immobilized enzyme reactors. Alternatively, our demonstration that whole cells containing this enzyme can convert chlorinated cinnamic acids to the corresponding epoxides offers the possibility that cellular TfdA can be used to expand the repertoire of available synthetic building blocks. Whole cells can be grown in an inexpensive minimal media and used to carry out the desired reaction. Since whole cells do not utilize either 2,4-dichlorocinnamic acid or 2,4-dichlorophenylglycidic acid, substrate is directly converted to product stoichiometrically. We have not characterized the stereospecificity of the TfdA oxygenase reaction for substituted cinnamic acids. However, we expect enantiopure epoxidation. TfdA has been shown previously to hydroxylate only (*S*)-dichloroprop, a phenoxypropionate herbicide [14]. Additionally, a fellow member of the superfamily of α-kG-dependent dioxygenases, EpoA, catalyzes the enantiomer specific production of fosfomycin ((−)-*cis*-1,2-epoxypropylphosphonic acid) from *cis*-propenylphosphonic acid in *Penicillium decumbens* [25]. Likewise, styrene monooxygenase (from a different superfamily of oxygenases) is capable of carrying out the enantioselective conversion of cinnamic acid methyl ester to phenylglycidic acid methyl ester [16]. The studies described here lay the groundwork for future efforts to develop this potential biocatalyst.

Acknowledgements

This work was supported by the National Science Foundation (MCB9603520) and the Michigan State University Agricultural Experiment Station. We thank Raghavakaimal Padmakumar for initiating several of these studies and performing NMR analysis of the benzofuran-2-carboxylate product, the staff of the NMR facility for their assistance, and members of the laboratory for their comments.

References

[1] R.P. Hausinger, F. Fukumori, D.A. Hogan, T.M. Sassanella, Y. Kamagata, H. Takami, R.E. Saari, in: K. Horikoshi, M. Fukuda, T. Kudo (Eds.), Microbial Diversity and Genetics of Biodegradation, Karger, New York, 1997, p. 35.
[2] R.H. Don, J.M. Pemberton, J. Bacteriol. 145 (1981) 681.
[3] F. Fukumori, R.P. Hausinger, J. Bacteriol. 175 (1993) 2083.
[4] E. Eichhorn, J.R. van der Ploeg, M.A. Kertesz, T. Leisinger, J. Biol. Chem. 272 (1997) 23031.
[5] D.A. Hogan, T.A. Auchtung, R.P. Hausinger, J. Bacteriol. 181 (1999) 5876.
[6] A. Kahnert, M.A. Kertesz, J. Biol. Chem. 275 (2000) 31661.
[7] D.A. Hogan, D.H. Buckley, C.H. Nakatsu, T.M. Schmidt, R.P. Hausinger, Microb. Ecol. 34 (1997) 90.
[8] J.-N. Denis, A.E. Greene, A.A. Serra, M.-J. Luche, J. Org. Chem. 1986 (1986) 46.
[9] A. Gentile, C. Giordano, J. Org. Chem. 57 (1992) 6635.
[10] F. Fukumori, R.P. Hausinger, J. Biol. Chem. 268 (1993) 24311.
[11] R.E. Saari, R.P. Hausinger, Biochemistry 37 (1998) 3035.
[12] E.M. Top, O.V. Maltseva, L.J. Forney, Appl. Environ. Microbiol. 62 (1996) 2470.
[13] R. Stanier, N. Palleroni, M. Doudoroff, J. Gen. Microbiol. 43 (1966) 159.
[14] R.E. Saari, D.A. Hogan, R.P. Hausinger, J. Molec. Catal. B: Enzymatic 6 (1999) 421.
[15] A.L. Feig, S.J. Lippard, Chem. Rev. 94 (1994) 759.
[16] P. DiGennaro, A. Colmegna, E. Galli, G. Sello, F. Pelizzoni, G. Bestetti, Appl. Environ. Microbiol. 65 (1999) 2794.
[17] J.L. Garraway, R.L. Wain, Drug Design 7 (1976) 115.
[18] S.N. Ege, Organic Chemistry: Structure and Reactivity, D. C. Heath and Company, Toronto, 1994, p. 1061.
[19] S.S. Que Hee, R.G. Sutherland, The Phenoxyalkanoic Herbicides, Vol. I, Chemistry, Analysis, and Environmental Pollution, CRC Press, Boca Raton, FL, 1981.
[20] L.J. Audus, J. Sci. Food Agric. 3 (1952) 268.
[21] T. Katase, R. Kondo, Soil Sci. 148 (1989) 258.
[22] D. Strack, in: P.M. Dey, J.B. Harbone (Eds.), Plant Biochemistry, Academic Press, San Diego, 1997, p. 387.
[23] J.P. Rosazza, Z. Huang, L. Dostal, T. Volm, B. Rousseau, J. Ind. Microbiol. 15 (1995) 457.
[24] A. Archelas, R. Furstoss, Ann. Rev. Microbiol. 51 (1997) 491.
[25] M. Watanabe, N. Sumida, S. Murakami, H. Anzai, C.J. Thompson, Y. Tateno, T. Murakami, Appl. Environ. Microbiol. 65 (1999) 1036.

COR-AAD1-00001369

Microbiol. Res. (2002) **157,** 317–322
http://www.urbanfischer.de/journals/microbiolres

© Urban & Fischer Verlag

The two enantiospecific dichlorprop/α-ketoglutarate-dioxygenases from *Delftia acidovorans* MC1 – protein and sequence data of RdpA and SdpA

Anne Westendorf*, Dirk Benndorf, Roland H. Müller, Wolfgang Babel

UFZ Centre for Environmental Research, Department of Environmental Microbiology, Permoserstr. 15, D-04318 Leipzig, Germany

Accepted: August 26, 2002

Abstract

Two α-ketoglutarate-dependent dioxygenases carrying enantiospecific activity for the etherolytic cleavage of racemic phenoxypropionate herbicides [(*RS*)-2-(2,4-dichlorophenoxy)propionate and (*RS*)-2-(4-chloro-2-methylphenoxy)propionate] from *Delftia acidovorans* MC1 were characterized with respect to protein and sequence data. The (*S*)-phenoxypropionate/α-ketoglutarate-dioxygenase (SdpA) appeared as a monomeric enzyme with a molecular weight of 32 kDa in the presence of SDS. N-terminal sequences revealed relationship to α-ketoglutarate-dependent taurine dioxygenase (TauD) and to 2,4-dichlorophenoxyacetate/α-ketoglutarate-dioxygenase (TfdA). The (*R*)-phenoxypropionate/α-ketoglutarate-dioxygenase (RdpA) referred to 36 kDa in the presence of SDS and to 108 kDa under native conditions. Internal sequences of fragments obtained after digestion made evident relationship to TfdA and TauD. Two-dimensional electrophoretic separation resulted in the resolution of up to 3 individual spots with almost identical molecular weights but different isoelectric points with both RdpA and SdpA. The structural differences of these isoenzyme forms are not yet clear.

Key words: chiral herbicides – (*S*)-phenoxypropionate/α-ketoglutarate-dioxygenase (SdpA) – (*R*)-phenoxypropionate/α-ketoglutarate-dioxygenase (RdpA) – partial amino acid sequences – *Delftia acidovorans* MC1

Corresponding author: Anne Westendorf
e-mail: annewest@umb.ufz.de

Introduction

A few strains are known which are capable of completely degrading the chiral herbicides (*RS*)-2-(2,4-dichlorophenoxy)propionate (2,4-DP; dichlorprop) and (*RS*)-2-(4-chloro-2-methylphenoxy)propionate (MCPP; mecoprop) in addition to phenoxyacetate derivatives (2,4-dichlorophenoxyacetate; 2,4-D) (Zipper *et al.* 1996; Ehrig *et al.* 1997; Müller *et al.* 1999). This contrasts with the experience gained with most of the phenoxy herbicide-degrading strains, which have been shown, based on the activity of 2,4-D/α-ketoglutarate-dioxygenase (TfdA) (Fukumori and Hausinger 1993a, b; but cf. Saari *et al.* 1999), to be highly specific for the productive degradation of phenoxyacetates (Häggblom 1992; Hoffmann *et al.* 1996). The enzymes of one of the above strains, *Delftia acidovorans* MC1, have been purified and characterized (Müller and Babel 1999; Westendorf *et al.* 2003). Both enzymes were found to be α-ketoglutarate-dependent dioxygenases (Müller *et al.* 2001). This matches the reaction mechanism described for the cleavage of 2,4-D by the respective enzyme, i.e. TfdA, from the canonical strain *Ralstonia eutropha* JMP134 (Fukumori and Hausinger 1993a, b). Kinetic investigations revealed that one of the enzymes of strain MC1 is highly specific for the cleavage of the R-configuration of 2,4-DP and MCPP, i.e. (*R*)-phenoxypropionate/α-ketoglutarate-dioxygenase, consecutively designed RdpA. The other enzyme, i.e. (*S*)-phenoxypropionate/α-ketoglutarate-dioxygenase and designed SdpA, prefers the S-configuration but also accepts phenoxyacetate derivatives as substrate (Westendorf *et al.* 2003). Sequence data of *tfdA* are known

COR-AAD1-00001370

(Streber *et al.* 1987; Suwa *et al.* 1996; Mäe *et al.* 1993; Ka *et al.* 1994; Vallaeys *et al.* 1996; Hoffmann *et al.* 2001) revealing high similarity in the gene of the various strains of β- and γ-*Proteobacteria* (but form a separate cluster, *tfdA*α, with α-*Proteobacteria*; Itoh *et al.* 2002). The application of primers derived from conserved regions of *tfdA* (Vallaeys *et al.* 1996), however, failed to result in amplification products in PCR with genomic DNA from strain MC1 (Müller *et al.* 2001). No data are available on a molecular level with respect to the phenoxypropionate-degrading strains. The present investigation is designed to elucidate the protein characteristics of the two different dioxygenases of strain MC1. Special attention is directed to sequence data whose elucidation should enable the assignment of the two enzymes to known proteins by databank comparison. Moreover, these data should favor targeting the genetic basis.

Materials and methods

Enzyme purification. The (*R*)-phenoxypropionate/α-ketoglutarate dioxygenase (RdpA) and the (*S*)-phenoxypropionate/α-ketoglutarate dioxygenase (SdpA) were purified as described by Westendorf *et al.* (2003). RdpA and SdpA were enriched by this procedure by a factor of 20 and 30, respectively, resulting in 71% and >98% purity.

Electrophoretic and chromatographic separations. For one-dimensional electrophoresis, a PowerEase 500 system was used (Novex, San Diego, USA). SDS PAGE of the enzyme preparation was carried out on 12% Tris glycine pre cast acrylamide gels (Novex). The samples contained 1–2 µg of protein. The running buffer was composed of 2.9 g/l Tris base, 14.4 g/l glycine and 1.0 g/l SDS. The sampling buffer contained 2 ml glycerol, 4 ml 10% SDS (w/v), 0.5 ml 1% bromophenol blue and 2.5 ml 0.5 M Tris/HCl, pH 6.8 in 10 ml distilled water. The gels were treated at 125 V. Mark 12™ Wide Range Standard (Novex) was applied as a reference.

Peptides obtained from the digested enzymes were separated by peptide electrophoresis with Tris/Tricin buffer as described by Schägger and von Jagow (1987) in Mini Protean II Cell (Bio-Rad, Hercules, CA, USA). The following concentrations of acrylamide were used: separation gel (16% T and 3% C), spacer gel (10% T and 3% C) and stacking gel (4% T and 3% C). Electrophoresis was run with a constant current of 10 mA/gel for 0.5 h and a constant current of 20 mA/gel for 4 h. Ultra low range proteins (Sigma) served as a standard.

Two dimensional electrophoresis was performed as described by Benndorf and Babel (2002).

Gels were stained by following a slightly modified procedure described by Blum et. al. (1987), or by using Coomassie brilliant blue.

Gel exclusion chromatography was performed by HPLC at 23 °C using a Bio-Sil SEC-250 column (300 mm × 7.8 mm, Biorad) and a mobile phase containing 5 mM Na_2HPO_4, 5 mM NaH_2PO_4, 150 mM NaCl (pH 6,8). The flow rate was 1 ml/min. The molecular weight was determined by using a gel filtration standard (Biorad).

Protein digestion. Before enzymatic and chemical in-gel digestion, bands containing protein separated by PAGE were cut out and the gel slices washed three times with 50% methanol and 5% acetic acid in water for at least 3 h. The gel slices were reduced and alkylated by adding and removing the following solutions: 200 µl acetonitrile (5 min), 30 µl 1,4-dithioerythrole in 100 mM ammonium hydrogencarbonate (30 min), 30 µl 100 mM iodoacetamide in 100 mM ammonium hydrogencarbonate (30 min), 200 µl acetonitrile (5 min), 200 µl 100 mM ammonium hydrogencarbonate (10 min) and 200 µl acetonitrile (5 min). After drying the gel slices by vacuum centrifugation, enzymatic digestion was performed by adding 30 µl LysC buffer (25 mM Tris/HCl pH 8.5) containing 1.5 µl LysC (Lys-C, enzyme preparation from *Lysobacter enzymogenes*, Merck, Germany), 1 mM EDTA and 0.05% SDS, and incubating the sample overnight. Chemical digestion was performed by adding 30 µl BNPS-scatole solution (500 µg BNPS in 500 µl 75% acetic acid; Crimmins *et al.*, 1990) and incubating the sample for 1 h at 47 °C by shaking. To recover the peptides, the following solution were added and the respective supernatants collected: twice 30 µl of 50 mM ammonium hydrogencarbonate for 10 min and twice 30 µl extraction buffer (10 ml acetonitrile, 5 ml dist. water, 1.176 ml 85% formic acid) for 10 min. The unified supernatants were centrifuged under vacuum until a dry pellet was obtained. The latter was solubilized in sample buffer for electrophoretic separation.

Amino acid sequencing. Peptides separated by PAGE were blotted onto PVDF membranes (Jin and Cerletti 1992). Gels were washed for 5 min in distilled water and for 15 min in 10 mM CAPS buffer pH 11. Blotting was performed for 2 h at 3 mA/cm² with 10 mM CAPS buffer pH 11 (Trans Blot, Semi-Dry Transfer Cell, Biorad, USA). After staining with Coomassie brilliant blue, the bands were cut and applied to N-terminal sequencing (PROCISE® 491 cLC Protein Sequencer or PROCISE® 473A Protein Sequencer, Applied Biosystems; Forster City, CA, USA).

COR-AAD1-00001371

Results

One-dimensional SDS PAGE of the enzyme preparation resulted in distinct protein bands corresponding to molecular weights of about 32 kDa for SdpA and to 36 kDa for RdpA. Gel exclusion chromatography of the preparations of RdpA under native conditions resulted in a protein fraction corresponding to about 108 kDa which showed cleavage activity for (*R*)-2,4-DP. In the presence of SDS, the 108 kDa band disappeared in favor of a band with an apparent molecular weight of 36 kDa. This pattern indicates a trimeric structure of the R-specific enzyme. By contrast, with SdpA no fraction could be detected which corresponded to a protein larger than the monomeric form in gel exclusion chromatography. The protein appeared at a retention time corresponding to a molecular weight of about 23.7 kDa under native conditions. The enzyme separated under these conditions showed specific activity corresponding to that found before. Applying this fraction to SDS PAGE or performing gel exclusion chromatography in the presence of SDS resulted in separation patterns which correlated this protein to a mass of about 32 kDa.

Table 1. Isoelectric points (pI) and molecular weight (MW) of the isoforms of RdpA (R) and SdpA (S)

Spot	pI	MW [kDa]
R1	5.9	36.3
R2	5.86	35.9
R3	5.8	35.7
S1	6.33	32.8
S2	6.16	32.5

Data were obtained from electropherograms by software analysis (Phoretix 2D, 5.01, NonLinear Dynamics Ltd., G.B.)

The purified enzymes were treated by 2D-electrophoresis which resulted in the resolution of the respective enzymes into 2–3 individual protein spots. This is convincingly shown for SdpA in Fig. 1. Apparently, the isoforms were characterized by an almost identical molecular weight but differed in their charges (Table 1). A similar pattern was observed for RdpA (Table 1). Although the latter preparation contained significant impurities which could not be eliminated by the purification protocol applied, three spots were clearly identified to correspond to RdpA by admixture of crude extracts of pyruvate grown cells (not shown). The individual protein spots with respect to RdpA and SpdA preparations were used to determine N-terminal amino acid sequences. The results are shown in Table 2. The sequences were clearly identical with respect to the isoforms of RdpA and SdpA except position 14 in spot R3 which may be attributed to analytical uncertainties in this late cycle of sequencing.

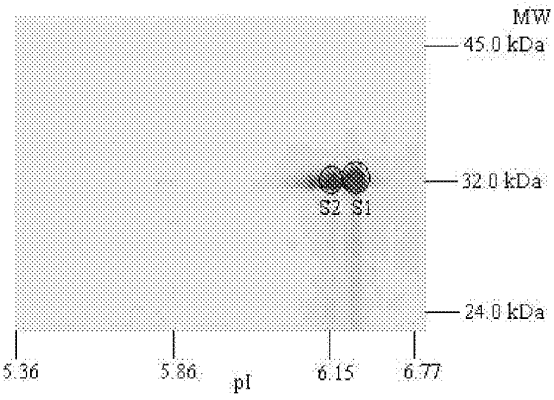


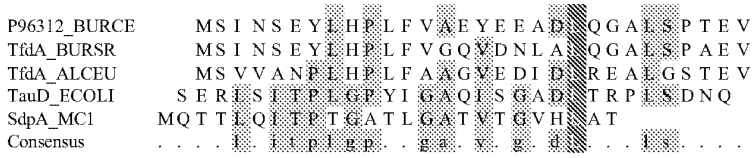
Fig. 1. Part of the 2D-gel of SdpA. Isoelectric focusing from right to left, SDS PAGE from up to down.

Table 2. N-terminal sequences of the isoforms of RdpA and SdpA and internal sequences from selected peptide fragments of the two proteins

Position		Sequence	Accession No
N-terminal	R-Spot1	MHAALSPLSQRFERI	P 83310
	R-Spot2	MHAALSPLSQRFERI	
	R-Spot3	MHAALSPLSQRFEEIAVQPLTGGV	
	S-Spot1	MQTTLQITPTGATLG	P 83309
	S-Spot2	MQTTLQITPTGATLGATVTGVHLAT	
Internal	R-Frac1E	SIEGYPEVQMIRREANESGRVIGDDXHTXII	P 83310
	R-Frac5C	ETLSPTMQATIEGLN	
	R-Frac8C	NEILDAFHTYQVIYFPGQAITNEQHIAFSR	
	S-Frac1E	LGHVQQAGSAYI	P 83309
	S-Frac2E	VIVGNTAG	

R, S refers to RdpA and SdpA; Spot designates the respective proteins in the 2D gels (Fig. 1 for SdpA; RdpA not shown); Frac refers to the protein fractions after enzymatic (E) and chemical (C) digestion

COR-AAD1-00001372



P96312_BURCE, 2.4-D-dioxygenase [*Burkholderia cepacia*]; TFDA_Bursr, *Burkholderia* sp. RASC; TfdA_Alceu, α-ketoglutarate-dependent 2,4-dichlorophenoxyacetate dioxygenase [*Ralstonia eutropha*]; TauD_Ecoli, taurine dioxygenase [*Escherichia coli* O157:H7];

Fig. 2. Alignment of the N-terminus of SdpA (SdpA_MC1) with TfdA and TauD.

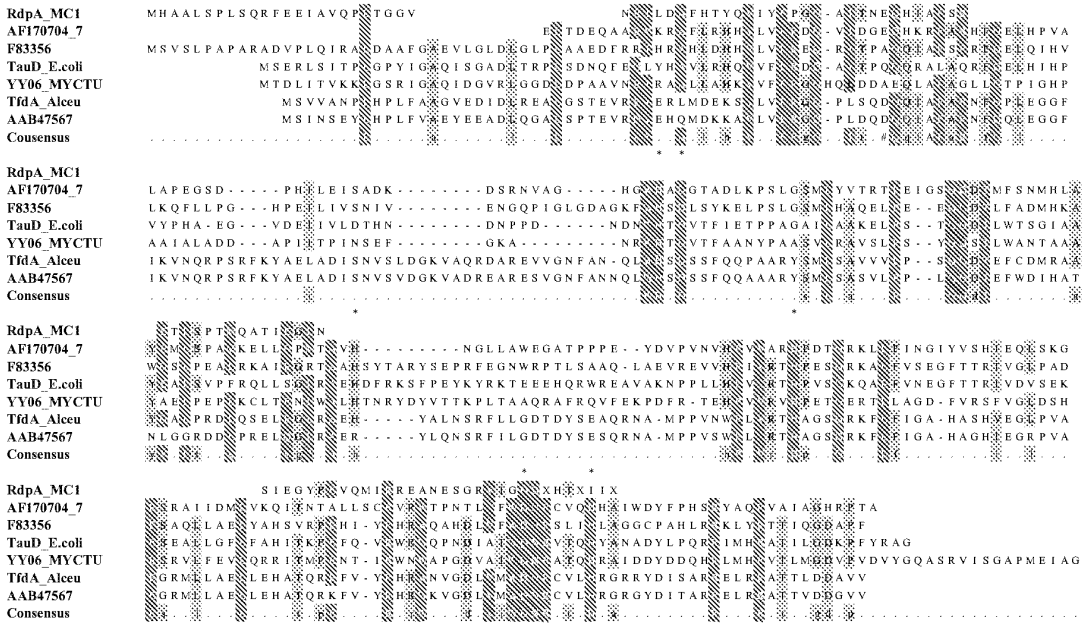


Fig. 3. Alignment of the N-terminal and internal amino acid sequences of RdpA (RdpA_MC1) with different proteins (dioxygenases).

N-terminal sequence data were applied for databank comparison using NCBI BLAST, Matrix Blossom 80. The data from SdpA revealed similarity to other α-ketoglutarate-dependent dioxygenases, i.e. taurine dioxygenase (TauD) and 2,4-D/α-ketoglutarate-dioxygenase (TfdA) (Fig. 2). This was supported further by data derived from internal sequences (see Discussion) the alignment of which is not shown in detail. With respect to RdpA, however, the N-terminal sequence did not exhibit significant homology with any known protein. Elucidation was expected from determining internal sequences. Three fragments obtained by electrophoretic separation after chemical and enzymatic proteolysis gave unequivocal sequence data for their individual N-termini (Table 2), in other cases the results were obviously disrupted by the presence of different fragments in the respective fractions. The sequence data from the positive results were again applied to databank comparison, which now resulted in significant correlation to known proteins. Accordingly, homology of the individual fragments to sequences of TfdA and TauD became evident (Fig. 3). The results of alignment also highlight that the N-terminal sequence of RdpA deviates significantly from the other cluster of proteins, which explains the failure to obtain any positive correlation in the initial attempts.

Discussion

The results make evident a monomeric structure with the (*S*)-phenoxypropionate/a-ketoglutarate dioxygenase (SdpA) and a trimeric structure with the (*R*)-phenoxypropionate/α-ketoglutarate dioxygenase (RdpA) from

COR-AAD1-00001373

D. acidovorans MC1. The molecular weights of the monomers, 32 and 36 kDa respectively, are in the range found for analog enzymes from other sources (Fukumori and Hausinger 1993b; Nickel *et al.* 1997). With TfdA from *R. eutropha* JMP134 a dimeric structure was deduced (Fukumori and Hausinger 1993b). Surprisingly, with the two enzymes of strain MC1, isomeric enzyme forms were detected being separated from each other by only slight differences in their isoelectric points. The structural reasons are not yet apparent. The existence of isoenzymes is a likely explanation of the complex, multiphasic characteristics observed in estimating the kinetic properties with preparations of both RdpA and SdpA (Westendorf *et al.* 2003).

The two dioxygenases from *D. acidovorans* MC1 were shown to be related to other α-ketoglutarate-dependent dioxygenases. According to the N-terminal and two internal sequences, SdpA showed closest homology to taurine dioxygenase (TauD) from *E. coli* (Eichhorn *et al.* 1997) and *Pseudomonas aeruginosa* (36%/43% identical/similar to *E. coli* enzyme) and to TfdA from strains JMP134 and RASC (18%/20% identical/similar to JMP134) based on 44 amino acids aligned. With respect to RdpA, homology with other α-ketoglutarate-dependent dioxygenases was detected from sequence data of various internal fragments, according to which highest homology was attributed to TfdA (32%/37% identical/similar; strain JMP134), followed by TauD (21%/30% identical/similar; *E. coli*) on the basis of 101 amino acids aligned. The present sequence data are considered an appropriate basis to target the genetic background and elucidate the whole structure of these enzymes.

Acknowledgements

This investigation was kindly supported by grant no. 13-8811.61/94 from the Saxon Ministry of the Environment and Agriculture (AW) and by E.C. (HERBICBIOREM, QLK3-CT-1999-00041) (DB).

References

Benndorf, D., Babel, W. (2002): Assimilarory detoxification of herbicides by *Delftia acidovorans* MC1: induction of two chlorocatechol 1,2-dioxygenases as a response to chemostress. Microbiology **148**, 2883–2888.

Blum, H., Beier, H., Gross, H. J. (1987): Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis **8,** 93–99.

Crimmins, D. L., McCourt, D. W., Thoma, R. S., Scott, M. G., Macke, K., Schwartz, B. D. (1990): In situ chemical cleavage of proteins immobilized to glass-fiber and polyvinylidenedifluoride membranes: cleavage at tryptophan residues with 2-(2'-nitrophenylsulfenyl)-3-methyl-3'-bromoindolenine to obtain internal amino acid sequence. Anal. Biochem. **187,** 27–38.

Ehrig, A., Müller, R. H., Babel, W. (1997): Isolation of phenoxy herbicide-degrading *Rhodoferax* species from contaminated building material. Acta Biotechnol. **17,** 351–356.

Eichhorn, E., Ploeg, J. R., Kertesz, M. A., Leisinger, T. (1997): Characterisation of α-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. J. Biochem. **37,** 23031–23036.

Fukumori, F., Hausinger, R. P. (1993a): *Alcaligenes eutrophus* JMP 134 "2,4-dichlorophenoxyacetate monooxygenase" is an α-ketogluterate-dependent dioxygenase. J. Bacteriol. **175,** 2083–2086.

Fukumori, F., Hausinger, R. P. (1993b): Purification and characterization of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. J. Biol. Chem. **268,** 24311–24317.

Häggblom, M. M. (1992): Microbial breakdown of halogenated aromatic pesticides and related compounds. FEMS Microbiol. Rev. **103,** 29–72.

Hoffmann, D., Müller, R. H., Kiesel, B., Babel, W. (1996): Isolation and characterization of an alkaliphilic bacterium capable of growing on 2,4-dichlorophenoxyacetic acid and 4-chloro-2-methylphenoxyacetic acid. Acta Biotechnol. **16,** 121–131.

Hoffmann, D., Kleinsteuber, S., Müller, R. H., Babel, W. (2001): Development and application of PCR primers for the detection of the *tfd* genes in *Delftia acidovorans* P4a involved in the degradation of 2,4-D. Acta Biotechnol. **21,** 321–331.

Jin, Y., Cerletti, N. (1992): Western blotting of transforming growth factor β-2. Optimization of the electrophoretic transfer. Appl. Theor. Electrophoresis **3,** 85–90.

Ka, J. O., Holben, W. E., Tiedje, J. M. (1994): Genetic and phenotypic diversity of 2,4-dichlorphenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D-treated field soils. Appl. Environ. Microbol. **60,** 1106–1115.

Itoh, K., Kanda, R., Sumita, Y., Kim, H., Kamagata, Y., Suyama, K., Yamamoto, H., Hausinger, R. P., Tiedja, J. M. (2002): *tdfA*-like genes in 2,4-dichlorophenoxyacetic acid-degrading bacteria belonging to the *Bradyrhizobium-Agromonas-Nitrobacter-Afipia* cluster in α-Proteobacteria. Appl. Environ. Microbiol. **68,** 3449–3454.

Mäe, A. A., Marits, R. O., Ausmess, N. R., Koiv, V. M., Heinaru, A. L. (1993): Characterization of a new 2,4-dichlorophenocyacetic acid degrading plasmid pEST4011: physical map and localization of catabolic genes. J. Gen. Microbiol. **139,** 3165–3170.

Müller, R. H., Jorks, S., Kleinsteuber, S., Babel, W. (1999): *Comamonas acidovorans* strain MC1: a new isolate capable of degrading the chiral herbicides dichlorprop and mecoprop and the herbicides 2,4-D and MCPA. Microbiol. Res. **154,** 241–246.

Müller, R. H., Babel, W. (1999): Separation of two dichloprop/α-ketoglutarate dioxygenases with enantiospecific properties from *Comamonas acidovorans* MC1. Acta Biotechnol. **19,** 349–355.

Müller, R. H., Kleinsteuber, S., Babel, W. (2001): Physiological and genetic characteristics of two bacterial strains utilizing phenoxypropionate and phenoxyacetate herbicides Microbiol. Res. **156,** 121–131.

COR-AAD1-00001374

Nickel, K., Suter, M. J. F., Kohler, H. P. E. (1997): Involvement of two α-ketoglutarate-dependent dioxygenases in the enantioselective degradation of (R)- and (S)-mecoprop by *Sphingomonas herbicidovorans* MH. J. Bacteriol. **179,** 6674–6679.

Saari, R. E., Hogan, D. A, Hausinger, R. P. (1999): Stereospecific degradation of the phenoxypropionate herbicide dichlorprop. J. Mol. Catal. B: Enzymatic **6,** 421–428.

Schägger, H., von Jagow, G. (1987): Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. **166,** 368–379.

Streber, W. R., Timmis, K. N., Zenk, M. H. (1987): Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. **169,** 2950–2955.

Suwa, Y., Wright, A. D., Fukumori, F., Nummy, K. A., Hausinger, R. P. (1996): Characterization of a chromosomally encoded 2,4-dichlorophenoxyacetic acid/α-ketoglutarate-dioxygenase from *Burkholderia* sp. strain RASC. Appl. Environ. Microbiol. **62,** 2464–2469.

Vallaeys, T., Fulthorpe, R. R., Wright, A. M., Soulas, G. (1996): The metabolic pathway of 2,4-dichlorophenoxyacetic acid degradation involves different families of *tfdA* and *tfdB* genes according to PCR-RFLP analysis. FEMS Microbiol. Ecol. **20,** 163–172.

Westendorf, A., Müller, R. H., Babel, W. (2003): Purification and characterization of the enantiospecific dioxygenases from *Delftia acidovorans* MC1 initiating the degradation of phenoxypropionate and phenoxyacetate herbicides. Acta Biotechnol. **23,** 11–24.

Zipper, C, Nickel, K., Angst, W., Kohler, H. P. E. (1996): Complete microbial degradation of both enantiomeres of the chiral herbicide mecoprop [(*R,S*)-2-(4-chloro-2-methylphenoxy)propionic acid] in an enantioselective manner by *Sphingomonas herbicidovorans* sp. nov. Appl. Environ. Microbiol. **62,** 4318–4322.

COR-AAD1-00001375

J. Microbiol. Biotechnol. (2003), **13**(2), 243–250

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

© The Korean Society for Microbiology and Biotechnology

Isolation and Characterization of 4-(2,4-Dichlorophenoxy)Butyric Acid-Degrading Bacteria from Agricultural Soils

PARK, IN-HYUN AND JONG-OK KA*

School of Agricultural Biotechnology, Seoul National University, Suwon 441-744, Korea

Received: September 9, 2002
Accepted: January 2, 2003

Abstract Eight numerically dominant 4-(2,4-dichlorophenoxy) butyric acid (2,4-DB)-degrading bacteria and three pairs of bacteria showing syntrophic metabolism of 2,4-DB were isolated from soils, and their phylogenetic and phenotypic characteristics were investigated. The isolates were able to utilize 2,4-DB as a sole source of carbon and energy, and their 2,4-DB degradative enzymes were induced by the presence of 2,4-DB. Analysis of 16S rDNA sequences indicated that the isolates were related to members of the genera, *Variovorax*, *Sphingomonas*, *Bradyrhizobium*, and *Pseudomonas*. The chromosomal DNA patterns of the isolates obtained by polymerase-chain-reaction (PCR) amplification of repetitive extragenic palindromic (REP) sequences were distinct from each other. Four of the isolates had plasmids, but only one strain, DB1, had a transmissible 2,4-D degradative plasmid. When analyzed with PCR using primers targeted to the *tfdA*, *B*, and *C* genes, only strains DB2 and DB9a produced DNA bands of the expected sizes with the *tfdA* and *C* primers, respectively. All of the isolates were able to degrade 2,4-D as well as 2,4-DB, suggesting that the degradation pathways of these compounds were closely related to each other, but respiratory activities of many isolates adapted to 2,4-DB metabolism were quite low with 2,4-D.

Key words: 4-(2,4-Dichlorophenoxy)butyric acid, bacteria, biodegradation

Large amounts of man-made chlorinated organic chemicals have been used as pesticides in agriculture. Among them, chlorinated phenoxyalkanoates are widely used, and thus many studies have focused on their persistence and fate in the natural environments [32, 33, 34]. Some phenoxyacid herbicides, such as 2,4-dichlorophenoxyacetic acid (2,4-D) and 2-methyl-4-chlorophenoxyacetic acid (MCPA), are readily mineralized by a number of soil microorganisms and stimulate substantial growth of the corresponding microbial population [5, 12, 27]. In contrast, other closely related pesticides, such as 2,4,5-trichlorophenoxyacetic acid (2,4,5-T), are not easily degraded by soil microorganisms, and organisms able to degrade them are very rare in the environment [1].

2,4-DB is one of the chlorinated phenoxy alkanoic herbicides and is closely related to 2,4-D in its structure and action mechanism. 2,4-DB was developed for use in agriculture as a selective herbicide against broad-leaved weeds in several leguminous crops that are sensitive to 2,4-D [37]. 2,4-DB itself has no effect as a herbicide, but when it is exposed to the plant, which is capable of beta-oxidation, it is converted to phytotoxic 2,4-D and kills the plant.

Soil microorganisms are the most active agents in detoxifying pesticide residues and industrial chemicals exposed to the environment, thus many studies have focused on the isolation and application of microorganisms capable of degrading pesticides and industrial organic pollutants [7, 17, 22, 26]. 2,4-DB reaches the soil mainly during application process and by transfer from plants to the soil by rainwater. In soil, 2,4-DB is suggested to undergo beta-oxidation to 2,4-D, which is further degraded by other soil microflora [14]. However, unlike other phenoxyalkanoate herbicides such as 2,4-D and MCPA, very few microorganisms able to degrade 2,4-DB have been isolated from soils and thus little information is available on their metabolic, genetic, and physiological properties. It is presumed that 2,4-DB could be degraded by 2,4-D-degrading bacteria either containing beta-oxidation systems or that are able to cleave the ether linkage. However, since most of the previously reported 2,4-D-degrading microorganisms were not able to attack 2,4-DB [8, 19], it appears that a new class of microorganisms different from the previous 2,4-D-degrading bacteria isolated under 2,4-D selection are expected to be involved in 2,4-DB degradation in natural soils.

*Corresponding author
Phone: 82-31-290-2475; Fax: 82-31-294-8533;
E-mail: kagil@plaza.snu.ac.kr

COR-AAD1-00001376

In this study, eight 2,4-DB-degrading bacteria and three pairs of syntrophic bacteria capable of co-mineralizing 2,4-DB were isolated from 64 agricultural soils. Their species diversity was investigated by 16S rDNA sequence analysis and REP-PCR patterns of chromosome, and their physiological and genetic properties related to 2,4-DB biodegradation are described.

MATERIALS AND METHODS

Media and Culture Conditions

Peptone-tryptone-yeast extract-glucose (PTYG) medium contained (in grams per liter) 0.25 g of peptone (Difco Laboratories, Detroit, MI, U.S.A.), 0.25 g of tryptone (Difco), 0.5 g of yeast extract (Difco), glucose (0.5), magnesium sulfate (0.03), calcium chloride (0.003), and agar (for plates only, 15). The 2,4-DB mineral medium used for enrichment, isolation, and maintenance of isolates contained (in grams per liter) 2,4-DB (0.3), KH_2PO_4 (1.36), Na_2HPO_4 (1.41), $(NH_4)_2SO_4$ (0.3), $MgSO_4 \cdot 7H_2O$ (0.05), $CaCl_2 \cdot H_2O$ (0.0058), trace metal solution (5 ml per liter), and agar (for plates only, 15). The trace metal solution contained (in grams per liter) $FeSO_4 \cdot 7H_2O$ (0.55), $ZnSO_4 \cdot 7H_2O$ (0.23), $MnSO_4 \cdot 7H_2O$ (0.34), $Co(NO_3)_2 \cdot 6H_2O$ (0.075), $CuSO_4 \cdot 5H_2O$ (0.047), and $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ (0.025). Each of the bacteria showing syntrophic degradation of 2,4-DB was maintained on PTYG medium. PTYG medium was also used for strain purification and colony production for the repetitive extragenic palindromic PCR (REP-PCR). All cultures were incubated at 30°C and liquid cultures were aerated by shaking at 200 rpm on a rotary shaker (Vision Co., Bucheon, Korea).

Chemicals

4-(2,4-Dichlorophenoxy)butyric acid (2,4-DB), 4-(2-methyl-4-chlorophenoxy)butyric acid (MCPB), 4-chlorophenoxyacetic acid (4-CPA), 4-chlorocatechol (4-CC), and chlorohydroquinone (CHQ) were obtained from Aldrich Co., Milwaukee, U.S.A., and analytical grade 2,4-dichlorophenoxyacetic acid (2,4-D), 2,4-dichlorophenol (2,4-DCP), and 2-methyl-4-chlorophenoxyacetic acid (MCPA) were obtained from Sigma Co., St. Louis, U.S.A. 3,5-Dichlorocatechol (3,5-DCC) was kindly donated by Dr. W. Reineke (Bergische Universitat-Gesamthochschul Wuppertal, Germany).

Enumeration and Isolation of 2,4-DB-Degraders

Agricultural soil samples were taken from upland fields in various locations in South Korea. The agricultural fields selected in this study have been under normal agricultural practices, growing dry crops, such as legumes and barley, for more than 10 years. Samples from the top 15 cm of soil were taken, sifted through a 2-mm-pore-size sieve, and kept at 4°C prior to use. A 10-g soil sample from each site was homogenized with 95 ml of sterile 0.85% saline solution by shaking the preparation at 200 rpm on a rotary shaker. Samples (0.1 ml) of appropriate 10-fold dilutions were inoculated into five replicate MPN (most-probable-number) tubes containing 3 ml of 2,4-DB medium (300 ppm). The tubes were incubated at 30°C for 4 weeks and degradation of 2,4-DB was analyzed by a Shimadzu LC-10 series high-performance liquid chromatography (Shimadzu Co., Kyoto, Japan) on μBondapack C18 column (3.9 by 300 mm; Waters, Milford, PA, U.S.A.) and a UV detector set at 278 nm; methanol-0.1% phosphoric acid (70:30) was used as the eluant. The culture of the terminal positive tube showing substantial cell growth and less than 20% of the 2,4-DB remaining was enriched by two additional transfers into fresh medium. Each enriched culture was streaked onto PTYG agar medium and single colonies were then tested for 2,4-DB degradation in fresh 2,4-DB medium before strain purification. The population density of 2,4-DB degraders was estimated from MPN tables [2].

16S rDNA Sequence Analysis

Total genomic DNA was isolated from the isolates and PCR amplification of 16S rRNA genes was performed with 27f and 1492r as previously described [24]. The amplified 16S rRNA genes were sequenced using an ABI Prism BigDye Terminator Cycle Sequencing Ready Kit (Applied Biosystems, Foster City, WI, U.S.A.) according to the manufacturer's instructions with the sequencing primers 27f and 519r [16]. Approximately 400 unambiguous nucleotide positions were used for comparison to the data in GenBank using Basic Local Alignment Search Tool (BLAST) [3]. Sequences from nearest relatives were identified from the Ribosome Database Project (RDP) using the SIMILARITY-RANK program of the RDP [30].

Colony REP-PCR

The colony REP-PCR was performed using BOXA1R as a primer as described previously [9, 36]. Each isolate was grown on the PTYG agar medium for 24 to 48 h, and then a small amount of cells was resuspended in 25 μl of PCR mixture. After PCR amplification, 10 μl samples of the REP-PCR products were separated by electrophoresis on horizontal 1% agarose gels.

Degradative Phenotype Analysis

Each strain was grown in PTYG medium. Cells were then harvested by centrifugation at 10,000 ×g for 10 min at 4°C, washed twice with an equal volume of 15 mM phosphate buffer (pH 7.0), and resuspended in the same buffer. Aliquots of suspended cells were inoculated into culture tubes, each of which contained the mineral medium supplemented with one of the structural analogs at a concentration of 300 μg/ml. After 4 week's incubation, the

COR-AAD1-00001377

cultures were centrifuged to remove the cellular material, and the UV absorption was measured to determine the degradation of phenoxyacetates.

Axenic Culture Experiment

After growth in 2,4-DB medium or PTYG medium, cells were harvested, washed, and resuspended in sodium phosphate buffer. Aliquots of suspended cells were inoculated into duplicate flasks containing 200 ml of 2,4-DB medium at a final density of OD_{550}=0.005. All cultures were incubated at 30°C and were aerated by shaking at 200 rpm on a rotary shaker. Aliquots of the cultures were regularly removed to determine cell growth and degradation of 2,4-DB.

Plasmid Detection and Conjugation

For detection of plasmid DNA, cells were lysed using a modified form [18] of the procedure of Kado and Liu [20]. To analyze the transferability of the 2,4-DB degradative phenotype of the isolates, matings were performed on membrane filters as described by Willetts [39], using *Alcaligenes eutrophus* JMP228 and *Pseudomonas cepacia* DBO1 as the recipients. Transconjugants were selected on 2,4-DB or 2,4-D mineral medium containing appropriate antibiotics, 1.5% Noble agar, and 300 μg/ml of 2,4-DB or 2,4-D, respectively.

PCR Amplification of the *tfd* Genes

The partial gene sequences specific to the 2,4-D degradation pathway were amplified by PCR, with specific primers targeted for the *tfdA*, *B*, and *C* genes of the 2,4-D degradative plasmid pJP4. The PCR cycles and primers for the *tfd* genes were previously reported [4, 21, 35]. The amplification of the *tfdA*, *B*, and *C* genes with the corresponding primers is expected to produce a 362 bp, 205 bp, and 361 bp DNA fragment, respectively.

Respirometric Experiments

Each strain was grown in 2,4-DB medium. Cells were then harvested at 4°C, washed twice, and resuspended in mineral medium. The resuspended cells were kept in ice prior to being added to the respiratory chambers. The initial rates of oxygen uptake were measured with Clark type polarographic oxygen probes of the YSI Model 5300 Biological Oxygen Monitor (Yellow Springs Instrument Co., Yellow Springs, OH, U.S.A.) according to the manufacturer's manual. The sample chamber contained 100–300 ppm substrate and the equivalent of about 0.4 mg dry wt. bacteria in a total volume of 3.0 ml. The experiments were carried out at 28°C in duplicates, and the indicated rates of oxygen uptake have been corrected for indigenous respiratory rates measured in the absence of any added substrates. Protein content was measured by the method of Lowry *et al.* [28], using bovine serum albumin as a standard and activities were expressed as μl O_2/h/mg protein.

RESULTS AND DISCUSSION

Distribution and Isolation of 2,4-DB-Degrading Bacteria

The numbers of 2,4-DB-degrading microorganisms in different agricultural soils were estimated with the MPN procedure. The 2,4-DB-degrading populations ranged from 0 to 18.8 cells/g soil in 64 different agricultural soils, and 82% of the soil samples did not show any detectable degradation of 2,4-DB during the incubation time.

Eight 2,4-DB-degrading bacteria were isolated and purified from the positive MPN tubes (Table 1). 2,4-DB in mineral medium was completely degraded by the isolates, and no dead-end products were accumulated during biodegradation of 2,4-DB, when analyzed with high-performance liquid chromatography. Since 2,4-DB is highly water-soluble and not volatile, its disappearance is due to microbial degradation. Figure 1 shows typical growth and degradation curves of strain DB3 on 2,4-DB medium. However, some enriched cultures failed to produce purified strains able to degrade 2,4-DB. When the different colony types from each of these cultures were combined on 2,4-DB medium, some mixed cultures were able to grow and mineralize 2,4-DB. From these mixed cultures, three pairs of presumably syntrophic bacteria (denoted by a and b) capable of degrading 2,4-DB were isolated from three different soils (Table 1). For each of three syntrophic pairs, both organisms were able to grow together on 2,4-DB medium, but neither organism alone was able to grow on 2,4-DB (Fig. 2). It has been suggested that synergistic degradation of agricultural pesticides in soil could be the important mechanism and pathway of detoxification of environmental contaminants. In the case of organophosphate insecticides, an *Arthrobacter* sp. and a *Streptomyces* sp. have been shown to degrade diazinon synergistically [13]

Table 1. Nearest relatives of the 2,4-DB-degrading isolates based upon 16S rDNA sequence analysis.

Isolate	Nearest relative	% Similarity*
DB1	*Variovorax* sp. WFF52	94
DB2	*Variovorax* sp. WFF52	96
DB3	*Sphingomonas macrogoltabidus*	94
DB4	*Sphingomonas macrogoltabidus*	96
DB5	*Bradyrhizobium japonicum*	98
DB6	*Bradyrhizobium japonicum*	97
DB7	*Bradyrhizobium japonicum*	95
DB8	*Bradyrhizobium japonicum*	96
DB9a	*Sphingomonas* sp. JSS-54	98
DB9b	*Pseudomonas* sp. Isolate CRE 12	99
DB10a	*Variovorax* sp. WFF52	97
DB10b	*Bradyrhizobium japonicum*	97
DB11a	*Variovorax* sp. WFF52	95
DB11b	*Pseudomonas* sp. Isolate CRE 12	96

*Based upon approx. 400 nucleotide positions at the 5' end of the 16S rRNA gene.

COR-AAD1-00001378

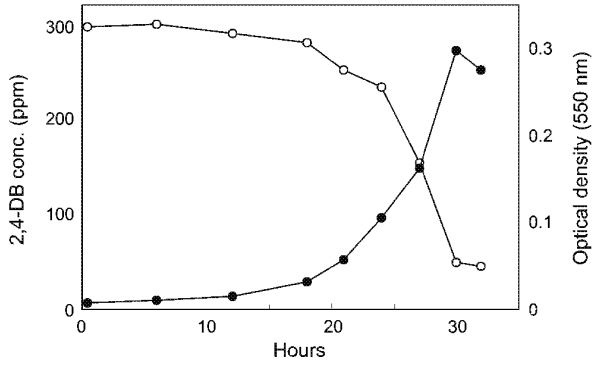


Fig. 1. Disappearance of 2,4-DB (○) and growth of strain DB3 (●) during its growth in mineral medium.
Each point is the mean for two replicate liquid cultures.

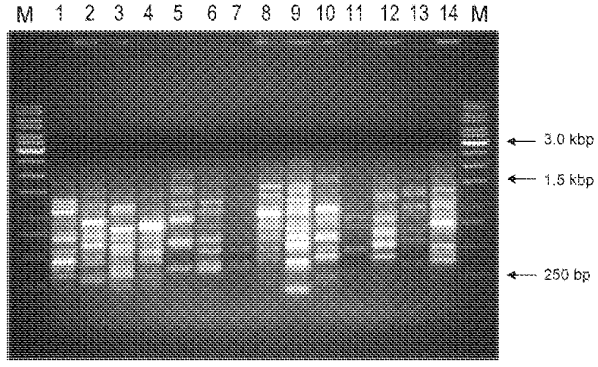


Fig. 3. REP-PCR band patterns of the isolates.
Lanes : 1, DB1; 2, DB2; 3, DB3; 4, DB4; 5, DB5; 6, DB6; 7, DB7; 8, DB8; 9, DB9a; 10, DB9b; 11, DB10a; 12, DB10b; 13, DB11a; 14, DB11b; M, DNA size marker.

and mixed cultures were able to grow on parathion [6, 31]. For the phenoxyalkanoate pesticides, complex enriched cultures have been reported to be able to degrade mecoprop [15, 25]. The results showed that 2,4-DB could also be syntrophically mineralized by two-component cultures.

16S rDNA Sequence and REP-PCR Analyses

When analyzed by 16S rDNA sequences, the isolates were found to be related to members of the genera *Variovorax*, *Sphingomonas*, *Bradyrhizobium*, and *Pseudomonas* (Table 1). All of the isolates were Gram-negative and belonged to the alpha, beta, and gamma subgroups of *Proteobacteria*. Although the 2,4-DB-degrading strains were isolated from different soil locations, some isolates were most closely related to the same species, such as *Variovorax* sp. WFF52 and *Bradyrhizobium japonicum*. A REP-PCR experiment was performed to study the genomic relatedness among the closely related isolates by 16S rDNA sequence analysis. It was revealed that not only did the eight isolates capable of degrading 2,4-DB alone exhibit distinct REP-PCR DNA patterns, but the three pairs of syntrophic bacteria also produced six different DNA fingerprint patterns (Fig. 3), suggesting that each isolate is a distinct strain.

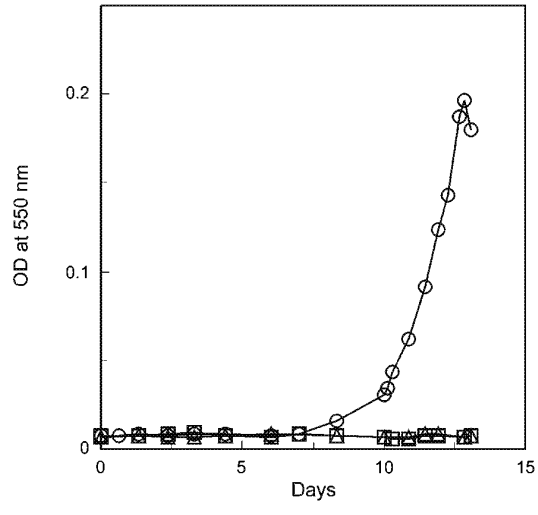


Fig. 2. Syntrophic growth of strains DB11a and DB11b in 2,4-DB mineral medium.
Symbols: ○, DB11a and DB11b; □, DB11a; △, DB11b. OD, optical density.

Growth of 2,4-DB-Degrading Bacteria in Axenic Cultures

To understand axenic growth patterns of the 2,4-DB degraders, the strains were inoculated into 2,4-DB medium after growth on 2,4-DB (inducing) or on PTYG (non-inducing). The growth curves of some representative strains are shown in Fig. 4. Among the isolates capable of degrading 2,4-DB alone, strains DB1 to DB4 grew quickly on 2,4-DB medium, showing doubling times ranging from 3.8 h (strain DB3) to 11.5 h (strain DB1), but strains DB5 to DB8 grew very slowly, taking about 20–30 days for degrading 300 μg/ml of 2,4-DB completely. Under an uninduced condition, the strains DB3 and DB4 exhibited a short lag period (ca. 2 days). Strains DB1 and DB2 exhibited relatively longer lag periods (ca. 5–7 days). By contrast, when adapted to 2,4-DB metabolism, the isolates degraded 2,4-DB much more quickly, suggesting that the 2,4-DB degradative enzymes were inducible by the presence of 2,4-DB. The inducibility of 2,4-DB degradative enzymes of the isolates by 2,4-DB was further confirmed by respirometry experiments conducted with adapted or non-adapted cells: Respiratory activities on 2,4-DB ranged from 8.0 (strain DB6) to 259.1 (strain DB3) μl O_2/h/mg protein for the isolates grown on 2,4-DB (Table 3), but the cells cultivated on glucose as the sole carbon source oxidized the herbicide very slowly, or not at all (data not shown).

COR-AAD1-00001379

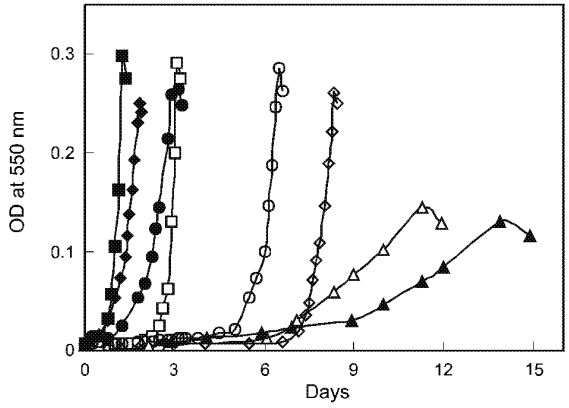


Fig. 4. Growth of 2,4-DB-degrading bacteria on 2,4-DB.
Symbols: ○ and ●, strain DB1; ◇ and ◆, strain DB2; □ and ■, strain DB3; △ and ▲, strain DB11ab. The inoculated bacteria were either adapted (solid symbol) or not adapted (open symbol) to 2,4-DB metabolism. Each point is the mean for two replicate liquid cultures. OD, optical density.

For the syntrophic bacteria, DB9ab showed a growth pattern similar to the strain DB3, but the other syntrophic groups exhibited longer lag periods (ca. 7 days) under uninduced conditions (Fig. 4). When induced with 2,4-DB metabolism, the syntrophic groups (DB10ab and DB11ab) still exhibited slow growth on 2,4-DB medium. From plate counts of each member of the syntrophic pair DB11ab, it was observed that the density of DB11b, which was initially inoculated at the ratio of 1:1 from the non-inducing PTYG medium, was about 4-fold higher than DB11a at the end of exponential growth on 2,4-DB medium. When these adapted cells were washed and transferred into fresh 2,4-DB medium, the density of DB11b declined slightly (ca. half-fold) at the early phase (for the first 9 days) before it resumed its growth, while the other member, DB11a, showed steady increase in plate counts during growth on 2,4-DB. These results may explain the observed slow growth of the adapted syntrophic pair DB11ab on 2,4-DB medium and suggest that DB11a attacks 2,4-DB initially, followed by DB11b which begins to degrade the accumulated intermediate from the middle phase of 2,4-DB degradation.

Degradative Diversity Analysis

The isolates were grown on PTYG medium, and then examined for their ability to degrade other compounds related to 2,4-DB. All of the isolates capable of degrading 2,4-DB independently were also able to mineralize 2,4-D (Table 2), suggesting that the degradation pathways for these two compounds were closely related to each other in these isolates. Strains DB2 and DB3 could utilize MCPB, in addition to 2,4-DB and 2,4-D, and strain DB1 was the most versatile in substrate utilization, using 2,4-DB, 2,4-D, MCPB, MCPA, and 4-CPA as sole carbon sources. By contrast, none of the isolates were able to degrade mecoprop or dichloroprop, which has a propionic acid side chain.

Table 2. Patterns of utilization of substrates by the isolates.

Isolate	Substrate*				
	2,4-DB	2,4-D	MCPB	MCPA	4-CPA
DB1	+	+	+	+	+
DB2	+	+	+	-	-
DB3	+	+	+	-	-
DB4	+	+	+	-	-
DB5	+	+	-	-	+
DB6	+	+	-	-	+
DB7	+	+	-	-	+
DB8	+	+	-	-	-
DB9a	-	-	-	-	-
DB9b	-	-	-	-	-
DB9ab	+	+	+	+	-
DB10a	-	-	-	-	-
DB10b	-	-	-	-	-
DB10ab	+	+	-	-	-
DB11a	-	-	-	-	-
DB11b	-	-	-	-	-
DB11ab	+	+	-	-	-

*Test for the utilization of substrates. +, >80% reduction in peak height as determined by UV scanning and substantial growth (optical density at 550 nm >0.13); -, <15% reduction in peak height and very scant growth (optical density at 550 nm <0.01). 2,4-DB, 4-(2,4-dichlorophenoxy)butyric acid; 2,4-D, 2,4-dichlorophenoxyacetic acid; MCPB, 4-(2-methyl-4-chlorophenoxy)butyric acid; MCPA, 2-methyl-4-chlorophenoxyacetic acid; 4-CPA, 4-chlorophenoxyacetic acid.

Each of the two-member communities was also able to degrade 2,4-DB and 2,4-D syntrophically, while none of those pure cultures could degrade any of the compounds (Table 2). Interestingly, the DB9a and DB9b pair was capable of synergistically mineralizing MCPB and MCPA, in addition to 2,4-DB and 2,4-D. It is not yet clear which strain attacked the herbicide initially and what kinds of intermediates were produced during syntrophic metabolisms. In future studies, more attention will be placed on understanding the detailed mechanisms of 2,4-DB as well as other herbicides and co-degradation by these syntrophic bacteria.

Respirometric Experiments

To investigate the diversity of 2,4-DB degradation pathways among the isolates, each strain and the syntrophic pairs were grown in 2,4-DB medium. They were then analyzed for oxygen uptake in the presence of various substrates, including the assumed intermediates of 2,4-DB, such as 2,4-D, 2,4-DCP, 3,5-DCC, and chlorohydroquinone (Table 3). 2,4-DB-grown cells of strain DB1 rapidly oxidized 2,4-DB, 2,4-D, 2,4-DCP, and 3,5-DCC, suggesting that this strain initially beta-oxidized 2,4-DB to 2,4-D, which was

COR-AAD1-00001380

Table 3. Respiratory activities (μl O_2/h/mg protein) of whole cells of 2,4-DB-degrading bacteria grown on 2,4-DB*.

Substrate†	Bacterial cells										
	DB1	DB2	DB3	DB4	DB5	DB6	DB7	DB8	DB9ab	DB10ab	DB11ab
2,4-DB	171.6	109.1	259.1	211.4	72.1	8.0	31.8	111.0	151.7	55.4	171.3
2,4-D	387.2	18.2	39.5	36.9	55.0	10.8	<1.0	30.7	43.5	7.0	171.7
2,4-DCP	495.8	160.1	95.4	71.5	113.1	<1.0	28.3	114.8	92.4	66.0	30.2
3,5-DCC	367.8	38.4	67.8	44.2	<1.0	<1.0	13.2	102.9	52.3	81.3	70.2
CHQ	24.7	24.4	<1.0	<1.0	37.5	9.6	2.8	74.9	30.9	<1.0	12.9
Glucose	36.4	21.5	<1.0	7.0	<1.0	19.3	<1.0	<1.0	20.6	3.1	10.6

*The values are the mean of two replicate assays of two independent cultures.
†2,4-(2,4-dichlorophenoxy)butyric acid; 2,4-D, 2,4-dichlorophenoxyacetic acid; 2,4-DCP, 2,4-dichlorophenol; 3,5-DCC, 3,5-dichlorocatechol; CHQ, chlorohydroxyquinone; 4-CC, 4-chlorocatechol.

then further mineralized through the previously described 2,4-D degradative pathway of *Alcaligenes eutrophus* JMP134(pJP4) [10]. In contrast, 2,4-DB-grown cells of strain DB2 had much lower oxygen consumption with 2,4-D, while this strain rapidly oxidized 2,4-DCP. These results suggest that strain DB2 might initially have attacked 2,4-DB by cleavage of the ether linkage of the side chain, which has been reported as an alternative pathway for 2,4-DB in a *Flavobacterium* species [29]. Strain DB7 also showed similar patterns to strain DB2 in oxygen consumption with 2,4-D and 2,4-DCP. Among the syntrophic bacteria, group DB10ab showed much higher oxygen consumption with 2,4-DCP than 2,4-D, while group DB11ab had higher oxygen consumption with 2,4-D than 2,4-DCP. It has been suggested that 2,4-DCP could be degraded through chlorohydroxyquinone in *Streptomyces* species [40]. Among the isolates, strains DB5 and DB6 consumed oxygen at a faster rate in the presence of chlorohydroxyquinone as compared to 3,5-DCC, suggesting that chlorohydroxyquinone rather than 3,5-DCC may be the prefered route for these strains. All of the other strains, including three syntrophic pairs, exhibited lower oxygen consumption with chlorohydroxyquinone than with 3,5-DCC, indicating that these isolates degrade 2,4-DB via the formation of 3,5-DCC instead of chlorohydroxyquinone. Most of the cells grown on 2,4-DB showed relatively low oxygen consumption with glucose.

Plasmid Detection and Transferability of 2,4-DB Degradation Phenotype

When the isolates were subjected to Kado's plasmid detection procedure, four strains of the 14 isolates exhibited one or two plasmid DNA bands (Fig. 5). These plasmids were stably maintained in cells cultivated for two months with repeated transfers in Luria-Bertani medium and were not cured with sodium dodecyl sulfate. To investigate whether the 2,4-DB degradative genes are on the plasmid and transmissible to other bacteria, the isolates containing plasmid were mated with two antibiotic-resistant recipients, *Pseudomonas cepacia* DBO1 and *Alcaligenes eutrophus* JMP228. The 2,4-DB degradation phenotype was not transferred to any of the two recipients at a detectable frequency ($<10^{-9}$) from any of the isolates, except strain DB1 which was identified as *Varivorax* sp. by 16S rDNA sequence analysis. The plasmid of DB1 was transferred to *Alcaligenes eutrophus* JMP228, but not to *Pseudomonas cepacia* DBO1, at an average frequency of 1.5×10^{-5}. The transconjugant was able to mineralize 2,4-D, but it could not degrade other phenoxyalkanoates, such as 2,4-DB, MCPB, MCPA, and 4-CPA, while the donor strain DB1 utilized them as sole carbon sources. Such results suggest that in DB1, chromosomal beta-oxidation systems convert 2,4-DB to 2,4-D, which is further degraded by 2,4-D degradative enzymes encoded from the plasmid, and that its chromosome contains indispensable genes for degradation of the other herbicides.

It has been reported that the aliphatic side chains of phenoxyalkyl carboxylic acids, such as 4-(4-chlorophenoxy)-butyric acid and 4-CPA, are metabolized by beta oxidation in certain pure cultures [38]. In the case of 2,4-DB, some plants and soil microflora have been shown to possess beta oxidation systems to convert 2,4-DB to 2,4-D [11, 14], but no pure cultures of bacteria have been reported to carry out beta oxidation for the biodegradation of 2,4-DB. Thus, strain DB1 may use the chromosomal linkage of beta

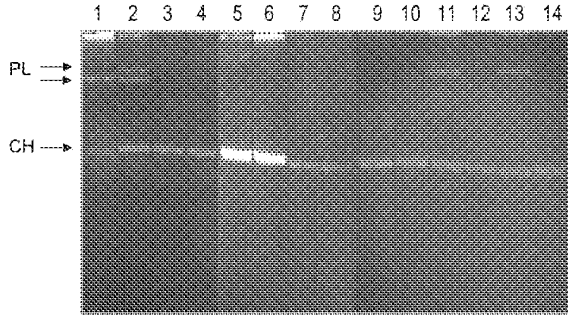


Fig. 5. Plasmid profiles of the isolates.
Lanes: 1, DB1; 2, DB2; 3, DB3; 4, DB4; 5, DB5; 6, DB6; 7, DB7; 8, DB8; 9, DB9a; 10, DB9b; 11, DB10a; 12, DB10b; 13, DB11a; 14, DB11b. PL, plasmid band; CH, chromosomal band.

COR-AAD1-00001381

oxidation systems to 2,4-D degradative plasmid, leading to its complete mineralization.

PCR Amplification of the *tfd* Genes

To investigate whether the 2,4-DB-degrading isolates have any *tfd*-like genes of the 2,4-D degradative plasmid pJP4, PCR amplification was performed for every isolate using specific primers selected from the internal sequences of the *tfdA*, *B*, and *C* genes. Among the 14 isolates, strain DB2 produced an amplified DNA band of the expected size only from the *tfdA* gene, and strain DB9a produced an expected DNA band only from the *tfdC* gene (data not shown). The other isolates did not give any DNA bands with any of the primers of the *tfdA*, *B*, and *C* genes. It has been reported that about 25% of the 47 2,4-D-degrading bacteria isolated under 2,4-D selection had all of the *tfdA*, *B*, *C*, and *D* genes [19]. All of the isolates, which were enriched under 2,4-DB selection, were able to degrade 2,4-D as well as 2,4-DB, but the occurrence frequency of the *tfd* genes in these strains was very low. The results suggest that the isolates have degradative genes dissimilar from those of the previously described 2,4-D-degrading strains, and that different groups of microorganisms are enriched depending on the applied selection, though the enriched strains share a common property of degrading the same compound.

Acknowledgments

We thank William W. Mohn for his critical reading of this manuscript. This study was supported by the Korean Ministry of Education through Research Fund.

References

1. Alexander, M. 1965. Biodegradation: Problems of molecular recalcitrance and microbial fallibility. *Adv. Appl. Microbiol.* **7:** 35–80.
2. Alexander, M. 1982. Most probable number methods for microbial populations, pp. 815–820. *In* A. L. Page (ed.), *Methods of Soil analysis, Part 2: Chemical and Microbiological Properties*, Soil Science Society of America, Wisconsin, U.S.A.
3. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment tool. *J. Mol. Biol.* **215:** 403–410.
4. Berthelet, M., L. G. Whyte, and C. W. Greer. 1996. Rapid, direct extraction of DNA from soils for PCR analysis using polyvinylpolypyrrolidone spin columns. *FEMS Microbiol. Lett.* **138:** 17–22.
5. Biederbeck, V. O., C. A. Campbell, and A. E. Smith. 1987. Effects of long-term 2,4-D field applications on soil biochemical processes. *J. Environ. Qual.* **16:** 257–262.
6. Chaudhry, G. R. and W. B. Wheeler. 1988. Isolation of a methyl parathion-degrading *Pseudomonas* sp. that possesses DNA homologous to the *opd* gene from a *Flavobacterium* sp. *Appl. Environ. Microbiol.* **54:** 288–293.
7. Chang, B. V., L. N. Yeh, and S. Y. Yuan. 1996. Effect of a dichlorophenol-adapted consortium on the dechlorination of 2,4,6-trichlorophenol and pentachlorophenol in soil. *Chemosphere* **33:** 303–311.
8. Chung, M. J. and J. O. Ka. 1998. Isolation and characterization of 2,4-dichlorophenoxyacetic acid-degrading bacteria from paddy soils. *J. Microbiol.* **36:** 256–261.
9. de Bruijn, F. J. 1992. Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. *Appl. Environ. Microbiol.* **58:** 2180–2187.
10. Don, R. H., A. J. Weightman, H. J. Knackmuss, and K. N. Timmis. 1985. Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4). *J. Bacteriol.* **161:** 85–90.
11. Fawcett, C. H., R. M. Pascal, M. B. Pybus, H. F. Taylor, R. L. Wain, and F. Wightman. 1959. Plant growth-regulating activity in homologous series of ω-phenoxyalkane carboxylic acids and the influence of ring substitution on their breakdown by β-oxidation within plant tissues. *Proc. R. Soc.* (London) **B150:** 95–119.
12. Foster, R. K. and R. B. Mckercher. 1973. Laboratory incubation studies of chlorophenoxyacetic acids in chernozemic soils. *Soil Biol. Biochem.* **5:** 333–337.
13. Gunner, H. B. and B. M. Zuckerman. 1968. Degradation of 'diazinon' by synergistic microbial action. *Nature* **217:** 1183–1184.
14. Gutenmann, W. H. and D. J. Lisk. 1964. Conversion of 4-(2,4-DB) to 2,4-dichlorophenoxycrotonic acid (2,4-DC) and production of 2,4-D from 2,4-DC in soil. *J. Agric. Food Chem.* **12:** 322–323.
15. Hallberg, K. B., M. P. Kelly, and O. H. Tuovinen. 1991. Simultaneous degradation of the herbicides 2,4-dichlorophenoxyacetic acid and 2-(2-methyl-4-chlorophenoxy) propionic acid by mixed bacterial cultures. *Curr. Microbiol.* **23:** 65–69.
16. Hugenholtz, P., C. Pitulle, K. L. Hershberger, and N. R. Pace. 1998. Novel division level bacterial diversity in a Yellowstone hot spring. *J. Bacteriol.* **180:** 366–376.
17. Joo, C. S., Y. S. Oh, and W. J. Chung. 2001. Evaluation of bioremediation effectiveness by resolving rate-limiting parameters in diesel-contaminated soil. *J. Microbiol. Biotechnol.* **11:** 607–613.
18. Ka, J. O. and J. M. Tiedje. 1994. Integration and excision of a 2,4-dichlorophenoxyacetic acid-degradative plasmid in *Alcaligenes paradoxus* and evidence of its natural intergeneric transfer. *J. Bacteriol.* **176:** 5284–5289.
19. Ka, J. O., W. E. Holben, and J. M. Tiedje. 1994. Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D-treated field soils. *Appl. Environ. Microbiol.* **60:** 1106–1115.

COR-AAD1-00001382

20. Kado, C. I. and S. T. Liu. 1981. Rapid procedure for detection and isolation of large and small plasmids. *J. Bacteriol.* **145:** 1365–1373.
21. Kamagata, Y., R. R. Fulthorpe, K. Tamura, H. Takami, L. J. Forney, and J. M. Tiedje. 1997. Pristine environments harbor a new group of oligotrophic 2,4-dichlorophenoxyacetic acid-degrading bacteria. *Appl. Environ. Microbiol.* **63:** 2266–2272.
22. Kilbane, J. J., D. K. Chatterjee, A. M. Chakrabarty. 1983. Detoxification of 2,4,5-trichlorophenoxyacetic acid from contaminated soil by *Pseudomonas cepacia*. *Appl. Environ. Microbiol.* **45:** 1697–1700.
23. Kobayashi, H. and B. E. Rittman. 1982. Microbial removal of hazardous organic compounds. *Environ. Sci. Technol.* **16:** 170–183.
24. Lane, D. J. 1991. 16S/23S rRNA sequencing, pp. 115–148. *In* E. Stackebrandt and M. Goodfellow (eds.), *Nucleic Acid Techniques in Bacterial Systematics*, John Wiley and Sons, Chichester, England.
25. Lappin, H. M., M. P. Greaves, and J. H. Slater. 1985. Degradation of the herbicide mecoprop [2-(2-methyl-4-chlorophenoxy)propionic acid] by a synergistic microbial community. *Appl. Environ. Microbiol.* **49:** 429–433.
26. Lee, H. S. and K. Lee. 2001. Bioremediation of diesel-contaminated soil by bacterial cells transported by electrokinetics. *J. Microbiol. Biotechnol.* **11:** 1038–1045.
27. Loos, M. A., I. F. Schlosser, and W. R. Mapham. 1979. Phenoxy herbicide degradation in soils: Quantitative studies of 2,4-D and MCPA-degrading microbial populations. *Soil Biol. Biochem.* **1:** 377–385.
28. Lowry, O. H., H. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193:** 265–275.
29. Macrae, I. C., M. Alexander, and A. D. Rovira. 1963. The decomposition of 4-(2,4-dichlorophenoxy)butyric acid by *Flavobacterium* sp. *J. Gen. Microbiol.* **32:** 69–76.
30. Maidak, B. L., J. R. Cole, T. G. Lilburn, C. T. Parker Jr, P. R. Saxman, J. M. Stredwick, G. M. Garrity, B. Li, G. J. Olsen, S. Pramanik, T. M. Schmidt, and J. M. Tiedje. 2000. The RDP (Ribosomal Database Project) continues. *Nucleic Acids Res.* **28:** 173–174.
31. Munnecke, D. M. and D. P. H. Hsi. 1976. Pathways of microbial metabolism of parathion. *Appl. Environ. Microbiol.* **31:** 63–69.
32. Smith, A. E. and A. J. Aubin. 1991. Effects of long-term 2,4-D and MCPA field applications on the soil breakdown of 2,4-D, MCPA, mecoprop, and 2,4,5-T. *J. Environ. Qual.* **20:** 436–438.
33. Smith, A. E. and B. J. Hayden. 1981. Relative persistence of MCPA, MCPB and mecoprop in Saskatchewan soils, and the identification of MCPA in MCPB-treated soils. *Weed Res.* **21:** 179–183.
34. Torstensson, N. T. L., J. Stark, and B. Goransson. 1975. The effect of repeated applications of 2,4-D and MCPA on their breakdown in soil. *Weed Res.* **15:** 159–164.
35. Vallaeys, T., R. R. Fulthorpe, A. M. Wright, and G. Soulas. 1996. The metabolic pathway of 2,4-dichlorophenoxyacetic acid degradation involves different families of *tfdA* and *tfdB* genes according to PCR-RFLP analysis. *FEMS Microbiol. Ecol.* **20:** 163–172.
36. Versalovic, J., M. Schneider, F. J. de Bruijn, and J. R. Lupski. 1994. Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol. Cell Biol.* **5:** 25–40.
37. Wain, R. L. 1955. Herbicidal selectivity through specific action of plants on compounds applied. *J. Agric. Food. Chem.* **3:** 128–130.
38. Webley, D. M., R. B. Duff, and V. C. Farmer. 1957. Formation of β-hydroxy-acid as an intermediate in the microbiological conversion of monochlorophenoxybutyric acid to the corresponding substituted acetic acids. *Nature* (London) **179:** 1130.
39. Willetts, N. S. 1988. Conjugation. *Methods Microbiol.* **21:** 49–77.
40. Zaborina, O., M. Latus, J. Eberspacher, L. A. Golovleva, and F. Lingens. 1995. Purification and characterization of 6-chlorohydroxyquinol 1,2-dioxygenase from *Streptomyces rochei* 303: comparison with an analogous enzyme from *Azotobacter* sp. strain GP1. *J. Bacteriol.* **177:** 229–234.

COR-AAD1-00001383

Mazur, E. 1984. "Annual Rates of Deposition of Polycyclic Aromatic Hydrocarbons in Different Forest Ecosystems." *Water, Air Soil Pollut.* 21(0): 425-434.

McMahon, C. K. and S. M. Tomkins. 1975. "Polycyclic Aromatic Hydrocarbons in Forest Fire Smoke." *Environmental Letters* 8 (1): 61-73.

Menzie, C. A., B. B. Potocki, and J. Santodonato. 1992. "Exposure to Carcinogenic PAHs in the Environment." *Environ. Sci. Technol.* 26(0): 1278-1284.

Neff, M. J. 1979. "The Environmental Load and Cycle of Polycyclic Aromatic Hydrocarbons." *Sci. Tot. Environ.* 6 (1): 239-350.

Wakeham, S. G., C. Schaffner, and W. Giger. 1980. "Polycyclic Aromatic Hydrocarbons in Recent Lake Sediments-I Compounds Having Anthropogenic Origin." *Geochim. Cosmochim. Acta* 44 (0): 403-413.

Ward, D. E., J. Peterson, and W. M. Hao. 1993. "An Inventory of Particulate Matter and Air Toxic Emissions from Prescribed Fires in the USA for 1989." In *Proceedings of the Air and Waste Management Association, 1993 Annual Meeting and Exhibition*, Denver, CO.

Wanless, H. J. G. and R. A. Hites 1980. "Polycyclic Aromatic Hydrocarbons in Gulf of Maine Sediments and Nova Scotia Soils." *Geochim. Cosmochim. Acta* 44(0): 37-43.

COR-AAD1-00001384

DEGRADABILITY AND RECALCITRANCE OF PHENOXYALKANOIC ACID HERBICIDES – TRAITS OF THE MICROBIAL METABOLISM

Roland H. Mueller and Wolfgang Babel

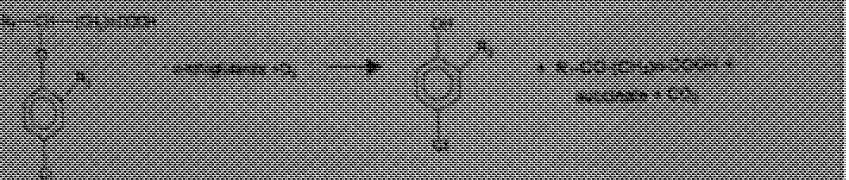
UFZ Centre for Environmental Research, D-04318 Leipzig, Germany

ABSTRACT: Cleavage of phenoxyalkanoic herbicides via an α-ketoglutarate-dependent dioxygenase mechanism necessitates the regeneration of this oxygen acceptor from the cleavage products if degradation is to be continuous. Although no limitation occurs when degradation is complete, this is not necessarily the case in the case of incomplete degradation. Theoretical considerations indicate that the regeneration of α-ketoglutarate is not possible under these conditions with 2,4-D as the substrate. For stoichiometric reasons regeneration is possible when phenylpropionate and lactate derivatives are used, although competitive reactions mean that cleavage ought to gradually cease. This thesis was experimentally verified with *Ralstonia eutropha* MC1, a strain which exhibits constitutive expression of α-ketoglutarate-dependent phenoxy cleavage activities for phenoxyacetate and enantiomeric phenoxypropionate herbicides. Under conditions of incomplete substrate degradation leaving the phenolic intermediate as a dead-end product, almost no cleavage of 2,4-dichlorophenoxyacetate (2,4-D) was observed, whereas the two enantiomers of 2-(2,4-dichlorophenoxy)-propionate (DCPs- and (S)-2,4-DP) were cleaved to a limited extent. This seems to correlate with the availability of α-ketoglutarate. Education of the pools of metabolites decreased the degree of cleavage, this reaction being reversed and supported by supplying extra metabolites such as α-ketoglutarate. A shortage of α-ketoglutarate is considered to be a reason for the mass recalcitrance of phenoxyalkanoic herbicides under defined metabolic conditions.

INTRODUCTION

The first enzymatic step in the degradation of 2,4-D, encoded by the gene *tfdA*, has been shown to be catalyzed by an α-ketoglutarate-dependent dioxygenase: oxidative decarboxylation of the electron acceptor results in the formation of succinate, whereas 2,4-D is oxidatively cleaved to 2,4-dichlorophenol (DCP) and glyoxylate (Fukumori and Hausinger 1993). This type of reaction, as studied with the enzyme of *Ralstonia eutropha* JMP134, has been shown to be widespread in nature, showing tfdA-homologous genes in various phylogenetic branches (Hogan et al. 1997; McGowan et al. 1998). Besides 2,4-D, phenoxypropionate herbicides, as shown with *Sphingomonas herbicidovorans* (Zipple et al. 1993), *Rhodoferax* sp. P230 (Lang et al. 1997) and *Ralstonia eutropha* MC1 (Müller et al. 1999).

A general equation of the cleavage of phenoxyalkanoates by this type of reaction reads as follows with R1 = H, CH_3, R2 = Cl, CH_3, n = 0, 2, in respective combinations for the usual herbicides.



As the initiation of the degradation of phenoxyalkanoates via this step requires α-ketoglutarate for continual substrate utilization, it must be regenerated. This has consequences for substrate consumption and growth, especially with strains exhibiting only partial substrate degradation by excreting the phenolic moiety as a dead-end product. The released alkanoic acids are the only sources in these cases for both regeneration of the electron acceptor from succinate and for growth (Müller and Babel 2000). From a global survey on the distribution of 2,4-D-utilizing bacteria, it became evident that this metabolic performance is not so much restricted to individual strains in natural biotopes as bound to consortia (Tiedje et al. 1987).

Objective. Availability of α-ketoglutarate is a prerequisite for continual substrate consumption. On the basis of the carbon flows it is shown in which manner this is impaired or corrected respectively with the various phenoxyalkanoic acid herbicides under different metabolic conditions especially referring to incomplete substrate degradation. The theoretically based considerations will be substantiated by experimental results showing the effect of the regeneration of α-ketoglutarate and its additional supply on the cleavage of these herbicides.

MATERIALS AND METHODS

The theoretical considerations for balancing substrate conversion and biomass synthesis were performed on the basis of the Y$_{ATP}$-concept (Stouthamer 1973, Babel and Müller 1985), as shown in detail with respect to phenolic and chlorophenolic compounds (Müller and Babel 1994, 2000).

The experimental investigations were performed with *C. pinidorans* AC1 (Müller et al. 1994). This strain exhibits constitutive expression of α-ketoglutarate-dependent dioxygenase activity with respect to (R) 2,4-DP with one enzyme, and with respect to (S) 2,4-DP and 2,4-D with another enzyme (Müller and Babel 1996). The influence of extra metabolites added to the cell suspension upon the cleavage of phenoxyalkanoates was investigated in the presence of acetylene (56 mg/l) in order to avoid induction of the whole degradative pathway in the model.

RESULTS AND DISCUSSION

Theoretical Considerations. The metabolism of phenoxyalkanoates and the formation of 3-phosphoglycerate (PGA) as a central metabolite of biomass synthesis including enzymatic cleavage of the herbicides via α-ketoglutarate-

dependent dioxygenase are illustrated in Fig. 1 using the example of 2,4-DP and taking into account both complete and incomplete degradation. Evidently and empirical, cleavage by an α-ketoglutarate-dependent mechanism makes available metabolites which may be used as a source of carbon and energy for growth on respective herbicides as it holds true also by taking into account a monooxygenase-mediated cleavage reaction (Müller and Babel 1994). By contrast, in the case of incomplete degradation by forming DCP as a dead-end product, the potential degradation of 2,4-DP can at best be measured according to stoichiometric factors. Under defined metabolic conditions the carbon made available from the metabolites liberated by the cleavage reaction will be completely lost as CO_2 thus preventing growth. As side reactions will make use of these metabolites for the regeneration of α-ketoglutarate should be limited and the cleavage reaction will gradually cease, as has been observed with certain strains known under the latter conditions. I.e. net synthesis of metabolites (pyruvate), would only be possible if extra carbon was fixed as CO_2 by carboxylase reactions. The maximum carbon available for biomass synthesis (C_{max}) under the different metabolic conditions considered is shown in Table 1.

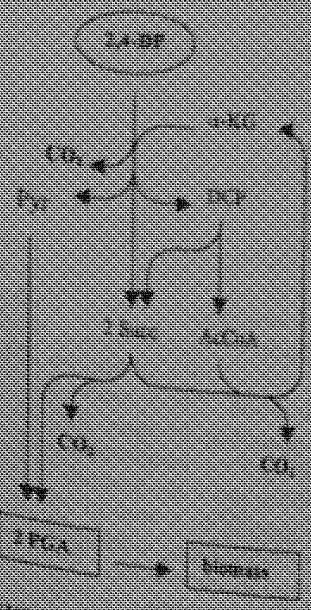


Figure 1. Fate of 2,4-DP and regeneration of α-ketoglutarate by complete (left) and incomplete (right) degradation of the substrate.

Abbreviations: Pyr, pyruvate; AcCoA, acetyl-coenzyme A; Succ, succinate.

Even more markedly, the cleavage reaction would not work with 2,4-D in the case of the incomplete degradation of this substrate due to the inability to regenerate α-ketoglutarate from the cleavage products succinate and glyoxylate.

COR-AAD1-00001387

Thus the substrate may prove to be apparently inert despite the presence of THA (Table 1). The carbon balance of the primary assimilatory pathway of various phenoxyalkanoic up to the synthesis of a central carbon precursor (PGA) for mureine synthesis are listed in Table 1.

TABLE 1. Carbon balance and maximum carbon conversion efficiency of phenoxyalkanoic acid utilization.

Substrate	Primary assimilation	Precursor synthesis	CCE_{max}
Complete degradation			
2,4-D	1 Succ $+ 4\,CO_2$	1 PGA $+ 3\,CO_2$	0.375
2,4-DP	1 Succ + 1 Pyr $+ 2\,CO_2$	2 PGA $+ 3\,CO_2$	0.667
2,4-DB	2 Succ $+ 2\,CO_2$	2 PGA $+ 4\,CO_2$	0.6
Incomplete degradation			
2,4-D	(no oxidative cleavage)		
$2,4\text{-DP}^{pyr}$		$3\,CO_2$ + [DCP]	no PGA 0
$2,4\text{-DP}^{pyr,ox}$	1 Glyox $+ 1\,CO_2$	+ [DCP]	0.5 PGA $+ 1.5\,CO_2$ 0.167
$2,4\text{-DP}^{ox}$	$+ 4\,CO_2$	+ [DCP]	no PGA 0
$2,4\text{-DB}^{ox}$	1 Glyox $+ 2\,CO_2$	+ [DCP]	0.5 PGA $+ 2.5\,CO_2$ 0.15

Assimilation via pyruvate dehydrogenase (PDH), pyruvate carboxylase (PC), malate lyase (ML) or malic enzyme (ME), respectively.

Experimental verification. Incubation of strain MC1 in the presence of streptomycin resulted in the cleavage of the various herbicides while excreting DCP (Fig. 3). The extent of cleavage differs whereas (R)-2,4-DP was completely cleaved, 2,4-D remained almost unaffected, matching the predictions. With the S-enantiomer of 2,4-DP, cleavage proceeded more slowly and later on stopped. Apparently the rate of cleavage influences its extent. This is shown in a more detail in Fig. 5 using various substrates. According to these data a linear dependency becomes evident which means that a fixed proportion between the time constants exists of the various reactions taking part in this process.

When correlating a given amount of bacteria to a pool of intracellular metabolites equivalent to a given potential for regenerating α-ketoglutarate, the extent of cleavage ought to correspond to the biomass concentration. This holds as it is shown in Fig. 4. The amount of substrate cleaved was about 7-fold higher with (R)-2,4-DP than with (S)-2,4-DP. As both enantiomers lead to the same reaction products, this differing degree of cleavage is probably caused by lacking intermediates due to general maintenance requirements, which should become more pronounced the lower the cleavage rate. After the exhaustion of suitable metabolites, cleavage ceased. However, the interference of secondary effects in this process (e.g. due to the accumulation of DCP) cannot be ruled out. The extent

of cleavage of 2,4-D is more than 30-fold lower than R,2,4-CP. This low possibility of regenerating α-ketoglutarate seems to be rather limited (see Eq. 1)

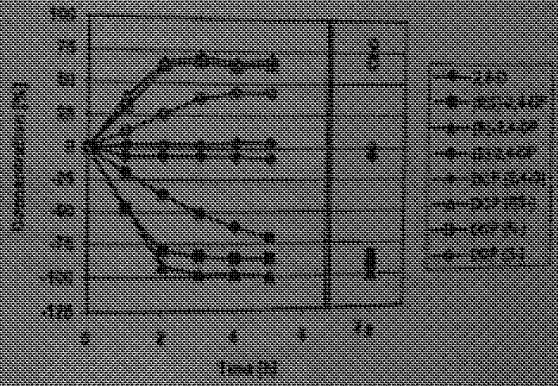


FIGURE 2. Cleavage of various phenoxyalkanoates (S₀ = 0.05 mM) and excretion of DCP by *C. acidovorans* MC1 in the presence of oxoglutarate.

The supply of additional metabolites should improve cleavage. This is shown in Fig. 5. Incubation of strain MC1 in the absence of substrate after growth on a complex carbon source resulted in a reduction in the extent of cleavage, which was almost zero after 4 h. The supply of α-oxoglutarate after one lag restored cleavage activity in an almost additive manner. Pre-incubation in the presence of α-ketoglutarate clearly improved the extent of cleavage. The plot holds for 2,4-D.

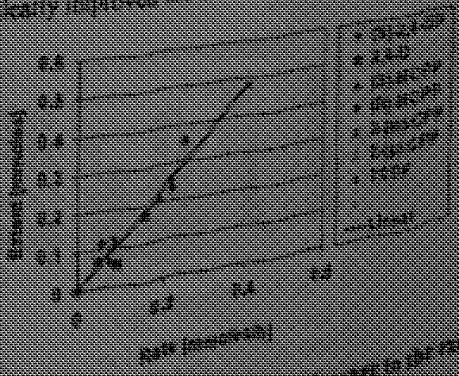


FIGURE 3. Correlation of the rate of cleavage to the rate of oxoglutarate excretion using various phenoxyalkanoates. 2,4-D, 2,4-dichlorophenoxyacetate; 2-(*o*)-CPP and 2-(*p*)-CPP, 2-(*o*-chlorophenoxy)- and 2-(*p*-chlorophenoxy)propionate, respectively; TCPP, 2,4,5-trichlorophenoxypropionate; for the other abbreviations see text.

COR-AAD1-00001389

the degradation of which was improved from zero in the absence of α-ketoglutarate to a significant extent after adding this metabolite with both cells obtained both after growth on complex medium (Table 2) and cells from stationary growth phase after cultivation on α-ketoglutarate (Table 2, values in brackets).

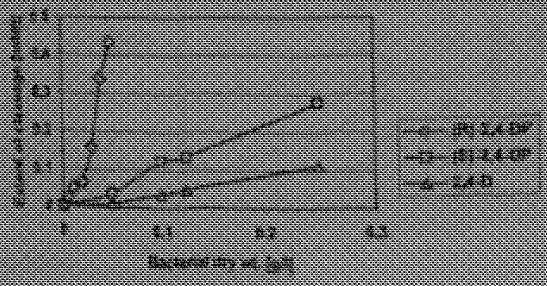


FIGURE 4. Correlation of the extent of cleavage to biomass concentration.

The results show that strains exhibiting TfdA-like activity without further metabolizing the phenolic intermediate may behave inertly with respect to phenylurea/amide herbicides. Thus any negative results in enrichment cultures concerning degradation might not exclusively be due to the absence of respective cleavage activity, but might instead be attributable to metabolic limitations.

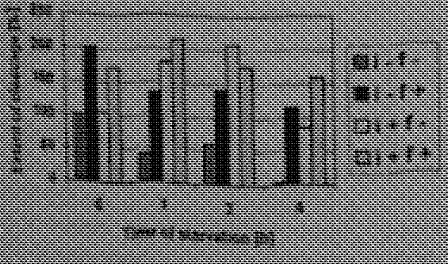


FIGURE 5. Effect of pretreatment and addition of α-ketoglutarate on the extent of cleavage of S-2,4-DP.

(-) addition of (S)-2,4-DP after time of starvation without (-) or with (+) 0.5 mM α-KG; (+) pre-incubation for 4 h on 0.5 mM α-KG after time of starvation before adding (S)-2,4-DP without (-) or with (+) a second quantity of 0.5 mM α-ketoglutarate.

COR-AAD1-00001390

Patterns in which limited degradation was occasionally observed when characterizing respective isolates correlates with this behavior. Vice versa, the sudden onset of degradation of phenoxy herbicides after several weeks' acclimation might possibly originate from genetic rearrangements in soil consortia by completing the pathway within one strain. Moreover, as this cleavage reaction may be regarded as a sink of the cell's pool of metabolites, maintenance failures may be regarded as a sink of the cell's pool of metabolites, maintenance failures may negatively be influenced targeting the survival of a given strain which may thus not only be caused by the potential accumulation of the toxic chlorophenolic intermediates.

TABLE 2. Influence of α-ketoglutarate on the cleavage of phenoxyalkanoates

Substrate	α-ketoglutarate	Rate (nanomol x g^{-1} dry wt. x min^{-1} (nanomol)	Extent
(R)-2,4-DP	–	0.71 (0)	> 0.45 (0)
(R)-2,4-DP	+	0.975 (0.275)	> 0.45 (0.11)
(S)-2,4-DP	–	0.15	0.065
(S)-2,4-DP	+	0.13	0.18
2,4-D	–	0 (0)	0 (0)
2,4-D	+	0.18 (0.14)	0.10 (0.07)

REFERENCES

Rawn, W. and R.H. Muller 1985. Correlation between cell composition and carbon conversion efficiency in microbial growth a theoretical study. *Appl. Microbiol. Biotechnol.* 22: 201-207.

Fukumori, F. and R.P. Hausinger (in) *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetate monooxygenase" is an α-ketoglutarate dependent dioxygenase. *J. Bacteriol.* 175: 2083-2086.

Hogan, D.A., Buckley, D.H., Nakatsu, C.H., Schmidt, T.M., and R.P. Hausinger 1997. Distribution of the *tfdA* gene in soil bacteria that do not degrade 2,4-dichlorophenoxyacetic acid (2,4-D). *Microb. Ecol.* 34: 60-68.

McGowan, C., Fulthorpe, R., Wright, A., and J.M. Tiedje 1998. Evidence for interspecies gene transfer in the evolution of 2,4-dichlorophenoxyacetic acid degraders. *Appl. Environ. Microbiol.* 64: 4089-4092.

Müller, R.H., and W. Babel. 1994. Phenol and its derivatives as heterotrophic substrates for microbial growth - an energetic comparison. *Appl. Microbiol. Biotechnol.* 42: 445-451.

Müller, R.H., and W. Babel. 1996. Separation of two dichlorprop-(2-oxoglutarate) dioxygenases with enantiospecific properties from *Delftia acidovorans* MC1. *Acta Biotechnol.* 16: 345-355.

Müller, R.H., S. Jorks, S. Kleinsteuber, and W. Babel. 1999. *Delftia acidovorans* MC1: a new isolate capable of degrading the chiral herbicides dichlorprop and mecoprop and the herbicides 2,4-D and MCPA. *Microbiol. Res.* 154: 241-246.

Müller, R.H., and W. Babel. 2000. A theoretical study on the metabolic requirements due to α-ketoglutarate-dependent cleavage of phenoxyalkanoates. *Appl. Environ. Microbiol.* in: 139-144.

Nickel, K., M.H. Suss, and H.P.E. Kohler. 1997. Involvement of two α-ketoglutarate dependent dioxygenases in enantioselective degradation of (R)- and (S)-mecoprop by *Sphingomonas herbicidovorans* MH. *J. Bacteriol.* 179: 6674-6679.

Stouthamer, A.H. 1973. A theoretical study on the amount of ATP required for the synthesis of microbial cell material. *Antonie van Leeuwenhoek* 39: 545-565.

Tiedje, J.M., R. Fulthorpe, V. Kanagawa, C. McGowan, A. Rhodes, and H. Tarsen. 1997. What is the global pattern of chloroaromatic degrading microbial populations? In K. Horikoshi, M. Fukuda and T. Kudo (Eds). *Microbial diversity and genetics of biodegradation* pp 63-74. S. Karger, Basel, Switzerland.

COR-AAD1-00001392

Biodegradation **8:** 43–52, 1997.
© 1997 *Kluwer Academic Publishers. Printed in the Netherlands.*

Biodegradation of the chlorophenoxy herbicide (*R*)-(+)-mecoprop by *Alcaligenes denitrificans*

Vanessa A. Tett, Andrew J. Willetts[1] & Hilary M. Lappin-Scott[1]
Research School of Biosciences, University of Kent at Canterbury, Canterbury, CT2 7NJ, Kent, UK; [1] *Department of Biological Sciences, Exeter University, Hatherly Laboratories, Exeter, EX4 4PS, Devon, UK*

Accepted 31 January 1997

Key words: Alcaligenes denitrificans, biodegradation, chloroaromatic, mecoprop (*R*)-(+)-2(2-methyl-4-chlorophenoxy)propionic acid

Abstract

An *Alcaligenes denitrificans* strain capable of utilizing the herbicide (*R*)-(+)-2(2-methyl-4-chlorophenoxy)propionic acid (mecoprop) as a sole carbon source was isolated from soil and cultured in liquid medium. Crude cell extracts of the bacterium were utilized in spectrophotometric assays to elucidate a biochemical pathway for degradation of mecoprop. Results indicated a reaction sequence analogous to the degradation of 2,4-dichlorophenoxyacetic acid (2,4-D). GC-MS analysis provided direct evidence for the biotransformation of mecoprop to the transient metabolite 4-chloro-2-methylphenol (MCP). No NADPH-dependent activity was observed during this reaction. Pyruvate was verified as the second product derived from the aliphatic side chain of mecoprop. MCP was subsequently transformed to a substituted catechol by an NADPH-dependent monooxygenase. When grown on mecoprop, *A. denitrificans* was adapted to oxidize catechol and its 4- and 3-methylated derivatives indicating the broad substrate specificity of catechol dioxygenase. The microorganism was demonstrated to adopt the *ortho* mechanism of aromatic cleavage which resulted in the formation of 2-methyl-4-carboxymethylene but-2-en-4-olide, a reaction intermediate of the β-ketoadipate pathway.

Introduction

The structurally related chlorinated phenoxy alkanoics, including 2,4-dichlorophenoxyacetic acid (2,4-D), 2-methyl 4-chlorophenoxyacetic acid (MCPA) and 2(2-methyl 4-chlorophenoxy) propionic acid (mecoprop) are among the most widely used herbicides to control broad leaved weeds in cereal crops throughout the world (Anonymous 1992; Worthing & Walker 1991). Many soil bacteria have been found to be able to grow on 2,4-D and MCPA as the sole source of organic carbon, and this has contributed to the successful elucidation of biodegradation pathways for the complete mineralization of these xenobiotics (Bollag et al. 1968a, b; Gamar & Gaunt 1971; Gaunt & Evans 1971a, b; Loos et al. 1967a, b; Tiedje & Alexander 1969).

It is widely accepted that the microbiological decomposition of herbicides in top soils occurs largely via aerobic metabolism (Sinton et al. 1986). Enzyme activities of an *Arthrobacter* species suggested that 2,4-D degradation was initiated by side chain cleavage of 2,4-dichlorophenol (Loos et al. 1967a, b). The phenol was hydroxylated to 3,5-dichlorocatechol by the *Arthrobacter* species and the catechol was oxidized to *cis, cis*-2,4-dichloromuconic acid, a step that effects an *ortho* cleavage of the aromatic ring. *Cis, cis*-2,4-dichloromuconic acid then underwent lactonization to 2-chloro-4-carboxymethylene but-2-en-4-olide with release of the second chloride ion and further conversion to chloromaleylacetate. Finally, chloromaleylacetate was transformed into succinyl-CoA enabling the ring carbon to be assimilated for microbial growth. Essentially the same pathway was proposed for MCPA degradation by a *Pseudomonas* species (Evans et al. 1971a).

On account of its chemical structure, mecoprop is considered likely to be a more recalcitrant compound than 2,4-D or MCPA, since the propionic acid

COR-AAD1-00001393

side chain is linked to the oxygen atom of the phenoxyacid moiety via the α-carbon atom adjacent to the carboxyl group (Kilpi 1980). Several mechanisms have been proposed for the degradation of mecoprop based on the precedents of 2,4-D and MCPA degradation but no pathway has yet been elucidated (Oh & Tuovinen 1991a, b, 1994). Lindholm and coworkers (1982) suggested decarboxylation of the propionic acid side chain of mecoprop to form MCPA which could then be degraded by the previously proposed pathway for the phenoxyacetic herbicide (Lindholm et al. 1982). Confirmation that phenoxyalkanoic acids with side chains longer than two carbon atoms can be degraded by a pathway involving some element of side-chain cleavage was provided by the observation that a *Flavobacterium* species isolated from soil degraded 4-(2,4-dichlorophenoxy)butyric acid with formation of 2,4-dichlorophenol (MacRae et al. 1963).

Research into the biodegradative pathway of mecoprop has been hindered by the difficulty of obtaining pure cultures of mecoprop-degrading bacteria. We have recently reported the stereoselective degradation of (*R*)-(+)-mecoprop by *Alcaligenes denitrificans* (Tett et al. 1994). The bacterium is capable of growth on (+)- and racemic mecoprop as the sole carbon and energy source, and can also degrade 2,4-D and MCPA. In this paper we report initial investigations to elucidate the biodegradative pathway of mecoprop.

Materials and methods

Bacteria and culture conditions

Alcaligenes denitrificans, able to utilize mecoprop as the sole carbon and energy source, was isolated from soil as described by Tett et al. 1994. The bacterium was grown in minimal medium containing (*R*)-(+)-mecoprop as the organic carbon source (Tett et al. 1994) unless indicated otherwise.

Chemicals

Gibbs reagent (2,6-dichloroquinone-4-chloroimide) and the reduced form of β-nicotinamide adenine dinucleotide were obtained from Sigma Chemical Company Ltd., (Dorset, UK), MCP was obtained from Aldrich Chemicals (Gillingham, Dorset, UK). Catechol, 4-methylcatechol, 3-methylcatechol and 4-carboxymethylbut-2-en-4-olide were kindly donated by Prof. D.W. Ribbons; 5-chloro-3-methylcatechol and 2-methyl-4-carboxymethylene but-2-en-4-olide were donated by Prof. H. Knackmuss. *Cis, cis*-muconic acid was synthesized using the amended method of Boeseken and Engelberts (Pandell 1976) and the identity of the product confirmed by $[H]^1$-NMR. *Cis, cis*-muconic acid was also produced through the peracetic acid oxidation of catechol. The (*R*)-(+)-isomer of mecoprop was used throughout this study and was obtained as a donation from Pennine Chemical Services Ltd. (Huddersfield, West Yorkshire, UK).

Measurement of gross mecoprop degradation

The degradation of mecoprop was measured directly by monitoring the change in the maximum absorption peak of mecoprop (A_{279}) in a scanning UV spectrophotometer. The culture supernatant was scanned from 200–500 nm to determine both the breakdown of mecoprop and the accumulation of biotransformation products.

Enzyme assays

A 24 h-culture of *A. denitrificans* was used to inoculate mecoprop medium at a concentration of 1 g l^{-1}. The batch culture was incubated overnight at 30 °C and sparged with air (4 l min^{-1}). When the microbial population had reached mid log phase (16–21 h), the cells were harvested by centrifugation at 5000 g for 15 min. The cell pellet was washed once with 50 mM sodium phosphate buffer pH 7.0 and respun under the same conditions as above. The cells were resuspended in a minimal amount of buffer to produce a concentrated cell suspension of at least 250 mg cells ml^{-1}. Cell lysis was achieved by sonication on ice at 22 × 10^{-1} nm for 4 × 30 seconds. The cell debris was removed by centrifugation at 12,000 *g* at 4 °C for 15 min, then further centrifuged at 300,000 *g* at 4 °C for 90 min. The supernatant which was used immediately for measurements of enzyme activities had a protein concentration of 11 mg ml^{-1}. All enzymes were assayed spectrophotometrically at 25 °C using an equivalent no-substrate control.

The activity of mecoprop monooxygenase in the cell extract was measured by the decrease in absorbance at 340 nm due to oxidation of the added cofactor NAD(P)H. The reaction mixture contained: 800 μl of 50 mM phosphate buffer pH 7.0, 20 μl of 10 mM NAD(P)H, 50–200 μl of enzyme extract, 20 μl of 50 mM mecoprop and distilled water to a total volume of 1 ml. The activity of MCP monooxygenase

COR-AAD1-00001394

was determined in an identical manner but substituting mecoprop with 20 μl of 50 mM MCP.

Catechol dioxygenase was assayed by a modification of the method described by Hegeman (Hegeman 1966). Each cuvette contained: 20 μl of 25 mM an aromatic cis-diol (catechol, 4-methylcatechol or 3-methylcatechol), 20 μl of crude cell extract, 900 μl of 50 mM phosphate buffer pH 7.0 and distilled water to a total volume of 1 ml.

The activity of muconate lactonizing enzyme was followed by the rate of decrease in muconic acid as assayed at its λmax. The assay mixture contained: 60 μl of 2.5 μM *cis, cis*-muconic acid, 10 μl of cell extract, 20 μl of 3 mM $MnCl_2$ solution and 50 mM Tris-Cl buffer pH 8.0 to a total volume of 1 ml.

The delactonizing enzyme that is responsible for the metabolism of 2-methyl-4-carboxymethylene but-2-en-4-olide was assayed by following the decrease in the maximum absorbance of the analogue 4-carboxymethylbut-2-en-4-olide (230 nm). The reaction mixture contained: 60 μl of 25 mM 4-carboxymethylbut-2-en-4-olide, 10 μl crude cell extract and Tris-Cl buffer pH 8.0 to a total volume of 1 ml.

Detection of phenols using Gibbs reagent

The putative production of phenolic compounds as early intermediates in the metabolism of mecoprop by *A. denitrificans* was monitored using the Gibbs reagent. The reaction mixture (100 ml) contained basal salts medium and 5 g l^{-1} mecoprop. In order to promote phenol accumulation and hence visualization, 2, 2′-bipyridyl (1 mM) was added to the medium of some cultures to inhibit ring cleavage enzymes requiring Fe^{2+}. The reaction was initiated with a 10% inoculum of *A. denitrificans* grown on mecoprop. At regular time intervals, a 5 ml sample was removed from the reaction and the cells were precipitated by centrifugation. The supernatant was acidified to pH 3.0 with 1 M HCl and extracted into ethyl acetate. The extracts were then separated using TLC (dichloromethane : methanol : acetic acid; 9:1:0.1). Phenolic compounds were visualized using Gibbs reagent (2% (w/v) 2,6 dichloroquinone chloroimide in absolute ethanol).

Diagnostic test for aromatic ring cleavage enzymes

A. denitrificans was spread onto protocatechuate agar of the following composition (g l^{-1} distilled water): protocatechuate, 0.2; KH_2PO_4, 0.5; K_2HPO_4, 0.5; $(NH_4)_2SO_4$, 0.2; $MgSO_4 . 7H_20$, 0.5; yeast extract, 0.01; agar, 15.0. The plates were incubated at 30 °C for 7 days. The cleavage pathway in the resulting colonies was determined colorimetrically (Stanier et al. 1966; Kilby 1948). The surfaces of the protocatechuate agar plates were flooded with 2 cm^3 10% (w/v) aqueous $FeCl_3$ which forms a black precipitate with protocatechuate and allowed to stand for 30 min. After this time interval, excess reagent was tipped off and the plates examined for clear haloes surrounding the colonies indicating utilization of protocatechuate.

Respirometric assays

The ability of the bacterium to oxidize mecoprop was studied by the measurement of oxygen uptake by whole cells using a standardized oxygen electrode. Cells were grown overnight in 500 ml of minimal medium containing 1 g l^{-1} mecoprop, harvested by centrifugation at 4 °C, washed once with phosphate buffer (50 mM, pH 7.0), and suspended in the buffer to a protein concentration of 440 μgml^{-1}. For assay of oxygen-dependent mecoprop metabolizing activity, the culture was suspended in phosphate buffer to achieve an oxygen uptake of 1–5 mg O_2 l^{-1} min^{-1}. The reaction was started by the addition of substrate at a final concentration of 1 mM.

Gas liquid chromatography (GLC) analysis

Samples were withdrawn from mecoprop-grown cultures (5 g l^{-1}) and centrifuged for 15 min at 3000 *g*. Aliquots of medium (20 ml) were acidified to pH 3.0 with 0.1 M HCl and extracted into ethyl acetate. Samples were analysed isothermally by gas liquid chromatography (Schimadzu GC-14A with FID; column: cyanopropylphenyldimethyl siloxane, 25m, i.d. = 0.2 mm, helium as carrier gas). Mecoprop and intermediates in the biodegradation pathway were identified by comparing their retention times with those of equivalent chemically-synthesized reference standards using naphthalene (RT 4.44 min) as the internal standard.

Volatile fatty acid analysis

At 2 h intervals throughout growth of *A. denitrificans* on mecoprop (1 g l^{-1}), three separate 1 ml aliquots were aseptically removed and to each sample, was added 100 μg ml^{-1} lysozyme. The tubes were incubated on ice for 10 min and then placed in a boiling

COR-AAD1-00001395

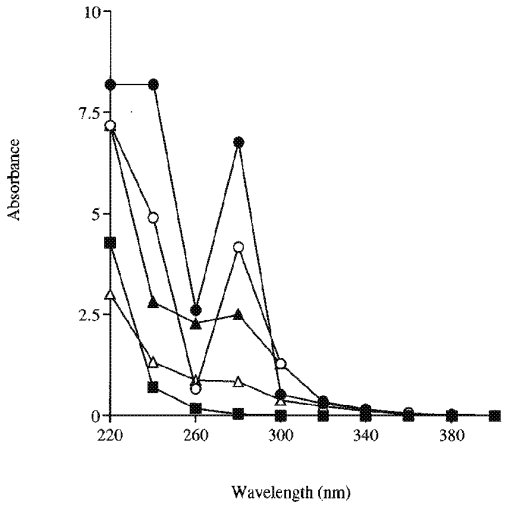


Figure 1. Change in UV absorption spectrum with time during batch culture growth of *A. denitrificans* on the (*R*)-(+)-enantiomer of mecoprop. At intervals during growth of the culture, samples of the medium were centrifuged to remove bacterial cells, diluted fivefold with water and the UV spectrum of the supernatant recorded between 220 and 400 nm. Symbols: ●, starting spectrum at inoculation; ○, spectrum 24 h after inoculation; ▲, spectrum after 32 h; △, spectrum after 46 h and ■, spectrum after 48 h.

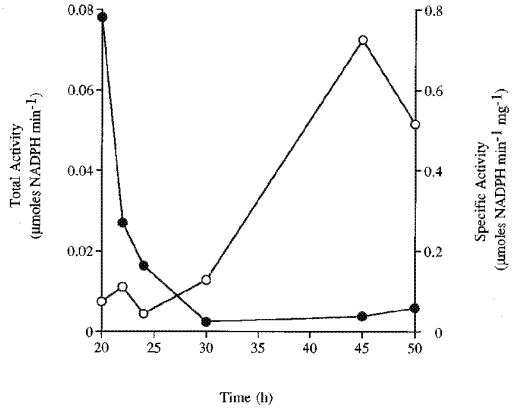


Figure 2. Change in MCP monooxygenase activity during growth of *A. denitrificans* on (*R*)-(+)-mecoprop. At various intervals during growth of the culture, bacterial cells were harvested, sonicated and the resulting enzyme activity of the crude cell extracts was determined as described in the text. The concentration of protein increased in crude cell extracts from 2.0 mg ml^{-1} at 20 h to 18.6 mg ml^{-1} at 50 h. Symbols: ●, specific activity of MCP monooxygenase and ○, total activity of MCP monooxygenase.

water bath for 60–90 seconds. The samples were diluted 1:100 in deionised water prior to injection.

Samples were analysed for volatile fatty acid concentrations by ion-exclusion HPLC. An ICE (Ion Chromatography Exclusion, type AS1, Dionex Ltd.) column was used, with eluent (5 mM heptafluorobutyric acid) flow rate 0.8 ml min^{-1}; regenerant (5 mM tertabutylammoniumhydroxyde) flow rate 2 ml min^{-1}. Peaks were identified by a comparison of their retention times to those of authentic standards of putative reaction products as follows: mecoprop (6.76 min), pyruvate (7.11 min), propionate (16.99 min), acetate (13.99 min), lactate (11.62 min).

Results

The changes in the UV spectra of culture supernatants of *A. denitrificans* while growing on (*R*)-(+)-mecoprop is shown in Figure 1. The sequence of changes with time may be summarized as follows:

- in the first 24 h of growth there was a decrease in A_{279} indicating the disappearance of mecoprop;
- at 32 h, there was a further decrease in A_{279} accompanied by a shift in maximum absorption (λmax) towards 260 nm, probably representing the accumulation of an intermediate;
- at 46 h, there was a decline in absorbance at 260 nm, the λmax occurring at 230 nm, this is more evident at 48 h.

In an attempt to elucidate a biodegradative pathway for (*R*)-(+)-mecoprop by *A. denitrificans* we postulated a pathway then undertook enzyme assays to test the hypothetical route. Crude enzyme extracts were produced from (*R*)-(+)-mecoprop-grown *A. denitrificans* to assay for a range of enzymes. The first step in the biodegradation may involve a mecoprop monooxygenase. However, spectrophotometric tests for this enzyme based on the oxydation of NAD(P)H failed to detect its presence, and evidence for this step was drawn from whole-cell methods using GLC analysis and respirometry.

Assays for activity of MCP were performed by monitoring the consumption of NADPH by mecoprop-grown cells. The specific and total activities for MCP monooxygenase are plotted in Figure 2, the maximum specific activity of the monooxygenase was 0.78 μmoles NADPH min^{-1} mg^{-1} protein. The reaction was specific for NADPH and there was no conversion of MCP to 5-chloro-3-methylcatechol in the

COR-AAD1-00001396

Table 1. Specific activities of catechol dioxygenase on three catechol analogues of 5-chloro-3-methylcatechol. The extinction coefficients were measured at the wavelengths cited in 50 mM phosphate buffer (pH 7.0). The protein concentration of the crude cell extract was 12.0 mg ml^{-1}. One enzyme unit is defined as that amount of enzyme that catalyses the formation of 1 μmol of the product or the disappearance of 1 μmol of the substrate per minute under the assay conditions

Substrate	Wavelength λ max (nm)	Molar extinction coefficient ($\times 10^3$ M)	Specific activity (units min^{-1} mg^{-1} protein)
Catechol	274	2.59	0.07
3-methylcatechol	272	1.73	0.23
4-methylcatechol	278	2.91	0.26

Table 2. Spectrophotometric investigation of the effect of substrate concentration on the specific activity of the muconate lactonizing and delactonizing enzymes of (*R*)-(+)-mecoprop-grown *A. denitrificans*. The assays were conducted at 260 nm and 230 nm respectively with crude cell extracts containing 12.4 mg ml^{-1} protein and started by the addition of substrate. The data represent the change in specific activity as a function of concentration of *cis, cis*-muconic acid and 4-carboxymethylbut-2-en-4-olide, structural analogues of 4-chloro-2-methyl muconate and 2-methyl-4-carboxymethylene but-2-en-4-olide

[*cis, cis*-muconic acid] (μM)	Specific activity (μmols *cis, cis*-muconic acid utilized min^{-1} mg^{-1} protein)	[4-carboxymethyl-but-2-en-4-olide] (μM)	Specific activity (μmols lactone (utilized min^{-1} mg^{-1} protein)
0.0	0	0.0	0
0.03	0.016	0.13	0.08
0.09	0.017	0.25	0.15
0.18	0.029	0.38	0.27
0.27	0.033	0.63	0.40
0.36	0.034	1.38	0.55

presence of NADH. Similarly, there was no detectable monooxygenase activity in extracts obtained from cells grown on nutrient broth. The monooxygenase was found to be intracellular and not membrane-bound as the activity was almost exclusively present in the supernatant and not in the pellet of cellular debris.

Utilization of catechol was monitored by the production of the corresponding muconic acid which absorbs maximally at 265 nm. The rate of change of optical density is directly proportional to the enzyme concentration within a 10-fold range of enzyme dilutions. Catechol dioxygenase activity toward a number of catechols was induced in cells of *A. denitrificans* grown on mecoprop, but not on cells grown on nutrient broth. The bacterium possessed an enzyme capable of utilizing catechol, 3-methyl and 4-methylcatechol. The specific activity of catechol dioxygenase was calculated using the molar extinction coefficients of the tested substrates (Table 1).

The spectrum of *cis, cis*-muconic acid has a λmax at 257 nm. Muconate cycloisomerase activity was assayed by following the decrease in absorbance at 260 nm without interference from the lactone at this wavelength. Crude cell extracts prepared from cells grown on both (*R*)-(+)-mecoprop and benzoic acid utilized muconic acid in the presence of Mn (II). In order to demonstrate lactonizing enzyme activity, the assay required very specific reaction conditions. The dependence of the rate of lactonizing enzyme activity on substrate concentration with *cis, cis*-muconic acid is demonstrated in Table 2. Maximum absorption of 4-carboxymethylbut-2-en-4-olide occurs at 220 nm. Disappearance of the lactone was followed at 230 nm since absorbance of β-ketoadipate is neglible at this wavelength. 4-carboxymethylbut-2-en-4-olide utilization was observed in cell extracts produced separately from cells grown on benzoate and mecoprop. The rate of delactonizing enzyme activity as a function of 4-carboxymethylbut-2-en-4-olide concentration is shown in Table 2.

The Gibbs reaction was used to detect the presence of phenol-containing compounds within the biodegra-

COR-AAD1-00001397

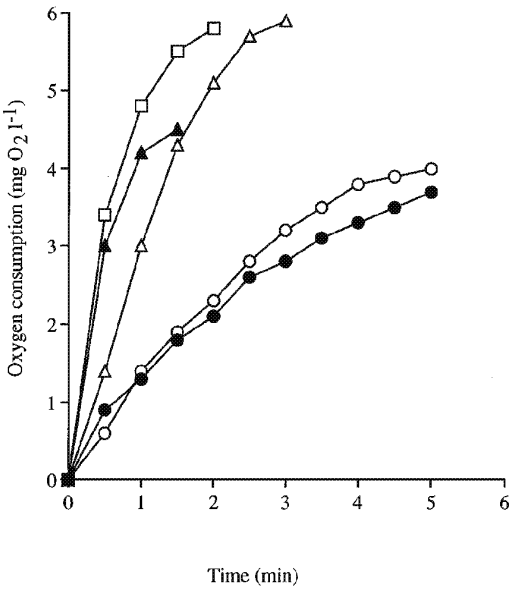


Figure 3. Oxygen uptake by (*R*)-(+)-mecoprop grown *A. denitrificans*. Substrate stimulated oxygen consumption was determined by subtracting the endogenous rate of oxygen uptake in the absence of substrate from the total oxygen consumption. The data represent the mean of triplicate values. The substrates indicated were added at zero time. Symbols: ●, mecoprop stimulated oxydation; ○, MCP fed reaction; ▲, oxygen consumption from oxidation of catechol; △, addition of 3-methylcatechol and ■, addition of 4-methylcatechol.

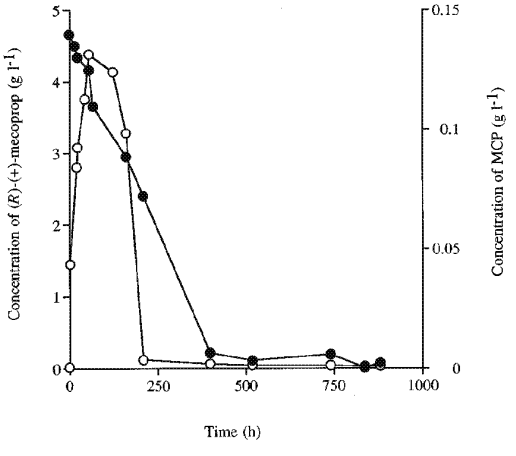


Figure 4. Changes in concentration of (*R*)-(+)-mecoprop and MCP during batch culture growth of *A. denitrificans* using gas liquid chromatography. Culture supernatants were analysed for the presence of mecoprop and the hypothetical metabolite MCP in media containing an initial substrate concentration of 5 g l^{-1} (*R*)-(+)-mecoprop. Symbols: ●, (*R*)-(+)-mecoprop and ○, MCP.

dation pathway. Although reported to be specific for 4-alkoxy-substituted phenolic compounds, mecoprop, 4-chloro-2-methylphenol and muconate were all detected using the reaction giving respective Rf values of 0.47, 0.71 and 0.08. There was no positive Gibbs response of phenol formation observed when mecoprop biodegradation was monitored in the presence of 2,2′-bipyridyl.

Using diagnostic tests to determine *ortho* and *meta* cleavage, mecoprop-grown *A. denitrificans* gave a negative test for *meta* cleavage but a positive Rothera reaction demonstrating that the degradation of mecoprop by *A. denitrificans* proceeds via *ortho* cleavage.

Oxygen uptake was measured and compared when (*R*)-(+)-mecoprop grown cells were supplied with mecoprop, MCP, catechol, 3-methyl and 4-methylcatechol. All substrates were rapidly oxidized without a lag phase (Figure 3). The rate of oxygen consumption during oxidation of each catechol was greater than that observed for oxidation of mecoprop or MCP.

GLC analysis was used to follow the breakdown of (*R*)-(+)-mecoprop and to detect and identify specific intermediates of the mecoprop biodegradation pathway (Figure 4). Results from GC-MS confirmed that MCP was an intermediate soon after the cells were exposed to mecoprop. Other intermediates were identified within the culture supernatants, occurring in minor quantities throughout growth of the bacterium and identified by their retention times against equivalent chemically-synthesized reference standards as follows: 5-chloro-3-methylcatechol RT 11.54 min, 2-methyl-4-carboxymethylene but-2-en-4-olide (lactone) RT 9.76 min. Upon disappearance of MCP after approximately 200 h growth on mecoprop, the formation of a more hydrophobic peak with a retention time of 5.8–5.9 min was observed. The compound increased steadily in the supernatants and was still present after 900 h. This peak could not be identified by GLC or GC-MS. After 900 h, a total amount of 1 mg of mecoprop remained at a 93% extraction efficiency (Figure 4). The muconic acid was not detected in the extracts from *A. denitrificans* and the chemically-synthesized standard was not able to dissolve in the ethyl acetate extraction solvent. Muconic acid was considered too polar for the extraction and partitioning characteristics required in the gas chromatographic procedure to isolate mecoprop.

COR-AAD1-00001398

Discussion

The 2,4-dichlorophenoxyacetic acid (2,4-D) biodegradation pathway is a paradigm for microbial metabolism of chloroaromatic compounds (Chaudhry & Chapalamadugu 1991; Häggblom 1992). The scheme for mineralization of 2,4-D was first established in an *Arthrobacter* species but variations of this pathway have been reported (Bollag et al. 1968a, b; Evans et al. 1971b; Gaunt & Evans 1971a; Loos et al. 1967a; Loos 1975; Tiedje et al. 1969). Unlike 2,4-D, mecoprop possesses an alkyl substituent on the aromatic ring in addition to a chlorine atom. As indicated previously, this has a radical influence on how the compound is degraded (Dagley 1987).

On the basis of the reaction sequence for 2,4-D degradation, the first metabolite in the metabolism of mecoprop would be MCP (Figure 5). This compound is also formed in conjunction with glyoxylate as the primary intermediates of MCPA metabolism (Gamar & Gaunt 1971; Gaunt & Evans 1971a). The biotransformation reaction may be initiated by the action of a monooxygenase(s) in the presence of a hydrogen donor (NADH or NADPH) and molecular oxygen. However, it was not possible to demonstrate mecoprop stimulated oxidation of NAD(P)H in crude cell extracts of *A. denitrificans*. Conversely, the enzyme responsible for the initial biotransformation may not be a monooxygenase with NAD(P)H-dependent activity but rather an α-ketoglutarate-coupled dioxygenase which catalyses an entirely different reaction mechanism for cleavage of the aliphatic side chain (Abbott & Udenfriend 1974). Dioxygenase activity is dependent on the presence of α-ketoglutarate, Fe^{2+} and a reducing agent (Fukumori & Hausinger 1993). The inability of the microorganism to transform mecoprop in the presence of 2,2′-bipyridyl suggests that the initial reaction step itself was inhibited which is consistent with the concept that the enzyme responsible for phenol formation is iron dependent.

As the initial enzyme proved to be unstable in cell extracts, experiments were carried out with whole cells. Disappearance of mecoprop was monitored on GLC and GC-MS. Analysis of supernatants revealed that cells produced MCP when incubated with mecoprop.

Oxygen consumption data supported the view that MCP is an early metabolite in the degradative pathway of (*R*)-(+)-mecoprop. MCP was rapidly oxidized without a lag by *A. denitrificans* implicating side-chain loss at the first metabolic step and demonstrating the oxygen dependency of the reaction. Although it was expected that mecoprop would consume twice the amount of oxygen of MCP and three times the amount of catechol, it is possible that the amount of oxygen consumption observed was due to different rates of transport and/or biooxidation of the substrates at comparable protein concentrations.

In addition to the production of MCP, a second primary metabolite was generated as the product of the three carbon aliphatic side chain. After 4 h batch culture growth, pyruvate (3 ppm) was formed as a result of cleavage of the ether link and could be fed directly into the TCA cycle.

MCP is subsequently converted to the corresponding catechol, a common central intermediary metabolite of aromatic compounds (Sandmann & Loos 1988). The introduction of an additional polar group into the ring is achieved by the employment of a monooxygenase and enables the compound to be further degraded by the β-keto adipate pathway. Trace amounts of 5-chloro-3-methylcatechol were detected in mecoprop grown cultures. Experiments with cell extracts showed that molecular oxygen and NADPH were required for MCP utilization. This implicates a monooxygenase catalysed reaction and is consistent with work by Bollag (1967) who partially purified an enzyme that could hydroxylate 2,4-dichlorophenol to 3,5-dichlorocatechol with NADPH (Bollag et al. 1967).

In respirometric experiments, catechol, 4-methylcatechol and 3-methylcatechol were oxidized without a lag by mecoprop grown cells. This further suggests that a catechol is formed as an intermediate when cells grow on mecoprop. The phenomenom illustrates a degree of nonspecificity of the mecoprop degrading enzyme systems of *A. denitrificans* towards methylated catechols and catechol intermediates.

There are two different metabolic pathways which can be used by microorganisms to metabolize gem-diols such as catechol, the *ortho* cleavage pathway (initiated by catechol 1, 2-oxygenase or protocatechuate 3,4-oxygenase) and the *meta* cleavage pathway (initiated by catechol 2, 3-oxygenase or protocatechuate 4, 5-oxygenase). The mechanism of aromatic cleavage is of great significance because it determines a long series of subsequent biochemical steps differing in many enzymic respects. Most microorganisms able to metabolize aromatic substrates possess either the *ortho* or the *meta* pathway however, there are microorganisms which exhibit both mechanisms of ring cleavage (Worsey & Williams 1975). Many investigations of the ring cleavage mechanism have been conducted with structures simpler than those of

COR-AAD1-00001399

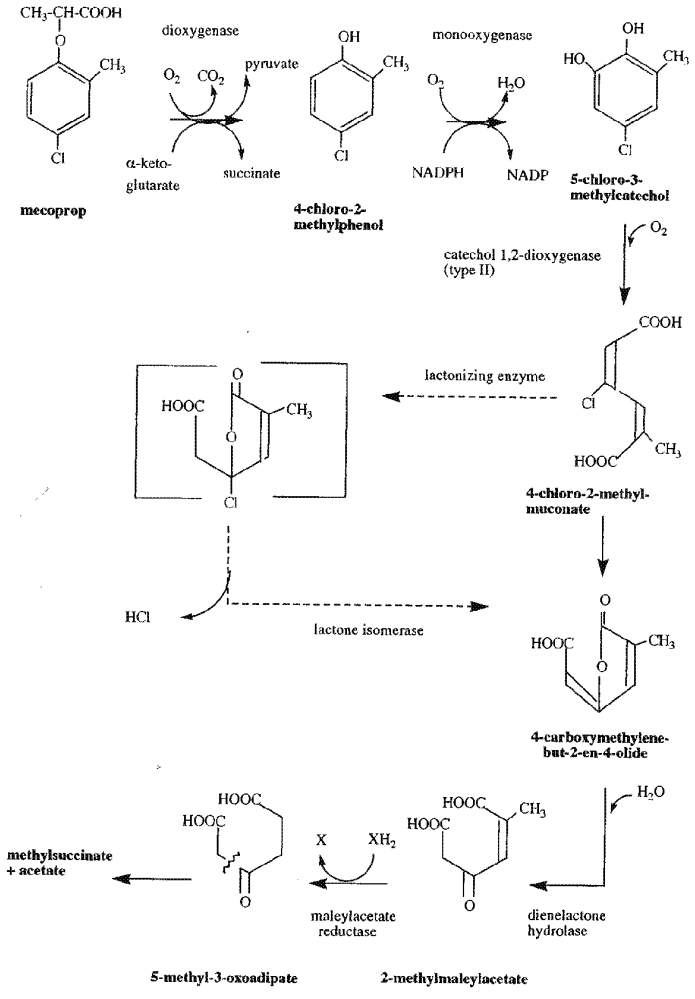


Figure 5. Proposed pathway for the degradation of (*R*)-(+)-mecoprop by *A. denitrificans*. Direct evidence was obtained for transformation of the herbicide as far as 2-methyl-4-carboxymethylene but-2-en-4-olide (—>), by detection of the metabolites using GLC, but the remaining conversion of the lactone to the putative products acetate and methyl succinate remains speculative (- - ->).

most pesticides (Ornston & Stanier 1966). In general, *ortho*- and *meta*-substituted methyl phenols seem to be preferentially metabolized by a *meta*-cleavage pathway starting with hydroxylation of methyl phenol to form 3-methylcatechol. In contrast, halogenated aromatics are generally degraded by *ortho* cleavage. It is not possible to predetermine the mechanism of aromatic cleavage of compounds possessing both methyl and halogen ring substituents. *Pseudomonas cepacia* MB2, for example, adopted the *meta* cleavage pathway for utilisation of 3-chloro-2-methylbenzoic acid via 4-chloro-3-methylcatechol (Higson & Focht 1992). In contrast, Pieper and coworkers described degradation of 2-chloro-4-methylphenoxyacetic acid via 5-chloro-3-methylcatechol (Pieper et al. 1993).

Results from the *ortho* and *meta* cleavage diagnostic tests indicated that mecoprop-grown *A. denitrificans* adopts the *ortho* cleavage mechanism of aromatic ring

COR-AAD1-00001400

cleavage for growth on protocatechuate. Further evidence for the metabolism of this aromatic substrate was obtained by the observation of clear haloes surrounding colonies of *A. denitrificans* indicating utilization of protocatechuate.

If mecoprop is degraded by *A. denitrificans* by *ortho* cleavage, catechol 1, 2-dioxygenase attack on 5-chloro-3-methylcatechol would produce 4-chloro-2-methylmuconate. In contrast to MCP monooxygenase, the ring cleavage enzyme is unstable in the presence of oxygen. The increase in absorbance at 265 nm in spectrophotometric assays of cell extracts fed with analogues of 5-chloro-3-methylcatechol is consistent with the production of a muconic acid. The ability of the crude cell extract to transform both catechol and the two methylated derivatives provides further evidence for the relative non-specificity of catechol dioxygenase. This is consistent with halobenzoate degradation in which both *Pseudomonas* and *Alcaligenes* species have been reported to possess catechol-1, 2-dioxygenases with high activities towards substituted catechols (Commandeur & Parsons 1990; Haigler et al. 1988; Tiedje et al. 1969). The relaxed substrate specificity of catechol 1,2-dioxygenase is characteristic of the modified *ortho* cleavage pathway for degradation of chloroaromatic compounds (Schlömann 1994). The formation of this ring cleavage product at the expense of 4-methylcatechol can be followed by monitoring the change in maximum absorbance with time over a range of wavelengths. The maximum absorption peak initially occurred at 278 nm which is consistent with that of 4-methylcatechol. With time, the peak became broader and there was a gradual shift in absorbance towards 250 nm indicative of formation of a muconic acid. There was no increase in absorbance at 375 nm corresponding to the production 2-hydroxymuconic acid semialdehyde of the *meta* cleavage pathway. The absence of catechol dioxygenase activity in uninduced cells indicates that mecoprop or intermediates formed from it are required for induction of catechol dioxygense and subsequent enzymes in the pathway. Expression of the aromatic degradation has been repressed in the presence of a non-aromatic growth substrate.

Results from the scanning spectrophotometric study following the transformation of 4-methylcatechol, reveal that after 50 min incubation of the crude cell extract with the substrate, the maximum absorption peak shifts back to a higher wavelength of 270–272 nm. This maybe due to the production of the derivatised butenolide (λmax 278 nm) occurring either biologically or through the spontaneous lactonization of 4-chloro-2-methylmuconate if it is unstable (Gaunt & Evans 1971b). Spectrophotometric assays revealed that cell extracts harvested after independent growth on both mecoprop and benzoic acid were indeed able to utilize 4-carboxymethylbut-2-en-4-olide, the lactone derivative formed during metabolism of benzoate. This indicates the broad substrate specificity of the dienelactone hydrolase induced by growth on mecoprop being able to metabolize an intermediate other than that derived from the mecoprop substrate. This is in contrast to the modified *ortho* cleavage pathway of chlorocatechols where the dienelactone hydrolases are more specific (Schlömann 1994). Both the lactonizing and delactonizing enzymes are stable under the above extraction conditions for several weeks. GLC also revealed the presence of a compound with a retention time equivalent to that of 2-methyl-4-carboxymethylene but-2-en-4-olide in culture supernatants during growth of *A. denitrificans* on (*R*)-(+)-mecoprop. Although it is considered that muconic acids and muconlactones are likely intermediates in the oxidative biotransformation of catechols, it is yet to be established whether this occurs via the *ortho* or modified *ortho* cleavage pathways. As far as it is known, this is the most comprehensive series of investigations to determine the biodegradative pathway of mecoprop.

Acknowledgments

We are most grateful to Professor Doug Ribbons for his help regarding the synthesis of *cis, cis*-muconic acid and his kind donations of authentic standards of catechol, 4-methylcatechol, 3-methylcatechol and carboxymethylbut-2-en-4-olide. We thank Prof. H. Knackmuss for his donations of standards of 5-chloro-3-methylcatechol and 2-methyl-4-carboxymethylene but-2-en-4-olide and Tony Davey for technical assistance. This work was supported by the award of a studentship from the Natural Environment Research Council.

References

Abbott MT & Udenfriend S (1974) α-Ketoglutarate-coupled dioxygenases. In: Hayaishi O (Ed) Molecular Mechanisms of Oxygen Activation (pp 167–214). Academic Press, London

Anonymous (1992) Produce Farmstat Survey

COR-AAD1-00001401

Bollag JM, Helling CS & Alexander M (1967) Metabolism of 4-chloro-2-methylphenoxyacetic acid by soil bacteria. Appl. Microbiol. 15: 1393–1398

— (1968a) 2,4-D metabolism: enzymatic hydroxylation of chlorinated phenols. J. Agr. Food Chem. 16: 826–828

Bollag JM, Briggs GG, Dawson JE & Alexander M (1968b) 2,4-D metabolism: enzymatic degradation of chlorocatechols. J. Agr. Food Chem. 16: 829–833

Chaudhry GR & Chapalamadugu S (1991) Biodegradation of halogenated organic compounds. Microbiol. Rev. 55: 59–79

Commandeur LCM & Parsons JR (1990) Degradation of halogenated aromatic compounds. Biodegradation 1: 207–220

Dagley S (1987) Lessons from biodegradation. Annu. Rev. Microbiol. 41: 1–23

Evans WC, Smith BSW, Moss P & Fernley HN (1971a) Bacterial metabolism of 4-chlorophenoxyacetate. Biochem. J. 122: 509–517

Evans WC, Smith BSW, Fernley HN & Davies JI (1971b) Bacterial metabolism of 2,4-dichlorophenoxyacetate. Biochem. J. 122: 543–552

Fukumori F & Hausinger RP (1993) *Alcaligenes eutrophus* JMP134 '2,4-dichlorophenoxyacetate monooxygenase' is an α-ketoglutarate-dependent dioxygenase. J. Bacteriol. 175: 2083–2086

Gamar Y & Gaunt JK (1971) Bacterial metabolism of 4-chloro-2-methylphenoxyacetate (MCPA): formation of glyoxylate by side-chain cleavage. Biochem. J. 122: 527–531

Gaunt JK & Evans WC (1971a) Metabolism of 4-chloro-2-methylphenoxyacetate by a soil pseudomonad: Preliminary evidence for the metabolic pathway. Biochem. J. 122: 519–526

— (1971b) Metabolism of 4-chloro-2-methylphenoxyacetate by a soil pseudomonad. Ring-fission, lactonizing and delactonizing enzymes. Biochem. J. 122: 533–542

Häggblom MM (1992) Microbial breakdown of halogenated aromatic pesticides and related compounds. FEMS Microbiol. Rev. 103: 29–72

Haigler BE, Nishino SF & Spain JC (1988) Degradation of 1, 2-dichlorobenzene by a *Pseudomonas* sp. Appl. Environ. Microbiol. 54: 294–300

Hegeman GD (1966) Synthesis of the enzymes of the mandelate pathway by *Pseudomonas putida*. J. Bacteriol. 91: 1140–1154

Higson FK & Focht DD (1992) Utilization of 3-chloro-2-methylbenzoic acid by *Pseudomonas cepacia* MB2 through the *meta* fission pathway. Appl. Environ. Microbiol. 58: 2501–2504

Kilby (1948) The bacterial oxidation of phenol to β-keto-adipic acid. Biochem. J. 43: V-VI

Kilpi S (1980) Degradation of some phenoxy acid herbicides by mixed cultures of bacteria from soil treated with 2-(2-methyl-4-chloro)phenoxypropionic acid. Microbial Ecol. 6: 261–270

Lindholm L, Backström V & Kilpi S (1982) Degradation of mecoprop in soil. Acta Agr. Scandinavica 32: 429–432

Loos MA, Roberts RN & Alexander M (1967a) Phenols as intermediates in the decomposition of phenoxyacetates by an *Arthrobacter* sp. Canadian J. Microbiol. 13: 679–690

— (1967b) Formation of 2,4-dichlorophenol and 2,4-dichloroanisole from 2,4-dichlorophenoxyacetate by *Arthrobacter* sp. Canadian J. Microbiol. 13: 691–699

Loos MA (1975) Phenoxyalkanoic acids. In: Kearney PC & Kaufman DD (Eds) Herbicides: Their Chemistry, Degradation and Mode of Action, Vol 1 (pp 34–57). Marcel Dekker, New York

MacRae IC, Alexander M & Rovira AD (1963) The decomposition of 4-(2, 4-dichlorophenoxy)butyric acid by *Flavobacterium* sp. J. Gen. Microbiol. 32: 69–76

Oh K & Tuovinen OH (1991a) Bacterial degradation of phenoxy herbicide mixtures 2,4-D and MCPP. Bull. Environ. Contam. Toxicol. 47: 1–8

— (1991b) Detection and identification of substituted phenols as intermediates of concurrent bacterial degradation of the phenoxy herbicides MCPP and 2,4-D. Microbiol. Lett. 79: 141–146

— (1994) Biodegradation of the phenoxy herbicides MCPP and 2,4-D in fixed-film column reactors. Int. Biodet. Biodeg. 33: 93–99

Ornston LN & Stanier RY (1966) The conversion of catechol and protocatechuate to β-ketoadipate by *Pseudomonas putida*. J. Biol. Chem. 241: 3776–3785

Pandell AJ (1976) Enzyme-like aromatic oxidations. Metal-catalysed peracetic acid oxidation of phenol and catechol to *cis, cis* muconic acid. J. Org. Chem. 41: 3992–3996

Pieper DH, Stadlerfritzsche K, Engesser KH, Knackmuss HJ (1993) Metabolism of 2-chloro-4-methylphenoxyacetate by *Alcaligenes eutrophus* JMP-134. Arch. Microbiol. 160: 169–178

Sandmann E & Loos MA (1988) Aromatic metabolism of a 2,4-D degrading *Arthrobacter* sp. Canadian J. Microbiol. 34: 125–130

Schlömann M (1994) Evolution of chlorocatechol catabolic pathways. Biodegradation 5: 301–321

Sinton GL, Fan LT, Erickson LE & Lee SM (1986) Biodegradation of 2,4-D and related xenobiotic compounds. Enzyme Microbial Technol. 8: 395–403

Stanier RY, Palleroni NJ & Doudoroff M (1966) The aerobic pseudomonads: a taxonomic study. J. Gen. Microbiol. 43: 159–271

Tett VA, Willetts AJ & Lappin-Scott HM (1994) Enantioselective degradation of the herbicide mecoprop [2-(2-methyl-4-chlorophenoxy)propionic acid] by mixed and pure bacterial cultures. FEMS Microbiol. Ecol. 14: 191–200

Tiedje JM & Alexander M (1969) Enzymatic cleavage of the ether bond of 2,4-dichlorophenoxyacetate. J. Agr. Food Chem. 17: 1080–1084

Tiedje JM, Duxbury JM, Alexander M & Dawson JE (1969) 2,4-D metabolism: pathway of degradation of chlorocatechols by *Arthrobacter* sp. J. Agr. Food Chem. 17: 1021–1026

Worsey MJ & Williams PA (1975) Metabolism of toluene and xylenes by *Pseudomonas putida* (arvilla) mt-2: evidence for a new function of the TOL plasmid. J. Bacteriol. 124: 7–13

Worthing CR & Walker SB (1991) The Pesticide Manual: A World Compendium. BCPC Publications, Bracknell, UK

COR-AAD1-00001402

Chapter 10

GENETIC ENGINEERING OF CROPS FOR TOLERANCE TO 2,4-D

Danny Llewellyn
David Last

CONTENTS

1. Phenoxyacetic Acid Herbicides 159
2. Current 2,4-D Usage Patterns in Cotton Cropping 161
3. Genetic Engineering Options for 2,4-D Resistance 162
4. A Microbial Source of 2,4-D Degradation Genes 163
5. Genetic Engineering of the *Alcaligenes tfdA* Gene for Expression in Plants 164
6. Expression of the 2,4-D Dioxygenase in Transgenic Tobacco 165
7. Transgenic Cotton Showing Tolerance to 2,4-D 168
8. 2,4-D-Tolerant Cotton Plants — An Ethical Dilemma? 170
9. Conclusions 171

References 172

1. PHENOXYACETIC ACID HERBICIDES

2,4-Dichlorophenoxyacetic acid (2,4-D) is a member of a class of herbicides with activity against many dicotyledonous weeds. Although it is an older-generation herbicide, introduced in the 1940s, it is still in general use in many countries. The chemical structure of 2,4-D and some of its more common relatives such as MCPA (2-methyl-4-chlorophenoxyacetic acid) are shown in Figure 1. The group includes butyric acid derivates such as 2,4-DB that are metabolized to active 2,4-D by ß-oxidation, giving them a degree of selectivity different from that of 2,4-D. Because many legumes lack the ß-oxidation activity, 2,4-DB can be used for weed control on legume-containing pastures or crops. The most notable member of the group is 2,4,5-T, which, as a mixture with 2,4-D, formed the notorious "Agent Orange" herbicide used for massive forest defoliation during the Vietnam War. The phenoxy herbicides such as 2,4-D are active as plant growth regulators or hormones with auxin activity. They are absorbed through the leaves and are translocated initially for short distances by

1-56670-045-0/96/$0.00+$.50
© 1996 by CRC Press Inc.

COR-AAD1-00001403

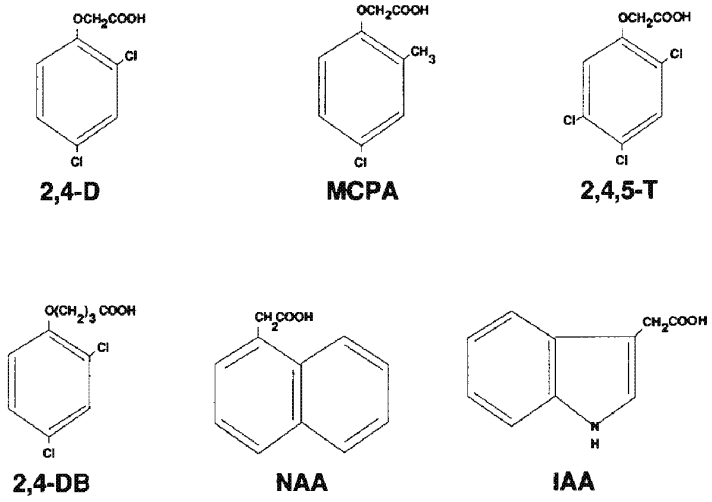


FIGURE 1 Chemical structures of phenoxyacetic acid herbicides and some natural plant auxins. 2,4-D, 2,4-dichlorophenoxyacetic acid; MCPA, 2-methyl,4-chlorophenoxyacetic acid; 2,4,5,-T, 2,4,5,-trichlorophenoxyacetic acid; 2,4-DB, 4-(2,4-dichlorophenoxy)butyric acid; IAA, indole-3-acetic acid; NAA, 1-naphthaleneacetic acid.

intercellular movement to the phloem and then over long distances in the phloem to the shoot meristems. Phloem-translocated herbicides such as the growth regulators move through the sieve tube system similar to sugars, from a source (region of sugar synthesis) to a sink (region of sugar utilization) (Singh and Muller 1979), although the exact mechanisms are still poorly understood. Stimulation of sink activity has been observed with 2,4-D (Kirkwood 1979), and this presumably potentiates its herbicidal activity. At the meristem 2,4-D and other phenoxy herbicides interfere with cell development, presumably through saturation effects arising from mimicry of auxin action on cell elongation (Dodge 1989). Characteristic damage to plants by low levels of 2,4-D is severe twisting and spindly growth of the vegetative and floral tips and organs (Figure 2), followed by the death of the meristems and eventually the plant.

Most commercial formulations contain salts, amine salts, or esters of the phenoxyacetic herbicides rather than the free acids. The esters are believed to be active only after hydrolysis to the free acid in the plant (Loos 1969). En route to the site of action the phenoxy herbicides are also exposed to the potential for modification to less active metabolites or in some cases to the active herbicide, e.g., 2,4-DB. Detoxification of the phenoxy herbicides may play an important part in the selectivity of these chemicals for different plant species. 2,4-D, for example, can undergo a variety of chemical modifications in insensitive plants (reviewed by Pillmoor and Gaunt 1981 and Loos 1969), including side chain degradation, side chain elongation, ring hydroxylation, and conjugation with proteins, amino acids, and glucose (Ashton and Crafts 1981).

Monocotyledonous plants are considerably less sensitive to phenoxyacetic herbicides and hence these herbicides are of particular use for broadleaf weed control in monocot crops such as cereals and sugarcane. 2,4-D is an important auxin in the tissue culture of cereals and is an active growth regulator for these species. It is not surprising, therefore, that cereals are sensitive to 2,4-D at particular developmental stages and can be seriously affected if this herbicide is used inappropriately (Evans 1968). However, the major scope for the production of crops resistant to 2,4-D is among the dicotyledonous crops that are sensitive to low concentrations of phenoxyacetic acid herbicides, particularly in multicropping systems where they may be exposed inadvertently to 2,4-D as spray drift or from residues in the soil or irrigation water. Among the dicot crops, the main focus of genetic engineering research worldwide has been into 2,4-D-tolerant cotton, a crop of considerable economic value that is extremely sensitive to 2,4-D.

COR-AAD1-00001404



FIGURE 2 2,4-D spray drift damage on cotton. Note the twisted distorted leaves and bracts and dead growing points.

2. CURRENT 2,4-D USAGE PATTERNS IN COTTON CROPPING

Weed control remains an important component of cotton production, primarily because of the noncompetitive nature of cotton seedlings (Keeley et al. 1986). Weeds can reduce cotton yields by direct competition for light, nutrients, and water, but can also affect quality parameters through contamination of the harvested lint. Weed spectrum will be determined by geographical location and weed management practices, but the problems are likely to be similar in all highly mechanized production systems such as those found in the U.S. and Australia. Charles (1991) surveyed the weed species and herbicide usage patterns on Australian cotton farms and found that most of the serious weeds were dicotyledonous. Generally, these were weeds of pastures and are either native or naturalized species. *Xanthium occidentale* (Noogoora burr), *X. spinosum* (Bathurst burr), and *Datura* spp. (thornapples) are large plants that are effective competitors with cotton and have been a long-term problem in cotton production. *Cyperus* spp. are perennial grasses adapted to moist soils and can be a problem in fields and drainage channels. Like the dicot weeds *Haloragis glauca* and *Polymeria longifolia,* these grasses have rhizomatous roots and are resistant to cultivation as a form of control. All of these species, except *Cyperus,* are sensitive to 2,4-D, but cannot be controlled effectively with this herbicide, except in fallow fields, because of the serious risk of damaging the cotton crop. 2,4-D is used on about one third of the cotton farms from ground rigs at any time after defoliation through to sowing the next season, but its popularity is reduced by fears of damage that could result from residues in the soil (Charles 1991). Current weed control programs rely heavily

COR-AAD1-00001405

on preemergence cultivation and use of nonselective herbicides such as glyphosate before planting, residual herbicides such as trifluralin and atrazine, and hand removal of weeds from postemergence cotton. 2,4-D is more commonly used on cereal crops (oats, barley, sorghum, and maize) in the same valleys in which cotton is grown or on fallow fields in wheat/fallow/cotton rotations to control broadleaf summer weeds. The majority of spraying of 2,4-D on cereals is done from the air, so spray drift, even on relatively calm days, away from the cereals into the cotton growing areas is almost inevitable. There are also instances in which 2,4-D residues are not removed from the tanks and have caused serious damage to cotton in subsequent spraying of an unrelated herbicide, miticide, or insecticide. In recent years there have been some isolated incidences of 2,4-D damage appearing in cotton as a result of its use on fallows. Such damage occurs when weather conditions do not favor growth of 2,4-D-degrading soil microbes, such as when it is unusually cold or dry (Hammond 1992). 2,4-D tolerant cotton would be useful both in the prevention of spray drift damage and perhaps for the control of some of the dicotyledonous weeds that currently trouble the industry by using this relatively cheap compound as an effective postemergence herbicide.

3. GENETIC ENGINEERING OPTIONS FOR 2,4-D RESISTANCE

Genetically engineered herbicide resistances are no longer new. Effective resistances have been developed to a variety of potent herbicides such as glyphosate and the sulfonylureas (Comai et al. 1985, Shah et al. 1986, Haughn et al. 1988). In these cases the molecular target of the herbicide was well characterized and resistances were developed through the manipulation or replacement of the target gene. Comai et al. selected a strain of *Salmonella typhimurium* that was resistant to glyphosate because of a mutation in the gene for 5-enolpyruvylshikimate 3-phosphate synthase (EPSP synthase), the target enzyme for glyphosate. When expressed in plants, the mutant enzyme could substitute for the endogenous plant enzyme that had been inactivated by the herbicide. Shah et al. used a different approach to achieve the same result. They selected a petunia cell line that was resistant to glyphosate through the overproduction of the EPSP synthase enzyme. This petunia gene was cloned and then reintroduced into petunia with a new, much stronger promoter. The new chimeric gene resulted in the over-production of the enzyme, and hence in resistance at a whole plant level. Haughn et al. achieved a similar result with an acetolactate synthase gene from *Arabidopsis,* which confers sulfonylurea herbicide resistance on transgenic tobacco plants. While all of these resistance genes are certainly effective, they may have some deleterious effects, especially in the absence of the herbicide.

Principles similar to those described above cannot be applied to developing a resistance gene against 2,4-D because its primary target is still unknown. 2,4-D is an analog of the plant hormone group known as auxins and at low levels is often used as a substitute for natural auxins in plant tissue culture systems. Some of the symptoms of 2,4-D damage to cotton such as twisting of shoots are suggestive of a hormone-induced mechanism such as perturbed rates of cell division. However, it is likely that there are other, more drastic and undetermined phytotoxic effects.

A second approach used to generate herbicide resistant plants is the engineering of detoxification pathways in susceptible species. A number of herbicides are toxic to some plant species but not others. The difference in sensitivity is often not associated with a difference in susceptibility of the molecular target, but to an ability of the resistant plant to detoxify or break down the herbicide before it has a chance to inactivate the susceptible target gene product. Two examples now exist in which a microbial gene encoding an enzyme that detoxifies a herbicide was modified for expression in plants and introduced into a model species by transformation (De Block et al. 1987, Stalker et al. 1988). The novel detoxification enzymes were expressed and conferred resistance to relatively high levels of phosphinothricin (Basta® or glufosinate) and 3,5-dibromo-4-hydroxybenzonitrile (bromoxynil), respectively. These artificial systems mimic detoxification systems found in plants and are likely to place the least genetic, physiological, or yield burden on the plant, both in the presence or absence of the herbicide. Furthermore, the mode of action of the herbicide need not be

COR-AAD1-00001406

known, provided that the appropriate detoxification genes can be found. This second strategy was the most promising for 2,4-D because a variety of microorganisms are known that naturally degrade the herbicide in the environment.

4. A MICROBIAL SOURCE OF 2,4-D DEGRADATION GENES

2,4-D has a relatively short residual life in soils, decreasing by about one half every 14 d or so (Noble and Hargreaves 1992). Other phenoxyacetic acid herbicides such as MCPA and 2,4,5-T break down much more slowly in the soil and can cause serious residual problems (Loos 1969). The breakdown of phenoxyacetic acid herbicides in the soil is performed by a variety of microorganisms including bacteria, yeasts, and fungi from several taxonomic groups. Loos (1969) lists many of these and the particular herbicides that can be used as substrates: a variety of *Pseudomonas* spp., *Achromobacter*, *Flavobacterium*, *Corynebacterium*, and *Arthrobacter* spp. and some Actinomycetes, such as *Nocardia* spp. and *Streptomyces viridochromogenes*, which will all degrade 2,4-D as a carbon source after appropriate enrichment of this substrate. The most well-characterized organisms are strains of *Alcaligenes eutrophus*, which are relatively simple Gram-negative rod bacteria found in most aerated soils. Several different strains of *A. eutrophus* were isolated from soils in fields commonly sprayed with 2,4-D and all are able to grow on synthetic media with 2,4-D as the sole source of carbon (Don and Pemberton 1981) (Figure 3). The common denominator between all the strains is a large (75 kb) plasmid that was shown to encode many of the enzymes necessary for the breakdown of 2,4-D. The overall pathway of 2,4-D degradation (Don et al. 1985) in *A. eutrophus* is shown in Figure 4.

It may be considered impractical to engineer all six *tfd* genes to function in plants, but data on the phytotoxicity of the first intermediate in the pathway, 2,4-dichlorophenol (DCP) suggested that only the first gene in the pathway may be necessary to confer a useful degree of tolerance on transgenic crop plants. Perkins et al. (1987) showed that DCP was toxic to cell suspension cultures of *Nicotiana tabacum* with a 50% inhibition at 10 mg/l and inhibited leaf and root development of tobacco seedlings at the same concentration. 2,4-D at 1 mg/l blocked all plant development. Streber and Willmitzer (1989) also tested the toxicity of DCP to tobacco callus and showed that DCP at 10 mg/l caused no obvious toxic effects, while 2,4-D at 2 mg/l severely reduced growth rates and turned the callus brown. Seed germination studies carried out by Llewellyn et al. (1990) in the presence of 2,4-D or DCP indicated that, in tobacco at least, DCP was some 50- to 100-fold less phytotoxic than 2,4-D. These results were encouraging with respect to expressing just the first step in the pathway for protection against 2,4-D damage.

Streber et al. (1987) were able to clone the *tfdA* gene encoding the first enzymatic step of 2,4-D degradation by Tn5 mutagenesis of the degradative plasmid of *A. eutrophus*. From >1000 kanamycin-resistant clones that had Tn5 insertions, two were identified that lacked the ability to cleave the acetate side chain of 2,4-D. A large *Hin*dIII fragment containing the wild-type plasmid DNA in the region of the tagged *tfdA* gene both complemented the mutations and allowed another strain of *A. eutrophus* (JMP222) to grow on 2,4-D analogs such as phenoxyacetic acid. JMP222 is able to grow on phenol as a carbon source by using chromosomal genes for a meta-cleavage pathway, but cannot normally grow on phenoxyacetate. The *tfdA* gene product is able to accept different substituted and unsubstituted phenoxyacetates as a substrate and can cleave phenoxyacetate to phenol. This assay was used to localize the gene by successive deletions of the large fragment down to 2 kb that could be sequenced. The *tfdA* gene was originally thought to encode a 2,4-D monooxygenase but subsequent biochemical analysis has indicated that the gene product is a ferrous iron dependent dioxygenase that uses α-ketoglutarate as a cosubstrate (Fukomori and Hausinger 1993). The α-ketoglutarate is converted to succinate and carbon dioxide concomitant with the conversion of 2,4-D to DCP and glyoxylate.

COR-AAD1-00001407

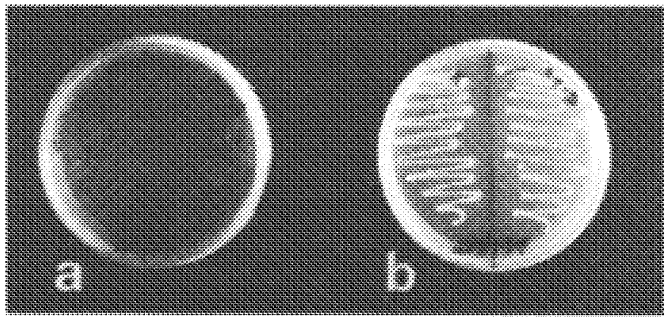


FIGURE 3 *Alcaligenes eutrophus* utilization of 2,4-D as a carbon source. Bacterial plates contain minimal medium supplemented with 125 µg/ml 2,4-D sodium salt as sole source of carbon. (a) JMP228R, a cured strain of *A. eutrophus* lacking the 2,4-D degradative plasmid is unable to grow on 2,4-D. (b) JMP134, a strain of *A. eutrophus* which has the degradative plasmid pJP4, grows on 2,4-D in the absence of any other carbon source.

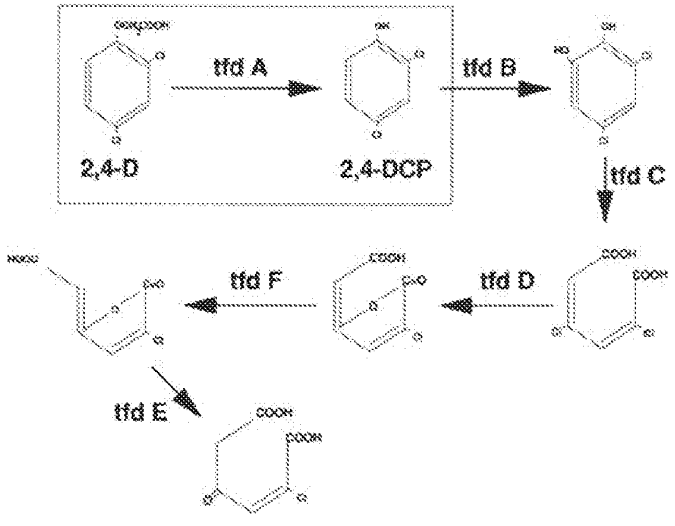


FIGURE 4 Biochemical pathway of 2,4-D degradation encoded by the degradative plasmid pJP4 of *Alcaligenes eutrophus*. 2,4-D degradative genes are designated *tfdA-F*, and the biochemical reaction catalyzed by their gene product is noted at each step. The 2,4-D dioxygenase reaction catalyzing the first breakdown step is boxed.

5. GENETIC ENGINEERING OF THE *ALCALIGENES TFDA* GENE FOR EXPRESSION IN PLANTS

Three research groups have engineered the *tfdA* gene for expression in plants (Streber and Willmitzer 1989, Lyon et al. 1989, Bayley et al. 1992). Sequencing of the *tfdA* gene revealed that it had an unusual translational start codon. This bacterial coding region began with a GTG codon rather than the usual ATG, a situation found in some other bacterial genes such as *tfdC* and *-D* (Perkins et al. 1990) and the *bar* gene from *Streptomyces hygrosopicus* (Thompson et al. 1987). The novel codon is not correctly recognized in plants (Lyon et al. 1989) and had to be modified to an ATG before reasonable expression levels could be attained. The modification of the translation start appeared to have no effect on 2,4-D dioxygenase enzyme activity in *Escherichia coli* (Lyon et al. 1989). Streber and Willmitzer and Lyon et al. also modified the region upstream of the ATG to more

COR-AAD1-00001408

closely resemble the consensus sequence compiled from a large number of plant genes (Joshi 1987). Bayley et al. (1992) used a *tfdA* gene derived from the *A. eutrophus* 2,4-D degradative plasmid pJP5 (Don and Pemberton 1981) rather than the pJP4 plasmid used by the earlier researchers. pJP5 has some restriction site polymorphisms relative to pJP4, but is very similar in size. No sequence data were presented for the pJP5 *tfdA* gene, but it is highly likely that it is nearly identical to the gene from pJP4, or at least had a common origin.

In order to express the 2,4-D dioxygenase gene in plants it was necessary to provide a strong plant promoter and transcription termination and polyadenylation signals. The latter were provided by a fragment of the nopaline synthase gene originally from the T-DNA of *Agrobacterium tumefaciens* (Herrera-Estrella et al. 1983) or a similar region from the T-DNA octopine synthase gene (Dhaese et al. 1983). All three groups used the promoter sequences from the 35S promoter of the cauliflower mosaic virus (CaMV), a strong and essentially constitutive promoter in many dicot plants (Odell et al. 1985). Two additional modified 35S promoters were used by Lyon et al. (1989) in an attempt to improve on the expression of the viral promoter. In the first, two additional copies of the so-called upstream enhancer region of the 35S promoter (the region between -90 and -340) were ligated onto the 35S promoter at -340 relative to the transcription start site. All three enhancer regions were in the same orientation. Kay et al. (1987) suggested that multiple copies of this enhancer increase the level of expression of the promoter more than arithmetically. In the second construction an additional 45 bp was added to the small leader region of the 35S promoter fragment (essentially only 2 bp). The addition consisted of the leader region of another viral gene, the coat protein gene of the alfalfa mosaic virus (AMV). This AMV leader has been suggested to improve the translational competitiveness of viral genes at least *in vitro* and perhaps also heterologous genes to which it is attached (Jobling and Gerhke 1987). Streber and Willmitzer also constructed a chimeric *tfdA* gene with the promoter from a light-inducible gene ST-LS1 (Stockhaus et al. 1987) that is expressed in all photosynthetically active tissues of a plant.

All chimeric *tfdA* genes were cloned either directly into binary plant transformation vectors such as pGA470, developed by An et al. (1985), or pBin19, developed by Bevan (1984), which was conjugally transferred into disarmed strains of *Agrobacterium tumefaciens*, or they were cloned into an intermediate vector that was transferred to *Agrobacterium* and recombined into a disarmed Ti plasmid. The binary plasmids are self-replicating and carry left and right T-DNA borders that direct integration into the plant genome, and a selectable kanamycin resistance gene that allows the recovery of transformed plant cells. The transfer functions are provided by modified Ti plasmids deleted for their own T-DNA. Lyon et al. used *Agrobacterium* strains LBA4404 (Hoekema et al. 1983) for transformation of tobacco and AGL1 (Lazo et al. 1991) for transformation of cotton. Bayley et al. used the octopine *Agrobacterium* strain GV3111 (containing a disarmed pTiB6S3) for both tobacco and cotton transformation. The system used by Streber and Willmitzer also used a kanamycin resistance gene as a selectable marker for tobacco transformation, but the chimeric *tfdA* gene, selectable marker, and T-DNA borders were present on the disarmed Ti plasmid GV3850Kan (Jones et al. 1985) rather than on a separate replicon in the *Agrobacterium* strain.

6. EXPRESSION OF THE 2,4-D DIOXYGENASE IN TRANSGENIC TOBACCO

Although none of the researchers were interested in 2,4-D-tolerant tobacco *per se*, tobacco provides a useful model system with which to test the function and specificity of introduced transgenes. Transgenic tobacco plants were produced by *Agrobacterium* infection of leaf pieces using relatively standard protocols with selection on toxic levels of kanamycin sulfate. Streber and Willmitzer (1989) screened 25 plants for production of nopaline (produced from an intact nopaline synthase gene present on the T-DNA) and for expression of the *tfdA* transcript in Northern blot

COR-AAD1-00001409

analyses. Of the 25 plants, 20 were producing detectable levels of the expected 1.3-kb chimeric *tfdA* transcript. Lyon et al. (1989) produced >30 transgenic tobacco from each of their three *tfdA* constructs and carried out preliminary screening with a radioactive enzyme assay using ^{14}C-labeled 2,4-D. While this assay worked well with bacterial extracts, it was only semiquantitative with plant extracts, but allowed the selection of the plants expressing the highest levels of the *tfdA* gene product. Since then the designation of the *tfdA* gene product has changed from a monooxygenase to a dioxygenase, new cofactor requirements have been discovered (Fukumori and Hausinger 1993), and it is likely that the original assay could be improved with the addition of ferrous ions and α-ketoglutarate. Bayley et al. (1992) also used a biochemical assay to measure expression of the *tfdA* gene product in transgenic tobacco, but they used whole cells (protoplasts) where endogenous cofactors would be present. They demonstrated an increase in $^{14}CO_2$ production from degraded ^{14}C-labeled 2,4-D in transgenic plants, ranging from 11 to 85 times higher than that from untransformed control tobacco tissues.

2,4-D has a drastic effect on plant growth and development, and when incorporated into tissue culture media can drastically perturb the effects of other hormones in the medium. When added to regeneration medium shoot development can be inhibited, and this provided a convenient assay procedure for ranking the 2,4-D tolerance of different transgenic plants. Streber and Willmitzer placed explanted leaf disks from their transgenic tobacco onto a medium containing 1 mg/l benzyl amino purine supplemented with varying amounts of 2,4-D from 0.1 to 4.0 mg/l. Nontransgenic plants would produce shoots on medium without 2,4-D, but shoot regeneration was inhibited by 0.2 mg/l 2,4-D. Depending on the individual transgenic plant, shoot regeneration was not inhibited by as much as 2 to 4 mg/l 2,4-D for plants transformed with a 35S-*tfdA* gene construct, indicating at least a 20-fold increase in tolerance to this herbicide. Plants transformed with the *tfdA* gene driven by the light-inducible promoter showed no more tolerance to 2,4-D than nontransformed tissues in this shoot regeneration assay, probably because the light-inducible promoter was not functional during the early stages of shoot morphogenesis, when growth is essentially nonphotosynthetic. In a similar assay Lyon et al. (1989) were able to produce plants with a tolerance to as high as 12.8 mg/l 2,4-D, representing a 32-fold increase in tolerance over a nontransformed control. The increased resistance in different transgenic plants varied from 10- to 30-fold over the control plants. In most cases the resistance in the leaf disk assay correlated well with the intensity of the DCP band in the radioactive enzyme assay. Genomic Southern analysis on some of the plants indicated that most contained single-copy, or at most two-copy, insertions of the 35S-AMV-*tfdA*-Nos chimeric gene described in this study. Bayley et al. (1992) used the shoot inhibition assay to select for the highest expressing tobacco plants and were able to recover seven primary transformants that could produce shoots on 6 mg/l 2,4-D and two others that could produce shoots on 4 mg/l 2,4-D.

In order to look at the segregation of the 2,4-D resistance gene in progeny Streber and Willmitzer tested the germination of 2,4-D-tolerant lines on medium containing 0 to 0.1 mg/l 2,4-D, but observed no significant difference in the sensitivity between transformed and nontransformed plants. Lyon et al. (1989), however, germinated F1 progeny seeds on a suitable germination medium containing from 0 to 3.2 mg/l 2,4-D and saw significant differences between controls and plants that showed high tolerance in the shoot inhibition assay. Seeds of control plants failed to germinate, or, if cotyledons emerged, failed to elongate on as little as 0.2 mg/l 2,4-D. Transgenic plants expressing 2,4-D dioxygenase activity, on the other hand, germinated and elongated on as much as 3.2 mg/l 2,4-D. In some progeny sets >90% of the seeds germinated and continued to grow, suggesting multiple gene insertions, while in others there seemed to be the expected three-to-one segregation of a single-copy insertion of a dominant character. At higher levels of 2,4-D there were some seeds obviously more resistant than the rest, and it seems likely that these were the homozygotes.

Selected transgenic tobacco lines or progeny from plants showing increased tolerance in the shoot inhibition assay were clonally or sexually propagated and tested in the glasshouse for increased tolerance to the herbicidal action of 2,4-D sprayed to simulate a field application. Streber

COR-AAD1-00001410

and Willmitzer (1989) sprayed clonally propagated plants after 2 weeks' growth in the soil with 2,4-D at doses equivalent to 1, 3, and 10 kg/ha (10 kg/ha was 4 ml of a 1% aqueous solution of 2,4-D sodium salt). Plant damage was assessed after 2 to 3 weeks. A dose of 1 kg/ha totally inhibited growth of nontransgenic control plants. A plant transformed with the 35S-*tfdA* construct was resistant to 10 kg/ha, while a plant transformed with the light-inducible promoter-driven *tfdA* gene was resistant to 1 kg/ha 2,4-D even though it showed no increased tolerance in the shoot inhibition assay. Lyon et al (1989) sprayed F1 progeny seed germinated in soil in small pots (four seedlings per pot) and grown in a containment glasshouse until they were 6 weeks old. Groups of 16 plants each were sprayed with a commercial preparation of 2,4-D isopropyl ester in 50% acetone at concentrations from 0 to 3000 ppm (Figure 5). Control plants sprayed with as little as 30 ppm began to show symptoms after only 3 d when they began to wilt and show abnormal twisting of the stems. Many of the transgenic plants only began to show these same symptoms at 3000 ppm. Some of the transgenic plants were affected more than others, presumably due to the segregation of the 2,4-D resistance gene among the F1 progeny. After about 3 weeks control plants sprayed with >100 ppm were dead or dying, while many of the clearly resistant transgenic plants had grown out of the effects of the herbicide even when sprayed at 3000 ppm. While it is difficult to estimate the equivalent field application rate corresponding to our experimental sprayings we estimate that resistance to 3000 ppm in our system corresponds to at least 3 to 6 kg/ha, which is at least 6- to 12-fold above normal field application rates for 2,4-D on cereals (around 0.5 kg/ha). Bayley et al. (1992) sprayed clones of their primary transformants with 10 ml of either 0.2, 0.5, and 1% aqueous solutions of 2,4-D with estimated dosages of 4, 10, and 20 kg/ha, although this may be on the high side. Many of their plants survived even the 20 kg/ha dose when assessed after 3 weeks.

A. eutrophus JMP134 is able to degrade other substituted phenoxyacetate compounds than 2,4-D, so Streber and Willmitzer (1989) tested their transgenic plants for cross-resistance to some other phenoxyacetate herbicides using the shoot inhibition tissue culture assay. They saw a tolerance similar to that seen with 2,4-D when 4-chlorophenoxyacetic acid was included in the medium and

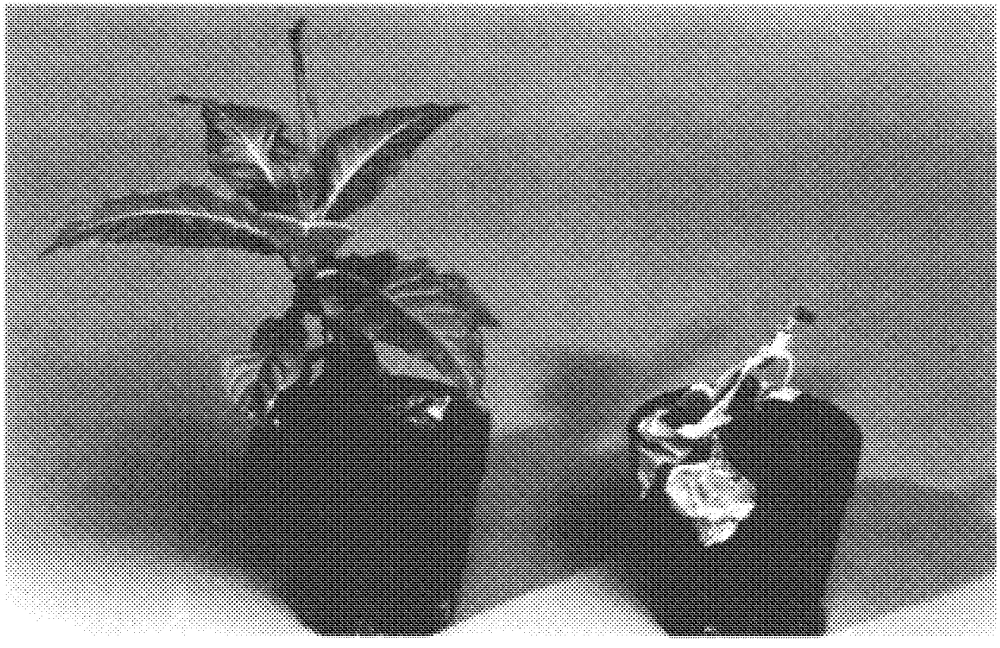


FIGURE 5 2,4-D dioxygenase gene expressed in transgenic tobacco confers tolerance to spraying with 2,4-D. Plants were sprayed with 1000 ppm 2,4-D isopropyl ester in 50% acetone until dripping wet. Damage was scored after 3 weeks. Transgenic 2,4-D-tolerant tobacco line is on the left and control sensitive tobacco on the right.

COR-AAD1-00001411

only a slight increase in tolerance with MCPA, but no increased tolerance to 2,4,5-T or any of the phenoxypropionic acid herbicides [2-(4-chlorophenoxy)propionic acid. 2-(2,4- dichlorophenoxy)-propionic acid, or 2-(2,4,5-trichlorophenoxy)propionic acid]. Similar results were reported by Lyon et al. (1989) using a seed germination assay. These data are consistent with the known specificity of the *tfd* degradative plasmid of *A. eutrophus* for these differently substituted phenoxyacetates. The enzyme specificity also means that the 2,4-D dioxygenase will not degrade natural plant auxins, such as indole acetic acid (Figure 1), and hence is not toxic to the plants.

7. TRANSGENIC COTTON SHOWING TOLERANCE TO 2,4-D

While there are several crops that could benefit from being made tolerant to the herbicide 2,4-D, little interest has been shown in any but cotton. Perkins et al. (1987) expressed interest in using a *tfdA* gene to confer 2,4-D tolerance on the legume cowpea (*Vigna sinensis*), but although they cloned a duplicated *tfdA* gene (*tfdAII*) from the *Alcaligenes* plasmid pJP4 (Perkins and Lurquin, 1988), they have not continued the project (Lurquin, personal communication), mainly because of concerns over the safety of 2,4-DCP. With this in mind, they cloned and sequenced the genes encoding later steps of the 2,4-D breakdown pathway, *tfdB* (encoding 2,4-DCP hydroxylase) and the chlorocatechol oxidative operon *tfdCDEF* (Perkins et al. 1990) with the idea of using these genes in transgenic plants.

Cotton has been the only commercial crop that has been engineered for 2,4-D tolerance, although it has been difficult, as has the regeneration of cotton plants from tissue culture, let alone transformation. Regenerability is highly genotype dependent and it was not until recently that reliable regeneration via somatic embryogenesis could be achieved (Gawel et al. 1986, Trolinder and Goodin 1987), and then predominantly with "Coker" varieties of cotton. Transformation was reported for Coker cottons using *Agrobacterium*-mediated gene transfer and regeneration from either seedling cotyledon (Firoozabady et al. 1987) or hypocotyl (Umbeck et al. 1987) explants. Transformed callus tissue was selected on media containing kanamycin at relatively low levels (25 to 35 μg/ml) and somatic embryos induced from the callus. These embryos were germinated on appropriate media and grown to maturity in soil. The frequency of transformation even with Coker cotton is still very low but managable. Cousins et al. (1991) developed a tissue culture system and transformation procedures for two commercial Australian cotton cultivars, offering the possibility for direct transformation of elite lines.

Bayley et al. (1992) transformed Coker 312 cotton with the *tfdA* gene from pJP5 driven by the 35S promoter, and regenerated 300 transgenic plantlets from 30 independent cell lines. They prescreened them for increased tolerance to 2,4-D by selection for rooting on medium containing 0.05 mg/l 2,4-D or by treating meristems with 0.01 mg/l 2,4-D. These levels inhibited new root and shoot development on nontransformed control plants. Three transformed lines that survived these treatments were transferred to soil for further analysis, and vegetatively propagated clones also tested for increased tolerance to commercial formulations of 2,4-D. Many of the primary transformants were unaffected by a dosage (1.5 kg/ha) of three times the normal field application rate, whereas control plants were severely affected by a level 300 times less than the recommended dose. Selfed progeny of some lines were assayed for 2,4-D degradation using the protoplast whole cell assay for 2,4-D dioxygenase and were shown to have between 17- and 38-fold the activity present in nontransformed cotton protoplasts. Germination assays showed that these plants could form roots on medium containing 0.1 mg/l 2,4-D, a level that inhibited germination of nontransformed plants. This assay also was used to show three-to-one segregation of the resistance phenotype in one of the transformed lines. Seedlings from the same line when germinated in the absence of 2,4-D, were sprayed with 2,4-D at 1.5 kg/ha, and 83 of the 107 plants sprayed survived this treatment. Transgenic seeds are being distributed to public and private breeders in the U.S. for incorporation of the trait into other commercial germplasm (Trolinder, personal communication).

COR-AAD1-00001412

Lyon et al. (1993) transformed Coker 315 cotton with one of the *tfdA* gene constructs used to transform tobacco and regenerated 201 plantlets from 11 different tissue culture lines. Only four of these produced self-fertile plants when transferred to the glasshouse. Southern analysis and progeny testing indicated that only one of the four lines had a single T-DNA insert, while the others had two copies of the T-DNA. There was no direct correlation between copy number and the level of resistance shown by spraying plants with 2,4-D. 2,4-D dioxygenase expressing progeny from all four lines was undamaged by 300 ppm 2,4-D isopropyl ester in 50% acetone (Figure 6). The line containing a single copy of the chimeric gene was shown to be the most tolerant and was resistant to over 600 ppm 2,4-D. T2 and T3 plants shown to be either hemizygous or homozygous indicated that at discriminating doses the resistance phenotype is semidominant (Lyon et al. 1993). When sprayed with 600 ppm 2,4-D the hemizygous plants containing a single gene copy are slightly damaged, but eventually grow out of the damage and have a normal yield and appearance. Homozygous plants are completely unaffected by this level of 2,4-D, but show some minor damage at 900 ppm and more severe damage at 1200 ppm, although they survive this higher dose. The best transgenic plants were estimated to be about 100-fold more tolerant to 2,4-D than control nontransformed cotton and were tolerant to about twice the normal field application rate. CSIRO is currently backcrossing the 2,4-D tolerance developed by Lyon et al. (1993) into important commercial Australian cultivars to provide them with spray drift protection from 2,4-D.

None of the 2,4-D-tolerant plants have progressed to field trials. In Australia the research group at CSIRO is continuing to develop new gene constructs that may have better 2,4-D tolerance levels in cotton. Because the site of action of the herbicide is at the shoot meristem they are investigating



FIGURE 6 2,4-D-tolerant cotton cultivars engineered to express the *A. eutrophus* 2,4-D dioxygenase gene. Primary Coker transformants and controls were sprayed with 300 ppm 2,4-D isopropyl ester in 50% acetone. The plant on the left contains a single T-DNA insertion of the *tfdA* gene and was unaffected by 2,4-D at the concentration tested. The plant on the right is a nontransformed control cotton that was rapidly killed by the herbicide.

COR-AAD1-00001413

using gene promoters other than the 35S promoter, which has relatively poor expression in the meristem. Promoters from genes, such as histones and ubiquitin that are expressed at high levels in actively dividing cells, were linked to the *tfdA* gene and are currently being assessed in transgenic tobacco. Expression of the second gene in the 2,4-D degradation pathway, 2,4-D hydroxylase (*tfdB*) (Perkins et al. 1990), is also being considered to detoxify the DCP produced by the 2,4-D dioxygenase. Increased tolerance to 2,4-D will provide a greater protection margin and may allow 2,4-D to be used directly on cotton to control problem broadleaf weeds.

8. 2,4-D-TOLERANT COTTON PLANTS — AN ETHICAL DILEMMA?

The ability to produce 2,4-D-tolerant crops such as cotton raises the ethical dilemma of whether these plants should be used in agriculture. Opponents of herbicide-tolerant crops warn that herbicide-tolerant plants will promote indiscriminant usage of herbicides with concomitant human health and environmental impacts (Goldburg et al. 1990; Phelps 1990), but are these fears justified with 2,4-D?

As 2,4-D has been used in agriculture for almost 50 years it sometimes has been labeled as an "older chemical," and like several other older chemicals it has acquired a tainted image as a suspected environmental and human health hazard. The name 2,4-D has become associated in the public mind with less benign chlorinated hydrocarbons such as DDT (dichlorodiphenyltrichloroethane). Unlike DDT, which accumulates in body fat, 2,4-D is rapidly excreted in an unchanged form from the human body. The biological half-life of 2,4-D in humans is approximately 18 to 20 h (reviewed by Ibrahim et al. 1991). Similarly, 2,4-D is unlikely to accumulate in aquatic organisms, as exemplified by the rapid renal elimination of 2,4-D administered to channel catfish (Plakas et al. 1992). This contrasts markedly with DDT, which is concentrated manyfold in the tissues of aquatic organisms, especially shellfish. 2,4-D is also rapidly broken down in the soil by bacteria and fungi, typically becoming undetectable about 5 to 6 weeks after application (Nobel and Hargreaves 1992). Soils with a history of 2,4-D treatment were shown to have an increased population of 2,4-D-degrading microorganisms (Holben et al. 1992). The unpopularity of 2,4-D in some circles stems in part from its inclusion along with 2,4,5-T in Agent Orange. However, the human health hazard posed by Agent Orange was caused by a contaminant, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin, an extremely toxic compound which causes a severe form of acne and is suspected of causing birth defects and specific cancers. Modern 2,4-D formulations do not contain this dangerous contaminant nor do they contain significant quantities of any other known carcinogens (Ibrahim et al. 1991).

2,4-D has a low acute oral or dermal toxicity. The Western Australian Department of Agriculture lists 2,4-D as "slightly toxic" (poison schedule S5), placing it in the same category as glyphosate, sodium chlorate, and kerosene. The LD_{50} values for 2,4-D (reviewed in Ibrahim et al. 1991) range from 100 mg/kg in dogs (which are known to excrete organic acids more slowly than rodents or humans) to 1000 mg/kg in guinea pigs. Chronic exposure to 2,4-D in rats did not produce any observable detrimental effects at levels of up to 30 mg/kg/d. Studies reviewed by Mullison (1988) concluded that 2,4-D is not carcinogenic in mice at up to 45 mg/kg/d, nor is it teratogenic in rats even at a dosage that caused toxic symptoms in pregnant rats (75 mg/kg/d). There has been a suggestion, based on epidemiological evidence, that exposure to 2,4-D may be linked with the occurrence of the human cancer non-Hodgkin's lymphoma (Woods et al. 1987; Zahm et al. 1990), and this has been the subject of much debate. A panel of 13 scientists reviewing the literature were unable to reach a single conclusion as to whether 2,4-D is a human carcinogen (Ibrahim et al. 1991). Of the 13, five thought that it was a possible but unlikely human carcinogen, while one considered

COR-AAD1-00001414

that it was a possible carcinogen. Two panelists thought that it was unlikely that 2,4-D can cause cancer in humans. Following a report of a possible link between 2,4-D and canine malignant lymphoma (Hayes et al. 1991), a scientific review panel examined the available evidence and concluded that the use of 2,4-D by dog owners did not show an association with the occurrence of canine malignant lymphoma (Carlo et al. 1992). In other words, there is no conclusive evidence to implicate 2,4-D as a carcinogen.

The *Agrochemicals Handbook* of the Royal Society of Chemistry (Hartley and Kidd 1983) lists 2,4-D as nontoxic to fish (average lethal concentration of sodium salt, 1160 mg/l/24 h) and nontoxic to bees. 2,4-D is only toxic to waterfowl at very high concentrations (LD_{50} = 2000 mg/kg; Weiss 1986), which would never be approached by the use of it in agriculture.

2,4-D has many of the desirable qualities sought in a herbicide, most notably its low toxicity to animal life and high biodegradability. Furthermore, it is very inexpensive and effective. Its use over many years has been relatively untroubled by problems of resistance development in weeds. There are only six sites in the world where 2,4-D-resistant weeds are thought to have evolved as a result of spraying, compared to over 1000 sites for herbicides such as the triazines and acetolactate synthase inhibitors (Holt et al. 1993). 2,4-D does, however, have a drawback in that it does not adsorb readily to some soils and can leach into groundwater. Precautions should be taken to ensure that pollution of ground water does not occur.

Like any agricultural chemical, 2,4-D should not be used carelessly, but the large body of toxicological data gathered over the past 40 or so years suggests that it can be used according to label directions without endangering humans or the environment.

9. CONCLUSIONS

The herbicide detoxification systems developed for the protection of cotton against 2,4-D have proven to be extremely effective in at least two plant species, and we expect that it also will be useful in other crops that are sensitive to 2,4-D. The plants containing the engineered resistance gene appear to grow and flower normally, producing just as many viable seeds as nontransformed plants, and there appears to be no yield penalty in the absence of any applied herbicide, at least in glasshouse trials. When sprayed at relatively high levels of 2,4-D there is a slight retardation of growth, but the plants recover quickly. Because most commonly used herbicides are broken down by soil microorganisms it would seem feasible to apply the same principles to those herbicides and to quickly develop a whole range of herbicide resistance genes that could, once transformation procedures become routine, revolutionize farming practices in high-input crops such as cotton.

Herbicide-tolerant crop projects, including the 2,4-D-resistant cotton projects described here, have been criticized by environmental groups, such as the Australian Conservation Foundation (Phelps 1990), on the grounds that they could lead to an increase in the use of synthetic chemicals in agriculture. This fear should be balanced with the environmental benefits of being able to extend the spectrum of use of relatively benign herbicides such as 2,4-D and thereby reduce our reliance on more environmentally damaging herbicides. Additionally, control of weeds within the crop with 2,4-D could potentially lead to a no-till cropping system which would have great benefits in reducing soil erosion, providing a mulch, and returning organic nitrogen to the soil.

Although widely held public concern about the use of 2,4-D-resistant cotton are not supported by scientific data, it is nevertheless essential that careful studies in the field are carried out in order to evaluate any possible environmental implications. This will enable cropping systems to be designed so that the new technology can be applied responsibly in a way that benefits the farmer, the consumer, and the environment.

COR-AAD1-00001415

REFERENCES

An, G., Watson, B., Stachel, S., Gordon, M., and Nester, E. 1985. New cloning vehicles for transformation of higher plants. *EMBO J.* 4: 277.

Ashton, F.M. and Crafts, A.S. 1981. *Mode of Action of Herbicides,* 2nd ed. John Wiley & Sons, New York.

Bayley, C., Trolinder, C., Morgan, M., Quisenberry, J.E., and Ow, D.W. 1992. Engineering 2,4-D resistance into cotton. *Theor. Appl. Genet.* 83: 645.

Bevan, M. 1984. Binary *Agrobacterium* vectors for plant transformation. *Nucl. Acids Res.* 12: 8711.

Carlo, G.L., Cole, P., Miller, A.B., Munro, I.C., Solomon, K.R., and Squire, R.A. 1992. Review of a study reporting an association between 2,4-dichlorophenoxyacetic acid and canine malignant lymphoma: report of an expert panel. *Regul. Toxicol. Pharmacol.* 16: 245.

Charles, G.W. 1991. A grower survey of weeds and herbicide use in the New South Wales cotton industry. *Aust. J. Exp. Agric.* 31: 387.

Comai, L., Facciotti, D., Hiatt, W., Thompson, G., Rose, R., and Stalker, D. 1985. Expression in plants of a mutant *aroA* gene from *Salmonella typhimurium* confers tolerance to glyphosate. *Nature* 317: 741.

Cousins, Y.L., Lyon, B.R., and Llewellyn, D.J. 1991. Transformation of an Australian cotton cultivar: prospects for cotton improvement through genetic engineering. *Aust. J. Plant Physiol.* 18: 481.

DeBlock, M., Botterman, J., Vandewiele, M., Dockyx, J., Thoen, A., Gossele, V., Rao Movva, N., Thompson, C., Van Montagu, M., and Leemans, J. 1987. Engineering herbicide resistance in plants by expression of a detoxifying enzyme. *EMBO J.* 6: 2513.

Dhaese, P., De Greve, H., Gielen, J., Seurinck, J., Van Montagu, M., and Schell, J. 1983. Identification of sequences involved in the polyadenylation of higher plant nuclear transcripts using *Agrobacterium* T-DNA genes as models. *EMBO J.* 2: 419.

Dodge, A.D. 1989. Principles. In *Weed Control Handbook.* Ed. by Hance, R.J. and Holly, K. Blackwell Scientific, Oxford, p. 201.

Don, R. and Pemberton, J. 1981. Properties of six pesticide degradation plasmids isolated from *Alcaligenes eutrophus* and *Alcaligenes paradoxus. J. Bacteriol.* 145: 681.

Don, R., Weightman, A., Knackmuss, H., and Timmis, K. 1985. Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4). *J. Bacteriol.* 161: 85.

Evans, S.A. 1968. Application of herbicides. In *Weed Control Handbook,* Vol. 1, 5th ed. Ed. by Fryer, J D. and Evans, S.A. Blackwell Scientific, Oxford, p. 149.

Firoozabady, E., DeBoer, D., Merlo, D., Halk, E., Amerson, L., Rashka, K., and Murray, E. 1987. Transformation of cotton (*Gossypium hirsutum* L.) by *Agrobacterium tumefaciens* and regeneration of transgenic plants. *Plant Mol. Biol.* 10: 105.

Fukumori, F. and Hausinger, R.P. 1993. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetate monooxygenase" is an α-ketoglutarate-dependent dioxygenase. *J. Bacteriol.* 175: 2083.

Gawel, N., Rao, A., and Robacker, C. 1986. Somatic embryogenesis from leaf and petiole callus cultures of *Gossypium hirsutum. Plant Cell Rep.* 5: 457.

Goldburg, R., Rissler, J., Shand, H., and Hasselbrook, C. 1990. Biotechnology's Bitter Harvest: Herbicide Tolerant Crops and the Threat to Sustainable Agriculture. Biotechnology Working Group, Washington, D.C., p. 53.

Hammond, D. 1992. Managing 2,4-D in cotton fallows. *Aust. Cottongrower 13*: 67.

Hartley, D. and Kidd, H. 1983. *Agrochemicals Handbook.* Royal Society of Chemistry, London.

Haughn, G., Smith, J., Mazur, B., and Sommerville, C. 1988. Transformation with a mutant *Arabidopsis* acetolactate synthase gene renders tobacco resistant to sulfonylurea herbicides. *Mol. Gen. Genet.* 211: 266.

Hayes, H.M., Tarone, R.E., Cantor, K.P., Jessen, C.R., McCurnin, D.M., and Richardson, R.C. 1991. Case-control study of canine malignant lymphoma: positive association with dog owner's use of 2,4-dichlorophenoxyacetic acid herbicides. *J. Nat. Cancer Inst.* 83: 1226.

Herrera-Estrella, L., De Block, M., Messens, E., Hernalsteens, J., Van Montagu, M., and Schell, J. 1983. Chimeric genes as dominant selectable markers in plant cells. *EMBO J.* 2: 987.

Hoekema, A., Hirsch, P., Hooykaas, P., and Schilperoort, R. 1983. A binary plant vector strategy based on separation of *vir*- and T-region of *Agrobacterium tumefaciens* Ti-plasmid. *Nature* 303: 179.

Holben, W.E., Schroeter, B.M., Calbrese, V.G.M., Olsen, R.H., Kukor, J.K., Diederbeck, V.O., Smith, A.E., and Tiedje, J.M. 1992. Gene probe analysis of soil microbial populations selected by amendment with 2,4-dichlorophenoxyacetic acid. *Appl. Environ. Microbiol.* 58: 3941.

Holt, J. S. 1993. Mechanisms and agronomic aspects of herbicide resistance. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 44: 203.

Ibrahim, M.A., Bond, G.G., Burke, T.A., Cole, P., Dost, F.N., Enterline, M., Gough, R.S., Greenberg, W., Halperin, E., McConnell, M., Munro, I.C., Swenberg, J.A., Zahm, S.H., and Graham, J.D. 1991. Weight of evidence on the human carcinogenicity of 2,4-D. *Environ. Health Perspect. 96*: 213.

COR-AAD1-00001416

Jobling, S. and Gerhke, L. 1987. Enhanced translation of chimeric mRNAs containing plant viral untranslated leader sequences. *Nature* 325: 622.

Jones, J., Dunsmuir, P., and Bedbrook, J. 1985. High level expression of introduced chimaeric genes in regenerated transformed plants. *EMBO J.* 4: 2411.

Joshi, C. 1987. An inspection of the domain between putative TATA box and translation start sites in 79 plant genes. *Nucl. Acids Res.* 15: 6643.

Kay, R., Chan, A., Daly, M., and McPherson, J. 1987. Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science* 236: 1299.

Keeley, P.E., Thullen, R.J., and Carter, C.H. 1986. Influence of planting date on growth of ivyleaf morningglory (*Ipomoea hederacea*) in cotton (*Gossypium hirsutum*). *Weed Sci.* 34: 906.

Kirkwood, R.C. 1979. The uptake and translocation of foliar applied herbicides using an explant system. In *Advances in Pesticide Science.* Ed. by Geissbuhler, H. Pergamon Press, Oxford, p. 410.

Lazo, G.R., Stein, P.A., and Ludwig, R.A. 1991. A DNA transformation competent *Arabidopsis* genomic library in *Agrobacterium. Bio/Technology* 9: 963.

Llewellyn, D., Lyon, B.R., Cousins, Y.L., Huppatz, J., Dennis, E.S., and Peacock, W.J. 1990. Genetic engineering of plants for resistance to the herbicide 2,4-D. In *Genetic Engineering of Crop Plants.* Ed. by Lycett, G.W. and Grierson, D. Butterworths, London, p. 67.

Loos, M.A. 1969. Phenoxyalkanoic acids. In *Degradation of Herbicides.* Ed. by Kearney, P.C. and Kaufman, D.D. Marcel Dekker, New York, p. 1.

Lyon, B.R., Cousins, Y.C., Llewellyn, D.J., and Dennis, E.S. 1993. Cotton plants transformed with a bacterial degradation gene are protected from accidental spray drift damage by the herbicide 2,4-dichlorophenoxyacetic acid. *Transgenic Res.* 2: 162.

Lyon, B.R., Llewellyn, D.J., Huppatz, J., Dennis, E.S., and Peacock, W.J. 1989. Expression of a bacterial gene in transgenic tobacco confers resistance to the herbicide 2,4-dichlorophenoxyacetic acid. *Plant Mol. Biol.* 13: 533.

Mullison, W.R. 1988. A 2,4-D toxicological update. *Proc. West Soc. Weed Sci.* 41: 20.

Noble, B. and Hargreaves, P. 1992. Are herbicide residues accumulating in our soils? *Austr. Grain* 2: 10.

Odell, J., Nagy, F., and Chua, N.H. 1985. Identification of DNA sequences required for activity of a plant promoter: the CaMV promoter. *Nature* 313: 810.

Perkins, E.J. and Lurquin, P.F. 1988. Duplication of a 2,4-dichlorophenoxyacetic acid monooxygenase gene in *Alcaligenes eutrophus* JMP134(pJP4). *J. Bacteriol.* 170: 5669.

Perkins, E.J., Stiff, C.M., and Lurquin, P.F. 1987. Use of *Alcaligenes eutrophus* as a source of genes for 2,4-D resistance in plants. *Weed Sci.* 35: 12.

Perkins, E.J., Gordon, M.P., Caceres, O., and Lurquin, P.F. 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. *J. Bacteriol.* 172: 2351.

Phelps, R. 1990. A leap into the unknown. *Consuming Interests* August, p. 5.

Pillmoor, J.B. and Gaunt, J.K. 1981. The behaviour and mode of action of phenoxyacetic acid herbicides. In *Progress in Pesticide Biochemistry,* Vol. 1. Ed. by Hutson, D.H. and Roberts, T.R. Wiley-Interscience, New York, p. 147.

Plakas, S.M., Khoo, L., and Barron, M.G. 1992. 2,4-Dichlorophenoxyacetic acid disposition after oral administration in channel catfish. *J. Agric. Food Chem.* 40: 1236.

Shah, D., Horsch, R., Klee, H., Kishore, G., Winter, J., Tumer, N., Hironaka, C., Sanders, P., Gasser, C., Aykent, S., Siegel, N., Rogers, S., and Fraley, R. 1986. Engineering herbicide tolerance in transgenic plants. *Science* 233: 478.

Singh, S.P. and Muller, F. 1979. Translocation of 2,4-D, asulam and amitrole in water hyacinth. *Weed Res.* 19: 171.

Stalker, D., McBride, K., and Malyj, L. 1988. Expression in plants of a bromoxynil-specific bacterial nitrilase that confers herbicide resistance. In *Genetic Improvements of Agriculturally Important Crops: Progress and Issues.* Ed. by Fraley, R.T., Frey, N.M., and Schell, J. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, p. 37.

Stockhaus, J., Eckes, P., Rocha-Sosa, M., Schell, J., and Willmitzer, L. 1987. Analysis of *cis* active sequence involved in leaf specific expression of a potato gene in transgenic plants. *Proc. Natl. Acad. Sci. U.S.A.* 84: 7943.

Streber, W., Timmis, K., and Zenk, M. 1987. Analysis, cloning and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. *J. Bacteriol.* 169: 2950.

Streber, W.R. and Willmitzer, L. 1989. Transgenic tobacco expressing a bacterial detoxifying gene are resistant to 2,4-D. *Bio/Technology* 7: 811.

Thompson, C.J., Movva, N.R., Tizard, R., Craneri, R., Davies, J.E., Lauwereys, M., and Botterman, J. 1987. Characterization of the herbicide resistance gene *bar* from *Streptomyces hygroscopicus. EMBO J.* 6: 2519.

Trolinder, N. and Goodin, J. 1987. Somatic embryogenesis and plant regeneration in cotton (*Gossypium hirsutum* L.). *Plant Cell Rep.* 6: 231.

Umbeck, P., Johnson, G., Barton, K., and Swain, W. 1987. Genetically transformed cotton (*Gossypium hirsutum* L.) plants. *Bio/Technology* 5: 263.

Weiss, G. (Ed.) 1986. *Hazardous Chemicals Data Book,* 2nd Ed. Noyes Data Corporation, Park Ridge, NJ, p. 363.

COR-AAD1-00001417

Woods, J.A., Polissar L., Severson, R.K., Heuser, L.S., and Kulander, B.G. 1987. Soft tissue sarcoma and non-Hodgkin's lymphoma in relation to phenoxyherbicide and chlorinated phenol exposure in western Washington. *JNCI* 78: 899.

Zahm, S.H., Weissenburger, D.D., Babbit, P.A., Saal, R.C., Vaught, J.B., Cantor, K.P., and Blair, A. 1990. A case-control study of non-Hodgkin's lymphoma and the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) in eastern Nebraska. *Epidemology* 1: 349.

COR-AAD1-00001418

Electronic Acknowledgement Receipt

EFS ID:	40617111
Application Number:	15288406
International Application Number:	
Confirmation Number:	5402
Title of Invention:	NOVEL HERBICIDE RESISTANCE GENES
First Named Inventor/Applicant Name:	Terry WRIGHT
Customer Number:	107540
Filer:	John P. Breen/Garla Twyman
Filer Authorized By:	John P. Breen
Attorney Docket Number:	14764-258662
Receipt Date:	22-SEP-2020
Filing Date:	07-OCT-2016
Time Stamp:	15:59:28
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Non Patent Literature	Poiger.pdf	1205653 65bb8d4aeacd8349d812174e5584cf6861787f39	no	8

Warnings:

Information:					
2	Non Patent Literature	Kaneko.pdf	301104 0ee0aa3a1a1eaa010c3605201286c127077e1948	no	9
Warnings:					
Information:					
3	Non Patent Literature	Rousseaux.pdf	468991 2835d16338da7e8fd2dbcea660c34c091c4ad338	no	12
Warnings:					
Information:					
4	Non Patent Literature	Saari.pdf	338362 73b907e3ef7fda9e090ee9cb02b7f23605d567ec	no	8
Warnings:					
Information:					
5	Non Patent Literature	SCHNEIDERHEINZE.pdf	157650 2f0d84bf93139fe00c39dca25439e6d6447fba3e	no	8
Warnings:					
Information:					
6	Non Patent Literature	Seifert.pdf	116691 bdaa6f899ded1aceaf7abd3b826e1680eb1f7734	no	9
Warnings:					
Information:					
7	Non Patent Literature	Shaw.pdf	418203 362f1c919d37ed84eeba6a0c93a7fd67f662571e	no	9
Warnings:					
Information:					
8	Non Patent Literature	Smejkal.pdf	247760 de6320468f20ba2e1fbc8e3f316f729ac8077723	no	7
Warnings:					
Information:					

9	Non Patent Literature	Streber_Transgenic_tobacco.pdf	1529677 670954e57ca902f0365dbc693f974ad55e1ef33c	no	6
Warnings:					
Information:					
10	Non Patent Literature	Swiegers.pdf	333574 2c19444444465911f7603b71503f33f61ce2a99ee	no	5
Warnings:					
Information:					
11	Non Patent Literature	Tett_Enantioselective.pdf	675968 828afabc56d7f66ddfbef18815fcf928d01ffe11	no	9
Warnings:					
Information:					
12	Non Patent Literature	Top_The_role.pdf	139029 4f78f706f0d53489b682bf70748d7ad371852918	no	8
Warnings:					
Information:					
13	Non Patent Literature	Top_Capture.pdf	274965 44134c7515271c37ff3c0e71ee8140fcca10ded0	no	7
Warnings:					
Information:					
14	Non Patent Literature	Top_Methane.pdf	299279 d547e09444ec1edb9d877b8aa4ab6dc91c4a67f5	no	11
Warnings:					
Information:					
15	Non Patent Literature	Travkin_Characterization.pdf	305721 82bcffa477e6b155e068fe53820007b088267df2	no	4
Warnings:					
Information:					

16	Non Patent Literature	Vallaeys_PCR-RFLP.pdf	3492232 10d7f012284a6cfb8575928c66fef0359c9bf9ce	no	10
Warnings:					
Information:					
17	Non Patent Literature	Vallaeys_Metabolic.pdf	828981 cb08d722ddd0a11d23d1f3cd888bc7e759f7667d	no	10
Warnings:					
Information:					
18	Non Patent Literature	Vedler_Analysis.pdf	219268 6f97aad220d6d8c70c94ee7aa2ff15c88daa8025	no	8
Warnings:					
Information:					
19	Non Patent Literature	Vedler_TfdR.pdf	312543 fde2a71350f4a35e0022682a2c37271d0328f6f9	no	8
Warnings:					
Information:					
20	Non Patent Literature	Velicer_Pleiotropic.pdf	120065 127e1f3936f1080bf2de78c22f0211bc73768396	no	6
Warnings:					
Information:					
21	Non Patent Literature	Whiting_Metal.pdf	260916 6abde09d7378260b75600eb85e8bbf9364379994	no	2
Warnings:					
Information:					
22	Non Patent Literature	Yi_Phylogenetic.pdf	344560 020ff877144fece8beb7d745ae8d900a52688b63	no	8
Warnings:					
Information:					

23	Non Patent Literature	Zhang_In_vitro.pdf	63605 4b42f8eb01a685607ca0e38769dab25d7968415e	no	5
Warnings:					
Information:					
24	Non Patent Literature	Zhang_Inheritance.pdf	69866 66556174337e14a0b51b03be853156f4b9050970	no	5
Warnings:					
Information:					
25	Non Patent Literature	Zipper_Complete.pdf	106309 4b60eb4b1eb01550eed604b599e612a58899e263	no	5
Warnings:					
Information:					
26	Non Patent Literature	Zipper_Enantioselective.pdf	149759 92ddf77211af1cecd89eff2b208cb780dfc8fec6	no	7
Warnings:					
Information:					
27	Non Patent Literature	Cho.pdf	993895 ea10d2fb423439fa551347d05df743b4c1ef5894	no	5
Warnings:					
Information:					
28	Non Patent Literature	Lim_Genetic.pdf	667500 5d3325b809a4f840e8b5a1afb6b51f944c88a964	no	7
Warnings:					
Information:					
29	Non Patent Literature	Tamura_Microbial.pdf	3554547 bd9daf99f199961b99dc97abe15c492d0be0ac7d	no	9
Warnings:					
Information:					

30	Non Patent Literature	Vallaeys_Isolation.pdf	90948 14e159e88274e1b5077b5e1d4e04edb02cabd862	no	4
Warnings:					
Information:					
31	Non Patent Literature	Fukamori_and_Hausinger_PURIFICATION.pdf	1933909 6da7f2d030722f71ad5e6f2e20c62e8a17696354	no	7
Warnings:					
Information:					
32	Non Patent Literature	Fukamori_Alkaligenes.pdf	743479 285351f606cb96a1644da93a189fd5847e290c51	no	4
Warnings:					
Information:					
33	Non Patent Literature	Ehrig.pdf	300271 6eba9dd696763469379e474da9661b25c30d77c2	no	6
Warnings:					
Information:					
34	Non Patent Literature	Hausinger.pdf	833045 3e8a30c46ced480be0e59cccdf1f54d0a45d86a4	no	48
Warnings:					
Information:					
35	Non Patent Literature	Kohler_Sphingomonas.pdf	1468107 3f189d83018f17d65abbf9357dcc8312886ffbb9	no	5
Warnings:					
Information:					
36	Non Patent Literature	Hausinger_Characterization.pdf	570979 b3e0d1f497b0a158c5cfeb0752c28ad4d1d2cc49	no	3
Warnings:					
Information:					

37	Non Patent Literature	Don_Transposon.pdf	1089812 0860bd3d95b981b6f02685136c7821ba756e5459	no	7
Warnings:					
Information:					
38	Non Patent Literature	Don_Genetic.pdf	512042 a89f491ad6abf4430ce340eab2b7a09f8530ad64	no	4
Warnings:					
Information:					
39	Non Patent Literature	Hotoff_Alternative.pdf	120580 05f34ae376abf51eaa462c75e96264b46c7fb043	no	8
Warnings:					
Information:					
40	Non Patent Literature	Westendorf_The_two_enantiospecfic.pdf	140777 95fe6487172f5b8605923665493f6ceb28081b27	no	6
Warnings:					
Information:					
41	Non Patent Literature	Park_Isolation.pdf	529400 405670535d003fc74f3973c684084a91decf5cda	no	8
Warnings:					
Information:					
42	Non Patent Literature	Muller_Degradability.pdf	2487522 4c1ac7d099f756d66b76c0d7c6f43130cd265279	no	9
Warnings:					
Information:					
43	Non Patent Literature	Tett_Biodegradation.pdf	284390 34adf5de43d582f121d5666a0d5cd8e04797da3d	no	10
Warnings:					
Information:					

44	Non Patent Literature	Llewellyn.pdf	1215712 a1faedd58863bf19875823cade980e2f727f6bda	no	16
Warnings:					
Information:					
			Total Files Size (in bytes):	30317646	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer:
(11) Publication number:
(11) Numéro de publication:

1 025 250

Internationale Anmeldung veröffentlicht durch die Weltorganisation für geistiges Eigentum unter der Nummer:

WO 99/23231 (art.158 des EPÜ).

International application published by the World Intellectual Property Organisation under number:

WO 99/23231 (art.158 of the EPC).

Demande internationale publiée par l'Organisation Mondiale de la Propriété sous le numéro:

WO 99/23231 (art.158 de la CBE).

COR-AAD1-00001427

PCT WELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales Büro

INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) **Internationale Patentklassifikation** 6 :
C12N 15/82, 15/52, 15/53, 9/00, 9/02, 5/10, G01N 33/50, A01H 5/00

A2

(11) **Internationale Veröffentlichungsnummer:** **WO 99/23231**

(43) **Internationales Veröffentlichungsdatum:** 14. Mai 1999 (14.05.99)

(21) **Internationales Aktenzeichen:** PCT/EP98/06851

(22) **Internationales Anmeldedatum:** 29. Oktober 1998 (29.10.98)

(30) **Prioritätsdaten:**
197 47 739.9 29. Oktober 1997 (29.10.97) DE

(71) **Anmelder** *(für alle Bestimmungsstaaten ausser US)*: INSTITUT FÜR PFLANZENGENETIK UND KULTURPFLANZENFORSCHUNG [DE/DE]; Corrensstrasse 3, D–06466 Gatersleben (DE).

(72) **Erfinder; und**
(75) **Erfinder/Anmelder** *(nur für US)*: GRIMM, Bernhard [DE/DE]; Corrensstrasse 2, D–06466 Gatersleben (DE). TANAKA, Ryouichi [JP/JP]; The Institute of Low Temperature Science, Hokkaido University, N19 W8, Sapporo 060–0819 (JP).

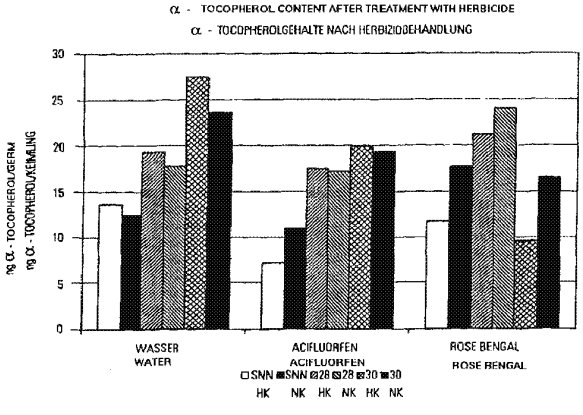
(74) **Anwalt:** MAIWALD, Walter; Maiwald GmbH, Elisenhof, Elisenstrasse 3, D–80335 München (DE).

(81) **Bestimmungsstaaten:** AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht
Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.

(54) **Title:** MODIFICATION OF THE TOCOPHEROL CONTENT OF TRANSGENIC PLANTS

(54) **Bezeichnung:** BEEINFLUSSUNG DES TOCOPHEROLGEHALTES IN TRANSGENEN PFLANZEN



(57) **Abstract**

The invention relates to novel nucleic acid sequences which provide the coding for a geranyl geranyl reductase, to a method for producing novel plants which contain a novel nucleic acid sequence and whose tocopherol and/or chlorophyl content is modified in relation to wild type plants, to these plants and to parts, products and plant cells thereof and to the use of the nucleic acid sequences for influencing the tocopherol, chlorophyl and/or vitamin K_1 content of transgenic plants and parts, products and plant cells thereof.

COR-AAD1-00001428

(57) Zusammenfassung

Die vorliegende Erfindung betrifft neue Nukleinsäuresequenzen, die für eine Geranylgeranyl–Reduktase kodieren, eine Methode zur Erzeugung neuer Pflanzen, die eine neue Nukleinsäuresequenz enthalten und deren Gehalt an Tocopherol und/oder Chlorophyll im Vergleich zu Wildtyppflanzen verändert ist, diese neuen Pflanzen, deren Teile und Produkte und Pflanzenzellen sowie die Verwendung der Nukleinsäuresequenzen zur Beeinflussung des Tocopherol–, Chlorophyll– und/oder Vitamin K_1–Gehaltes in transgenen Pflanzen, deren Teilen und Produkten und Pflanzenzellen.

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	**ES**	Spanien	**LS**	Lesotho	**SI**	Slowenien
AM	Armenien	**FI**	Finnland	**LT**	Litauen	**SK**	Slowakei
AT	Österreich	**FR**	Frankreich	**LU**	Luxemburg	**SN**	Senegal
AU	Australien	**GA**	Gabun	**LV**	Lettland	**SZ**	Swasiland
AZ	Aserbaidschan	**GB**	Vereinigtes Königreich	**MC**	Monaco	**TD**	Tschad
BA	Bosnien-Herzegowina	**GE**	Georgien	**MD**	Republik Moldau	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagaskar	**TJ**	Tadschikistan
BE	Belgien	**GN**	Guinea	**MK**	Die ehemalige jugoslawische Republik Mazedonien	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Griechenland			**TR**	Türkei
BG	Bulgarien	**HU**	Ungarn	**ML**	Mali	**TT**	Trinidad und Tobago
BJ	Benin	**IE**	Irland	**MN**	Mongolei	**UA**	Ukraine
BR	Brasilien	**IL**	Israel	**MR**	Mauretanien	**UG**	Uganda
BY	Belarus	**IS**	Island	**MW**	Malawi	**US**	Vereinigte Staaten von Amerika
CA	Kanada	**IT**	Italien	**MX**	Mexiko		
CF	Zentralafrikanische Republik	**JP**	Japan	**NE**	Niger	**UZ**	Usbekistan
CG	Kongo	**KE**	Kenia	**NL**	Niederlande	**VN**	Vietnam
CH	Schweiz	**KG**	Kirgisistan	**NO**	Norwegen	**YU**	Jugoslawien
CI	Côte d'Ivoire	**KP**	Demokratische Volksrepublik Korea	**NZ**	Neuseeland	**ZW**	Zimbabwe
CM	Kamerun			**PL**	Polen		
CN	China	**KR**	Republik Korea	**PT**	Portugal		
CU	Kuba	**KZ**	Kasachstan	**RO**	Rumänien		
CZ	Tschechische Republik	**LC**	St. Lucia	**RU**	Russische Föderation		
DE	Deutschland	**LI**	Liechtenstein	**SD**	Sudan		
DK	Dänemark	**LK**	Sri Lanka	**SE**	Schweden		
EE	Estland	**LR**	Liberia	**SG**	Singapur		

Beeinflussung des Tocopherolgehaltes in transgenen Pflanzen

Die vorliegende Erfindung betrifft neue Nukleinsäuresequenzen, die für eine Geranylgeranyl-Reduktase kodieren, eine Methode zur Erzeugung neuer Pflanzen, die eine neue Nukleinsäuresequenz enthalten und deren Gehalt an Tocopherol und/oder Chlorophyll im Vergleich zu Wildtyppflanzen verändert ist, diese neuen Pflanzen, deren Teile und Produkte und Pflanzenzellen sowie die Verwendung der Nukleinsäuresequenzen zur Beeinflussung des Tocopherol-, Chlorophyll- und/oder Vitamin K_1-Gehaltes in transgenen Pflanzen, deren Teilen und Produkten und Pflanzenzellen.

Das Diterpen Geranylgeranyl-pyrophosphat (GGPP) entsteht als 20-C-Zwischenprodukt im pflanzlichen Isoprenoidstoffwechsel. Es resultiert aus der Addition einer Einheit Isopentenyl-pyrophosphat (IPP) an Farnesyl-pyrophosphat, ein 15-C-Sesquiterpen. GGPP fließt in diverse Synthesewege des pflanzlichen Sekundärstoffwechsels. So können beispielsweise zwei Moleküle GGPP "Schwanz an Schwanz" zu 40-C-Körpern zusammengefügt werden, die Tetraterpene, die allgemein als Carotinoide bezeichnet werden und zu denen beispielsweise das β-Carotin gehört. Durch die Addition weiterer Moleküle IPP fließt GGPP des weiteren in die Biosynthese von Polyterpenen, wie Kautschuk und Guttapercha, ein.

GGPP kann darüber hinaus in weitere Diterpene, wie z.B. Phytyl-pyrophosphat (PPP), überführt werden. Der 20-C-Körper Phytol ist ein obligatorisches Intermediat in der Biosynthese der Tocopherole (Soll and Schulz (1981) Biochem. Biophys. Res. Commun. 99, 907-912) sowie der Synthese der Chlorophylle (Beale and Weinstein (1990) In: Biosynthesis of Heme and Chlorophyll (Dailey H.A., ed.) McGraw Hill, NY, 287-391). Während die Grundstruktur aller Chlorophylle

(Chlorophyll a, b, c, usw.) ein aus 4 Pyrrolringen aufgebautes Porphyrinsystem ist, an welches das Phytol esterartig über den Pyrrolring IV gebunden ist, zeichnen sich die Tocopherole durch eine aus Homogentisat und einem Phytol-Schwanz bestehende Struktur aus.

Die Gruppe der Tocopherole, die allgemein als Vitamin E bezeichnet werden, umfaßt eine Reihe strukturell nahe verwandter fettlöslicher Vitamine, nämlich α-, β-, γ-, δ- und ε-Tocopherol, wovon α-Tocopherol biologisch am wichtigsten ist. Die Tocopherole kommen in vielen Pflanzenölen vor, besonders reich an Tocopherolen sind die Samenöle von Soja, Weizen, Mais, Reis, Baumwolle, Luzerne und Nüssen. Auch Früchte und Gemüse, z.B. Himbeeren, Bohnen, Erbsen, Fenchel, Paprika, etc. enthalten Tocopherole. Soweit bisher bekannt werden Tocopherole ausschließlich in Pflanzen bzw. photosynthetisch aktiven Organismen synthetisiert.

Tocopherole tragen aufgrund ihres Redoxpotentials dazu bei, die Oxidation ungesättigter Fettsäuren durch Luftsauerstoff zu vermeiden; im Menschen ist α-Tocopherol das wichtigste fettlösliche Antioxidans. Es wird angenommen, daß die Tocopherole durch diese Funktion als Antioxidantien zur Stabilisierung biologischer Membranen beitragen, da durch den Schutz der ungesättigten Fettsäuren der Membranlipide die Membran-Fluidität aufrechterhalten wird. Jüngeren Erkenntnissen zufolge kann darüber hinaus mit der regelmäßigen Zufuhr relativ hoher Tocopherol-Dosen der Ausbildung der Arteriosklerose entgegengewirkt werden. Als weitere günstige physiologische Eigenschaft der Tocopherole wurde die Verzögerung Diabetes bedingter Spätschäden, die Verminderung des Risikos der Kataraktbildung, Verminderung des oxidativen Stresses bei Rauchern, anticarcinogene Effekte, protektive Wirkung gegen Hautschäden wie Erythreme und Hautalterung, etc. beschrieben.

Wegen ihrer oxidationshemmenden Eigenschaften werden die Tocopherole nicht nur lebensmitteltechnologisch genutzt, sondern auch in auf natürlichen Ölen basierenden Anstrichfarben, in Desodorantien und anderen Kosmetika, beispielsweise Sonnenschutzmittel, Hautpflegemitteln, Lippenstiften etc. eingesetzt. Dabei sind Tocopherolverbindungen wie Tocopherylacetat und -succinat die üblichen Applikationsformen für die Anwendung als Vitamin E, in durchblutungsfördernden und Lipid-senkenden Mitteln und veterinärmedizinisch als Futtermittelzusatz.

In der Biosythese der Tocopherole, insbesondere des α-Tocopherols, scheint Phytylpyrophosphat ein limitierender Faktor zu sein. Bisherige Untersuchungen deuten darauf hin, daß PPP durch sequentielle Hydrogenierung der Isoprenoidgruppe aus GGPP hervorgeht, wobei als Intermediate Dihydro-GGPP und Tetrahydro-GGPP entstehen (GGPP -> Dihydro-GGPP -> Tetrahydro-GGPP -> PPP; vgl. bspw. Bollivar *et al.* (1994) Biochemistry 33, 12763-12768).

Die schrittweise Hydrogenierung von GGPP zu PPP wird, so wird gegenwärtig angenommen, durch das Enzym Geranylgeranyl-Reduktase (GGPP-Reduktase, auch als Geranylgeranyl-pyrophosphat-Hydrogenase bzw. GGPP-Hydrogenase bezeichnet) katalysiert, das in Pflanzen im Gen *Chl P* kodiert wird. Das Enzym Geranylgeranyl-Reduktase gehört zum Isoprenoidstoffwechsel und fungiert für zwei Stoffwechselwege: die Tocopherolsynthese und die Chlorophyllsynthese.

Die essentielle Bedeutung dieses Enzyms ist erstmals für die Chlorophyllbiosynthese gezeigt worden (Benz *et al.* (1980) Plant Sci. Lett. 19, 225-230; Soll and Schultz (1981) Biochem. Biophys. Res. Commun. 99, 907-912; Schoch *et al.* (1977) Z. Pflanzenphysiol. 83, 427-436). Der letzte Schritt der Chlorophyllbiosynthese ist die Veresterung von Chlorophyllid, die sowohl mit Phytylpyrophosphat als auch mit Geranylgeranyl-pyrophosphat erfolgen kann. Systema-

tische Untersuchungen von *Rhodobacter capsulatus*-Mutanten haben zeigen können, daß zuerst Bacteriochlorophyllid mit GGPP verestert und nachfolgend esterifiziertes Chlorophyll-GG hydrogeniert wird (Katz *et al.* (1972) J. Am. Chem. Soc. 94, 7938-7939). In höheren Pflanzen wird größtenteils Phytyl-Chlorophyll (Chlorophyll-P) nachgewiesen (Rüdiger and Schoch (1991) in Chlorophylls (Scheer, H., Ed.) pp. 451-464, CRC Press, Boca Raton, Florida, USA). Es ist bisher nicht geklärt, mit welchen Substraten die Reduktase-Reaktion in Pflanzen erfolgt. Gegenwärtig wird angenommen, daß die pflanzliche Geranylgeranyl-Reduktase in der Lage ist, sowohl Chlorophyll-GG in Chlorophyll-P umzuwandeln (Schoch *et al.* (1978) Z. Pflanzenphysiol. 83, 427-436) als auch GGPP zu PPP zu hydrogenieren, das dann anschließend mit Chlorophyllid verbunden wird (Soll *et al.* (1983) Plant Physiol. 71, 849-854).

GGPP dient als Substrat für die Synthesewege des Tocopherols und des Phyllochinons in Chloroplastenhüllmembranen und für die Chlorophyllsynthese in den Thylakoidmembranen. Die Reduktion von GGPP zu PPP wurde erstmals 1983 von Soll *et al.* beschrieben (1983, Plant. Physiol. 71, 849-854). Bisher ist allerdings die Isolierung und Charakterisierung von Nukleinsäuresequenzen, die für das pflanzliche Enzym kodieren und für die Beeinflussung des Tocopherol-gehaltes in transgenen Pflanzen eingesetzt werden können, nicht gelungen.

Die essentielle Rolle der Geranylgeranyl-Reduktase im Tocopherol- und Chlorophyllstoffwechsel macht dieses Enzym zu einem besonders wertvollen Instrument für die molekulare Biotechnologie. Mit Hilfe molekularbiologischer Techniken wie der Übertragung von DNA-Sequenzen, die für Geranylgeranyl-Reduktase kodieren, könnten in Pflanzen Änderungen in der Tocopherol- und/oder Chlorophyllbiosyntheseleistung erzielt werden. Auf diese Weise wäre z.B. die Erzeugung von transgenen Pflanzen mit erhöhtem oder verringertem Tocopherol-gehalt möglich. Solche transgenen Pflanzen bzw. deren Teile, Zellen und/oder

Produkte könnten anschließend als Nahrungs- und Futtermittel bzw. allgemein als Produktionsstätte für Tocopherol, das in der chemischen, pharmazeutischen und kosmetischen Industrie Anwendung findet, eingesetzt werden.

Darüber hinaus ist davon auszugehen, daß Pflanzen, die einen gegenüber Wildtyppflanzen erhöhten Gehalt an antioxidativ wirksamen Tocopherolen aufweisen, auch eine erhöhte Resistenz gegenüber Streßbedingungen, insbesondere gegenüber oxidativem Streß, aufweisen.

Es ist daher Aufgabe der Erfindung, neue Nukleinsäuresequenzen bereitzustellen, mit deren Hilfe in Pflanzen, Pflanzenzellen, -teilen und/oder -produkten der Gehalt an Tocopherol beeinflußt werden kann.

Des weiteren ist es eine wichtige Aufgabe der Erfindung, transgene Pflanzen, -zellen, -produkte und -teile mit gegenüber Wildtyppflanzen verändertem Tocopherolgehalt bereitzustellen.

Eine weitere Aufgabe der Erfindung besteht darin, Möglichkeiten der Verwendung der erfindungsgemäßen DNA-Sequenzen, deren Genprodukte sowie der erfindungsgemäßen transgenen Pflanzen für die pflanzenzüchterische Praxis aufzuzeigen.

Weitere Aufgaben der Erfindung ergeben sich aus der folgenden Beschreibung. Diese Aufgaben werden durch die Gegenstände der unabhängigen Ansprüche, insbesondere basierend auf der Bereitstellung der erfindungsgemäßen DNA-Sequenzen, deren Genprodukte unmittelbar an der Tocopherolbiosynthese beteiligt sind, und der in einem veränderten Tocopherolgehalt resultierenden Übertragung dieser DNA-Sequenzen auf Pflanzen, gelöst.

Die vorliegende Erfindung betrifft somit DNA-Sequenzen, die für Proteine mit der biologischen Aktivität einer Geranylgeranyl-Reduktase (auch Geranylgeranyl-pyrophosphat-Hydrogenase genannt) oder ein biologisch aktives Fragment davon kodieren. Biologisch aktives Fragment bedeutet im Zusammenhang mit dieser Erfindung, daß die vermittelte biologische Aktivität zu einer Beeinflussung des Tocopherolgehalts ausreicht. Insbesondere betrifft die Erfindung pflanzliche DNA-Sequenzen, die für Proteine mit der biologischen Aktivität einer Geranylgeranyl-Reduktase oder ein biologisch aktives Fragment davon kodieren, besonders bevorzugt betrifft die Erfindung die in SEQ:ID NO. 1 angegebene DNA-Sequenz (s. auch Abb. 1).

Des weiteren betrifft die Erfindung Allele und Derivate der erfindungsgemäßen DNA-Sequenzen, die für ein Protein mit der biologischen Aktivität einer Geranylgeranyl-Reduktase kodieren, insbesondere Nukleinsäuremoleküle, deren Sequenzen sich aufgrund der Degeneration des genetischen Codes von den erfindungsgemäßen DNA-Sequenzen unterscheiden und die für ein Protein oder ein Fragment davon kodieren, das die biologische Aktivität einer Geranylgeranyl-Reduktase aufweist.

Des weiteren betrifft die Erfindung Nukleinsäuremoleküle, die die erfindungsgemäßen DNA-Sequenzen enthalten oder durch natürlich vorkommende oder durch gentechnische oder chemische Prozesse und Syntheseverfahren aus diesen entstanden sind bzw. von diesen abgeleitet wurden. Hierbei kann es sich beispielsweise um DNA- oder RNA-Moleküle, cDNA, genomische DNA, mRNA etc. handeln.

Die Erfindung betrifft auch solche Nukleinsäuremoleküle, in denen die erfindungsgemäßen DNA-Sequenzen mit regulatorischen Elementen verknüpft sind, die die Transkription und, falls erwünscht, Translation in der Pflanzenzelle gewährleisten.

So können die erfindungsgemäßen DNA-Sequenzen beispielsweise unter Kontrolle konstitutiver, aber auch induzierbarer oder gewebe- bzw. enwicklungsspezifischer Regulationselemente, insbesondere Promotoren, in Pflanzenzellen exprimiert werden. Während beispielsweise die Verwendung eines induzierbaren Promotors die gezielt ausgelöste Expression der erfindungsgemäßen DNA-Sequenzen in Pflanzenzellen ermöglicht, bietet z.B. der Einsatz von gewebespezifischen, beispielsweise samenspezifischen, Promotoren die Möglichkeit, den Tocopherolgehalt in bestimmten Gewebe, beispielsweise in Samengewebe zu verändern. So liegen die erfindungsgemäßen DNA-Sequenzen in einer bevorzugten Ausführungsform in Kombination mit gewebespezifischen Promotoren, insbesondere samenspezifischen Promotoren, vor.

Die Erfindung betrifft weiterhin Proteine mit der biologischen Aktivität einer Geranylgeranyl-Reduktase oder aktive Fragmente davon, die durch eine erfindungsgemäße DNA-Sequenz oder ein erfindungsgemäßes Nukleinsäuremolekül kodiert werden. Vorzugsweise handelt es sich um eine pflanzliche Geranylgeranyl-Reduktase, bevorzugt aus *Nicotiana tabacum*, besonders bevorzugt um ein Protein mit der in SEQ:ID No. 2 (s. auch Abb. 2) gezeigten Aminosäuresequenz oder ein aktives Fragement davon.

Eine weitere Aufgabe der Erfindung besteht darin, Vektoren und Mikroorganismen zu liefern, deren Verwendung die Herstellung neuer Pflanzen, in denen ein veränderter Tocopherolgehalt erzielt werden kann, ermöglicht. Diese Aufgabe wird durch die Bereitstellung der erfindungsgemäßen Vektoren und Mikroorganismen gelöst, die für Enzyme mit der Aktivität einer Geranylgeranyl-Reduktase kodierende Nukleinsäuresequenzen enthalten.

Die vorliegende Erfindung betrifft somit auch Vektoren, insbesondere Plasmide, Cosmide, Viren, Bakteriophagen und andere in der Gentechnik gängige Vektoren,

die die oben beschriebenen erfindungsgemäßen Nukleinsäuremoleküle enthalten und gegebenfalls für den Transfer der erfindungsgemäßen Nukleinsäuremoleküle auf Pflanzen bzw. Pflanzenzellen eingesetzt werden können.

Ebenso betrifft die Erfindung transformierte Mikroorganismen, wie Bakterien, Viren, Pilze, Hefen etc., die die erfindungsgemäßen Nukleinsäuresequenzen enthalten.

In einer bevorzugten Ausführungsform sind die in den Vektoren enthaltenen Nukleinsäuremoleküle verknüpft mit regulatorischen Elementen, die die Transkription und ggf. Translation in prokaryontischen und eukaryontischen Zellen gewährleisten.

Gegebenenfalls können die Nukleinsäuresequenzen der Erfindung durch Enhancer-Sequenzen oder andere regulatorische Sequenzen ergänzt sein. Diese regulatorischen Sequenzen beinhalten z.B. auch Signalsequenzen, die für den Transport des Genprodukts zu einem bestimmten Kompartiment sorgen.

Es ist ebenso eine Aufgabe der Erfindung neue Pflanzen, Pflanzenzellen, -teile oder -produkte bereitzustellen, die sich durch einen veränderten Tocopherolgehalt auszeichnen, der ggf. mit einer gegenüber Wildtyppflanzen veränderten Chlorophyllbiosyntheseleistung gekoppelt sein kann.

Diese Aufgaben werden durch die Übertragung der erfindungsgemäßen Nukleinsäuremoleküle und ihre Expression in Pflanzen gelöst. Durch die Bereitstellung der erfindungsgemäßen Nukleinsäuremoleküle besteht nun die Möglichkeit, pflanzliche Zellen mittels gentechnischer Methoden dahingehend zu verändern, daß sie im Vergleich zu Wildtypzellen eine neue oder veränderte Geranylgeranyl-

Reduktase-Aktivität aufweisen und es als Folge dazu zu einer veränderten Tocopherolbiosyntheseleistung und zu einem veränderten Tocopherolgehalt kommt.

In einer bevorzugten Ausführungsform handelt es sich um Pflanzen bzw. deren Zellen und Teile, in denen der Tocopherolgehalt aufgrund der Gegenwart und Expression der erfindungsgemäßen Nukleinsäuremoleküle gegenüber Wildtyp-pflanzen erhöht ist.

Die Erfindung betrifft aber auch solche Pflanzen, in denen die Übertragung der erfindungsgemäßen Nukleinsäuremoleküle zu einer Verringerung des Tocopherol- und/oder Chlorophyllgehalts führt. Eine reduzierte Tocopherol- und/oder Chlorophyllbiosyntheseleistung kann beispielsweise durch den Transfer von antisense-Konstrukten oder andere Suppressionsmechanismen, wie beispielsweise Cosuppression, erreicht werden.

Gegenstand der Erfindung sind weiterhin transgene Pflanzenzellen bzw. solche Pflanzenzellen umfassende Pflanzen und deren Teile und Produkte, in denen die neuen Nukleinsäuremolelüle integriert in das pflanzliche Genom vorliegen. Ebenfalls Gegenstand der Erfindung sind Pflanzen, in deren Zellen die erfindungsgemäße Nukleinsäuresequenz in selbstreplizierender Form vorliegt, d.h. die Pflanzenzelle enthält die fremde DNA auf einem eigenständigen Nukleinsäure-molekül.

Bei den Pflanzen, die mit den erfindungsgemäßen Nukleinsäure-molekülen transformiert sind und in denen aufgrund der Einführung eines solchen Moleküls eine veränderte Menge an Tocopherol und/oder Chlorophyll synthetisiert wird, kann es sich im Prinzip um jede beliebige Pflanze handeln. Vorzugsweise ist es eine monokotyle oder dikotyle Nutzpflanze. Beispiele für monokotyle Pflanzen

sind die Pflanzen, die zu den Gattungen Avena (Hafer), Triticum (Weizen), Secale (Roggen), Hordeum (Gerste), Oryza (Reis), Panicum, Pennisetum, Setaria, Sorghum (Hirse), Zea (Mais) gehören. Bei den dicotylen Nutzpflanzen sind u.a. zu nennen Leguminosen, wie Hülsenfrüchte und insgesondere Alfalfa, Sojabohne, Raps, Tomate, Zuckerrübe, Kartoffel, Zierpflanzen, Bäume. Weitere Nutzpflanzen können beispielsweise Obst (insbesondere Äpfel, Birnen, Kirschen, Weintrauben, Citrus, Ananas und Bananen), Ölpalmen, Tee-, Kakao- und Kaffeesträucher, Tabak, Sisal, Baumwolle, Lein, Sonnenblume sowie Heilpflanzen und Weidegräser und Futterpflanzen sein. Besonders bevorzugt sind die Getreide, Weizen, Roggen, Hafer, Gerste, Reis, Mais und Hirse, Futtergetreide, Zuckerrübe, Raps, Soja, Tomate, Kartoffel, Süßgräser, Futtergräser, und Klee. Es ergibt sich von selbst, daß die Erfindung insbesondere übliche Nahrungs- bzw. Futterpflanzen betrifft. Hier sind neben den bereits erwähnten Pflanzen zusätzlich Erdnuß, Linse, Ackerbohne, Runkelrübe, Buchweizen, Möhre, Sonnenblume, Topinambur, Rübsen, Weißer Senf, Kohlrübe und Stoppelrübe zu nennen.

Gegenstand der Erfindung sind ferner Vermehrungsmaterial von erfindungsgemäßen Pflanzen, beispielsweise Samen, Früchte, Stecklinge, Knollen, Wurzelstöcke etc., wobei dieses Vermehrungsmaterial gegebenenfalls oben beschriebene transgene Pflanzenzellen enthält, sowie Teile dieser Pflanzen wie Protoplasten, Pflanzenzellen und Kalli.

Die Erfindung betrifft ebenfalls Pflanzenzellen, die aufgrund der Gegenwart und ggf. Expression der erfindungsgemäßen Nukleinsäuremoleküle einen im Vergleich zu Pflanzenzellen, die die Nukleinsäuremoleküle nicht enthalten, veränderten Gehalt an Vitamin K_1 aufweisen. Das fettlösliche Vitamin K_1, das insbesondere in Pflanzen enthalten ist, hat eine wichtige Funktion bei der Bildung von Gerinnungsfaktoren, Mangel an Vitamin K_1 führt zu einer Verringerung der Blutgerinnung, weshalb es auch als antihämorhagisches oder Koagulationsvitamin

bezeichnet wird. Da die Expression der erfindungsgemäßen Nukleinsäuremoleküle in einer veränderten Geranylgeranyl-Reduktase-Aktivität und demzufolge veränderten PPP-Sytheseleistung resultiert und das als Vitamin K_1 bezeichnete Phyllochinon, wie die Tocopherole, eine Einheit Phytol umfaßt, sind auch solche Pflanzenzellen bzw. Pflanzen Gegenstand der Erfindung, die einen veränderte Vitamin K_1-Gehalt alleine oder in Kombination mit einem veränderten Tocopherolgehalt aufweisen.

In einer bevorzugten Ausführungsform handelt es sich um transgene Pflanzenzellen bzw. Pflanzen und deren Teile und Produkte, die aufgrund der Gegenwart und ggf. Expression einer für eine Geranylgeranyl-Reduktase aus Pflanzen kodierenden DNA-Sequenz einen gegenüber nicht-transformierten Zellen veränderten Tocopherolgehalt aufweisen. Bevorzugt handelt es sich bei der in den Pflanzenzellen enthaltenen DNA-Sequenz um eine für Geranylgeranyl-Reduktase kodierende Sequenz, die aus Tabak stammt. Besonders bevorzugt handelt es sich um die in SEQ:ID NO. 1 dargestellte DNA-Sequenz (siehe auch Abb. 1). In einer besonders bevorzugten Ausführungsform kodieren die erfindungsgemäßen DNA-Sequenzen für ein Geranylgeranyl-Reduktase-Vorstufenenzym, das eine Transitsequenz für die Translokation in Plastiden umfaßt.

Die Erfindung betrifft des weiteren Pflanzen, in denen neben dem chl P-Gen zusätzlich ein Gen für Hydroxyphenylpyruvat-Dioxygenase (HPD) exprimiert wird. Das Enzym HPD katalysiert die Umsetzung von 4-Hydroxyphenylpyruvat in Homogentisat, das wie oben erwähnt neben dem Phytol den zweiten Baustein der Tocopherole darstellt. Das Enzym HPD sowie seine Stellung im pflanzlichen Isoprenoid-Stoffwechsel sind *inter alia* in Norris *et al.* (1995) The Plant Cell 7, 2139-2149, beschrieben.

Durch die gemeinsame Expression, vorzugsweise Überexpression, von Sequenzen, die für Geranylgeranyl-Reduktase und HPD kodieren, kann der Tocopherolgehalt in transgenen Pflanzen gegenüber solchen Pflanzen, die nur die erfindungsgemäßen, für CHL P kodierenden Sequenzen enthalten, zusätzlich gesteigert werden.

In einer weiteren Ausführungsform betrifft die Erfindung Wirtszellen, insbesondere prokaryontische und eukaryontische Zellen, die mit einem oben beschriebenen Nukleinsäuremolekül oder einem Vektor transformiert bzw. infiziert wurden, und Zellen, die von derartigen Wirtszellen abstammen und die beschriebenen Nukleinsäuremoleküle oder Vektoren enthalten. Die Wirtszellen können Bakterien, Algen, Hefe- und Pilzzellen sowie pflanzliche oder tierische Zellen sein. Gegenstand der Erfindung sind auch solche Wirtszellen, die neben den erfindungsgemäßen Nukleinsäuremolekülen ein oder mehrere auf gentechnologischem oder natürlichem Weg übertragene Nukleinsäuremoleküle enthalten, die die genetische Information für an der Tocopherol-, Chlorophyll- und/oder Vitamin K_1-Biosynthese beteiligte Enzyme tragen.

Der vorliegenden Erfindung liegt außerdem die Aufgabe zugrunde, Verfahren zur Herstellung von Pflanzenzellen und Pflanzen, die sich durch einen veränderten Tocopherolgehalt auszeichnen, bereitzustellen.

Diese Aufgabe wird durch Verfahren gelöst, mit deren Hilfe die Erzeugung neuer Pflanzenzellen und Pflanzen, die aufgrund der Übertragung von für Geranylgeranyl-Reduktase kodierenden Nukleinsäuremolekülen einen veränderten Tocopherolgehalt aufweisen, möglich ist.

Des weiteren wird diese Aufgabe durch Verfahren gelöst, mit deren Hilfe die Erzeugung neuer Pflanzenzellen und Pflanzen möglich ist, die aufgrund der

gemeinsamen Übertragung von für Geranylgeranyl-Reduktase kodierenden Nukleinsäurenmolekülen und für HPD kodierenden Nukleinsäuremolekülen oder der Übertragung von für Geranylgeranyl-Reduktase und für HPD kodierenden Nukleinsäuremolekülen einen gegenüber Wildtyppflanzen veränderten Tocopherolgehalt aufweisen.

Zur Erzeugung solcher neuer Pflanzenzellen und Pflanzen bieten sich verschiedene Methoden an. Zum einen können Pflanzen bzw. Pflanzenzellen mit Hilfe herkömmlicher gentechnologischer Transformationsmethoden derart verändert werden, daß die neuen Nukleinsäuremoleküle in das pflanzliche Genom integriert werden, d.h. daß stabile Transformanten erzeugt werden. Zum anderen kann ein erfindungsgemäßes Nukleinsäuremolekül, dessen Anwesenheit und ggf. Expression in der Pflanzenzelle eine veränderte Tocopherolbiosyntheseleistung bewirkt, in der Pflanzenzelle bzw. der Pflanze als selbstreplizierendes System enthalten sein. So können die erfindungsgemäßen Nukleinsäuremoleküle z.B. in einem Virus enthalten sein, mit dem die Pflanze bzw. Pflanzenzelle in Kontakt kommt.

Erfindungsgemäß werden Pflanzenzellen, die aufgrund der Expression einer erfindungsgemäßen Nukleinsäuresequenz einen veränderten Tocopherolgehalt aufweisen, durch ein Verfahren hergestellt, das folgende Schritte umfaßt:

a) Herstellung einer Expressionskassette, die folgende DNA-Sequenzen umfaßt:
   - einen Promotor, der die Transkription in pflanzlichen Zellen gewährleistet;
   - mindestens eine erfindungsgemäße Nukleinsäuresequenz, die für ein Protein oder ein Fragment mit der enzymatischen Aktivität einer Geranylgeranyl-Reduktase kodiert, wobei die Nukleinsäuresequenz in Sense-Orientierung an das 3'-Ende des

Promotors gekoppelt ist; und

- gegebenenfalls ein Terminationssignal für die Termination der Transkription und die Addition eines poly-A-Schwanzes an das entsprechende Transkript, das an das 3'-Ende der kodierenden Region gekoppelt ist.

b) Transformation pflanzlicher Zellen mit der in Schritt a) hergestellten Expressionskassette.

c) Regeneration transgener Pflanzen und gegebenenfalls die Vermehrung der Pflanzen.

Alternativ kann eine oder mehrere erfindungsgemäße Nukleinsäuresequenzen als selbstreplizierendes System in die Pflanzenzelle bzw. Pflanze eingebracht werden.

Als weitere Alternative kann Schritt a) des obigen Verfahrens dahingehend abgewandelt werden, daß die mindestens eine erfindungsgemäße Nukleinsäuresequenz, die für ein Protein oder ein Fragment mit der enzymatischen Aktivität einer Geranylgeranyl-Reduktase kodiert, in Antisense-Orientierung an das 3'-Ende des Promotors gekoppelt ist.

Eine weitere Aufgabe der Erfindung besteht darin, Verwendungen der erfindungsgemäßen Nukleinsäuresequenzen sowie der sie enthaltenen Nukleinsäuremoleküle aufzuzeigen.

Diese Aufgabe wird durch die erfindungsgemäßen Verwendungen der neuen DNA-Moleküle zur Erzeugung von Pflanzenzellen und Pflanzen, die sich durch

einen im Vergleich zu Wildtypzellen bzw. -pflanzen veränderten, vorzugsweise erhöhten, Tocopherolgehalt auszeichnen, gelöst.

Des weiteren betrifft die Erfindung die Verwendung der erfindungsgemäßen Nukleinsäuresequenzen zur Erzeugung von Pflanzen, die einen veränderten Chlorophyllgehalt aufweisen.

Weiterhin betrifft die Erfindung die Verwendung der erfindungsgemäßen Nukleinsäuresequenzen zur Erzeugung von Pflanzen, die sich durch einen veränderten, vorzugsweise erhöhten, Gehalt an Vitamin K_1 auszeichnen.

Eine weitere Aufgabe der Erfindung besteht darin, die Möglichkeiten der Verwendung der erfindungsgemäßen Pflanzen bzw. deren Zellen, Teile und Produkte aufzuzeichnen.

Gegenstand der Erfindung ist insbesondere die Verwendung der erfindungsgemäßen Pflanzen als Futter- und/oder Nahrungspflanze. In Abhängigkeit von der erzielten Erhöhung des Gehalts an Vitamin E und/oder Vitamin K_1 in der transgenen Nutzpflanze bzw. deren Produkte und Teile kann eine sonst allgemein übliche und oft auch erforderliche Zumischung der entsprechenden Vitamine, insbesondere von Vitamin E, mengenmäßig beschränkt bzw. vollkommen überflüssig werden. Unabhängig davon betrifft die Erfindung allgemein die Erhöhung des Nährwerts von Nutzpflanzen durch eine Steigerung des Gehalts an Tocopherolen und/oder Phyllochinon.

Des weiteren betrifft die Erfindung die Verwendung der erfindungsgemäßen Pflanzenzellen, Pflanzen, deren Teile und Produkte, als Produktionsstätten für Vitamin E und/oder Vitamin K_1. Neben ihrer Anwendung aufgrund ihres Vitamin-Charakters, beispielsweise in diätetischen und pharmazeutischen Produkten,

Kosmetika, Hautpflegeprodukten, allgemein zur Vitamin E-Supplementierung, etc., finden Tocopherole auch als Antioxidantien in chemischen Produkten wie beispielsweise Fetten und Ölen Anwendung. Die erfindungsgemäßen Pflanzen stellen somit eine wichtige Quelle für die Gewinnung von Tocopherolen und/oder Vitamin K_1 für ein breites Spektrum gewerblicher Zwecke dar.

Gegenstand der Erfindung ist ebenfalls die Verwendung der erfindungsgemäßen Nukleinsäuresequenzen in Kombination mit samenspezifischen Promotoren zur Erzeugung von Pflanzen, bei denen sich vor allem das Samengewebe durch einen veränderten, vorzugsweise erhöhten, Tocopherolgehalt auszeichnet. In einer bevorzugten Ausführungsform der Erfindung handelt es sich um die Verwendung der erfindungsgemäßen Nukleinsäuresequenzen in Kombination mit dem USP-(Bäumlein *et al.* (1991) Mol. Gen. Genet. 225, 459-467) oder Hordein-Promotor (Brandt *et al.* (1985) Carlsberg Res. Commun. 50, 333-345).

Diese genannten Promotoren, insbesondere die samenspezifischen Promotoren, eignen sich auch besonders für die gezielte Reduktion des Tocopherol- bzw. Chlorophyllgehalts in transgenen Samen unter Einsatz der erfindungsgemäßen DNA-Sequenzen im Zusammenhang mit der Antisense-Technik.

Weiter betrifft die Erfindung die Verwendung eines Geranylgeranyl-Reduktase-Gens zur Erzeugung eines veränderten Tocopherolgehalts in Pflanzen.

Weiter betrifft die Erfindung die Verwendung eines Proteins mit der enzymatischen Aktivität einer Geranylgeranyl-Reduktase, um in Pflanzen einen veränderten Tocopherolgehalt zu erzielen.

Des weiteren betrifft die Erfindung die Verwendung der erfindungsgemäßen Nukleinsäuremoleküle, der erfindungsgemäßen Proteine mit Geranylgeranyl-

Reduktase-Aktivität und/oder der erfindungsgemäßen transgenen Pflanzen bzw. Wirtszellen mit neuer bzw. veränderter Geranylgeranyl-Reduktase-Aktivität zur Identifizierung neuer herbizider Wirkstoffe für den Pflanzenschutz. Dank der Schlüsselstellung der Geranylgeranyl-Reduktase innerhalb der Chlorophyll- und Tocopherolbiosynthese stellen die erfindungsgemäßen DNA-Sequenzen und die durch sie kodierten Proteine ein äußerst wertvolles Target für die Herbizidforschung dar. So können beispielsweise die erfindungsgemäßen Proteine mit enzymatischer Geranylgeranyl-Reduktase-Aktivität für die Röntgenstrukturanalyse, NMR-Spektroskopie, molecular modeling, und drug design eingesetzt werden, um auf der Basis der aus diesen Verfahren gewonnenen Erkenntnisse Inhibitoren und/oder Effektoren der Geranylgeranyl-Reduktase und somit potentielle Herbizide zu identifizieren oder synthetisieren.

Die Erfindung betrifft weiterhin die Verwendung der erfindungsgemäßen Nukleinsäuresequenzen zur Herstellung herbizidtoleranter Pflanzen. So können für Geranylgeranyl-Reduktase kodierende Sequenzen mittels Standard-Methoden verändert oder um neue Sequenzelemente erweitert werden und anschließend auf Pflanzenzellen übertragen werden. Die Einbringung von aus den erfindungsgemäßen Sequenzen abgeleiteten Sequenzen kann z.B. dazu genutzt werden, die Eigenschaften der Pflanzen dahingehend zu verändern, daß in der transgenen Pflanze mehr oder weniger funktionell aktive Geranylgeranyl-Reduktase oder eine Variante der Geranylgeranyl-Reduktase mit veränderten Eigenschaften gebildet wird oder daß das Expressionsniveau des in der transgenen Pflanze vorhandenen chl P-Gens vermindert wird. So kann mittels einer Erhöhung der CHL P-Aktivität eine Erhöhung der Toleranz gegenüber Herbiziden, welche die Chlorophyllbiosynthese inhibieren, erreicht werden. Ebenso kann z.B. die Expression veränderter Geranylgeranyl-Reduktase-Gene in transgenen Pflanzenzellen mit einer Erhöhung der Herbizidtoleranz verbunden sein.

Des weiteren betrifft die Erfindung die Verwendung der erfindungsgemäßen Nukleinsäuresequenzen oder eines durch sie kodierten Proteins zur Herstellung von Antikörpern.

Die vorliegende Erfindung umfaßt somit jede mögliche Form des Einsatzes der erfindungsgemäßen Nukleinsäuremoleküle, deren Gegenwart und ggf. Expression in Pflanzen eine Veränderung des Tocopherolgehalts und/oder Chlorophyllgehalts bewirkt, sowie des Einsatzes der erfindungsgemäßen Proteine oder Fragmente davon, deren enzymatische Aktivität eine solche Veränderung herbeiführt.

Für den genannten Promotor kommt im Prinzip jeder in den für die Transformation gewählten Pflanzen funktionale Promotor in Betracht, der die Bedingung erfüllt, daß die von ihm regulierte Expression zu einer veränderten Tocopherolsyntheseleistung führt. Im Hinblick auf die Verwendung der transgenen Pflanzen als Nahrungs- bzw. Futterpflanzen erscheinen hierfür besonders solche Promotoren sinnvoll, die eine samenspezifische Expression gewährleisten. Beispiele für solche Promotoren sind der USP-Promotor, Hordein-Promotor und Napin-Promotor.

Falls solche Promotoren nicht bekannt sind oder nicht zur Verfügung stehen, ist auf jeden Fall das Konzept zur Isolierung solcher Promotoren dem Fachmann bekannt. Dabei wird in einem ersten Schritt aus Samengewebe die poly(A)$^+$ RNA isoliert und eine cDNA-Bank angelegt. In einem zweiten Schritt werden mit Hilfe von cDNA-Klonen, die auf poly(A)$^+$ RNA-Molekülen aus einem nicht aus Samen stammenden Gewebe basieren, aus der ersten Bank mittels Hybridisierung diejenigen Klone identifiziert, deren korrespondierende poly(A)$^+$ RNA-Moleküle lediglich im Samengewebe exprimiert werden. Anschließend werden mit Hilfe dieser so identifizierten cDNA's Promotoren isoliert, die sodann für die Expression der hier beschriebenen kodierenden Nukleinsäuresequenzen verwendet

werden können. Analog können andere gewebespezifische oder entwicklungsspezifische oder durch abiotische Stimuli induzierbare Promotoren isoliert und erfindungsgemäß eingesetzt werden.

Alternativ kann erstrebenswert sein, daß die Pflanze in vielen Bereichen aufgrund der Expression der erfindungsgemäßen Nukleinsäuremoleküle einen veränderten, vorzugsweise erhöhten Tocopherolgehalt aufweist. In diesem Fall bietet sich die Verwendung eines konstitutiven Promotors, beispielsweise des 35S RNA Promotors aus Cauliflower Mosaic Virus, an.

Gegenstand der Erfindung sind ebenfalls Nukleinsäuremoleküle, die für Proteine mit der biologischen Aktivität einer Geranylgeranyl-Reduktase kodieren oder biologisch aktive Fragmente davon und die mit einem der oben beschriebenen Nukleinsäuremoleküle hybridisieren. Der Begriff biologisch aktive Fragmente bezieht sich auf Fragmente, die eine Veränderung des Tocopherolgehalts bewirken können. Der Begriff "Hybridisierung" bedeutet im Rahmen dieser Erfindung eine Hybridisierung unter konventionellen Hybridisierungsbedingungen, vorzugsweise unter stringenten Bedingungen, wie sie beispielsweise in Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual, 2. Auflage, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, beschrieben sind.

Nukleinsäuremoleküle, die mit den erfindungsgemäßen Molekülen hybridisieren, können z.B. aus genomischen oder aus cDNA-Bibliotheken isoliert werden.

Die Identifizierung und Isolierung derartiger Nukleinsäuremoleküle kann dabei unter Verwendung der erfindungsgemäßen Nukleinsäuremoleküle oder Teile dieser Moleküle bzw. der reversen Komplemente dieser Moleküle erfolgen, z.B. mittels Hybridisierung nach Standardverfahren (siehe z.B. Sambrook *et al.*, supra). Zur Identifizierung und Isolierung derartiger Nukleinsäuremoleküle können auch von

den erfindungsgemäßen DNA-Sequenzen abgeleitete Sequenzfolgen, beispielsweise degenerierte Oligonukleotidprimer, verwendet werden.

Somit betrifft die Erfindung ebenfalls die Verwendung einer erfindungsgemäßen DNA-Sequenz oder Teile davon zur Identifizierung und Isolierung homologer Sequenzen aus Pflanzen oder anderen Organismen.

Als Hybridisierungssonde können z.B. Nukleinsäuremoleküle verwendet werden, die exakt oder im wesentlichen die oben aufgeführten Nukleotidsequenzen oder Teile dieser Sequenzen aufweisen. Bei den als Hybridisierungssonde verwendeten Fragmenten kann es sich auch um synthetische Fragmente handeln, die mit Hilfe der gängigen Synthesetechniken hergestellt wurden und deren Sequenz im wesentlichen mit der eines erfindungsgemäßen Nukleinsäuremoleküls übereinstimmt. Hat man Gene identifiziert und isoliert, die mit den erfindungsgemäßen Nukleinsäuresequenzen hybridisieren, ist eine Bestimmung der Sequenz und eine Analyse der Eigenschaften der von dieser Sequenz kodierten Proteine erforderlich. Hierzu stehen dem Fachmann eine Reihe von molekularbiologischen, biochemischen und biotechnologischen Standardverfahren zur Verfügung.

Die mit den erfindungsgemäßen Nukleinsäuremolekülen hybridisierenden Moleküle umfassen auch Fragmente, Derivate und allelische Varianten der oben beschriebenen DNA-Moleküle, die für eine Geranylgeranyl-Reduktase kodieren oder ein biologisch, d.h. enzymatisch aktives Fragment davon. Unter Fragmenten werden dabei Teile der Nukleinsäuremoleküle verstanden, die lang genug sind, um ein Polypeptid oder Protein mit der enzymatischen Aktivität einer Geranylgeranyl-Reduktase oder einer vergleichbaren enzymatischen Aktivität, die einen veränderten Tocopherolgehalt bedingt, zu kodieren. Der Ausdruck Derivat bedeutet in diesem Zusammenhang, daß die Sequenzen dieser Moleküle sich von den Sequenzen der oben beschriebenen Nukleinsäuremoleküle an einer oder

mehreren Positionen unterscheiden und einen hohen Grad an Homologie zu diesen Sequenzen aufweisen. Homologie bedeutet dabei eine Sequenzidentität von mindestens 40%, insbesondere eine Identität von mindestens 60%, vorzugsweise über 80% und besonders bevorzugt über 90%. Die Abweichungen zu den oben beschriebenen Nukleinsäuremolekülen können dabei durch Deletion, Addition, Substitution, Insertion oder Rekombination entstanden sein.

Homologie bedeutet ferner, daß funktionelle und/oder strukturelle Äquivalenz zwischen den betreffenden Nukleinsäuremolekülen oder den durch sie kodierten Proteinen besteht. Bei den Nukleinsäuremolekülen, die homolog zu den oben beschriebenen Molekülen sind und Derivate dieser Moleküle darstellen, handelt es sich in der Regel um Variationen dieser Moleküle, die Modifikationen darstellen, die dieselbe biologische Funktion ausüben. Es kann sich dabei sowohl um natürlicherweise auftretende Variationen handeln, beispielsweise um Sequenzen aus anderen Organismen, oder um Mutationen, wobei diese Modifikationen auf natürliche Weise aufgetreten sein können oder durch gezielte Mutagenese eingeführt wurden. Ferner kann es sich bei den Variationen um synthetisch hergestellte Sequenzen handeln. Bei den allelischen Varianten kann es sich sowohl um natürlich auftretende als auch um synthetisch hergestellte oder durch rekombinante DNA-Techniken erzeugte Varianten handeln.

Üblicherweise weisen die von den verschiedenen Varianten und Derivaten der erfindungsgemäßen Nukleinsäuremoleküle kodierten Proteine bestimmte gemeinsame Charakteristika auf. Dazu können z.B. Enzymaktivität, Molekulargewicht, immunologische Reaktivität, Konformation etc. gehören. Weitere gemeinsame Charakteristika können physikalische Eigenschaften wie z. B. das Laufverhalten in Gelelektrophoresen, chromatographisches Verhalten, Sedimentationskoeffizienten, Löslichkeit, spektroskopische Eigenschaften, Stabilität, pH-Optimum, Temperatur-

Optimum etc. darstellen. Des weiteren können natürlich die Produkte der von den Proteinen katalysierten Reaktionen gemeinsame oder ähnliche Merkmale aufweisen.

Zur Vorbereitung der Einführung fremder Gene in höhere Pflanzen stehen eine große Anzahl von Klonierungsvektoren zur Verfügung, die ein Replikationssignal für *E. coli* und ein Markergen zur Selektion transformierter Bakterienzellen enthalten. Beispiele für derartige Vektoren sind pBR322, pUC-Serien, M13mp-Serien, pACYC184 usw. Die gewünschte Sequenz kann an einer passenden Restriktionsschnittstelle in den Vektor eingeführt werden. Das erhaltene Plasmid wird für die Transformation von *E. coli*-Zellen verwendet. Transformierte *E. coli*-Zellen werden in einem geeigneten Medium gezüchtet und anschließend geerntet und lysiert. Das Plasmid wird wiedergewonnen. Als Analysemethode zur Charakterisierung der gewonnenen Plasmid-DNA werden im allgemeinen Restriktionsanalysen, Gelelektrophoresen und weitere biochemisch-molekularbiologische Methoden eingesetzt. Nach jeder Manipulation können die Plasmid-DNA gespalten und gewonnene DNA-Fragmente mit anderen DNA-Sequenzen verknüpft werden. Jede Plasmid-DNA-Sequenz kann in den gleichen oder anderen Plasmiden kloniert werden.

Für die Einführung von DNA in eine pflanzliche Wirtszelle stehen eine Vielzahl bekannter Techniken zur Verfügung, wobei der Fachmann die jeweils geeignete Methode ohne Schwierigkeiten ermitteln kann. Diese Techniken umfassen die Transformation pflanzlicher Zellen mit T-DNA unter Verwendung von *Agrobacterium tumefaciens* oder *Agrobacterium rhizogenes* als Transformationsmittel, die Fusion von Protoplasten, den direkten Gentransfer isolierter DNA in Protoplasten, die Mikroinjektion und Elektroporation von DNA, die Einbringung von DNA mittels der biolistischen Methode sowie weitere Möglichkeiten.

Bei der Injektion und Elektroporation von DNA in Pflanzenzellen werden *per se* keine speziellen Anforderungen an die verwendeten Plasmide gestellt. Ähnliches gilt für den direkten Gentransfer. Es können einfache Plasmide wie z.B. pUC-Derivate verwendet werden. Sollen aber aus derartig transformierten Zellen ganze Pflanzen regeneriert werden, ist in der Regel die Anwesenheit eines selektierbaren Markergens notwendig. Dem Fachmann sind die gängigen Selektionsmarker bekannt und es stellt für ihn kein Problem dar, einen geeigneten Marker auszuwählen.

Je nach Einführungsmethode kann neben dem gewünschten Gen bzw. Gene die Anwesenheit weiterer DNA-Sequenzen erforderlich sein. Werden z.B. für die Transformation der Pflanzenzelle das Ti- oder Ri-Plasmid verwendet, so muß mindestens die rechte Begrenzung, häufig jedoch die rechte und linke Begrenzung der im Ti- und Ri-Plasmid enthaltenen T-DNA als Flankenbereich mit den einzuführenden Genen verbunden werden.

Werden für die Transformation Agrobakterien verwendet, muß die einzuführende DNA in spezielle Plasmide kloniert werden, und zwar entweder in einen intermediären oder in einen binären Vektor. Die intermediären Vektoren können aufgrund von Sequenzen, die homolog zu Sequenzen in der T-DNA sind, durch homologe Rekombination in das Ti- oder Ri-Plasmid der Agrobakterien integriert werden. Dieses enthält außerdem die für den Transfer der T-DNA notwendige *vir*-Region. Intermediäre Vektoren können nicht in Agrobakterien replizieren. Mittels eines Helferplasmids kann der intermediäre Vektor auf *Agrobacterium tumefaciens* übertragen werden (Konjugation). Binäre Vektoren können sowohl in *E. coli* als auch in Agrobakterien replizieren. Sie enthalten ein Selektionsmarker-Gen und einen Linker oder Polylinker, welche von der rechten und linken T-DNA-Grenzregion eingerahmt werden. Sie können direkt in die Agrobakterien transformiert werden (Holsters *et al.* (1978) Molecular and General Genetics 163, 181-

187). Das als Wirtszelle dienende Agrobakterium soll ein Plasmid, das eine *vir*-Region trägt, enthalten. Die vir-Region ist für den Transfer der T-DNA in die Pflanzenzelle notwendig. Zusätzliche T-DNA kann vorhanden sein. Das derartig transformierte Agrobakterium wird zur Transformation von Pflanzenzellen verwendet.

Die Verwendung von T-DNA für die Transformation von Pflanzenzellen ist intensiv untersucht und ausreichend in EP 120 515; Hoekema in: The Binary Plant Vector System, Offsetdrokkerij Kanters B.V., Alblasserdam (1985) Chapter V; Fraley *et al.* (1993) Crit. Rev. Plant. Sci., 4, 1-46 und An *et al.* (1985) EMBO J. 4, 277-287 beschrieben worden.

Für den Transfer der DNA in die Pflanzenzelle können Pflanzen-Explantate zweckmäßigerweise mit *Agrobacterium tumefaciens* oder *Agrobacterium rhizogenes* kultiviert werden. Aus dem infizierten Pflanzenmaterial (z.B. Blätter, Blattstücke, Stengelsegmente, Wurzeln, aber auch Protoplasten oder Suspensions-kultivierte Pflanzenzellen) können dann in einem geeigneten Medium, welches Antibiotika oder Biozide zur Selektion transformierter Zellen enthalten können, wieder ganze Pflanzen regeneriert werden. Die Regeneration der Pflanzen erfolgt nach üblichen Regenerationsmethoden unter Verwendung bekannter Nährmedien. Die so erhaltenen Pflanzen können dann auf Anwesenheit der eingeführten DNA untersucht werden. Andere Möglichkeiten der Einführung fremder DNA unter Verwendung des biolistischen Verfahrens oder durch Protoplasten-Transformation sind bekannt (vgl. z.B. Wilmitzer L. (1993) Transgenic Plants, in: Biotechnology, A Multi-Volume Comprehensive Treatise (H.J. Rehm, G. Reed, A. Pühler, P. Stadler, eds.) Vol. 2, 627-659, V.C.H. Weinheim - New York - Basel - Cambridge).

Während die Transformation dikotyler Pflanzen über Ti-Plasmid-Vektorsysteme mit Hilfe von *Agrobacterium tumefaciens* wohl etabliert ist, weisen neuere Arbeiten darauf hin, daß auch monokotyle Pflanzen der Transformation mittels *Agrobacterium*-basierender Vektoren sehr wohl zugänglich sind (Chan *et al.* (1993) Plant Mol. Biol. 22, 491-506; Hiei *et al.* (1994) Plant J. 6, 271-282; Deng *et al.* (1990) Science in China 33, 28-34; Wilmink *et al.* (1992) Plant Cell Reports 11, 76-80; May *et al.* (1995) Bio/Technology 13, 486-492; Conner und Domiss (1992) Int. J. Plant Sci. 153, 550-555; Ritchie *et al.* (1993) Transgenic Res. 2, 252-265).

Alternative Systeme zur Transformation von monokotylen Pflanzen sind die Transformationen mittels des biolistischen Ansatzes (Wan and Lemaux (1994) Plant Physiol. 104, 37-48; Vasil *et al.* (1993) Bio/Technology 11, 1553-1558; Ritala *et al.* (1994) Plant Mol. Biol. 24, 317-325; Spencer *et al.* (1990) Theor. Appl. Genet. 79, 625-631; Altpeter *et al.* (1996) Plant Cell Reports 16, 12-17), die Protoplasten-Transformation, die Elektroporation von partiell permeabilisierten Zellen, die Einbringung von DNA mittels Glasfasern.

Spezifisch die Transformation von Mais wird in der Literatur verschiedentlich beschrieben (vgl. z.B. WO 95/06128, EP 0 513 849; EP 0 465 875; Fromm *et al.* (1990) Biotechnology 8, 833-844; Gordon-Kamm *et al.* (1990) Plant Cell 2, 603-618; Koziel *et al.* (1993) Biotechnology 11, 194-200). In EP 292 435 wird ein Verfahren beschrieben, mit Hilfe dessen, ausgehend von einem schleimlosen, weichen, granulösen Mais-Kallus, fertile Pflanzen erhalten werden können. Shillito *et al.* ((1989) Bio/Technology 7, 581) haben in diesem Zusammenhang beobachtet, daß es ferner für die Regenerierbarkeit zu fertilen Pflanzen notwendig ist, von Kallus-Suspensionskulturen auszugehen, aus denen eine sich teilende

Protoplastenkultur, mit der Fähigkeit zu Pflanzen zu regenerieren, herstellbar ist. Nach einer *in vitro* Kultivierungszeit von sieben bis acht Monaten erhalten Shillito *et al.* Pflanzen mit lebensfähigen Nachkommen.

Prioli und Söndahl ((1989) Bio/Technology 7, 589) beschreiben die Regeneration und die Gewinnung fertiler Pflanzen aus Mais-Protoplasten, der Cateto-Mais-Inzuchtlinie Cat 100-1. Die Autoren vermuten, daß die Protoplasten-Regeneration zu fertilen Pflanzen von einer Anzahl verschiedener Faktoren, wie z.B. vom Genotyp, vom physiologischen Zustand der Donor-Zellen und von den Kultivierungsbedingungen, abhängig ist.

Auch die erfolgreiche Transformation anderer Getreidearten wurde bereits beschrieben, z.B. für Gerste (Wan und Lemaux, supra; Ritala *et al.*, supra) und für Weizen (Nehra *et al.* (1994) Plant J. 5, 285-297; Altpeter *et al.*, supra).

Ist die eingeführte DNA einmal im Genom der Pflanzenzelle integriert, so ist sie dort in der Regel stabil und bleibt auch in den Nachkommen der ursprünglich transformierten Zelle erhalten. Sie enthält normalerweise einen Selektionsmarker, der den transformierten Pflanzenzellen Resistenz gegenüber einem Biozid oder einem Antibiotikum wie Kanamycin, G418, Bleomycin, Hygromycin, Methotrexat, Glyphosat, Streptomycin, Sulfonyl-Harnstoff, Gentamycin oder Phosphinotricin u.a. vermittelt. Der individuell gewählte Marker sollte daher die Selektion transformierter Zellen gegenüber Zellen, denen die eingeführte DNA fehlt, gestatten.

Die transformierten Zellen wachsen innerhalb der Pflanze in der üblichen Weise (siehe auch McCormick *et al.* (1986) Plant Cell Reports 5, 81-84). Die resultierenden Pflanzen können normal angezogen werden und mit Pflanzen, die die gleiche transformierte Erbanlage oder andere Erbanlagen besitzen, gekreuzt

werden. Die daraus entstehenden hybriden Individuen haben die entsprechenden phänotypischen Eigenschaften. Von den Pflanzenzellen können Samen gewonnen werden.

Es sollten zwei oder mehrere Generationen angezogen werden, um sicherzustellen, daß das phänotypische Merkmal stabil beibehalten und vererbt wird. Auch sollten Samen geerntet werden, um sicherzustellen, daß der entsprechende Phänotyp oder andere Eigenarten erhalten geblieben sind.

Ebenso können nach üblichen Methoden transgene Linien bestimmt werden, die für die neuen Nukleinsäuremoleküle homozygot sind und ihr phänotypisches Verhalten hinsichtlich eines veränderten Tocopherolgehalts untersucht und mit dem von hemizygoten Linien verglichen werden.

Die Expression der erfindungsmäßen Proteine mit Geranylgeranyl-Reduktase-Aktivität kann mit Hilfe herkömmlicher molekularbiologischer und biochemischer Methoden erfolgen. Dem Fachmann sind diese Techniken bekannt und er ist problemlos in der Lage, eine geeignete Nachweismethode zu wählen, beispielsweise eine Northern Blot-Analyse zum Nachweis Geranylgeranyl-Reduktase-spezifischer RNA bzw. zur Bestimmung der Höhe der Akkumulation von Geranylgeranyl-Reduktase-spezifischer RNA, eine Southern-Blot Analyse zur Identifizierung für Geranylgeranyl-Reduktase kodierender DNA-Sequenzen oder eine Western Blot-Analyse zum Nachweis des durch die erfindungsgemäßen DNA-Sequenzen kodierten Proteins, vorzugsweise CHL P. Der Nachweis der enzymatischen Aktivität der Geranylgeranyl-Reduktase kann beispielsweise mit dem von Soll und Schultz (1981) in Biochem. Biophys. Res. Commun. 99, 907-912 beschriebenen Enzymassay über die Bildung von Chlorophyll-Phytyl erfolgen.

Die Erfindung basiert auf der erfolgreichen Isolierung eines für eine Geranylgeranyl-Reduktase kodierenden cDNA-Klons aus einer cDNA-Library aus *Nicotiana tabacum* cv. Petit Havana SR1. Die Sequenz dieses cDNA-Klons, die einen vollständigen offenen Leserahmen umfaßt, ist in SEQ:ID NO. 1 dargestellt. Unter Verwendung der Sequenz gemäß SEQ:ID No. 1 gelang die Erzeugung transgener Pflanzen, die einen gegenüber Wildtyppflanzen veränderten Tocopherolgehalt aufweisen.

Der die DNA-Sequenz gemäß SEQ:ID NO. 1 enthaltende cDNA-Klon wurde in *Escherichia coli* transformiert und der entsprechende *E. coli*-Stamm am 16.10.1997 bei der Deutschen Sammlung für Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1b, D-38124 Braunschweig, unter der Hinterlegungsnummer DSM 11816 gemäß dem Budapester Vertrag hinterlegt.

Die nachfolgenden Beispiele dienen zur Erläuterung der Erfindung.

BEISPIELE

Beispiel 1:
Klonierung einer Tabak-cDNA, die eine Geranylgeranyl-Reduktase (CHL P) kodiert

Zur Identifizierung der Geranylgeranyl-Reduktase-cDNA aus Tabak wurde eine Lambda ZAP II cDNA-Bibliothek (*Nicotiana tabacum* SR1, Stratagene, USA) unter Verwendung eines für den Locus 4D9T7P kodierenden EST aus *Arabidopsis thaliana* nach Vorschrift gescreent. Die verwendete EST-Sequenz weist eine Ähnlichkeit zu den bekannten bch P/chl P-Sequenzen aus *Rhodobacter capsulatus* (Young *et al.* (1989) Mol. Gen. Genet. 218, 1-12; Bollivar *et al.* (1994) J. Mol.

Biol. 237, 622-640; Bollivar *et al.* (1994) Biochemistry 33, 12763-12768) und *Synechocystis PCC6803* (Addlesee *et al.* (1996) FEBS Lett. 389, 126-130) auf.

Die verwendete Hybridisierungssonde umfaßt den Bereich der in 4D9T7P (Accession No. T04791) dargestellten EST-Sequenz von Base 1 bis Base 364. Die Sonde wurde als NotI/SalI-Restriktionsfragment aus der PRL2-Library von *A. thaliana* (Vektor: λZipLox) (Newman *et al.* (1994) Plant Physiol. 106:1241-1255) isoliert und mit [α-^{32}P]dCTP mittels Nicktranslation (Life Technologies, Eggenstein) radioaktiv markiert.

Die Hybridisierung wurde nach folgendem Protokoll durchgeführt.

- 2 h Vorhybridisierung bei 55 °C mit Hybridisierungslösung folgender Zusammensetzung: 5 x SSC, 0,1 % SDS, 5 x Denhardtreagenz, 100 µg/ml denaturierte Salmonsperm-DNA;
- 12 h Haupthybridisierung bei 55 °C mit frischer Hybridisierungslösung der oben genannten Zusammensetzung plus radioaktiv markierte Sonde;
- Waschen: 2 x 10 min. bei 55 °C mit 2 x SSC und 0,1 % SDS, und 1 x 5 min. bei 55 °C mit 1 x SSC und 0,1 % SDS.

Die nach cDNA-Bank-Screening isolierte Plasmid-DNA wurde sequenziert. Die identifizierte und in SEQ:ID NO. 1 dargestellte chl P-cDNA-Sequenz umfaßt 1510 Nukleotide (ohne polyA-Schwanz), von denen die Nukleotide 1 bis 1392 für ein 52 kDa-Protein von 464 Aminosäuren (einschließlich des Start-Methionins, und ohne das Stopkodon (Nukleotide 1393 bis 1395) gerechnet) kodieren. Die abgeleitete Aminosäuresequenz des CHL P ist in SEQ:ID NO. 2 gezeigt. Die in SEQ:ID NO. 1 gezeigte Nukleotidsequenz umfaßt einen 3' untranslatierten Bereich von Nukleotid 1396 bis 1510.

Zur DNA-RNA-Isolierung, Sequenzanalye, Restriktion, Klonierung, Gelelektrophorese, radioaktive Markierung, Southern, Northen und Western Blot Analysen, Hybridisierung und dergleichen wurden gängige Methoden angewandt, wie sie in einschlägigen Laborhandbüchern, wie Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual, 2. Auflage, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, beschrieben sind.

Beispiel 2:

Transformation von Tabakpflanzen und Regeneration intakter Pflanzen

Für die Herstellung transgener Pflanzen, die CHL P überexprimieren und daher einen gegenüber nicht-transformierten Pflanzen erhöhten Tocopherolgehalt aufweisen, wurde die DNA-Sequenz gemäß SEQ:ID NO. 1 mittels der Restriktionsenzyme BamHI und SalI in der multiplen Klonierungsschnittstelle des pBluescript-Vektors aus dem Vektor herausgeschnitten und in Sense-Orientierung in den binären Vektor BinAR-TX (Höfgen and Willmitzer (1990) Plant Science 66, 221-230), einem pBIB-Abkömmling (Becker (1990) Nucleic Acid Res. 18, 203), der mit den selben Restriktionsendonukleasen verdaut wurde, hinter den CaMV 35S-Promotor hineinligiert. Zur Verdeutlichung ist eine Restriktionskarte des Vektors BinAR-TX als Abbildung 3 beigefügt.

Anstelle des genannten binären Vektors BinAR-TX kann jeder beliebige für die Pflanzentransformation geeignete Vektor für die Herstellung eines chimären Gens, bestehend aus einer Fusion des CaMV 35S-Promotors oder eines anderen Promotors, der die Transkription und Translation in Pflanzenzellen gewährleistet, und DNA-Sequenzen, die für CHL P kodieren, verwendet werden.

Der rekombinante Vektor pCHLPbin wurde sodann in Agrobacterium tumefaciens (Stamm GV2260; Horsch *et al.* (1985) Science 227, 1229-1231) transformiert und zur Transformation von Tabakpflanzen (SNN) mittels der Blattscheiben-Transformationstechnik (Horsch *et al.*, supra) eingesetzt.

Hierzu wurde eine Übernachtkultur des entsprechenden *Agrobacterium tumefaciens*-Klons für 10 Minuten bei 5000 rpm abzentrifugiert, und die Bakterien wurden in 2YT-Medium resuspendiert. Junge Tabakblätter einer Sterilkultur (*Nicotiana tabacum* cv. *Samsun* NN) wurden in kleine, ca. 1 cm^2 große Stücke zerschnitten und kurzzeitig in die Bakteriensuspension gelegt. Die Blattstücke wurden anschließend auf MS-Medium (Murashige und Skoog (1962) Physiol. Plant. 15, 473; 0,7 % Agar) gelegt und zwei Tage im Dunkeln inkubiert. Anschließend wurden die Blattstücke zur Sproßinduktion auf MS-Medium (0,7 % Agar) mit 1,6 % Glukose, 1 mg/l 6-Benzylaminopurin, 0,2 mg/l Naphthylessigsäure, 500 mg/l Claforan (Cefotaxim, Hoechst, Frankfut) und 50 mg/l Kanamycin gelegt. Das Medium wurde alle sieben bis zehn Tage gewechselt. Wenn sich Sprosse entwickelt hatten, wurden die Blattstücke in Glasgefäße, die dasselbe Medium enthielten, überführt. Entstehende Sprosse wurden abgeschnitten und auf MS-Medium mit 2 % Saccharose und 250 mg/l Claforan gegeben und zu ganzen Pflanzen regeneriert.

Beispiel 3:
Analyse transgener Tabakpflanzen, die den rekombinanten Vektor pCHLPbin enthalten

Transgene Tabakpflanzen wurden wie oben beschrieben transformiert, selektioniert und regeneriert. Nach Bewurzelung in Sterilkultur wurden ca. 100 unabhängige Transformanten im Gewächshaus in Erde überführt. Die Tabakpflanzen wurden im

Gewächshaus bei 60 % Luftfeuchtigkeit und 20-25 °C für 16 Stunden im Licht und 18-20 °C für 8 Stunden in Dunkelheit gehalten.

Die Transformanten mit normalen oder erhöhten Tocopherol- bzw. Chlorophyll-gehalten zeigten weder hinsichtlich ihres Phänotyps ein verändertes Aussehen noch eine im Vergleich zu Kontrollpflanzen veränderte Wachstumsrate.

Einige der Primärtransformanten zeigten einen gegenüber Wildtyppflanzen bis zu 4- bis 6fach-gesteigerten Tocopherolgehalt. Diese Steigerung des Tocopherol-gehalts konnte in Nachkommen der T1- und T2-Generation und in homozygoten Tochterpflanzen, die durch übliche Selbstbestäubung und anschließende Bestimmung des Aufspaltungsmusters der Samen auf Kanamycin-haltigem Medium gewonnen wurden, zusätzlich erhöht werden.

Des weiteren konnte beobachtet werden, daß die Tocopherolgehalte in transgenen Pflanzen unter Streßbedingungen, wie z.B. Anzucht unter tiefen und erhöhten Temperaturen bzw. Starklicht, und in seneszenten Blättern gegenüber Kontroll-pflanzen zusätzlich erhöht waren.

Die Bestimmung des Tocopherolgehalts erfolgte nach folgendem Protokoll: Blattscheiben wurden in flüssigem Stickstoff homogenisiert und dreimal in Methanol extrahiert. Die Extrakte wurden gesammelt und auf einer LiCrospher 100 HPLC RP-18-Säule (Merck, Darmstadt, Deutschland) bei einem Fluß von 1 ml/min mit folgendem Gradienten eluiert: 94 % Laufmittel B (100 % Methanol) / 6 % Laufmittel A (30 % Methanol, 10 % 0,1 M Ammoniumacetat, pH 5,1) für 7 min., für weitere 17 min. 99 % Laufmittel B / 1 % Laufmittel A, dann weitere 26 min. 94 % Laufmittel B / 6 % Laufmittel A.

Alternativ wurden die gesammelten Extrakte mittels HPLC in einem isokratischen Gradienten analysiert (Gradient besteht zu 2 % aus Lösung A [10 % Methanol und 10 % Essigsäure] und zu 98 % aus Methanol (Lösung B); Flußrate 1 ml/min.). Es wurde eine Waters LC-Modul-Anlage mit Shimadzu RF 551 Fluoreszenzdetektor (295 nm_{ex}, 325 nm_{em}) verwendet.

Das Ergebnis eines Tocopherolassays ist in Abbildung 4 in Form eines Balkendiagramms dargestellt. Der Vergleich der Blätter 6, 9, 12 (gezählt von der Pflanzenspitze) von den Transformanten 28 und 30 mit den entsprechenden Blättern der Kontrollpflanze (SNN) belegt einen bis zu 6-fach gesteigerten Gehalt an Tocopherol in den transgenen Linien.

Unabhängig von ihrer Fähigkeit, auf Kanamycin-haltigem Medium zu wachsen, wurden die transgenen Tabakpflanzen auch im Southern Blot analysiert. Hierbei ergaben sich nach Hybridisierung mit einem markierten cDNA-Fragment für CHL P zusätzliche radioaktiv-markierte Banden der mit Restriktionsenzymen verdauten genomischen DNA der Transformanten im Vergleich zu Kontrollpflanzen

Eine Northern Blot-Analyse ergab eine gegenüber den CHL P-RNA-Gehalten der Kontrollpflanzen erhöhte Menge an spezifischer RNA in den Transformanten.

Eine erhöhte Geranylgeranyl-Reduktase-Expression in den transgenen Pflanzen konnte auch im Western Blot nachgewiesen werden. Die Transformanten zeigten gegenüber Kontrollpflanzen eine erhöhte Menge an CHL P-Protein.

Zusätzlich konnte in Plastiden-Importexperimenten (durchgeführt nach Grimm *et al.* (1989) Plant Mol. Biol. 13, 583-593) bestätigt werden, daß das durch die Sequenz gemäß SEQ:ID NO. 1 kodierte CHL P-Vorstufenprotein nach *in vitro*-Transkription und -Translation in die Plastiden importiert wurde.

Beispiel 4:
Herstellung von CHL P-Antisense-Konstrukten und Übertragung auf Tabak

Während die in den Beispielen 2 und 3 erzeugten und analysierten transgenen Pflanzen einen aufgrund der Überexpression der erfindungsgemäßen DNA-Sequenzen erhöhten Tocopherolgehalt aufwiesen, wurde für die Herstellung von transgenen Tabakpflanzen, die eine reduzierte Aktivität von CHL P aufweisen, folgendes Antisense-Konstrukt erzeugt und auf Tabak übertragen.

Die cDNA-Sequenz gemäß SEQ:ID NO. 1 wurde mittels der Restriktionsenzyme KpnI und XbaI in der multiplen Klonierungsschnittstelle des pBluescript-Vektors aus dem Vektor herausgeschnitten und in Antisense-Orientierung in den binären Vektor BinAR-TX (siehe Beispiel 2), der mit den selben Restriktionsenzymen verdaut wurde, mit dem 35S-Promotor von Cauliflower Mosaic Virus fusioniert. Der hieraus resultierende rekombinante Vektor pCHLPASbin wurde wie in Beispiel 2 beschrieben mittels *Agrobacterium tumefaciens*-vermittelter Leaf Disc-Transformation auf Tabak übertragen. Anschließend wurden transgene Pflanzen regeneriert. Ungefähr 100 unabhängige transgene Linien wurden regeneriert und die Insertion von Kopien des Transgens mittels üblicher Verfahren (z.B. Southern Blot-Hybridisierung) bestätigt.

Die Transformanten zeigten einen im Vergleich zu Kontrollpflanzen verminderten Wuchs, einen ausgebleichten Phänotyp, reduzierte RNA- und Proteingehalte für CHL P, einen hohen Gehalt an Geranylgeranyl-Chlorophyll (bis zu 50% des Gesamt-Chlorophyllgehalts im Vergleich zu 100% Phytyl-Chlorophyll der Wildtyppflanzen) sowie einen verminderten Chlorophyll- und Tocopherolgehalt.

Beispiel 5:

Überexpression aktiver Geranylgeranyl-Reduktase in *Escherichia coli*

Für die Herstellung von Expressionsklonen, die das rekombinante CHL P in *E. coli* überexprimieren, wurde der offene Leserahmen für ein vermutlich reifes (prozessiertes) Protein mittels der Oligonukleotid-Primer

- CSYN 1 5'-cgc cat ggg ccg caa tct tcg tgt tgc ggt-3' und
- CSYN 2 5'-gca gat ctg tcc att tcc ctt ctt agt gca-3'

von der DNA-Sequenz gemäß SEQ:ID No. 1 mittels PCR amplifiziert (1 min. 94 °C; 2 min. 60 °C; 3 min. 72 °C für 25 Zyklen). Das amplifizierte PCR-Fragment wurde aufgereinigt und mit den Restriktionsenzymen NcoI und BglII geschnitten und in den mit den selben Enzymen verdauten Expressionsvektor pQE60 (Qiagen, Hilden) hineinligiert. Hinter dem Initiationskodon ATG (Bestandteil der Erkennungssequenz für NcoI) folgte die kodierende *Chl P*-Sequenz, beginnend mit dem Nukleotid Nr. 148 des offenen Leserahmens der CHL P-cDNA-Sequenz. Daraus resultiert nach dem Methionin der Einbau eines Glycins (Aminosäure Nr. 50 des aus der cDNA-Sequenz übersetzten Proteins).

Für die Expression des pflanzlichen CHL P wurden die *E. coli*-Stämme XL 1 Blue (Stratagene, LaJolla, CA, USA) oder SG 13009 (Gottesman *et al.* (1981) J. Bacteriol. 148, 265-273) mit dem rekombinanten Vektor transformiert. Nach Induktion der Transkription des rekombinanten Gens mittels IPTG wurde in den *E. coli*-Stämmen ein Protein mit einem Molekulargewicht von ca. 47 kDa exprimiert. Das Protein wurde in der pelletierbaren Fraktion des Bakterienextraktes nachgewiesen und aus dem Gesamtextrakt unter denaturierenden Bedingungen über eine Nickel-Affinitätssäule gemäß den Instruktionen des Anbieters (Qiagen, Hilden) aufgereinigt.

Das aufgereinigte Protein wurde zur Antikörperherstellung in Kaninchen injiziert.

Das Protein aus dem Gesamtextrakt besitzt eine Geranylgeranyl-Reduktase-Aktivität in einem kombinierten Enzymassay mit bakterieller Bacteriochlorophyllsynthase unter Verwendung von Chlorophyllid und GGPP. Der Enzymassay wurde gemäß dem Protokoll in Oster *et al.* (1997) J. Biol. Chem. 272, 9671-9676, durchgeführt.

Die Auftrennung von Chlorophyll-GG und Chlorophyll-Phytol erfolgte auf einer HPLC mit der RP 18-Säule unter Verwendung der Lösungsmittelgemische Lösung A (60% Aceton) und Lösung B (100% Aceton). Der Gradient an Lösungsmitteln wurde eingesetzt: t_o 75% Lösung A und 25% Lösung B, 2 min.;innerhalb t_{2-4} auf 45% Lösung A und 55% Lösung B; innerhalb t_{4-13} auf 30% Lösung A und 70% Lösung B; innerhalb t_{13-17} auf 100% Lösung B; t_{17-21} 100% Lösung B isokratisch; anschließend innerhalb von 5 min. auf 75% Lösung A und 25% Lösung B; dann weitere 5 min. 75% Lösung A und 25% Lösung B isokratisch. Zur Detektion der Tetrapyrrole wurde ein Fluoreszenzdetektor verwendet (λ_{ex} 425 nm, λ_{em} 665 nm).

Beispiel 6:

Co-Expression des CHL P-Gens und des HPD-Gens in *Nicotiana tabacum*

Mittels der Oligonukleotid-Primer

- hpdoli1 5'-tta ggt acc atg ggc cac caa aac gcc gcc gtt tca g-3'

und

- hpdoli2 5'-tga gtc gac cac aat cct tta gtt ggt tct tct tct tg-3'

wurde aus einer *Arabidopsis thaliana*-cDNA-Bibliothek die Sequenz der HPD-cDNA (Accession No.: AF 000228) zwischen Nukleotid 37 und 1404 amplifiziert,

kloniert und ansequenziert. Das Fragment wurde mit den Restriktionsendonukleasen KpnI und SalI geschnitten und in den KpnI/SalI-geschnittenen binären Vektor Bin-Hyg-TX hineinligiert. Bei dem Vektor Bin-Hyg-TX handelt es sich (wie bei dem in Beispiel 2 verwendeten Vektor BinAR-TX) um einen pBIB-Abkömmling (Becker, supra), der die Expression einer in die multiple Klonierungsstelle einligierten kodierenden Region unter Kontrolle des 35S RNA Promotors aus Cauliflower Mosaic Virus ermöglicht. Im Unterschied zu dem für die Expression des CHL P-Gens eingesetzten Vektors BinAR-TX, der in Pflanzenzellen eine Selektion auf Kanamycin ermöglicht, trägt der binäre Vektor Bin-Hyg-TX ein Hygromycinresistenzgen als Selektionsmarker für Pflanzenzellen. Zur Verdeutlichung ist eine Restriktionskarte des Vektors Bin-Hyg-TX als Abbildung 5 beigefügt.

Anstelle des genannten binären Vektors Bin-Hyg-TX kann jeder beliebige für die Pflanzentransformation geeignete Vektor für die Herstellung eines chimären Gens, bestehend aus einer Fusion des CaMV 35S-Promotors oder eines anderen Promotors, der die Transkription und Translation in Pflanzenzellen gewährleistet und DNA-Sequenzen, die für HPD kodieren, verwendet werden.

Dann erfolgte die Übertragung des erhaltenen, rekombinanten Vektor pBinHygHPD auf Tabak SNN mittels *Agrobacterium tumefaciens* wie in obigem Beispiel 2 beschrieben, wobei transgene Tabaksprosse auf Hygromycin-haltigem Medium selektiert wurden. Die nach Regeneration erhaltenen Pflanzen dienten als Kontrollpflanzen.

Zusätzlich wurden die in Beispiel 2 beschriebenen Transformanten 28 und 30 und wahlweise andere Transformanten, die das CHL P-Gen unter Kontrolle des 35S

RNA Promotors überexprimieren, mittels Leaf Disc Transformation mit dem rekombinanten Vektor pBinHygHPD transformiert und ganze Pflanzen unter Selektion auf Kanamycin und Hygromycin regeneriert.

In sämtlichen transgenen Pflanzen wurde die erfolgreiche Übertragung und Expression des chimären HPD-Gens durch geeignete Southern- und Northernblot-Experimente mit dem oben erwähnten PCR-Fragment für HPD als Sonde bestätigt.

Ein Vergleich der Tocopherolgehalte (Analyse wie in Beispiel 3) in Blättern von Pflanzen, die nur das CHL P-Gen exprimieren (siehe Beispiel 3, Transformanten 28 und 30), und solchen Pflanzen, die das HPD-Gen und das CHL P-Gen co-exprimieren (Transformanten 28+HPD und 30+HPD) zeigte, daß der Tocopherol-gehalt durch die gleichzeitige Expression des Hydroxyphenylpyruvat-Dioxygenase-Gens zusätzlich gesteigert werden konnte.

Alternativ können transgene Pflanzen, die sowohl das CHL P-Gen als auch das HPD-Gen exprimieren, auch durch Kreuzung homozygoter "CHL P-Linien" mit homozygoten "HPD-Linien" erhalten werden.

Eine zusätzliche Steigerung des Tocopherolgehalts in Doppeltransformanten (bzw. Doppelkreuzungen) ist unter Streßbedingungen (z.B. erhöhte Temperatur, Licht-streß u.ä.) zu erwarten.

Neben den vorangehend beschriebenen Doppeltransformanten, die HPD und CHL P exprimieren, wurden auch transgene HPD-Linien auf ihren Tocopherolgehalt und den Einfluß von Streßbedingungen auf den Tocopherolgehalt untersucht. Es konnte gezeigt werden, daß Überexpression von HPD in Tabakblättern eine 2- bis

3-fache Steigerung des Tocopherolgehalts zur Folge hat und daß insbesondere die Werte unter Streßbedingungen gegenüber nichttransgenen Kontrollpflanzen ansteigen.

Abbildung 6 zeigt Tocopherolgehalte in den Blättern 4, 7 und 10 von 12 Wochen-alten transgenen Takaklinien (# 2, 6 und 33), die das Arabidopsis-Enyzm Hydroxyphenyl-Pyruvat-Dioxygenase (HPD) überexprimieren, im Vergleich zu Kontrollpflanzen (SNN). Die Pflanzen wurden entweder bei 38 °C oder bei 10 °C und einer Lichtintensität von ca. 200 µmol Photonen/m^2/s abgezogen. Tocopherol wird mit zunehmendem Blattalter in den Pflanzen verstärkt angereichert. Insbesondere in den älteren Blättern steigt der Gehalt des Tocopherols in den unter 38 °C angezogenen Pflanzen sehr viel steiler an und erreicht die zweifache Menge gegenüber den Kontrollpflanzen. Dagegen sind die Tocopherolgehalte in den Transformanten in der Kälte nur leicht gegenüber denen der Wildtyppflanzen erhöht.

Die Ergebnisse deuten darauf hin, daß HPD bei erhöhtem Bedarf an Tocopherolen, also insbesondere unter Streßbedingungen, die Regeneration dieser Antioxidantien in transgenen Pflanzen besonders begünstigt.

Beispiel 7:
Erhöhte Resistenz der transgenen Pflanzen gegenüber oxidativem Streß

Nach Beispiel 2 und 3 hergestellte transgene Pflanzen wurden in einem Blattscheiben-Inkubationsversuch auf den Einfluß inhibitorischer und oxidativer Substanzen untersucht. 15 Keimlinge wurden in 20 mM Kaliumphosphat-Puffer (pH 7,1) für 10 h im Licht inkubiert. Neben den Kontrollansätzen (Wasser) wurden die Pflanzen in Ansätzen mit 3,3 µM (niedrige Konzentration = NK) oder

33 µM (hohe Konzentration = HK) Acifluorfen (BASF, Ludwigshafen, Deutschland), einem Hemmstoff der Protoporphyrinogen-Oxidase, und in anderen Ansätzen mit 1,7 µM (niedrige Konzentration = NK) oder 17 µM (hohe Konzentration = HK) Bengalrosa (4,5,6,7-Tetrachlor-2',4',5',7'-tetrajodfluorescein, Sigma-Aldrich, Deisenhofen, Deutschland), einem Erzeuger von reaktiven Sauerstoffspezies, inkubiert. Der Gehalt an Tocopherol ist in den untersuchten transgenen Linien (28, 30) sowohl in den Puffer-Kontrollansätzen wie auch unter oxidativen Streßbedingungen gegenüber Wildtyppflanzen (SNN) jeweils um das 2- bis 3-fache erhöht. Die Ergebnisse sind in Abbildung 7 veranschaulicht.

Die Ergebnisse deuten darauf hin, daß sich die erfindungsgemäßen Pflanzen dank ihres gegenüber Wildtyppflanzen erhöhten Tocopherolgehalts durch eine erhöhte Resistenz gegenüber oxidativem Streß auszeichnen. Ein erhöhter antioxidativer Schutz der zellulären Membranen vor reaktiven Sauerstoffspezies in den erfindungsgemäßen Pflanzen ist somit zu erwarten.

Beispiel 8:
Tocopherolgehalt in Samen von transgenen Pflanzen

Tocopherol wurde wie oben beschrieben aus Samen von transgenen Pflanzen extrahiert, die sich aufgrund der Expression von CHL P unter Kontrolle des CaMV 35S-Promotors (vgl. Beispiel 2) durch einen im Vergleich zu Wildtyppflanzen erhöhten Tocopherolgehalt in Blättern auszeichnen. Die Ergebnisse der mittels HPLC durchgeführten Tocopherolquantifizierung sind im Vergleich zu Tabakkontrollpflanzen in Abbildung 8 dargestellt.

Neben α-Tocopherol wurde auch γ-Tocopherol quantifiziert. Letzteres liegt grundsätzlich in Tabaksamen in größeren Mengen vor; das Verhältnis von γ-Tocopherol

zu α-Tocopherol beträgt in Tabaksamen ca. 10:1. Die Gehalte an beiden Formen der Tocopherole, insbesondere α-Tocopherol, sind im Vergleich zu den Kontrollpflanzen in den transgenen Pflanzen mit erhöhter Geranylgeranyl-Reduktase-Expression um das 2- bis 3-fache erhöht.

Diese Ergebnisse, die die Auswirkungen einer konstitutiven CHL P-Expression unter Kontrolle des 35S-Promotors auf den Tocopherolgehalt im Samen wiedergeben, zeigen deutlich, daß bei Einsatz eines samenspezifischen Promotors (bzw. eines in Samen induzierbaren Promotors) die Expression der Geranylgeranyl-Reduktase im Samen und hierdurch der Tocopherolgehalt in Samen von transgenen Pflanzen zusätzlich gesteigert werden kann.

Sollten in irgendeiner Weise molekularbiologische Arbeiten nicht hinreichend beschrieben worden sein, so wurden diese nach Standardmethoden, wie bei Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual, 2. Auflage, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, beschrieben, durchgeführt. Bezüglich der Transformation von Pflanzen wird auf allgemein bekannte Übersichtsartikel sowie auf die oben genannten Veröffentlichungen verwiesen.

BESCHREIBUNG DER ABBILDUNGEN

Abbildung 1: SEQ:ID NO. 1 zeigt die Nukleotid-Sequenz einer chl P-cDNA der Geranylgeranyl-Reduktase (CHL P) aus *Nicotiana tabacum*.

Abbildung 2: SEQ:ID NO. 2 zeigt eine Aminosäuresequenz des Enzyms CHL P aus *N. tabacum*, abgeleitet von der in Abbildung 1 gezeigten SEQ:ID NO.1.

Abbildung 3: Restriktionskarte des für die Pflanzentransformation eingesetzten binären Vektors BinAR. BinAR (Höfgen and Willmitzer (1990) Plant Science 66, 221) ist ein Bin19-Derivat (Bevan (1984) Nucl. Acids Res. 12, 8711), das eine Expressionskassette für die konstitutive Expression von chimären Genen in Pflanzen enthält, wobei die Expressionskassette über die EcoRI- und HindIII-Restriktionsschnittstellen von Bin19 kloniert ist. Die Kassette umfaßt ein 770 Bp.-EcoRI/HindIII-Fragment, das den CaMV 35S-Promotor, einen teilweisen pUC18-Polylinker und das Terminationssignal des Octopinsynthase-Gens (OCS) enthält. Für die Insertion kodierender Sequenzen eignen sich besonders die unikalen Schnittstellen des pUC18-Polylinkers, nämlich KpnI, SmaI, BamHI, XbaI und SalI. Als Pflanzenselektionsmarker trägt der binäre Vektor BinAR ein Kanamycinresistenzgen.

Abbildung 4: Balkendiagramm, das den Tocopherolgehalt in den Blättern 6, 9, 12 (gezählt von der Pflanzenspitze) von transgenen Tabakpflanzen (Linien 28 und 30) versus den entsprechenden Blättern von Kontrollpflanzen (SNN) zeigt.

Abbildung 5: Restriktionskarte des für die Pflanzentransformation eingesetzten binären Vektors Bin-Hyg-TX, bei dem es sich um ein pBIB-Derivat (Becker, supra; Bevan, supra) handelt, das eine Expressionskassette für die konstitutive Expression von chimären Genen in Pflanzen enthält. Für die Insertion kodierender Sequenzen eignen sich besonders die unikalen Schnittstellen des pUC18-Polylinkers, nämlich HpaI, KpnI, SmaI, XbaI und SalI. Als Pflanzenselektionsmarker trägt der binäre Vektor Bin-Hyg-TX ein Hygromycinresistenzgen.

Abbildung 6: Säulendiagramm der Tocopherolgehalte in den Blättern 4, 7 und 10 in 12 Wochen-alten transgenen Linien (# 2, 6, 33), die das Arabidopsis-Enzym HPD überexprimieren, und in Kontroll-pflanzen (SNN).

Abbildung 7: Tocopherolgehalt in 12 Tage-alten Keimlingen, die während einer 10stündigen Belichtung in 20 mM Kaliumphosphatpuffer (pH 7,1, Kontrolle = Wasser), mit 3,3 µM (NK) bzw. 33 µM (HK) Acifluorfen oder mit 1,7 µM (NK) bzw. 17 µM (HK) Bengalrosa inkubiert wurden (SNN = Wildtyp-Kontrollpflanzen, 28 und 30 = transgene Linien).

Abbildung 8: relative Werte an α-Tocopherol und γ-Tocopherol in Samen von transgenen Tabakpflanzen (# 7, 39, 67, 96) und Wildtyp-tabakpflanzen (SNN). 100% α-Tocopherol entspricht 6 ng/mg Samen; 100% γ-Tocopherol entspricht 62,8 ng/mg Samen.

ANSPRÜCHE

1. Nukleinsäuresequenz, dadurch gekennzeichnet, daß sie für ein Protein mit der Aktivität einer Geranylgeranyl-Reduktase oder ein aktives Fragment davon kodiert.

2. Nukleinsäuresequenz gemäß Anspruch 1, dadurch gekennzeichnet, daß sie für ein pflanzliches Protein kodiert.

3. Nukleinsäuresequenz gemäß Anspruch 1 oder 2, dadurch gekennzeichnet, daß sie für ein Protein aus Tabak kodiert.

4. Nukleinsäuresequenz gemäß SEQ:ID No. 1.

5. Nukleinsäuresequenz gemäß Anspruch 1, dadurch gekennzeichnet, daß sie in dem Klon DSM 11816 enthalten ist.

6. Allele und Derivate der Nukleinsäuressequenz gemäß einem der vorangehenden Ansprüche, dadurch gekennzeichnet, daß sie für ein Protein mit der Aktivität einer Geranylgeranyl-Reduktase oder ein aktives Fragment davon kodieren.

7. Nukleinsäuremolekül, dadurch gekennzeichnet, daß es eine Nukleinsäuresequenz gemäß einem der vorangehenden Ansprüche umfaßt.

8. Nukleinsäuremolekül gemäß Anspruch 7, dadurch gekennzeichnet, daß es eine Nukleinsäuressequenz gemäß einem der Ansprüche 1-6 in Kombination mit regulatorischen Elementen umfaßt, die die Transkription und Translation in prokaryontischen oder eukaryontischen Zellen gewährleisten.

9. Nukleinsäuremolekül gemäß Anspruch 7 oder 8, dadurch gekennzeichnet, daß es eine Nukleinsäuresequenz gemäß einem der Ansprüche 1-6 in Kombination mit Promotoren umfaßt, die die Transkription und Translation in Pflanzenzellen gewährleisten.

10. Nukleinsäuremolekül gemäß Anspruch 9, dadurch gekennzeichnet, daß es sich bei dem Promotor um einen konstitutiven, induzierbaren, gewebespezifischen oder entwicklungsspezifischen Promotor handelt.

11. Nukleinsäuremolekül gemäß Anspruch 10, dadurch gekennzeichnet, daß es sich bei dem Promotor um einen samenspezifischen Promotor handelt.

12. Nukleinsäuremolekül gemäß einem der Ansprüche 7-11, dadurch gekennzeichnet, daß es zusätzlich Enhancer-Sequenzen, für Signalpeptide kodierende Sequenzen oder andere regulatorische Sequenzen umfaßt.

13. Nukleinsäuremolekül gemäß einem der Ansprüche 7-12, in dem die kodierende Nukleinsäuresequenz in der Sense-Orientierung vorliegt.

14. Nukleinsäuremolekül gemäß einem der Ansprüche 7-12, in dem die kodierende Nukleinsäuresequenz in der Antisense-Orientierung vorliegt.

15. Protein mit der biologischen Aktivität einer Geranylgeranyl-Reduktase oder ein aktives Fragment davon, dadurch gekennzeichnet, daß es durch eine DNA-Sequenz oder ein Nukleinsäuremolekül gemäß einem der vorangehenden Ansprüche kodiert wird.

16. Protein mit der biologischen Aktivität einer Geranylgeranyl-Reduktase oder ein aktives Fragment davon, dadurch gekennzeichnet, daß es die in SEQ:ID NO. 2 gezeigte Aminosäuresequenz oder Bereiche davon aufweist.

17. Mikroorganismus, dadurch gekennzeichnet, daß er eine Nukleinsäuresequenz oder ein Nukleinsäuremolekül gemäß einem der Ansprüche 1-14 enthält.

18. Transgene Pflanzen, die eine Nukleinsäuresequenz, eine davon abgeleitete Nukleinsäuresequenz oder ein Nukleinsäuremolekül gemäß einem der Ansprüche 1-14 enthalten, sowie Teile dieser Pflanzen und deren Vermehrungsmaterial, wie Protoplasten, Pflanzenzellen, Kalli, Samen, Knollen oder Stecklinge, usw., sowie die Nachkommen dieser Pflanzen.

19. Pflanzen gemäß Anspruch 18, die einen gegenüber Wildtypflanzen veränderten Tocopherolgehalt aufweisen.

20. Pflanzen gemäß Anspruch 19, die einen gegenüber Wildtyppflanzen erhöhten Tocopherolgehalt aufweisen.

21. Pflanzen gemäß Anspruch 18, die einen gegenüber Wildtyppflanzen veränderten Gehalt an Vitamin K_1 aufweisen.

22. Pflanzen gemäß Anspruch 21, die einen gegenüber Wildtyppflanzen erhöhten Gehalt an Vitamin K_1 aufweisen.

23. Pflanzen gemäß Anspruch 18, die einen gegenüber Wildtyppflanzen veränderten Chlorophyllgehalt aufweisen.

24. Pflanzen gemäß Anspruch 23, die einen gegenüber Wildtyppflanzen erhöhten Chlorophyllgehalt aufweisen.

25. Dikotyle Pflanzen gemäß einem der Ansprüche 18-24.

26. Pflanzen gemäß Anspruch 25, bei denen es sich um Nutzpflanzen, Nahrungs- und/oder Futterpflanzen, insbesondere Raps, Soja, Tomate, Kartoffel, Zuckerrübe, Klee, handelt.

27. Monokotyle Pflanzen gemäß einem der Ansprüche 18-24.

28. Pflanzen gemäß Anspruch 27, bei denen es sich um Nutzpflanzen, Nahrungs- und/oder Futterpflanzen, insbesondere Getreide, wie Weizen, Gerste, Mais, Hafer, Roggen, Reis, Süß- und Weidegräser, handelt.

29. Pflanzen gemäß einem der Ansprüche 18-28, in denen eine Nukleinsäuresequenz gemäß einem der Ansprüche 1-6 integriert im Genom der Pflanze vorliegt, sowie Teile dieser Pflanzen und deren Vermehrungsmaterial, wie Protoplasten, Pflanzenzellen, Kalli, Samen, Knollen oder Stecklinge, usw., sowie die Nachkommen dieser Pflanzen.

30. Transgene Pflanzenzellen, einschließlich Protoplasten, die eine Nukleinsäuresequenz, eine davon abgeleitete Nukleinsäuresequenz oder ein Nukleinsäuremolekül gemäß einem der Ansprüche 1-14 enthalten.

31. Pflanzenzellen gemäß Anspruch 30, einschließlich Protoplasten, die einen gegenüber nicht-transformierten Pflanzenzellen veränderten, insbesondere erhöhten, Tocopherol-, Vitamin K_1- und/oder Chlorophyllgehalt aufweisen.

32. Pflanzen bzw. Pflanzenzellen gemäß einem der Ansprüche 18-31, die zusätzlich eine für Hydroxyphenylpyruvat-Dioxygenase kodierende Nukleinsäuresequenz enthalten.

33. Verfahren zur Herstellung von Pflanzen bzw. Pflanzenzellen gemäß einem der Ansprüche 18-32, folgende Schritte umfassend:

a) Herstellung einer Nukleinsäuresequenz, bestehend aus den folgenden Bestandteilen, die in der 5'-3'-Orientierung aneinandergereiht sind:
   - ein in Pflanzen funktionsfähiger, insbesondere samenspezifischer, Promotor,
   - mindestens eine Nukleinsäuresequenz, die für ein Protein mit der Aktivität einer Geranylgeranyl-Reduktase oder ein aktives Fragment davon kodiert und
   - gegebenenfalls ein Terminationssignal für die Termination der Transkription und die Addition eines poly-A-Schwanzes an das entsprechende Transkript, sowie ggf. davon abgeleitete DNA-Sequenzen;

b) Übertragung der Nukleinsäuresequenz aus a) auf pflanzliche Zellen und gegebenenfalls Integration der Nukleinsäuresequenz in das pflanzliche Genom,

c) falls erwünscht, Regeneration vollständig transformierter Pflanzen und gegebenenfalls Vermehrung dieser Pflanzen.

34. Verwendung einer Pflanze gemäß einem der Ansprüche 18-29 und 32 als Nahrungs- oder Futterpflanze.

35. Verwendung einer Pflanze gemäß einem der Ansprüche 18-29 und 32 als Produktionsstätte für Tocopherole und/oder Vitamin K_1.

36. Verwendung einer Nukleinsäuresequenz oder eines Nukleinsäuremoleküls gemäß einem der Ansprüche 1-14 zur Identifizierung, Isolierung oder Amplifizierung einer Nukleinsäuresequenz, die für ein Protein mit der Aktivität einer Geranylgeranyl-Reduktase oder ein aktives Fragment davon kodiert.

37. Verwendung einer Nukleinsäuresequenz, eines Nukleinsäuremoleküls oder eines Proteins gemäß einem der Ansprüche 1-16 zur Herstellung von Antikörpern.

38. Verwendung einer Nukleinsäuresequenz oder eines Nukleinsäuremoleküls gemäß einem der Ansprüche 1-14 zur Herstellung transgener Pflanzen bzw. Pflanzenzellen.

39. Verwendung gemäß Anspruch 38, worin die Pflanzen bzw. Pflanzenzellen einen veränderten Tocopherolgehalt aufweisen.

40. Verwendung gemäß Anspruch 38, worin die Pflanzen bzw. Pflanzenzellen einen veränderten Chlorophyllgehalt aufweisen.

41. Verwendung gemäß Anspruch 38, worin die Pflanzen bzw. Pflanzenzellen einen veränderten Vitamin K_1-Gehalt aufweisen.

42. Verwendung gemäß Anspruch 38, worin die Pflanzen bzw. Pflanzenzellen eine gegenüber Wildtyppflanzen erhöhte Herbizidtoleranz aufweisen.

43. Verwendung gemäß einem der Ansprüche 38-42, worin die Pflanzen bzw. Pflanzenzellen zusätzlich eine für Hydroxyphenylpyruvat-Dioxygenase kodierende Nukleinsäuresequenz enthalten.

44. Verwendung einer Nukleinsäuresequenz, eines Nukleinsäuremoleküls, eines Proteins oder eines Mikroorganismus gemäß einem der Ansprüche 1-17 zur Auffindung von Effektoren pflanzlicher Geranylgeranyl-Reduktase.

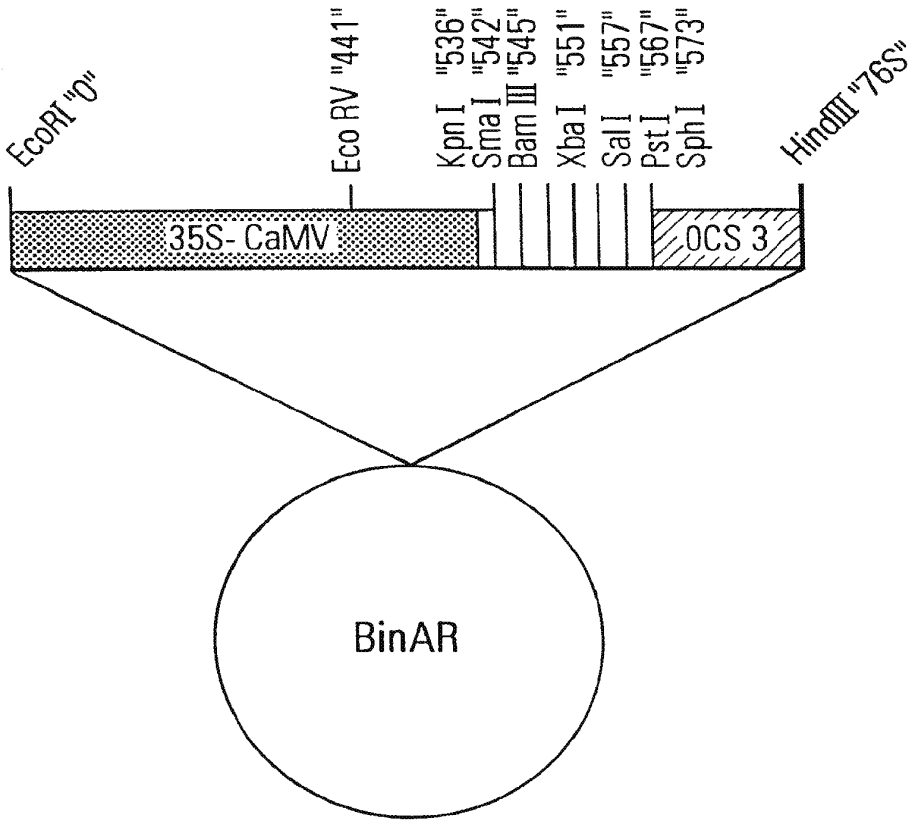
Abb. 1

ATGGCTTCCA TTGCTCTCAA AACTTTCACC GGCCTCCGTC AATCCTCGCC GGAAAACAAT
TCCATTACTC TTTCTAAATC CCTCCCCTTC ACCCAAACCC ACCGTAGGCT CCGAATCAAT
GCTTCCAAAT CCAGCCCAAG AGTCAACGGC CGCAATCTTC GTGTTGCGGT GGTGGGCGGT
GGTCCTGCTG GTGGCGCCGC CGCTGAAACA CTCGCCAAGG GAGGAATTGA AACCTTCTTA
ATCGAACGCA AAATGGACAA CTGCAAACCC TGCGGTGGGG CCATCCCACT TTGCATGGTG
GGAGAATTTG ACCTCCCTTT GGATATCATT GACCGGAAAG TTACAAAGAT GAAGATGATT
TCCCCATCCA ACGTTGCTGT TGATATTGGT CAGACTTTAA AGCCTCACGA GTACATCGGT
ATGGTGCGCC GCGAAGTACT CGATGCTTAC CTCCGTGACC GCGCTGCTGA AGCCGGAGCC
TCTGTTCTCA ACGGCTTGTT CCTCAAAATG GACATGCCCA AAGCTCCCAA CGCACCTTAC
GTCCTTCACT ACACAGCTTA CGACTCCAAA ACTAATGGCG CGGGGGAGAA GCGTACCCTG
GAAGTTGACG CCGTTATCGG CGCTGACGGT GCAAATTCCC GTGTCGCAAA ATCCATAAAC
GCCGGTGACT ACGAGTACGC TATTGCATTC CAAGAAAGGA TTAAAATTTC CGATGATAAA
ATGAAGTATT ACGAGAATTT AGCTGAAATG TACGTGGGTG ATGACGTGTC CCCTGATTTT
TACGGGTGGG TTTTCCCCAA ATGTGACCAC GTTGCCGTTG GCACTGGCAC AGTCACCCAC
AAAGCTGACA TCAAAAAATT CCAGCTAGCT ACAAGATTGA GAGCTGATTC CAAAATCACC
GGCGGAAAAA TTATCCGGGT CGAGGCCCAC CCGATTCCAG AACACCCAAG ACCCAGAAGA
TTACAAGACA GAGTTGCATT GGTTGGTGAT GCGGCAGGGT ACGTGACCAA ATGTTCGGGC
GAAGGGATTT ACTTCGCGGC AAAGAGTGGA CGTATGTGTG CTGAAGCAAT TGTTGAAGGG
TCAGAAATGG GAAAAAGAAT GGTGGACGAG AGTGATTTGA GGAAGTATTT GGAGAAATGG
GACAAGACTT ATTGGCCAAC GTACAAGGTG CTTGATATAT TGCAGAAGGT ATTTTACAGG
TCGAATCCGG CGAGGGAAGC ATTTGTTGAA ATGTGCGCAG ATGAGTATGT GCAGAAGATG
ACATTTGACA GCTATTTGTA CAAGAAAGTA GCACCAGGAA ACCCAATTGA AGACTTGAAG
CTTGCTGTGA ATACCATTGG AAGTTTGGTG AGAGCTAATG CACTAAGAAG GGAAATGGAC
AAGCTCAGTG TATAAGAAGA TTAACAGCAT TAATATTTTC TTGTAATTGA AGGATTTATT
TCTCAAATTA CTCTGTAAAC ACCTTTCATC CTGCCTTTAA TCGGATTTAT GTAACTTCAT
AATTTGAGCT

Abb. 2

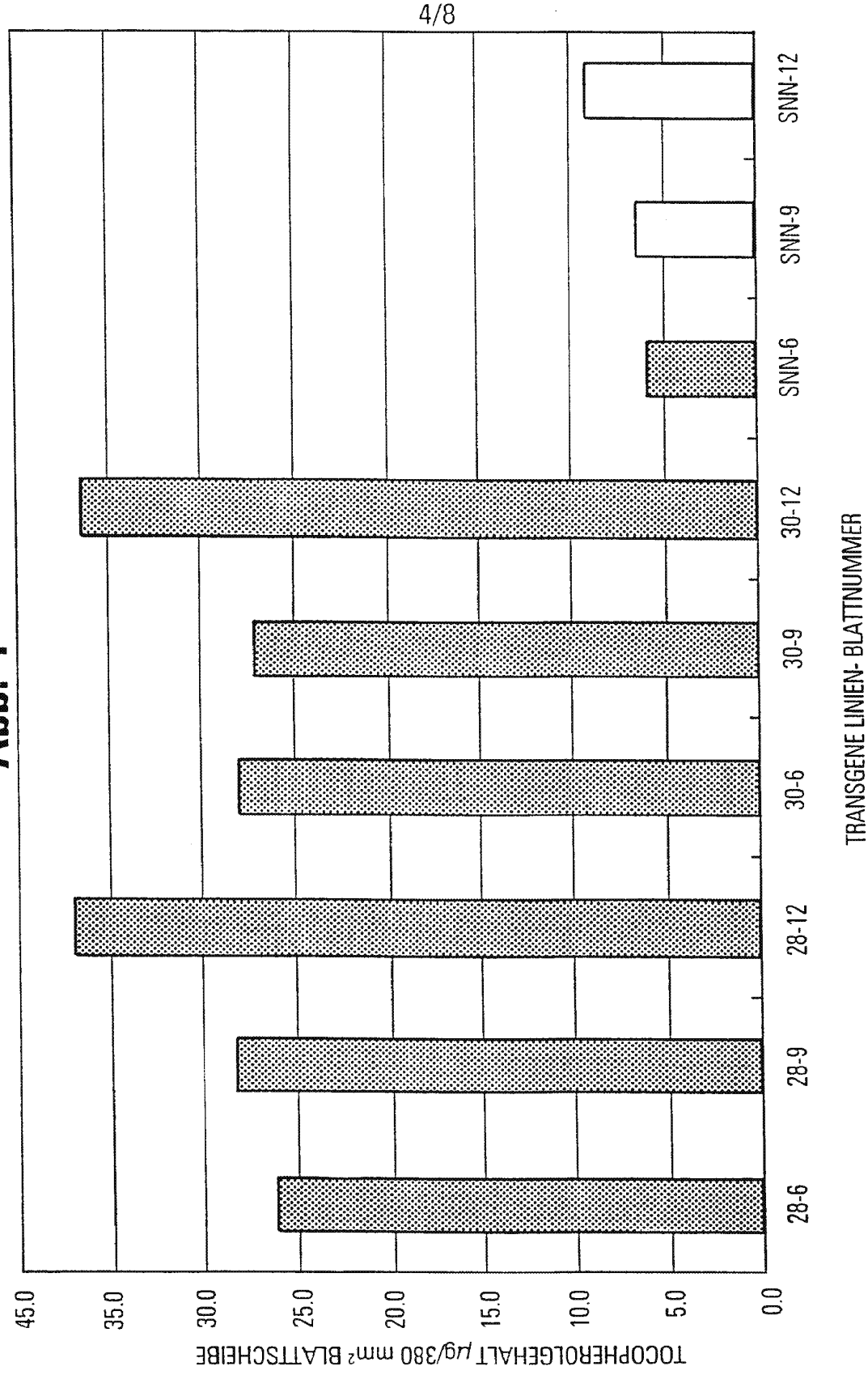
```
  1 Met Ala Ser Ile Ala Leu Lys Thr Phe Thr Gly Leu Arg Gln Ser
 16 Ser Pro Glu Asn Asn Ser Ile Thr Leu Ser Lys Ser Leu Pro Phe
 31 Thr Gln Thr His Arg Arg Leu Arg Ile Asn Ala Ser Lys Ser Ser
 46 Pro Arg Val Asn Gly Arg Asn Leu Arg Val Ala Val Val Gly Gly
 61 Gly Pro Ala Gly Gly Ala Ala Ala Glu Thr Leu Ala Lys Gly Gly
 76 Ile Glu Thr Phe Leu ILe Glu Arg Lys Met Asp Asn Cys Lys Pro
 91 Cys Gly Gly Ala Ile Pro Leu Cys Met Val Gly Glu Phe Asp Leu
106 Pro Leu Asp Ile Ile Asp Arg Lys Val Thr Lys Met Lys Met Ile
121 Ser Pro Ser Asn Val Ala Val Asp Ile Gly Gln Thr Leu Lys Pro
136 His Glu Tyr Ile Gly Met Val Arg Arg Glu Val Leu Asp Ala Tyr
151 Leu Arg Asp Arg Ala Ala Glu Ala Gly Ala Ser Val Leu Asn Gly
166 Leu Phe Leu Lys Met Asp Met Pro Lys Ala Pro Asn Ala Pro Tyr
181 Val Leu His Tyr Thr Ala Tyr Asp Ser Lys Thr Asn Gly Ala Gly
196 Glu Lys Arg Thr Leu Glu Val Asp Ala Val Ile Gly Ala Asp Gly
211 Ala Asn Ser Arg Val Ala Lys Ser Ile Asn Ala Gly Asp Tyr Glu
226 Tyr Ala Ile Ala Phe Gln Glu Arg Ile Lys Ile Ser Asp Asp Lys
241 Met Lys Tyr Tyr Glu Asn Leu Ala Glu Met Tyr Val Gly Asp Asp
256 Val Ser Pro Asp Phe Tyr Gly Trp Val Phe Pro Lys Cys Asp His
271 Val Ala Val Gly Thr Gly Thr Val THr His Lys Ala Asp Ile Lys
286 Lys Phe Gln Leu Ala Thr Arg Leu Arg Ala Asp Ser Lys Ile Thr
301 Gly Gly Lys Ile Ile Arg Val Glu Ala His Pro Ile Pro Glu His
316 Pro Arg Pro Arg Arg Leu Gln ASp Arg Val Ala Leu Val Gly Asp
331 Ala Ala Gly Tyr Val Thr Lys Cys Ser Gly Glu Gly Ile Tyr Phe
346 Ala Ala Lys Ser Gly Arg Met Cys Ala Glu Ala Ile Val Glu Gly
361 Ser Glu Met Gly Lys Arg Met Val Asp Glu Ser Asp Leu Arg Lys
376 Tyr Leu Glu Lys Trp Asp Lys Thr Tyr Trp Pro Thr Tyr Lys Val
391 Leu Asp Ile Leu Gln Lys Val Phe Tyr Arg Ser Asn Pro Ala Arg
406 Glu Ala Phe Val Glu Met Cys Ala ASp Glu Tyr Val Gln Lys Met
421 Thr Phe Asp Ser Tyr Leu Tyr Lys Lys Val Ala Pro Gly Asn Pro
436 Ile Glu Asp Leu Lys Leu Ala Val Asn Thr Ile Gly Ser Leu Val
451 Arg Ala Asn Ala Leu Arg Arg Glu Met Asp Lys Leu Ser Val
```

Abb. 3

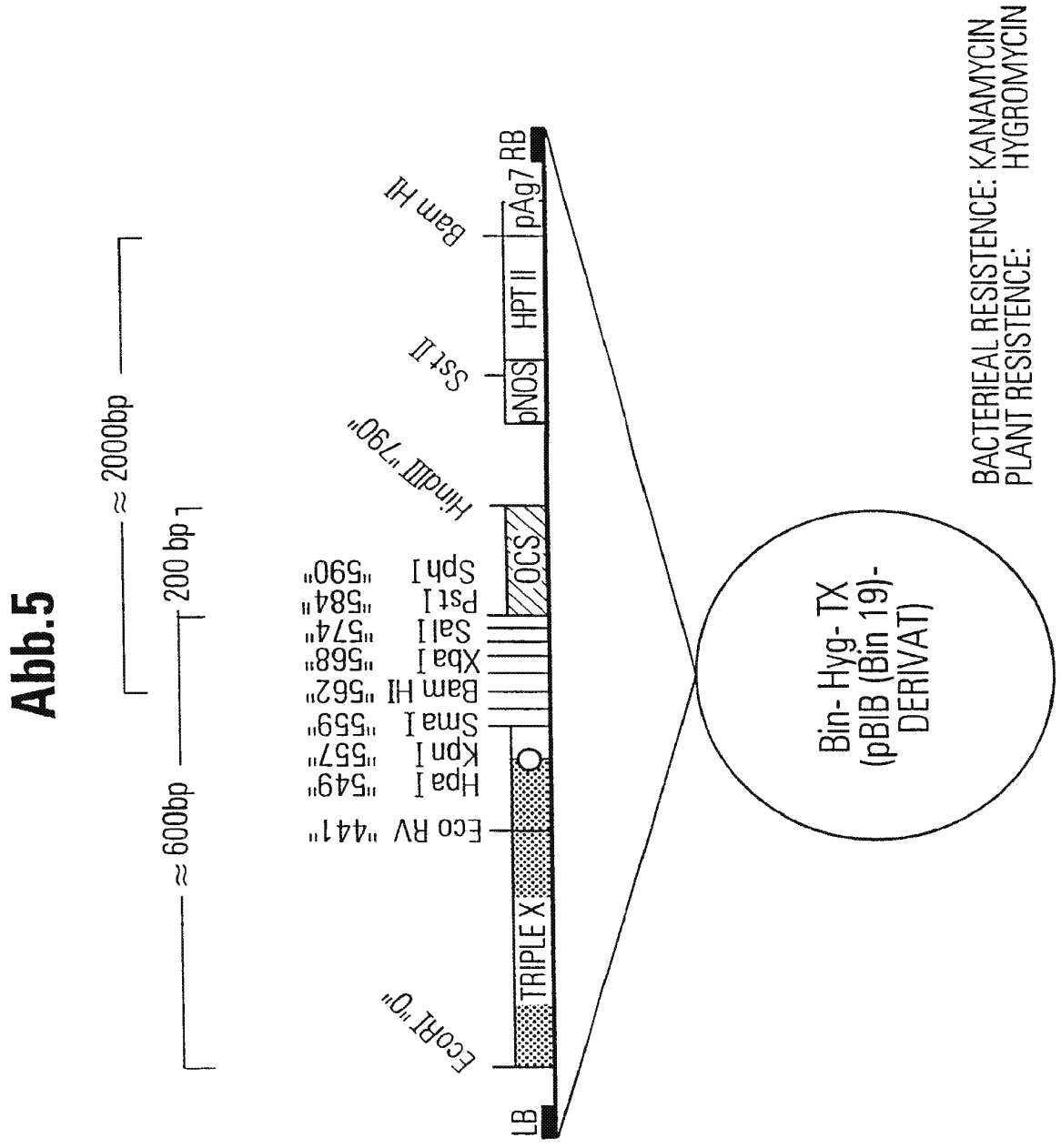


BACTERIAL RESISTANCE: KANAMYCINE
PLANT SELECTION: KANAMYCINE

Abb. 4

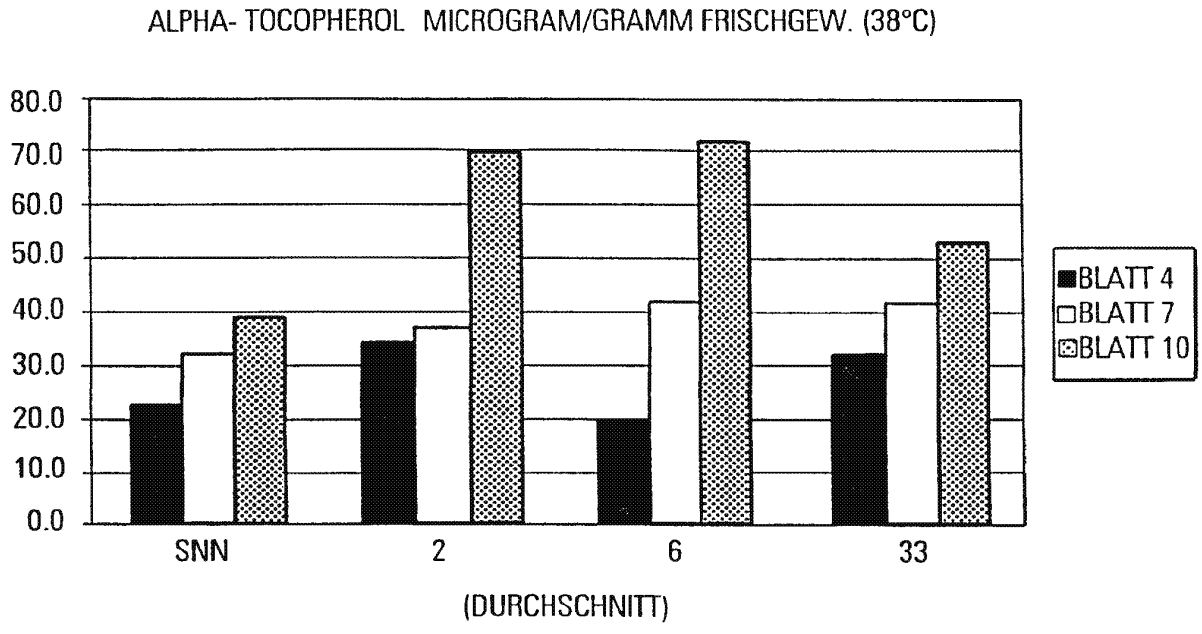


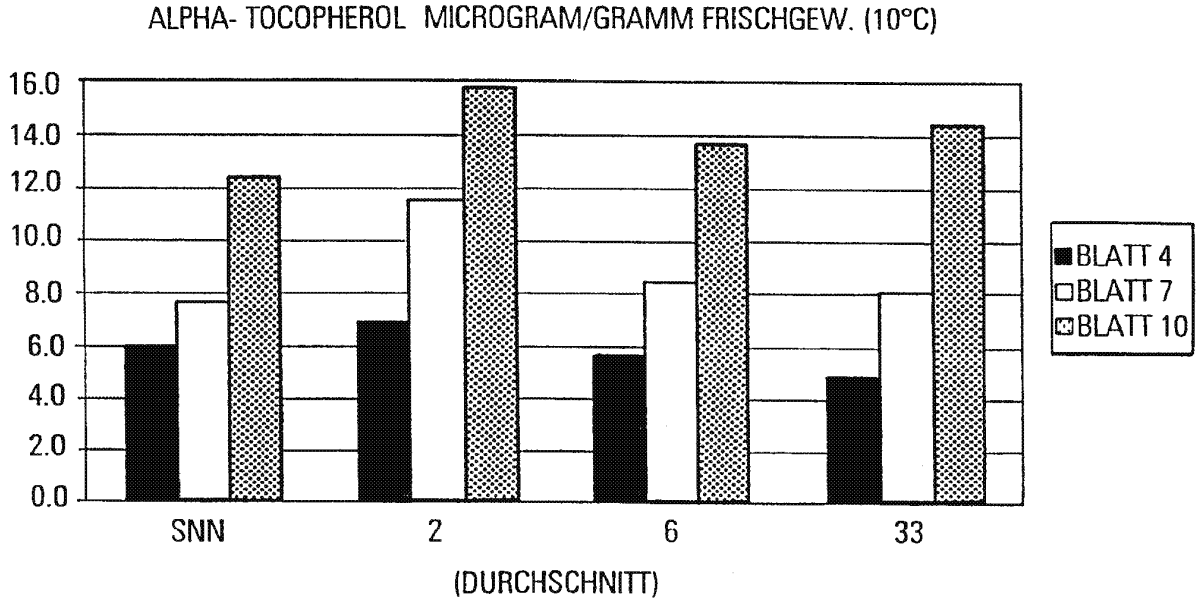
COR-AAD1-00001483



COR-AAD1-00001484

Abb.6





COR-AAD1-00001485

Abb.7

α - TOCOPHEROLGEHALTE NACH HERBIZIDBEHANDLUNG

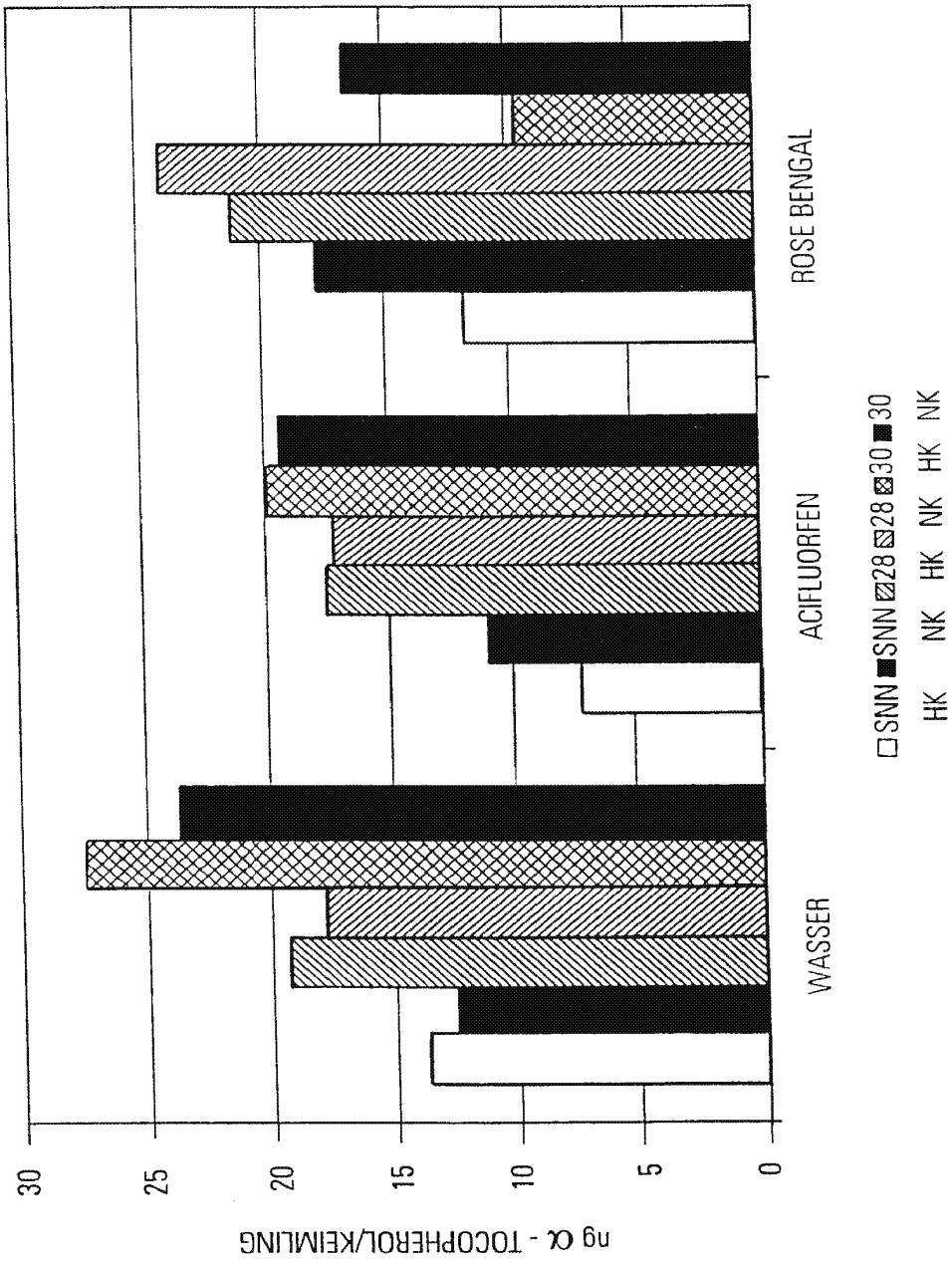
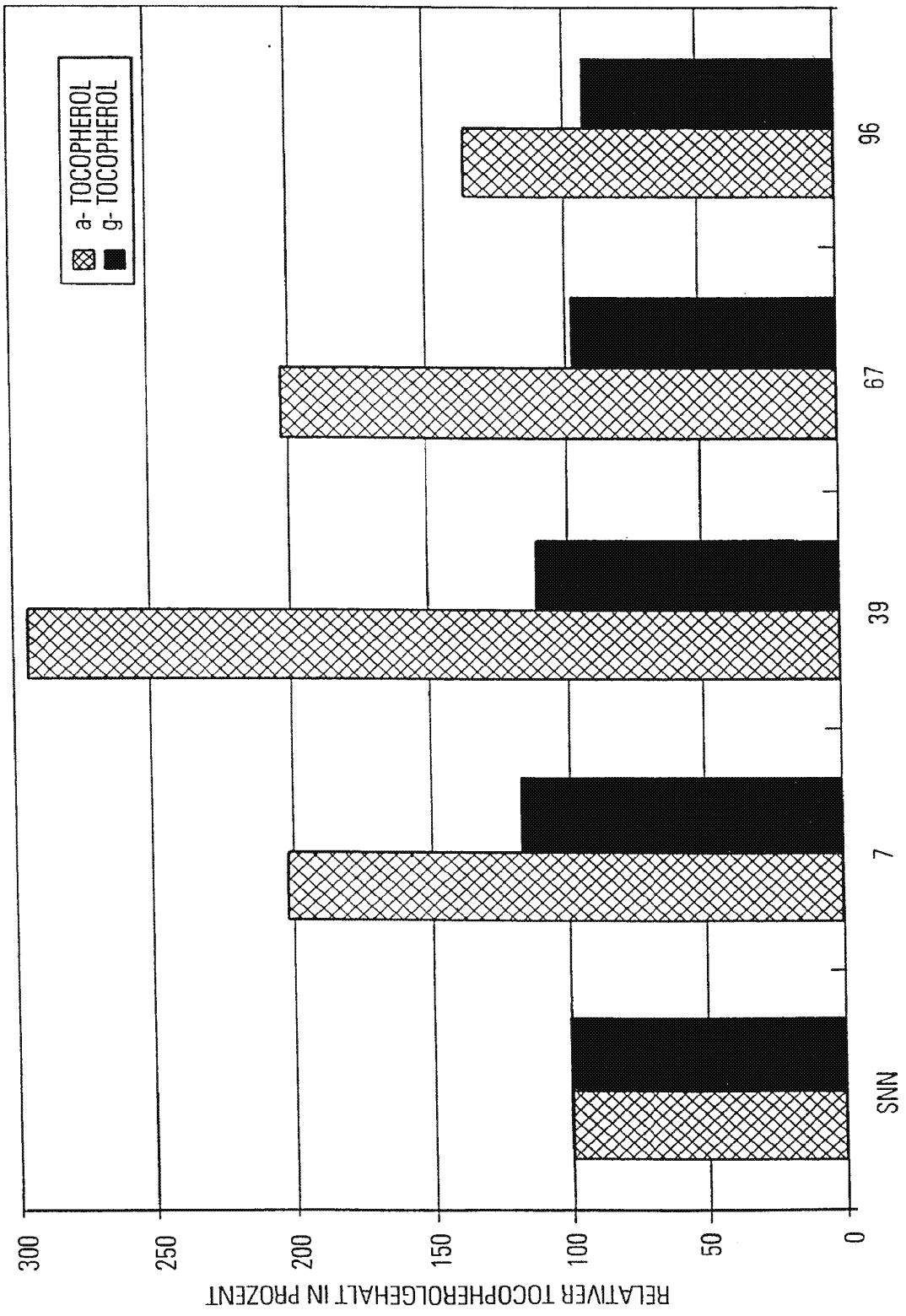


Abb.8

TOCOPHEROLGEHALT IN TABAKSAMEN



COR-AAD1-00001487

SEQUENZPROTOKOLL

(1) ALLGEMEINE ANGABEN:

(i) ANMELDER:
(A) NAME: Institut fuer Pflanzengenetik und Kulturpflanzenforschung
(B) STRASSE: Correnssstrasse 3
(C) ORT: Gatersleben
(E) LAND: Deutschland
(F) POSTLEITZAHL: 06466

(ii) BEZEICHNUNG DER ERFINDUNG: Beeinflussung des Tocopherolgehaltes in transgenen Pflanzen

(iii) ANZAHL DER SEQUENZEN: 3

(iv) COMPUTER-LESBARE FASSUNG:
(A) DATENTRAEGER: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

(2) ANGABEN ZU SEQ ID NO: 1:

(i) SEQUENZKENNZEICHEN:
(A) LAENGE: 1510 Basenpaare
(B) ART: Nucleotid
(C) STRANGFORM: Doppelstrang
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKUELS: cDNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

ATGGCTTCCA TTGCTCTCAA AACTTTCACC GGCCTCCGTC AATCCTCGCC GGAAAACAAT 60

TCCATTACTC TTTCTAAATC CCTCCCCTTC ACCCAAACCC ACCGTAGGCT CCGAATCAAT 120

GCTTCCAAAT CCAGCCCAAG AGTCAACGGC CGCAATCTTC GTGTTGCGGT GGTGGGCGGT 180

GGTCCTGCTG GTGGCGCCGC CGCTGAAACA CTCGCCAAGG GAGGAATTGA AACCTTCTTA 240

ATCGAACGCA AAATGGACAA CTGCAAACCC TGCGGTGGGG CCATCCCACT TTGCATGGTG 300

GGAGAATTTG ACCTCCCTTT GGATATCATT GACCGGAAAG TTACAAAGAT GAAGATGATT 360

TCCCCATCCA ACGTTGCTGT TGATATTGGT CAGACTTTAA AGCCTCACGA GTACATCGGT 420

ATGGTGCGCC GCGAAGTACT CGATGCTTAC CTCCGTGACC GCGCTGCTGA AGCCGGAGCC 480

```
TCTGTTCTCA ACGGCTTGTT CCTCAAAATG GACATGCCCA AAGCTCCCAA CGCACCTTAC      540

GTCCTTCACT ACACAGCTTA CGACTCCAAA ACTAATGGCG CGGGGGAGAA GCGTACCCTG      600

GAAGTTGACG CCGTTATCGG CGCTGACGGT GCAAATTCCC GTGTCGCAAA ATCCATAAAC      660

GCCGGTGACT ACGAGTACGC TATTGCATTC CAAGAAAGGA TTAAAATTTC CGATGATAAA      720

ATGAAGTATT ACGAGAATTT AGCTGAAATG TACGTGGGTG ATGACGTGTC CCCTGATTTT      780

TACGGGTGGG TTTTCCCCAA ATGTGACCAC GTTGCCGTTG GCACTGGCAC AGTCACCCAC      840

AAAGCTGACA TCAAAAAATT CCAGCTAGCT ACAAGATTGA GAGCTGATTC CAAAATCACC      900

GGCGGAAAAA TTATCCGGGT CGAGGCCCAC CCGATTCCAG AACACCCAAG ACCCAGAAGA      960

TTACAAGACA GAGTTGCATT GGTTGGTGAT GCGGCAGGGT ACGTGACCAA ATGTTCGGGC     1020

GAAGGGATTT ACTTCGCGGC AAAGAGTGGA CGTATGTGTG CTGAAGCAAT TGTTGAAGGG     1080

TCAGAAATGG GAAAAAGAAT GGTGGACGAG AGTGATTTGA GGAAGTATTT GGAGAAATGG     1140

GACAAGACTT ATTGGCCAAC GTACAAGGTG CTTGATATAT TGCAGAAGGT ATTTTACAGG     1200

TCGAATCCGG CGAGGGAAGC ATTTGTTGAA ATGTGCGCAG ATGAGTATGT GCAGAAGATG     1260

ACATTTGACA GCTATTTGTA CAAGAAAGTA GCACCAGGAA ACCCAATTGA AGACTTGAAG     1320

CTTGCTGTGA ATACCATTGG AAGTTTGGTG AGAGCTAATG CACTAAGAAG GGAAATGGAC     1380

AAGCTCAGTG TATAAGAAGA TTAACAGCAT TAATATTTTC TTGTAATTGA AGGATTTATT     1440

TCTCAAATTA CTCTGTAAAC ACCTTTCATC CTGCCTTTAA TCGGATTTAT GTAACTTCAT     1500

AATTTGAGCT                                                            1510
```

(2) ANGABEN ZU SEQ ID NO: 2:

(i) SEQUENZKENNZEICHEN:
(A) LAENGE: 1510 Basenpaare
(B) ART: Nucleotid
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKUELS: Protein

(ix) MERKMAL:
(A) NAME/SCHLUESSEL: CDS
(B) LAGE:1..1392

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

```
ATG GCT TCC ATT GCT CTC AAA ACT TTC ACC GGC CTC CGT CAA TCC TCG        48
Met Ala Ser Ile Ala Leu Lys Thr Phe Thr Gly Leu Arg Gln Ser Ser
  1               5                  10                  15

CCG GAA AAC AAT TCC ATT ACT CTT TCT AAA TCC CTC CCC TTC ACC CAA        96
```

```
Pro Glu Asn Asn Ser Ile Thr Leu Ser Lys Ser Leu Pro Phe Thr Gln
             20                  25                  30

ACC CAC CGT AGG CTC CGA ATC AAT GCT TCC AAA TCC AGC CCA AGA GTC      144
Thr His Arg Arg Leu Arg Ile Asn Ala Ser Lys Ser Ser Pro Arg Val
         35                  40                  45

AAC GGC CGC AAT CTT CGT GTT GCG GTG GTG GGC GGT GGT CCT GCT GGT      192
Asn Gly Arg Asn Leu Arg Val Ala Val Val Gly Gly Gly Pro Ala Gly
     50                  55                  60

GGC GCC GCC GCT GAA ACA CTC GCC AAG GGA GGA ATT GAA ACC TTC TTA      240
Gly Ala Ala Ala Glu Thr Leu Ala Lys Gly Gly Ile Glu Thr Phe Leu
 65                  70                  75                  80

ATC GAA CGC AAA ATG GAC AAC TGC AAA CCC TGC GGT GGG GCC ATC CCA      288
Ile Glu Arg Lys Met Asp Asn Cys Lys Pro Cys Gly Gly Ala Ile Pro
                 85                  90                  95

CTT TGC ATG GTG GGA GAA TTT GAC CTC CCT TTG GAT ATC ATT GAC CGG      336
Leu Cys Met Val Gly Glu Phe Asp Leu Pro Leu Asp Ile Ile Asp Arg
            100                 105                 110

AAA GTT ACA AAG ATG AAG ATG ATT TCC CCA TCC AAC GTT GCT GTT GAT      384
Lys Val Thr Lys Met Lys Met Ile Ser Pro Ser Asn Val Ala Val Asp
        115                 120                 125

ATT GGT CAG ACT TTA AAG CCT CAC GAG TAC ATC GGT ATG GTG CGC CGC      432
Ile Gly Gln Thr Leu Lys Pro His Glu Tyr Ile Gly Met Val Arg Arg
    130                 135                 140

GAA GTA CTC GAT GCT TAC CTC CGT GAC CGC GCT GCT GAA GCC GGA GCC      480
Glu Val Leu Asp Ala Tyr Leu Arg Asp Arg Ala Ala Glu Ala Gly Ala
145                 150                 155                 160

TCT GTT CTC AAC GGC TTG TTC CTC AAA ATG GAC ATG CCC AAA GCT CCC      528
Ser Val Leu Asn Gly Leu Phe Leu Lys Met Asp Met Pro Lys Ala Pro
                165                 170                 175

AAC GCA CCT TAC GTC CTT CAC TAC ACA GCT TAC GAC TCC AAA ACT AAT      576
Asn Ala Pro Tyr Val Leu His Tyr Thr Ala Tyr Asp Ser Lys Thr Asn
            180                 185                 190

GGC GCG GGG GAG AAG CGT ACC CTG GAA GTT GAC GCC GTT ATC GGC GCT      624
Gly Ala Gly Glu Lys Arg Thr Leu Glu Val Asp Ala Val Ile Gly Ala
        195                 200                 205

GAC GGT GCA AAT TCC CGT GTC GCA AAA TCC ATA AAC GCC GGT GAC TAC      672
Asp Gly Ala Asn Ser Arg Val Ala Lys Ser Ile Asn Ala Gly Asp Tyr
    210                 215                 220

GAG TAC GCT ATT GCA TTC CAA GAA AGG ATT AAA ATT TCC GAT GAT AAA      720
Glu Tyr Ala Ile Ala Phe Gln Glu Arg Ile Lys Ile Ser Asp Asp Lys
225                 230                 235                 240

ATG AAG TAT TAC GAG AAT TTA GCT GAA ATG TAC GTG GGT GAT GAC GTG      768
Met Lys Tyr Tyr Glu Asn Leu Ala Glu Met Tyr Val Gly Asp Asp Val
                245                 250                 255
```

```
TCC CCT GAT TTT TAC GGG TGG GTT TTC CCC AAA TGT GAC CAC GTT GCC        816
Ser Pro Asp Phe Tyr Gly Trp Val Phe Pro Lys Cys Asp His Val Ala
            260                 265                 270

GTT GGC ACT GGC ACA GTC ACC CAC AAA GCT GAC ATC AAA AAA TTC CAG        864
Val Gly Thr Gly Thr Val Thr His Lys Ala Asp Ile Lys Lys Phe Gln
        275                 280                 285

CTA GCT ACA AGA TTG AGA GCT GAT TCC AAA ATC ACC GGC GGA AAA ATT        912
Leu Ala Thr Arg Leu Arg Ala Asp Ser Lys Ile Thr Gly Gly Lys Ile
    290                 295                 300

ATC CGG GTC GAG GCC CAC CCG ATT CCA GAA CAC CCA AGA CCC AGA AGA        960
Ile Arg Val Glu Ala His Pro Ile Pro Glu His Pro Arg Pro Arg Arg
305                 310                 315                 320

TTA CAA GAC AGA GTT GCA TTG GTT GGT GAT GCG GCA GGG TAC GTG ACC       1008
Leu Gln Asp Arg Val Ala Leu Val Gly Asp Ala Ala Gly Tyr Val Thr
                325                 330                 335

AAA TGT TCG GGC GAA GGG ATT TAC TTC GCG GCA AAG AGT GGA CGT ATG       1056
Lys Cys Ser Gly Glu Gly Ile Tyr Phe Ala Ala Lys Ser Gly Arg Met
            340                 345                 350

TGT GCT GAA GCA ATT GTT GAA GGG TCA GAA ATG GGA AAA AGA ATG GTG       1104
Cys Ala Glu Ala Ile Val Glu Gly Ser Glu Met Gly Lys Arg Met Val
        355                 360                 365

GAC GAG AGT GAT TTG AGG AAG TAT TTG GAG AAA TGG GAC AAG ACT TAT       1152
Asp Glu Ser Asp Leu Arg Lys Tyr Leu Glu Lys Trp Asp Lys Thr Tyr
    370                 375                 380

TGG CCA ACG TAC AAG GTG CTT GAT ATA TTG CAG AAG GTA TTT TAC AGG       1200
Trp Pro Thr Tyr Lys Val Leu Asp Ile Leu Gln Lys Val Phe Tyr Arg
385                 390                 395                 400

TCG AAT CCG GCG AGG GAA GCA TTT GTT GAA ATG TGC GCA GAT GAG TAT       1248
Ser Asn Pro Ala Arg Glu Ala Phe Val Glu Met Cys Ala Asp Glu Tyr
                405                 410                 415

GTG CAG AAG ATG ACA TTT GAC AGC TAT TTG TAC AAG AAA GTA GCA CCA       1296
Val Gln Lys Met Thr Phe Asp Ser Tyr Leu Tyr Lys Lys Val Ala Pro
            420                 425                 430

GGA AAC CCA ATT GAA GAC TTG AAG CTT GCT GTG AAT ACC ATT GGA AGT       1344
Gly Asn Pro Ile Glu Asp Leu Lys Leu Ala Val Asn Thr Ile Gly Ser
        435                 440                 445

TTG GTG AGA GCT AAT GCA CTA AGA AGG GAA ATG GAC AAG CTC AGT GTA       1392
Leu Val Arg Ala Asn Ala Leu Arg Arg Glu Met Asp Lys Leu Ser Val
    450                 455                 460

TAAGAAGATT AACAGCATTA ATATTTTCTT GTAATTGAAG GATTTATTTC TCAAATTACT     1452

CTGTAAACAC CTTTCATCCT GCCTTTAATC GGATTTATGT AACTTCATAA TTTGAGCT       1510
```

(2) ANGABEN ZU SEQ ID NO: 3:

(i) SEQUENZKENNZEICHEN:
(A) LAENGE: 464 Aminosaeuren
(B) ART: Aminosaeure
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKUELS: Protein
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

```
Met Ala Ser Ile Ala Leu Lys Thr Phe Thr Gly Leu Arg Gln Ser Ser
 1               5                  10                  15

Pro Glu Asn Asn Ser Ile Thr Leu Ser Lys Ser Leu Pro Phe Thr Gln
            20                  25                  30

Thr His Arg Arg Leu Arg Ile Asn Ala Ser Lys Ser Ser Pro Arg Val
        35                  40                  45

Asn Gly Arg Asn Leu Arg Val Ala Val Val Gly Gly Gly Pro Ala Gly
    50                  55                  60

Gly Ala Ala Ala Glu Thr Leu Ala Lys Gly Gly Ile Glu Thr Phe Leu
65                  70                  75                  80

Ile Glu Arg Lys Met Asp Asn Cys Lys Pro Cys Gly Gly Ala Ile Pro
                85                  90                  95

Leu Cys Met Val Gly Glu Phe Asp Leu Pro Leu Asp Ile Ile Asp Arg
            100                 105                 110

Lys Val Thr Lys Met Lys Met Ile Ser Pro Ser Asn Val Ala Val Asp
        115                 120                 125

Ile Gly Gln Thr Leu Lys Pro His Glu Tyr Ile Gly Met Val Arg Arg
    130                 135                 140

Glu Val Leu Asp Ala Tyr Leu Arg Asp Arg Ala Ala Glu Ala Gly Ala
145                 150                 155                 160

Ser Val Leu Asn Gly Leu Phe Leu Lys Met Asp Met Pro Lys Ala Pro
                165                 170                 175

Asn Ala Pro Tyr Val Leu His Tyr Thr Ala Tyr Asp Ser Lys Thr Asn
            180                 185                 190

Gly Ala Gly Glu Lys Arg Thr Leu Glu Val Asp Ala Val Ile Gly Ala
        195                 200                 205

Asp Gly Ala Asn Ser Arg Val Ala Lys Ser Ile Asn Ala Gly Asp Tyr
    210                 215                 220

Glu Tyr Ala Ile Ala Phe Gln Glu Arg Ile Lys Ile Ser Asp Asp Lys
225                 230                 235                 240

Met Lys Tyr Tyr Glu Asn Leu Ala Glu Met Tyr Val Gly Asp Asp Val
                245                 250                 255

Ser Pro Asp Phe Tyr Gly Trp Val Phe Pro Lys Cys Asp His Val Ala
```

```
            260                 265                 270

Val Gly Thr Gly Thr Val Thr His Lys Ala Asp Ile Lys Lys Phe Gln
        275                 280                 285

Leu Ala Thr Arg Leu Arg Ala Asp Ser Lys Ile Thr Gly Gly Lys Ile
    290                 295                 300

Ile Arg Val Glu Ala His Pro Ile Pro Glu His Pro Arg Pro Arg Arg
305                 310                 315                 320

Leu Gln Asp Arg Val Ala Leu Val Gly Asp Ala Ala Gly Tyr Val Thr
                325                 330                 335

Lys Cys Ser Gly Glu Gly Ile Tyr Phe Ala Ala Lys Ser Gly Arg Met
            340                 345                 350

Cys Ala Glu Ala Ile Val Glu Gly Ser Glu Met Gly Lys Arg Met Val
        355                 360                 365

Asp Glu Ser Asp Leu Arg Lys Tyr Leu Glu Lys Trp Asp Lys Thr Tyr
    370                 375                 380

Trp Pro Thr Tyr Lys Val Leu Asp Ile Leu Gln Lys Val Phe Tyr Arg
385                 390                 395                 400

Ser Asn Pro Ala Arg Glu Ala Phe Val Glu Met Cys Ala Asp Glu Tyr
                405                 410                 415

Val Gln Lys Met Thr Phe Asp Ser Tyr Leu Tyr Lys Lys Val Ala Pro
            420                 425                 430

Gly Asn Pro Ile Glu Asp Leu Lys Leu Ala Val Asn Thr Ile Gly Ser
        435                 440                 445

Leu Val Arg Ala Asn Ala Leu Arg Arg Glu Met Asp Lys Leu Ser Val
    450                 455                 460
```

US006624342B1

(12) **United States Patent**
Grimm et al.

(10) **Patent No.: US 6,624,342 B1**
(45) **Date of Patent: Sep. 23, 2003**

(54) **MANIPULATION OF TOCOPHEROL CONTENT IN TRANSGENIC PLANTS**

(75) Inventors: **Bernhard Grimm**, Gatersleben (DE); **Ryouichi Tanaka**, Sapporo (JP)

(73) Assignee: **Institut fur Pflanzengenetik und Kulturpflanzenforschung** (DE)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/564,367**

(22) Filed: **Apr. 28, 2000**

Related U.S. Application Data

(63) Continuation of application No. PCT/EP98/06851, filed on Oct. 29, 1998.

(30) **Foreign Application Priority Data**

Oct. 29, 1997 (DE) 197 47 739

(51) **Int. Cl.**[7] **C12N 15/82**; C12N 5/04; A01H 5/00

(52) **U.S. Cl.** **800/278**; 800/298; 800/306; 800/312; 800/317.2; 800/317.4; 800/320; 800/320.1; 800/320.2; 800/320.3; 435/419

(58) **Field of Search** 435/419, 468, 435/320.1; 800/278, 281, 282, 298, 306, 12, 317.2, 317.4, 320, 320.1, 320.2, 320.3

(56) **References Cited**

PUBLICATIONS

Shewmaker et al, "Seed–specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects", 1999, The Plant Journal vol. 20 No. 4, pp. 401–412.*

Burkhardt et al, Transgenic rice (Oryza sativa) endosperm expressing daffodil (Narcissus pseudonarcissus) phytoene synthase . . . A bioysynthese, 1997, The Plant Journal vol. 11 No. 5, pp. 1071–1078.*

* cited by examiner

Primary Examiner—Elizabeth F. McElwain
Assistant Examiner—Cynthia Collins
(74) *Attorney, Agent, or Firm*—Schwegman, Lundberg, Woessner & Kluth, P.A.

(57) **ABSTRACT**

The present invention relates to new nucleic acid sequences, which code for a geranylgeranyl reductase, a method for producing new plants which contain the new DNA sequence and the content of tocopherol and/or chlorophyll of which is altered in comparison to wild-type plants, these new plants, parts and products thereof and plant cells as well as the use of the DNA sequences for the manipulation of tocopherol, chlorophyll and/or vitamin K_1 content in transgenic plants, parts and products thereof and plant cells.

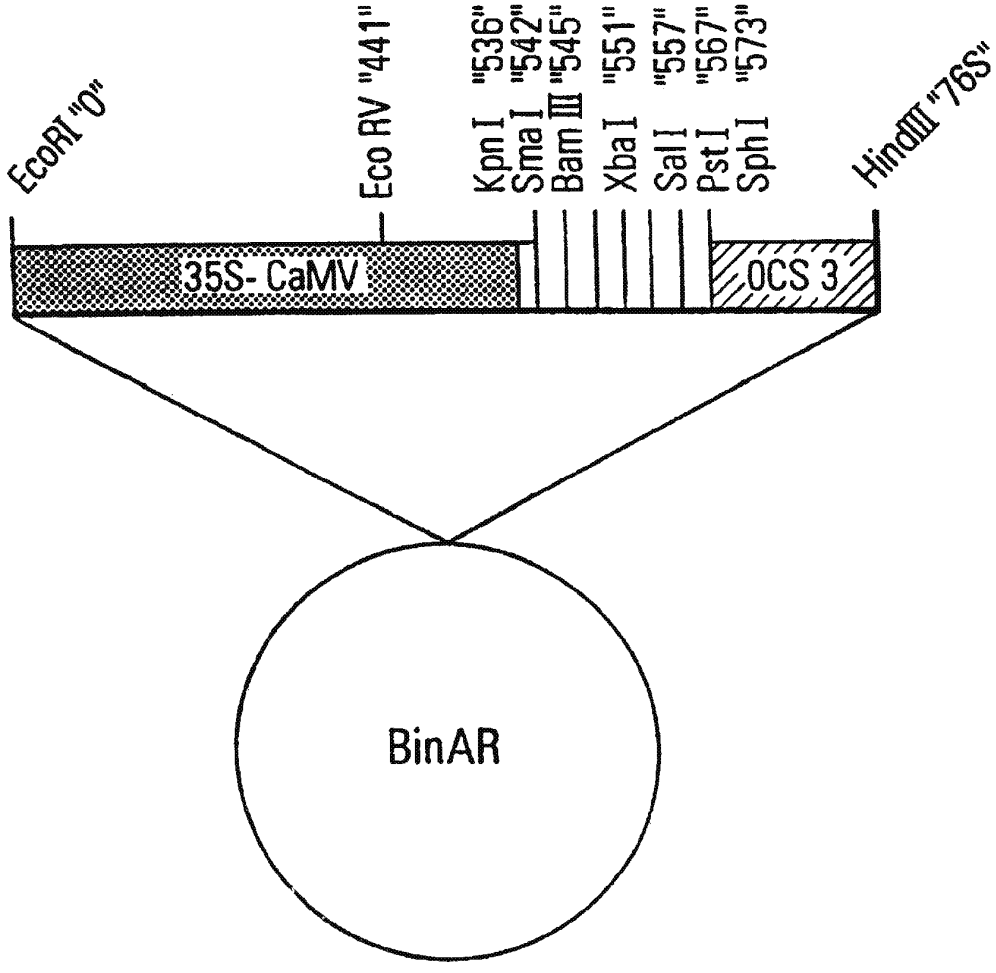
31 Claims, 8 Drawing Sheets

ATGGCTTCCA TTGCTCTCAA AACTTTCACC GGCCTCCGTC AATCCTCGCC GGAAAACAAT
TCCATTACTC TTTCTAAATC CCTCCCCTTC ACCCAAACCC ACCGTAGGCT CCGAATCAAT
GCTTCCAAAT CCAGCCCAAG AGTCAACGGC CGCAATCTTC GTGTTGCGGT GGTGGGCGGT
GGTCCTGCTG GTGGCGCCGC CGCTGAAACA CTCCCCAAGG GAGGAATTGA AACCTTCTTA
ATCGAACGCA AAATGGACAA CTGCAAACCC TGCGGTGGGG CCATCCCACT TTGCATGGTG
GGAGAATTTG ACCTCCCTTT GGATATCATT GACCGGAAAG TTACAAAGAT GAAGATGATT
TCCCCATCCA ACGTTGCTGT TGATATTGGT CAGACTTTAA AGCCTCACGA GTACATCGGT
ATGGTGCGCC GCGAAGTACT CGATGCTTAC CTCCGTGACC GCGCTGCTGA AGCCGGAGCC
TCTGTTCTCA ACGGCTTGTT CCTCAAAATG GACATGCCCA AAGCTCCCAA CGCACCTTAC
GTCCTTCACT ACACAGCTTA CGACTCCAAA ACTAATGGCG CGGGGGAGAA GCGTACCCTG
GAAGTTGACG CCGTTATCGG CGCTGACGGT GCAAATTCCC GTGTCGCAAA ATCCATAAAC
GCCGGTGACT ACGAGTACGC TATTGCATTC CAAGAAAGGA TTAAAATTTC CGATGATAAA
ATGAAGTATT ACGAGAATTT AGCTGAAATG TACGTGGGTG ATGACGTGTC CCCTGATTTT
TACGGGTGGG TTTTCCCCAA ATGTGACCAC GTTGCCGTTG GCACTGGCAC AGTCACCCAC
AAAGCTGACA TCAAAAAATT CCAGCTAGCT ACAAGATTGA GAGCTGATTC CAAAATCACC
GGCGGAAAAA TTATCCGGGT CGAGGCCCAC CCGATTCCAG AACACCCAAG ACCCAGAAGA
TTACAAGACA GAGTTGCATT GGTTGGTGAT GCGGCAGGGT ACGTGACCAA ATGTTCGGGC
GAAGGGATTT ACTTCGCGGC AAAGAGTGGA CGTATGTGTG CTGAAGCAAT TGTTGAAGGG
TCAGAAATGG GAAAAAGAAT GGTGGACGAG AGTGATTTGA GGAAGTATTT GGAGAAATGG
GACAAGACTT ATTGGCCAAC GTACAAGGTG CITCATATAT TGCAGAAGGT ATTTTACAGG
TCGAATCCGG CGAGGGAAGC ATTTGTTGAA ATGTGCGCAG ATGAGTATGT GCAGAAGATG
ACATTTGACA GCTATTTGTA CAAGAAAGTA GCACCAGGAA ACCCAATTGA AGACTTGAAG
CTTGCTGTGA ATACCATTGG AAGTTTGGTG AGAGCTAATG CACTAAGAAG GGAAATGGAC
AAGCTCAGTG TATAAGAAGA TTAACAGCAT TAATATTTTC TTGTAATTGA AGGATTTATT
TCTCAAATTA CTCTGTAAAC ACCTTTCATC CTGCCTTTAA TCGGATTTAT GTAACTTCAT
AATTTGAGCT

Fig.1

```
  1 Met Ala Ser Ile Ala Leu Lys Thr Phe Thr Gly Leu Arg Gln Ser
 16 Ser Pro Glu Asn Asn Ser Ile Thr Leu Ser Lys Ser Leu Pro Phe
 31 Thr Gln Thr His Arg Arg Leu Arg Ile Asn Ala Ser Lys Ser Ser
 46 Pro Arg Val Asn Gly Arg Asn Leu Arg Val Ala Val Val Gly Gly
 61 Gly Pro Ala Gly Gly Ala Ala Ala Glu Thr Leu Ala Lys Gly Gly
 76 Ile Glu Thr Phe Leu Ile Glu Arg Lys Met Asp Asn Cys Lys Pro
 91 Cys Gly Gly Ala Ile Pro Leu Cys Met Vat Gly Glu Phe Asp Leu
106 Pro Leu Asp Ile Ile Asp Arg Lys Val Thr Lys Met Lys Met Ile
121 Ser Pro Ser Asn Val Ala Vat Asp Ile Gly Gln Thr Leu Lys Pro
136 His Glu Tyr Ile Gly Met Val Arg Arg Glu Val Leu Asp Ala Tyr
151 Leu Arg Asp Arg Ala Ala Glu Ala Gly Ala Ser Val Leu Asn Gly
166 Leu Phe Leu Lys Met Asp Met Pro Lys Ala Pro Asn Ala Pro Tyr
181 Val Leu His Tyr Thr Ala Tyr Asp Ser Lys Thr Asn Gly Ala Gly
196 Glu Lys Arg Thr Leu Glu Val Asp Ala Val Ile Gly Ala Asp Gly
211 Ala Asn Ser Arg Val Ala Lys Ser Ile Asn Ala Gly Asp Tyr Glu
226 Tyr Ala Ile Ala Phe Gln Glu Arg Ile Lys Ile Ser Asp Asp Lys
241 Met Lys Tyr Tyr Glu Asn Leu Ala Glu Met Tyr Val Gly Asp Asp
256 Val Ser Pro Asp Phe Tyr Gly Trp Val Phe Pro Lys Cys Asp His
271 Val Ala Val Gly Thr Gly Thr Val THr His Lys Ala Asp Ile Lys
286 Lys Phe Gln Leu Ala Thr Arg Leu Arg Ala Asp Ser Lys Ile Thr
301 Gly Gly Lys Ile Ile Arg Val Glu Ala His Pro Ile Pro Glu His
316 Pro Arg Pro Arg Arg Leu Gln ASp Arg Val Ala Leu Val Gly Asp
331 Ala Ala Gly Tyr Val Thr Lys Cys Ser Gly Glu Gly Ile Tyr Phe
346 Ala Ala Lys Ser Gly Arg Met Cys Ala Glu Ala Ile Val Glu Gly
361 Ser Glu Met Gly Lys Arg Met Val Asp Glu Ser Asp Leu Arg Lys
376 Tyr Leu Glu Lys Trp Asp Lys Thr Tyr Trp Pro Thr Tyr Lys Val
391 Leu Asp Ile Leu Gln Lys Val Phe Tyr Arg Ser Asn Pro Ala Arg
406 Glu Ala Phe Val Glu Met Cys Ala ASp Glu Tyr Val Gln Lys Met
421 Thr Phe Asp Ser Tyr Leu Tyr Lys Lys Val Ala Pro Gly Asn Pro
436 Ile Glu Asp Leu Lys Leu Ala Val Asn Thr Ile Gly Ser Leu Val
451 Arg Ala Asn Ala Leu Arg Arg Glu Met Asp Lys Leu Ser Val
```

Fig.2



BACTERIAL RESISTANCE: KANAMYCINE
PLANT SELECTION: KANAMYCINE

Fig.3

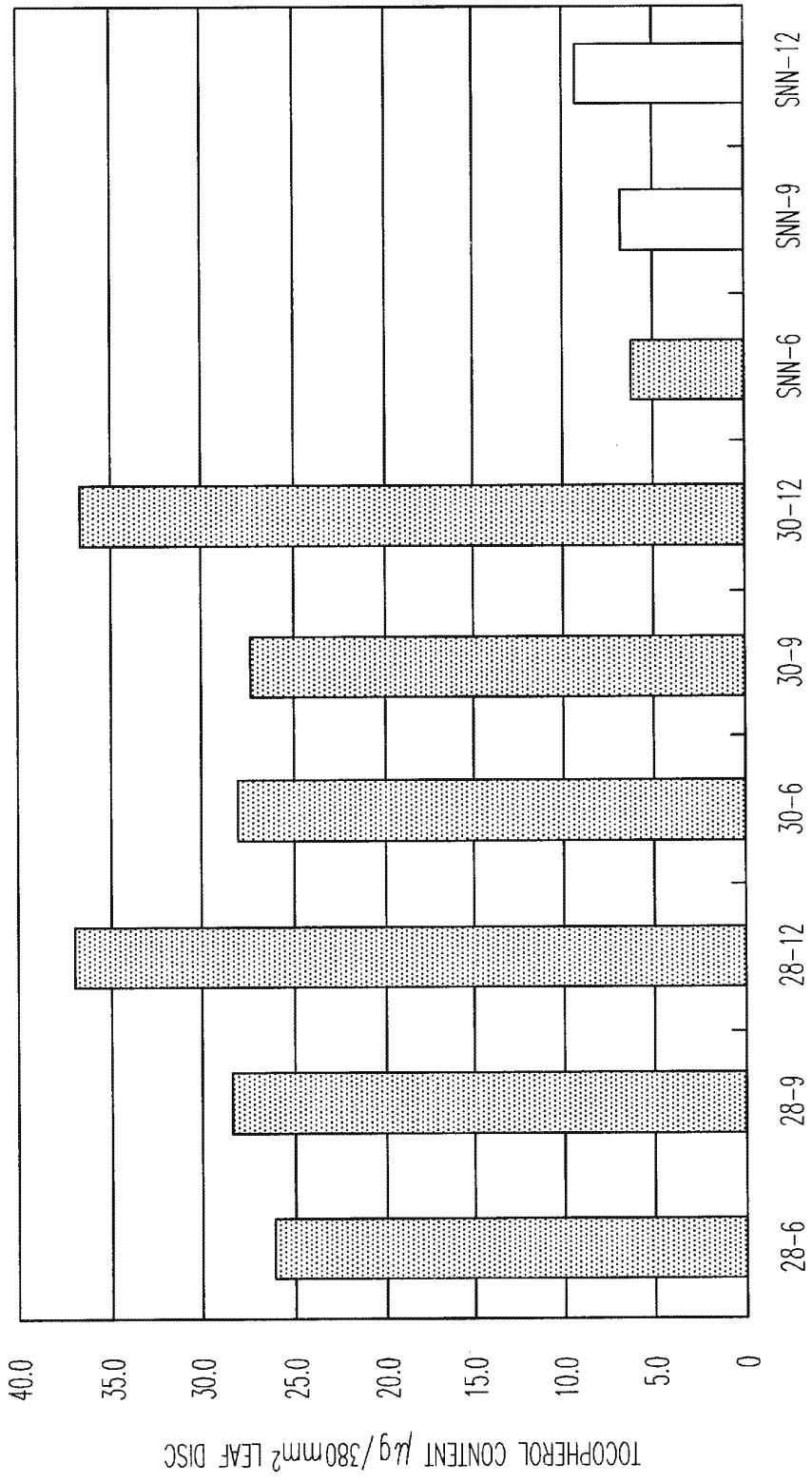

TOCOPHEROL CONTENT μg/380mm² LEAF DISC
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0
28-6
28-9
28-12
30-6
30-9
30-12
SNN-6
SNN-9
SNN-12
TRANSGENIC LINE-LEAF NUMBER

Fig. 4

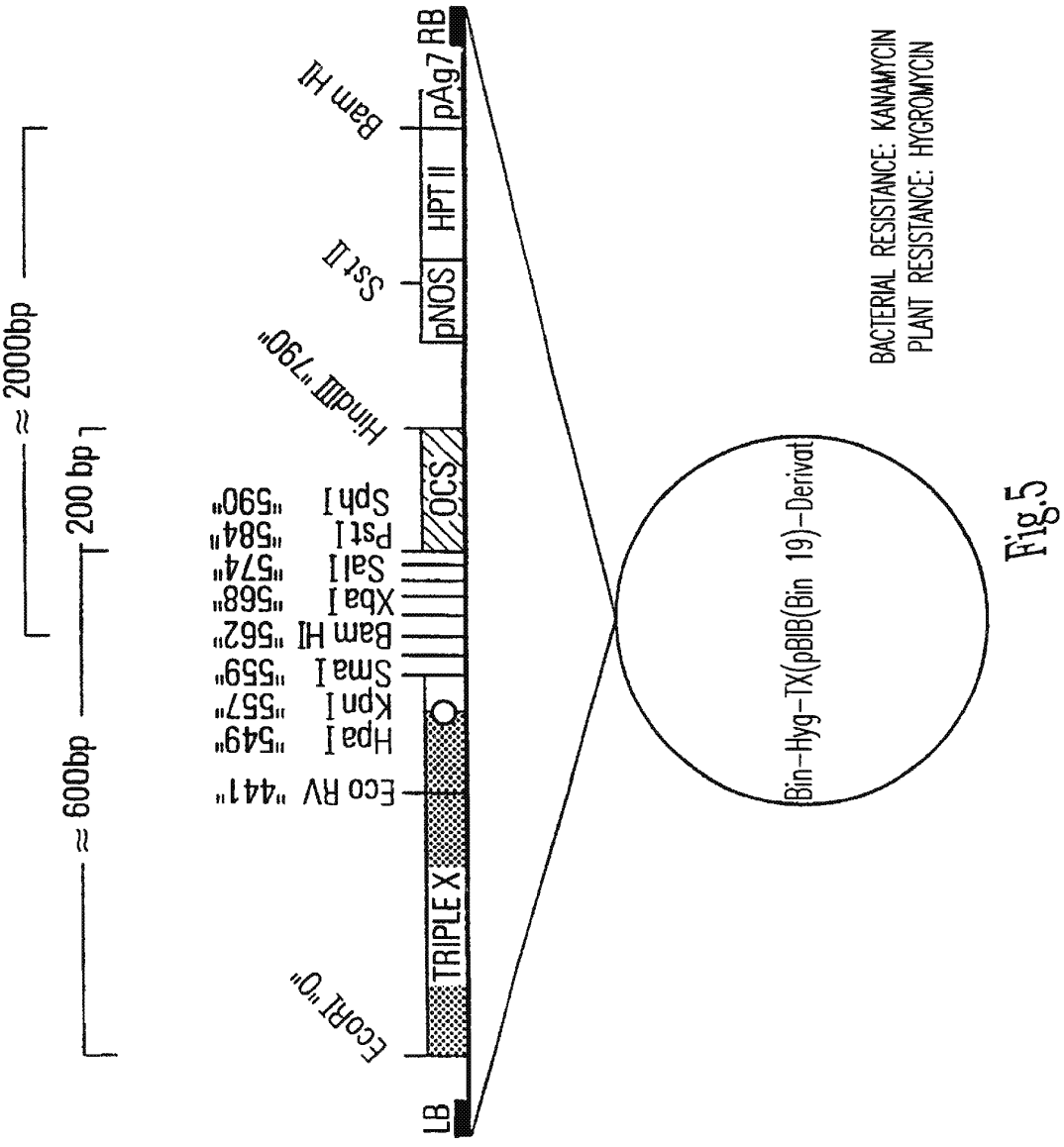
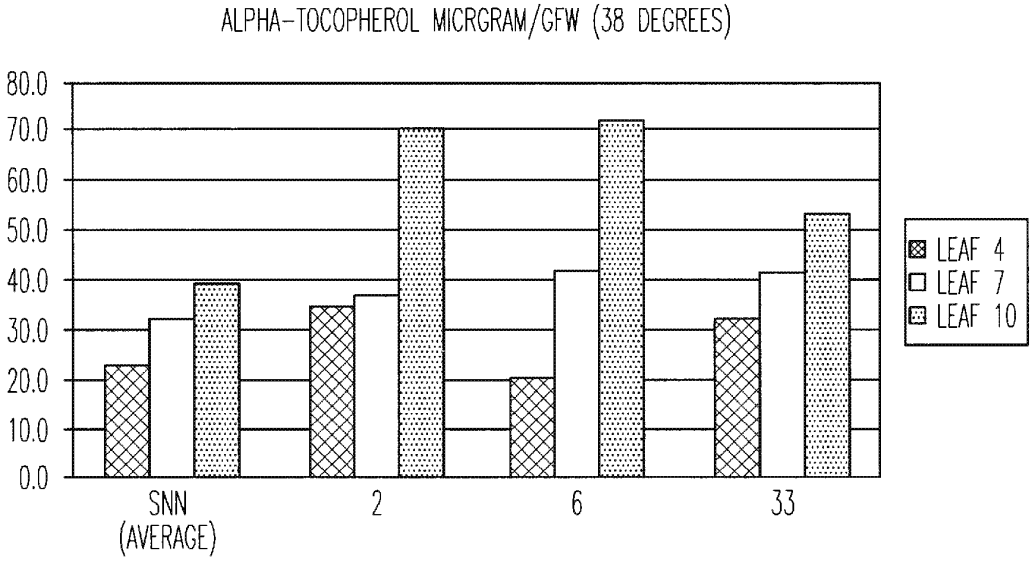

≈ 2000bp
200 bp
≈ 600bp
LB
EcoRI "0"
TRIPLE X
Eco RV "441"
Hpa I "549"
Kpn I "557"
Sma I "559"
Bam HI "562"
Xba I "568"
Sal I "574"
Pst I "584"
Sph I "590"
HindIII "790"
OCS'
SstII
pNOS
HPT II
BamHI
pAg7
RB
Bin-Hyg-TX(pBIB(Bin 19)-Derivat
BACTERIAL RESISTANCE: KANAMYCIN
PLANT RESISTANCE: HYGROMYCIN

Fig.5


ALPHA-TOCOPHEROL MICRGRAM/GFW (38 DEGREES)
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
SNN (AVERAGE)
2
6
33
LEAF 4
LEAF 7
LEAF 10

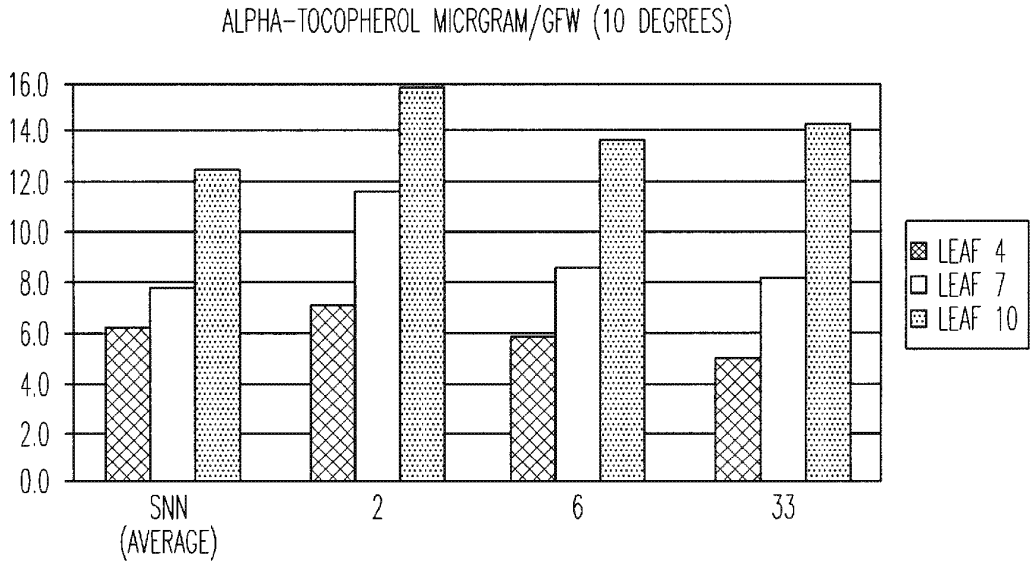

ALPHA-TOCOPHEROL MICRGRAM/GFW (10 DEGREES)
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
SNN (AVERAGE)
2
6
33
LEAF 4
LEAF 7
LEAF 10

Fig.6

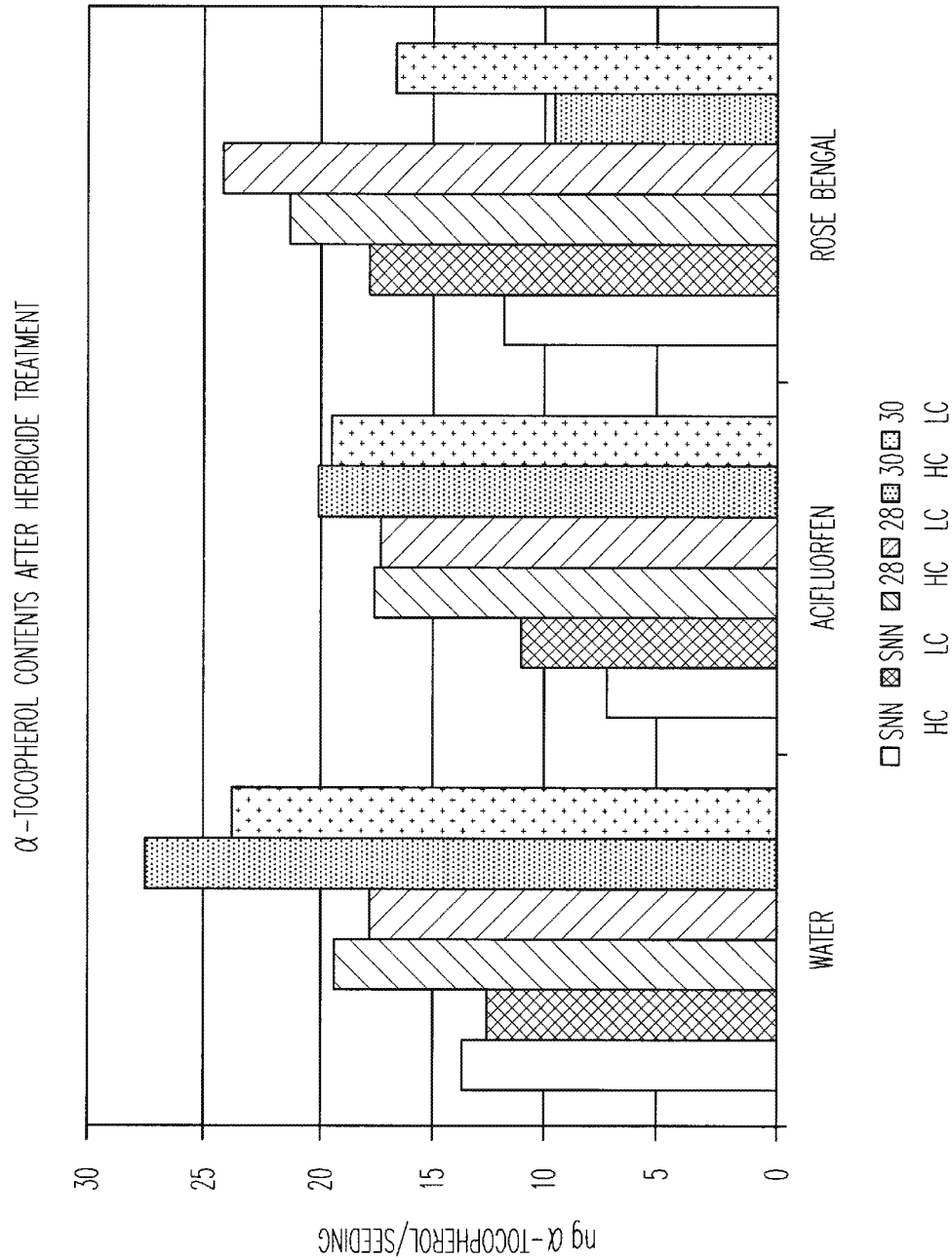

α-TOCOPHEROL CONTENTS AFTER HERBICIDE TREATMENT
ng α-TOCOPHEROL/SEEDING
0
5
10
15
20
25
30
WATER
ACIFLUORFEN
ROSE BENGAL
SNN 28 30 30
HC LC HC LC HC LC

Fig.7

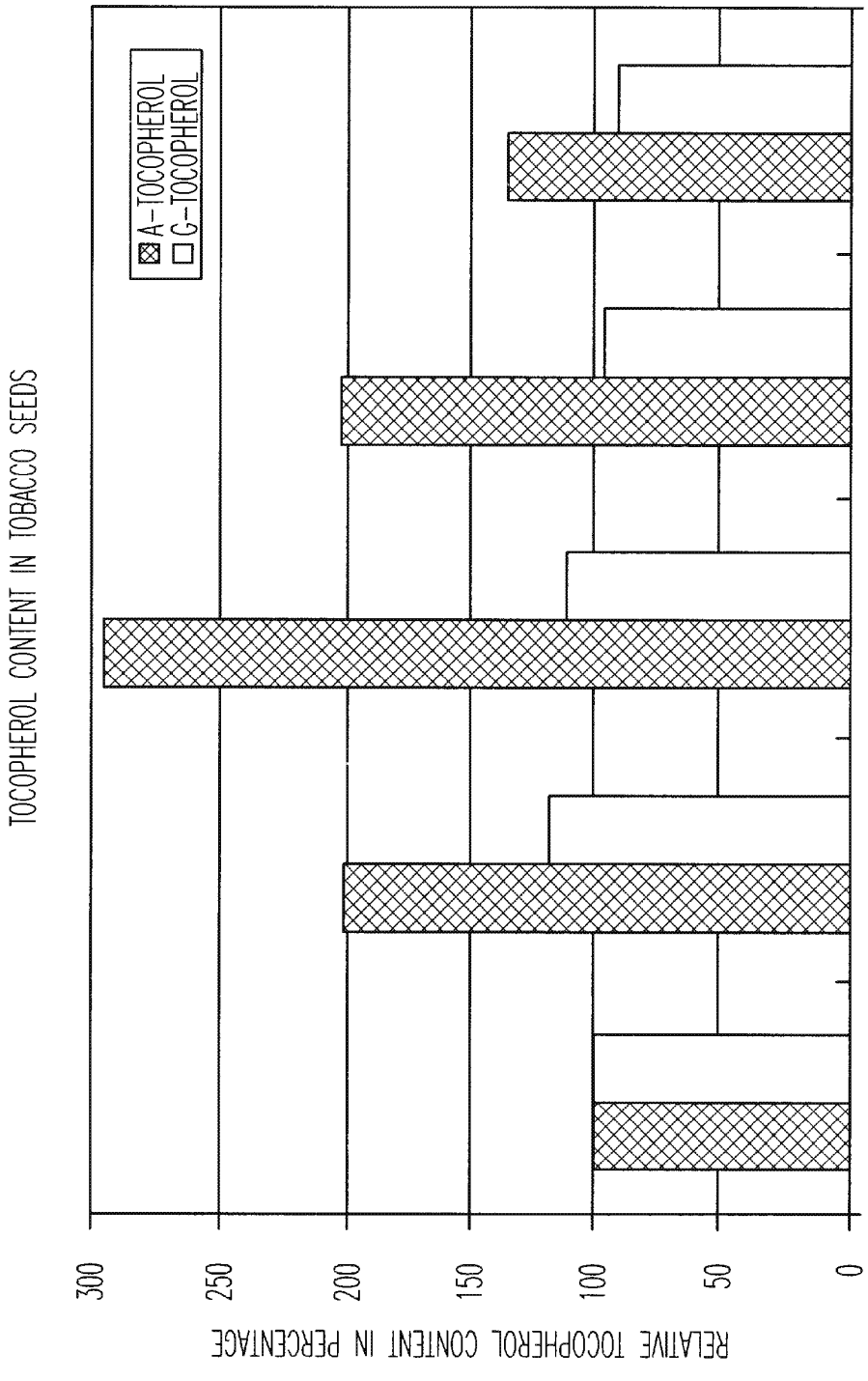

TOCOPHEROL CONTENT IN TOBACCO SEEDS
A-TOCOPHEROL
G-TOCOPHEROL
RELATIVE TOCOPHEROL CONTENT IN PERCENTAGE
300
250
200
150
100
50
0
SNN
7
39
67
96

Fig.8

MANIPULATION OF TOCOPHEROL CONTENT IN TRANSGENIC PLANTS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation of International Patent Application No. PCT/EP98/06851, filed on Oct. 29, 1998, which in turn is an international filing of German Patent Application No. 19747739.9 filed on Oct. 29, 1997, all of which are incorporated herein by reference.

The present invention relates to novel nucleic acid sequences, which code for a geranylgeranyl reductase, a method for producing novel plants, which contain a novel nucleic acid sequence and the tocopherol and/or chlorophyll content of which is altered in comparison with wild-type plants, these novel plants, parts and products thereof and plant cells as well as the use of the nucleic acid sequences for the manipulation of tocopherol, chlorophyll and/or vitamin K_1 content in transgenic plants, parts and products thereof and plant cells.

The diterpene geranylgeranyl pyrophosphate (GGPP) is formed as C_{20}-intermediate in the plant isoprenoid metabolism. It results from the addition of one unit isopentenyl pyrophosphate (IPP) to farnesyl pyrophosphate, a C_{15}-sesquiterpene. GGPP enters several synthesis pathways of plant secondary metabolism. For example, two molecules of GGPP can be joined "tail to tail" to give C_{40}-bodies, the tetraterpenes, generally called carotenoides and to which, for example, the β-carotene belongs. By the addition of further molecules of IPP, GGPP furthermore enters the biosynthesis of polyterpenes, such as rubber and guttapercha.

Further, GGPP can be transformed into other diterpenes, such as phytyl pyrophosphate (PPP). The C_{20}-body phytol is an obligatory intermediate in the biosynthesis of the tocopherols (Soll and Schulz (1981) Biochem. Biophys. Res. Commun. 99, 907–912) as well as the synthesis of the chlorophylls (Beale and Weinstein (1990) in: Biosynthesis of Heme and Chlorophyll (Daily H. A., ed.) McGraw Hill, N.Y., 287–391). While the basic structure of all chlorophylls (chlorophyll a, b, c, etc.) is a porphyrin system consisting of four pyrrole rings, to which system the phytol is bound by an ester-like bond through the pyrrole ring IV, the tocopherols are characterised by a structure consisting of homogentisate and a phytol tail.

The group of tocopherols, commonly described as vitamin E, comprises several structurally closely related lipophilic vitamins, viz α-, β-, γ-, δ- und ε-tocopherol, α-tocopherol being the most important in biological terms. The tocopherols can be found in many plant oils, specially rich in tocopherols are the seed oils of soybean, wheat, maize, rice, cotton, lucerne and nuts. Also fruits and vegetables, e.g. raspberries, beans, peas, fennel, pepper etc. contain tocophrols. As far as presently known, tocopherols are exclusively synthesized in plants and photosynthetically active organisms.

Due to their redox potential tocopherols contribute to avoid oxidation of unsaturated fatty acids by air oxygen; α-tocopherol is the most important lipophilic antioxidant in human. It is assumed that due to their function as antioxidative agents the tocopherols contribute to the stabilisation of biological membranes, because the fluidity of the membranes is maintained by the protection of the unsaturated fatty acids of the membrane lipids. Moreover, according to recent observations, regular uptake of relatively high tocopherol doses can counteract the development of arteriosclerosis. Further positive physiological properties and influences of tocopherols have been described, such as delay of late damages associated with diabetes, reduction of the risk of cataract development, reduction of oxidative stress in smokers, anticarcinogenic effects, protective effects against skin damages such as erythremes and skin aging etc.

Due to their oxidation inhibiting properties the tocopherols are not only used in food technology applications, but also employed in paintings based on natural oils, in deodorants and other cosmetics, such as sun protection agents, skin care agents, lip sticks etc. In such applications tocopherol compounds like tocopheryl acetate and succinate are usual application forms for the use as vitamin E, in circulation promoting and lipid reducing agents and as food additive in veterinary applications.

In the biosynthesis of tocopherols, in particular in the biosynthesis of α-tocopherol, phytyl pyrophosphate is believed to be the limiting factor. Previous studies indicate that PPP is formed from GGPP by sequential hydrogenation of the isoprenoid group, during which reaction dihydro-GGPP and tetrahydro-GGPP are formed as intermediates (GGPP->dihydro-GGPP->tetrahydro-GGPP->PPP; cf., for example, Bollivar et al. (1994) Biochemistry 33, 12763–12768).

The stepwise hydrogenation of GGPP to PPP is, as presently assumed, catalyzed by the enzyme geranylgeranyl reductase (GGPP reductase, also called geranylgeranyl pyrophosphate hydrogenase and GGPP hydrogenase), which is coded in plants in the gene Chl P. The enzyme geranylgeranyl reductase belongs to the isoprenoid metabolism and functions for two metabolic pathways: the tocoperol biosynthesis and the chlorophyll biosynthesis.

The essential role of this enzyme has been shown for the first time for the biosynthesis of chlorophyll (Benz et al. (1980) Plant Sci. Lett. 19, 225–230; Soll and Schultz (1981) Biochem. Biophys. Res. Commun. 99, 907–912; Schoch et al. (1977) Z. Pflanzenphysiol. 83, 427–436). The final step in chlorophyll biosynthesis is the esterification of chlorophyllide, which may take place with phytyl pyrophosphate as well as with geranylgeranyl pyrophosphate. In systematic studies of *Rhodobacter capsulatus* mutants it could be demonstrated that bacteriochlorophyllide is esterified with GGPP in a first step and that subsequently esterified chlorophyll-GG is hydrogenated (Katz et al. (1972) J. Am. Chem. Soc. 94, 7938–7939). In higher plants, phytyl chlorophyll (chlorophyll-P) can be found for the most part (Rüdiger and Schoch (1991) In: Chlorophylls (Scheer, H., Ed.) pp. 451–464, CRC Press, Boca Raton, Fla., USA). So far it has not been elucidated yet, which substrates are involved in the reductase reaction in plants. Presently, it is assumed that the plant enyzme geranylgeranyl reductase is able to transform chlorophyll-GG into chlorophyll-P (Schoch et al. (1978) Z. Pflanzenphysiol. 83, 427–436) as well as to hydrogenate GGPP to PPP, which is then subsequently joined to chlorophyllide (Soll et al. (1983) Plant Physiol. 71, 849–854).

GGPP serves as the substrate for the synthesis pathways of tocopherol and phyllochinone in the chloroplast outer membranes and for chlorophyll synthesis in the thylakoid membranes. The reduction of GGPP to PPP has been described for the first time 1983 by Soll et al. (Plant. Physiol. (1983) 71, 849–854). However, until now the isolation and characterisation of nuclic acid sequences, which code for the plant enzyme and which can be used for the manipulation of tocopherol content in transgenic plants, was unsuccessful.

The essential role of geranylgeranyl reductase in tocopherol and chlorophyll metabolism makes this enzyme a

particular valuable instrument for molecular biotechnology. By means of molecular biological techniques such as the transfer of DNA sequences coding for geranylgeranyl reductase, it should be possible to achieve alterations in tocopherol and/or chlorophyll biosynthesis performance in plants. By this way it would, for example, become possible to produce transgenic plants having an increased or reduced tocopherol content. Such transgenic plants and parts, cells and/or products thereof could subsequently be used as food and feed and in general as production center for tocopherol, for use in chemical, pharmaceutical and cosmetic industrial applications.

Further, there is reason to expect that plants which exhibit an increased content of antioxidative tocopherols, in comparison with wild-type plants, also show increased tolerance against stress conditions, in particular against oxidative stress.

It is therefore an object of the invention to provide new nucleic acid sequences, with the help of which the content of tocopherol can be manipulated in plants, plant cells, plant parts and/or plant products.

Further it is an important object of the invention to provide transgenic plants, plant cells, plant products and plant parts having an altered tocopherol content compared to wild-type plants.

It is a further object of the invention to show possible ways how to use the DNA sequences according to the invention, their gene products as well as the transgenic plants according to the invention for plant breeding practice.

Further objects of the invention will be seen from the following description. These problems are solved by the subject-matters of the independent claims, particularly based on the provision of the DNA sequences according to the invention, the gene products of which are directly involved in tocopherol biosynthesis, and the transfer of these DNA sequences to plants, which results in an altered tocopherol content.

The present invention thus relates to DNA sequences which code for proteins having biological activity of a geranylgeranyl reductase (also called geranylgeranyl pyrophosphate hydrogenase) or for a biologically active fragment thereof. In connection with this invention, biologically active fragment means that the mediated biological activity is sufficient to influence the tocopherol content. The invention relates in particular to DNA sequences which are isolated from plants and which code for proteins having enyzmatic activity of a geranylgeranyl reductase or a biologically active fragment thereof. Particularly preferred is the DNA sequence shown in SEQ ID NO:1 (see also FIG. **1**).

Further, the invention relates to alleles and derivatives of the DNA sequences according to the invention, which code for a protein having biological activity of a geranylgeranyl reductase, especially nuclic acid molecules, the sequences of which differ from the DNA sequences according to the invention due to the degeneracy of the genetic code and which code for a protein or a fragment thereof having the biological activity of a geranylgeranyl reductase.

Furthermore, the invention relates to nucleic acid molecules which comprise the DNA sequences according to the invention or which originate from the sequences according to the invention by naturally occurring or by gene technological or chemical processes and synthesis methods or which are deduced therefrom. The nucleic acid molecules can be any form of nucleic acid, such as DNA or RNA molecules, cDNA, genomic DNA, mRNA, etc.

The invention also relates to nucleic acid molecules wherein the DNA sequences according to the invention are combined with regulatory elements that provide transcription and, if desired, translation in the plant cell.

Thus, it is possible to express the DNA sequences according to the invention in plant cells, for example, under control of constitutive, but also under control of inducible or tissue-specific or developmental specific regulatory elements, particularly promoters. While, for example, the use of an inducible promoter makes it possible to achieve specifically triggered expression of the DNA sequences according to the invention in plant cells, the use of tissue-specific, for example seed-specific, promoters provides the possibility to modify the tocopherol content in specific tissues, for example, in seed tissue. Therefore, in a preferred embodiment of the invention, the DNA sequences according to the invention are in combination with tissue-specific promoters, particularly in combination with seed-specific promoters.

The invention further relates to proteins having the biological activity of a geranylgeranyl reductase or active fragments thereof, which are encoded by a DNA sequence according to the invention or a nucleic acid molecule according to the invention. Preferably, the protein is a plant geranylgeranyl reductase, preferably from *Nicotiana tabacum,* especially preferred is a protein having the amino acid sequence shown in SEQ ID NO:2 (cf. also FIG. **2**), or an active fragement thereof.

It is a further object of the invention to provide vectors and microorganisms, the use of which makes it possible to produce new plants wherein an altered tocopherol content can be achieved. This problem is solved by the provision of the vectors and microorganisms according to the invention, which comprise nucleic acid sequences that code for enzymes having the activity of a geranylgeranyl reductase.

The present invention, thus, also relates to vectors, in particular plasmids, cosmids, viruses, bacteriophages and other vectors customarily used in genetic engineering, which comprise the nucleic acid molecules according to the invention, as described above, and which, if desired, can be used for the transfer of the nucleic acid molecules according to the invention to plants and plant cells.

The invention also relates to transformed microorganisms, such as bacteria, viruses, fungi, yeasts, etc., which contain the nucleic acid sequences according to the invention.

In a preferred embodiment the nucleic acid molecules, contained in the vectors, are combined with regulatory elements that provide transcription and, if desired, translation in procaryotic and eucaryotic cells.

If desired, the nucleic acid sequences according to the invention may be supplemented by enhancer sequences or other regulatory sequences. These regulatory sequences comprise, e.g. also signal sequences which provide the transport of the gene product to a certain cell compartment.

It is also an object of the invention to provide new plants, plant cells, plant parts or plant products exhibiting altered tocopherol content, which may be linked to modified chlorophyll biosynthesis performance, compared to wild-type plants.

These problems are solved by the transfer of the nucleic acid molecules according to the invention and their expression in plants. By providing the nucleic acid molecules according to the invention, it is now possible to manipulate plant cells by gene technology methods in such a way that they exhibit new or altered geranylgeranyl reductase activity, in comparison with wild-type cells, and as a consequence show an altered tocopherol biosynthesis performance and modified tocopherol content.

COR-AAD1-00001504

In a preferred embodiment, the invention relates to plants and plant cells and parts thereof, wherein the tocopherol content is increased, in comparison with wild-type plants, due to the presence and expression of the nucleic acid molecules according to the invention.

The invention also relates to plants wherein the transfer of the nucleic acid molecules according to the invention leads to a reduction of tocopherol and/or chlorophyll content. A reduced tocopherol and/or chlorophyll biosynthesis productivity may, for example, be achieved by the transfer of antisense constructs or other suppression mechanisms, such as co-suppression.

Further, the invention relates to transgenic plant cells and plants comprising such plant cells, and parts and products thereof, wherein the new nucleic acid molecules are integrated into the plant genome. The invention also relates to plants, in the cells of which the nucleic acid sequence according to the invention is present in self-replicating form, i.e. the plant cell contains the foreign DNA on an autonomous nucleic acid molecule.

The plants, which are transformed with the nucleic acid molecules according to the invention and wherein an altered amount of tocopherol and/or chlorophyll is synthesised due to the transfer of such molecule, can, in principle, be any plant. Preferably, the plant is a monocotyle or dicotyle useful plant. Examples of monocotyl plants are plants which belong to the genus of avena (oat), triticum (wheat), secale (rye), hordeum (barley), oryza (rice), panicum, pennisetum, setaria, sorghum (millet), zea (maize). Dicotyl useful plants are, inter alia, leguminous plants, such as legumes and especially alfalfa, soy bean, rape, tomato, sugar beet, potato, ornamental plants, trees. Other useful plants can be, for example, fruit-bearing plants (particularly apples, pears, cherries, grapes, citrus fruits, pineapples and bananas), oil palms, tea, cocoa and coffee shrubs, tobacco, sisal, cotton, flax, sunflower as well as medical plants and pasture grasses, forage cereals and feed plants. Special preference is given to grains, cereals, wheat, rye, oat, barley, rice, maize and millet, forage cereals, sugar beet, rape, soy bean, tomato, potato, sweet grasses, feed grasses, forage grasses and clover. It is self-evident that the invention particularly relates to common food and forage plants. In this context, in addition to the plants already mentioned, peanut, lentil, forage bean (Ackerbohne), mangel, buckwheat, carrot, topinambur, Brassica (rapa, oleifera, napus, rapifera), white mustard and swede are to be mentioned.

Furthermore, the invention relates to propagation material of plants according to the invention, such as seeds, fruits, cuttings, tubers, root stocks, etc., whereby this propagation material may contain the above described transgenic plant cells, as well as parts of such plants, such as protoplasts, plant cells and calli.

The invention further relates to plant cells which, due to the presence and, if desired, expression of the nucleic acid molecules according to the invention, have an altered content of vitamin K_1 in comparison with plant cells which do not contain the nucleic acid molecules. The lipophilic vitamin K_1, which is present in particular in plants, plays an essential role in the formation of coagulation factors; lack of vitamin K_1 leads to a reduction in blood coagulation, which is why vitamin K_1 is also called anti-haemorrhagic or coagulation vitamin. Since the expression of the nucleic acid molecules according to the invention results in an altered geranylgeranyl reductase activity and, thus, in an altered PPP-synthesis performance, and in view of the fact that phylloquinone, called vitamin K_1, as the tocopherols, comprises one unit of phytol, the invention also relates to such plant cells and plants which exhibit an altered vitamin K_1 content, alone or in combination with an altered tocopherol content.

In a preferred embodiment, the invention relates to transgenic plant cells and plants and parts and products thereof, which have an altered tocopherol content, in comparison with non-transformed cells, due to the presence and, if desired, expression of a DNA sequence coding for a plant geranylgeranyl reductase. Preferably, the DNA sequence, contained within the plant cells, is a sequence coding for geranylgeranyl reductase which is isolated from tobacco. Specially preferred is a DNA sequence as shown in SEQ ID NO:1 (cf. also FIG. **1**). In a particularly preferred embodiment the DNA sequences according to the invention encode for a geranylgeranyl reductase pre-enzyme, comprising a transit sequence for translocation into plastids.

The invention further relates to plants wherein, in addition to the chl P gene, a gene for hydroxyphenyl pyruvate dioxygenase (HPD) is expressed. The enzyme HPD catalyses the reaction of 4-hydroxyphenylpyruvate into homogentisate, which, as mentioned above, represents the second precursor of the tocopherols, besides phytol. The enzyme HPD as well as its role within the plant isoprenoid metabolism are described, inter alia, in Norris et al. (1995) The Plant Cell 7, 2139–2149.

By co-expression, preferably over-expression, of sequences which code for geranlygeranyl reductase and HPD, respectively, the tocopherol content in transgenic plants can be further increased in comparison with plants which only contain the sequences according to the invention coding for chl P.

In a further embodiment, the invention relates to host cells, particularly procaryotic or eucaryotic cells, which have been transformed or infected with a nucleic acid molecule or a vector, as described above, and cells which originate from such host cells and which contain the described nucleic acid molecules or vectors. The host cells can, e.g., be bacteria, algae, yeast and fungus cells as well as plant or animal cells. The invention also relates to such host cells which not only contain the nucleic acid molecules according to the invention, but further contain one or more nucleic acid molecules, transferred by gene technology or naturally, which carry the genetic information for enzymes involved in the biosynthesis of tocopherol, chlorophyll and/or vitamin K_1.

It is a further object of the present invention to provide processes for producing plant cells and plants which exhibit altered tocopherol content.

This problem is solved through processes by means of which it is possible to produce new plants and plant cells which show an altered tocopherol content due to the transfer of nucleic acid molecules coding for geranylgeranyl reductase.

Furthermore, this problem is solved through processes by means of which it is possible to produce new plant cells and plants, which, due to co-transfer of nucleic acid molecules coding for geranylgeranyl reductase and nucleic acid molecules coding for HPD or the transfer of nucleic acid molecules coding for geranylgeranyl reductase and for HPD, show an altered tocopherol content in comparison with wild-type plants.

For the production of such new plant cells and plants several different methods can be applied. On the one hand, plants and plant cells can be modified by conventional gene technological transformation methods in such a way that the

new nucleic acid molecules are integrated into the plant genome, which means that stable transformants are produced. On the other hand, the nucleic acid molecules according to the invention, the presence and, if desired, expression of which results in altered tocopherol biosynthesis performance, can also be introduced into the plant cell or plant as self-replicating system. For instance, the nucleic acid molecules according to the invention may be contained in a virus that gets in contact with the plant or plant cell.

According to the invention, plant cells which, due to the expression of a nucleic acid sequence according to the invention, show an altered tocopherol content, are produced by a method which includes the following steps:

a) Manufacture of an expression cassette, comprising the following DNA sequences:
   a promoter that provides transcription in plant cells;
   at least one nucleic acid sequence that codes for a protein or a fragement having enzymatic activity of a geranylgeranyl reductase, whereby the nucleic acid sequence is linked to the 3' end of the promoter in sense orientation; and
   if desired, a termination signal for transcription termination and addition of a poly-A-tail to the respective transcript, wherein the termination signal is linked to the 3' end of the coding region;
b) transformation of plant cells with the expression cassette produced in step a);
c) regeneration of transgenic plants and, if desired, propagation of the plants.

As an alternative, the one or more nucleic acid sequences according to the invention can be introduced into the plant cell or plant as self-replicating system.

In a further alternative, step a) of the above method may be modified in such a way that the at least one nucleic acid sequence according to the invention, which codes for a protein or a fragment having enzymatic activity of a geranylgeranyl reductase, is linked to the 3' end of the promoter in antisense orientation.

It is a further object of the invention to show advantageous applications of the nucleic acid sequences according to the invention as well as of the nucleic acid molecules containing these nucleic acid sequences.

This problem is solved by the uses according to the invention of the new DNA molecules for the production of plant cells and plants which exhibit an altered, preferably increased, tocopherol content in comparison to wild-type cells and wild-type plants.

Further, the invention relates to the use of the nucleic acid sequences according to the invention for the production of plants which show an altered chlorophyll content.

Moreover, the invention relates to the use of the nucleic acid sequences according to the invention for the production of plants which show an altered, preferably increased, content of vitamin K_1.

It is a further object of the invention to show possible uses of the plants according to the invention and cells, parts and products thereof.

The invention particularly relates to the use of the plants according to the invention as forage and/or food plant. Depending on the achieved increase in vitamin E and/oder K_1 content in the transgenic useful plant and products and parts thereof, it may be possible to reduce the amount of respective vitamins, particularly of vitamin E, which otherwise is usually admixed to the feed/food and which is often also required. Under certain circumstances, conventional supplement with vitamins may become superfluous. Aside from this the invention relates in general to an enhancement of the nutritional value of useful plants by increasing the content of tocopherols and/or phyllochinone.

Further, the invention relates to the use of the plant cells, plants, parts and products thereof according to the invention as production sites for vitamin E and/or vitamin K_1. Apart from their application due to their vitamin characteristics, for example in dietetic and pharmaceutical products, cosmetics, skin care products, generally for vitamin E supplement, etc., tocopherols are also applied as antioxidants in chemical products such as lipids and oils. The plants according to the invention, thus, represent an important source for the production of tocopherols and/or vitamin K_1 in a broad spectrum of commercial purposes.

The invention further relates to the use of the nucleic acid sequences according to the invention in combination with seed-specific promoters for the production of plants, wherein particularly seed tissue exhibits an altered, preferably increased, tocopherol content. In a preferred embodiment of the invention, the nucleic acid sequences according to the invention are used in combination with the USP (B äumlein et al. (1991) Mol. Gen. Genet. 225, 459–467) or the hordein promoter (Brandt et al. (1985) Carlsberg Res. Commun. 50, 333–345).

The mentioned promoters, particularly seed-specific promoters, are especially useful for specific reduction of the tocopherol and chlorophyll content in transgenic seeds by use of the DNA sequences according to the invention in connection with the antisense approach.

Further, the invention relates to the use of a geranylgeranyl reductase gene for producing an altered tocopherol content in plants.

Moreover, the invention relates to the use of a protein having enzymatic activity of a geranylgeranyl reductase in order to achieve an altered tocopherol content in plants.

Further, the invention relates to the use of the nucleic acid molecules according to the invention, of the proteins according to the invention having geranylgeranyl reductase activity and/or of the transgenic plants and host cells according to the invention having new or altered geranylgeranyl reductase activity for the identification of new herbicidal substances for plant protection. Due to the key role of geranylgeranyl reductase within the chlorophyll and tocopherol biosynthesis, the DNA sequences according to the invention and the proteins encoded thereby are an extremely valuable target for herbicide research.

For example, the proteins according to the invention having enzymatic geranylgeranyl reductase activity can be used for X-ray structure analysis, NMR spectroscopy, molecular modelling and drug design, in order to identify or synthesise inhibitors or effectors of geranylgeranyl reductase and thus potential herbicides, on the basis of the data and knowledge obtained by means of these techniques.

The invention further relates to the use of the nucleic acid sequences according to the invention for the production of herbicide tolerant plants. Sequences coding for geranlygeranyl reductase can be modified by means of standard techniques, or can be supplemented by new sequence elements, and subsequently transferred to plant cells. The transfer of sequences deduced from the sequences according to the invention can, e.g. be used to modify the properties of plants in such a way that more or less functionally active geranylgeranyl reductase or a variant of the geranlygeranyl reductase having altered characteristics is synthesised in the transgenic plant or that the expression level of the chl P gene, present in the transgenic plant, is reduced. As a consequence, by increasing the CHL P activity an increase in the tolerance against herbicides which block chlorophyll

biosynthesis, can be achieved. Similarly, e.g. the expression of modified geranylgeranyl reductase genes in transgenic plant cells can be linked to an increase in the tolerance against herbicides.

Further, the invention relates to the use of the nucleic acid sequences according to the invention or of a protein encoded thereby for the production of antibodies.

Thus, the present invention comprises any possible use of the nucleic acid molecules according to the invention, the presence and, if desired, the expression of which in plants causes an alteration in tocopherol content and/or chlorophyll content, as any possible uses of the proteins according to the invention and fragments thereof, the enzymatic activity of which leads to such alteration.

In principle, any promoter functional in the plant of choice can be used, which fulfils the prerequisite that expression controlled by said promoter leads to an altered tocopherol synthesis capacity. In view of the use of the transgenic plants as food and/or forage plants, promoters which provide seed-specific expression are particularly useful in this respect. Examples for such promoters are the USP promoter, the hordein promoter and the napine promoter.

In case such promoters are not already known or not yet available, the strategy and methods for the isolation of such promoters are known to the person skilled in the art. In general, in a first step poly(A)$^+$ RNA is isolated from seed tissue and a cDNA library is established. In a second step, with the help of cDNA clones based on poly(A)$^+$ RNA molecules originating from a non-seed tissue, those clones are identified by hybridization from the first library, whose corresponding poly(A)$^+$ RNA molecules are expressed only in seed tissue. Subsequently, promoters are isolated with the help of cDNAs identified in this manner, which can then be used in order to control expression of the coding nucleic acid sequences described herein. Likewise, other tissue-specific or developmental specific promoters or promoters which can be induced by abiotic stimuli can be isolated and used according to the invention.

Alternatively it may be desired that the plant shows an altered, preferably increased, tocopherol content in several sections or organs, due to expression of the nucleic acid molecules according to the invention. In this case, the use of a constitutive promoter, for example, the use of the 35S RNA promoter from cauliflower mosaic virus may be desirable.

The invention also comprises nucleic acid molecules that code for proteins having biological activity of a geranylgeranyl reductase or biologically active fragments thereof, and which hybridise to the nucleic acid molecules described above. In the context of this invention "biologically active fragment" means in general that the fragment is sufficient for causing an alteration in tocopherol content. The term "hybridisation" means in the context of this invention a hybridisation under conventional hybridisation conditions, preferably under stringent conditions, as e.g. described in Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Nucleic acid molecules that hybridise with the molecules according to the invention, may be isolated, e.g. from genomic or cDNA libraries.

Identification and isolation of such nucleic acid molecules can be carried out using the nucleic acid molecules or parts of these molecules or the reverse complements of these molecules, e.g. by hybridisation according to standard techniques (see, for example, Sambrook et al., supra). For identification and isolation of such nucleic acid molecules, also such sequences can be used which are deduced from the DNA sequences according to the invention, for example degenerated oligonucleotide primers.

Thus, the invention also comprises the use of a DNA sequence according to the invention or fragments thereof for the identification and isolation of homologous sequences from plants or other organisms.

For instance, nucleic acid molecules which display exactly or essentially the nucleic acid sequences described above or fragments of these sequences can be used as hybridisation probe. The fragments used as hybridisation probe can also be synthetic fragments produced with the help of customary synthesis techniques and the sequence of which basically corresponds with that of a nucleic acid molecule according to the invention. Once genes that hybridise with the nucleic acid sequences according to the invention have been identified and isolated, it is necessary to determine their sequence and to analyse the properties of the proteins encoded by these sequences. To do so, a number of molecular biological, biochemical and biotechnological standard methods are available to the person skilled in the art.

The molecules that hybridise with the nucleic acid molecules according to the invention include also fragments, derivatives and allelic variants of the above-described DNA molecules that code for a geranylgeranyl reductase or a biologically, i.e. enzymatically, active fragment thereof. Fragment means in this respect fragments or regions of the nucleic acid molecules which are sufficiently long to code for a polypeptide or protein having enzymatic activity of a geranylgeranyl reductase or a comparable enzymatic activity, which is able to cause an altered tocopherol content. The term "derivative" means in this context that the sequences of these molecules are distinguishable from the sequences of the above-described nucleic acid molecules in one or several positions and show a great extent of homology with these sequences. Homology in this connection means a sequence identity of at least 40%, especially an identity of at least 60%, preferably above 80%, and especially preferably above 90%. The deviations from the above-described nucleic acid molecules can be due to deletion, addition, substitution, insertion or recombination.

Homology means further that there is functional and/or structural equivalence between the nucleic acid molecules in question or between the proteins encoded thereby. With respect to the nucleic acid molecules which are homologous with the above-described molecules and which represent derivatives of these molecules, it concerns usually variants of these molecules that represent modifications which perform the same biological function. These may concern naturally occurring variations, e.g. sequences from other organisms, or mutations, whereby these modifications can have occurred naturally or were introduced through specific mutagenesis. Furthermore, the variations may concern synthetically produced sequences. With respect to the allelic variants, these may occur naturally as well as be synthetically produced variants, or variants produced by recombinant DNA technology.

Usually the proteins encoded by the different variants and derivatives of the nucleic acid molecules according to the invention have common characteristics. Such characteristics are e.g. enzyme activity, molecular weight, immunological reactivity, conformation, etc. Further common characteristics may be physical properties, e.g. migration pattern in gel electrophoresis, chromatographic characteristics, sedimentation coefficients, solubility, spectroscopic properties, stability, pH optimum, temperature optimum, etc. Further, of

course, the reactions catalysed by the proteins may have common or similar features.

To prepare the introduction of foreign genes into higher plants, a variety of cloning vectors are available which contain an origin of replication active in *E. coli* and a marker gene for the selection of transformed bacteria cells. Examples of such vectors are pBR322, pUC series, M13mp series, pACYC184, etc. The desired sequence can be inserted into the vector at a suitable restriction site. The plasmid obtained is used for the transformation of *E. coli* cells. Transformed *E. coli* cells are cultured in a suitable medium and subsequently harvested and lysed. The plasmid is recovered. Usually, restriction mapping, gel electrophoresis and other biochemical, molecular biological methods are applied as analysing methods to characterise the recovered plasmid DNA. After each manipulation the plasmid DNA can be digested and the obtained DNA fragments can be linked with other DNA sequences. Each plasmid DNA sequence can be cloned in the same or other plasmids.

For introducing DNA into a plant host cell many well known methods are available and the skilled person can easily determine and select the respectively suitable procedure. These techniques include the transformation of plant cells with T-DNA by using *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes* as transformation means, fusion of protoplasts, direct gene transfer of isolated DNA into protoplasts, microinjection and electroporation of DNA, introduction of DNA by means of biolistic methods and other possibilities.

When injecting and electroporating DNA into plant cells no specific demands per se are made on the plasmids used. The same appplies to direct gene transfer. Here, simple plasmids such as pUC-derivatives, can often be used. If however, whole intact plants are to be regnerated from cells transformed in this way, the presence of a selectable marker gene is usually required. The skilled person is familiar with custamary selection markers, and he can easily select an appropriate marker.

Depending on the method chosen for the introduction of the gene(s) of interest, additional DNA sequences may be required. If, for example, the Ti or Ri plasmid is used for transformation of the plant cell, at least the right border, but frequently both the right and left boader, of the T-DNA contained in the Ti and Ri plasmid must be combined with the gene to be introduced as flanking regions.

If agrobacteria are used for transformation, the DNA to be introduced has to be clones into special plasmids, viz into an intermediary or a binary vector. Due to sequences which are homologous with sequences in the T-DNA, the intermediary vectors can be integrated in the Ti or Ri plasmid of agrobacteria by homologous recombination. In addition, the latter contains the vir region required for the transfer of the T-DNA. Intermediary vectors cannot replicate in agrobacteria. The intermediary vector can be transferred to agrobacteria *Agrobacterium tumefaciens* by means of a helper plasmids (conjugation). Binary vectors can replicate in *E. coli* as well as in agrobacteria. They contain a selection marker gene and a linker or polylinker framed by the right and left border regions of the T-DNA. These vectors can be directly transformed into agrobacteria (Holsters et al.(1978) Molecular and General Genetics 163, 181–187). The agrobacterium which serves as host cell shall contain a plasmid carrying the vir region. The vir region is required for the transfer of the T-DNA to the plant cell. Additional T-DNA may be present. The agrobacterium transformed in the manner described is used for the transformation of plant cells.

The use of T-DNA for the transformation of plant cells has been thoroughly studied and is adequately described in EP 120 515; Hoekema in: The Binary Plant Vector System, Offsetdrokkerij Kanters B. V., Alblasserdam (1985) Chapter V; Fraley et al. (1993) Crit. Rev. Plant. Sci., 4, 1–46 and An et al. (1985) EMBO J. 4, 277–287.

For the transfer of the DNA to the plant cell, plant explantates can be appropriately cultivated with *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*. Out of the infected plant material (e.g. leaves, leaf pieces, stem segments, rootes, but also protoplasts or plant cells cultivated in suspension cultures) whole plants can be regenerated in a suitable medium which can contain antibiotics or biocides for the selection of transformed cells. The regeneration of plants is carried out according to customary regeneration methods by using conventional nutrient media. Plants obtained in the above-described manner can then be examined for the presence of the introduced DNA. Other possibilities for introducing foreign DNA by applying the biolistic method or through protoplast transformation are known (see, e.g. Wilmitzer L (1993) Transgenic Plants, in: Biotechnology, A Multi-Volume Comprehensive Treatise (H. J. Rehm, G. Reed, A. Pühler, P. Stadler, eds.) Vol. 2, 627–659, V. C. H. Weinheim—New York, Basel—Cambridge).

While the transformation of dicotyle plants via Ti plasmid vector systems with the help of *Agrobacterium tumefaciens* is well established, recent studies indicate that also monocotyle plants can be transformed via Agrobacterium-based vectors (Chan et al. (1993) Plant Mol. Biol. 22, 491–506; Hiei et al. (1994) Plant J. 6, 271–282; Deng et al. (1990) Science in China 33, 28–34; Wilmink et al. (1992) Plant Cell Reports 11, 67–80; May et al. (1995) Bio/Technology 13, 486–492; Conner and Domiss (1992) Int. J. Plant Sci. 153, 550–555; Ritchie et al. (1993) Transgenic Res. 2, 252–265).

Alternative systems for the transformation of monocotyle plants are transformations by means of the biolistic approach (Wan and Lemaux (1994) Plant Physiol. 104, 37–48; Vasil et al. (1993) Bio/Technology 11, 1553–1558; Ritala et al. (1990) Tehor. Appl. Genet. 79, 625–631; Altpeter et al. (1996) Plant Cell Reports 16, 12–17), transformation of protoplasts, electroporation of partially permeabilised cells and the introduction of DNA by means of glass fibres.

The transformation of maize is specifically described several times in the literature (cf. e.g. WO 95/06128, EP 0 513 849; EP 0 465 875; Fromm et al. (1990) Biotechnology 8, 833–844; Gordon-Kamm et al. (1990) Plant Cell 2, 603–618; Koziel et al. (1993) Biotechnology 11, 194–200). EP 292 435 describes a process by means of which, starting from mucusless friable granulous maize callus, fertile plants can be obtained. Shillito et al. ((1989) Bio/Technology 7, 581) have observed in this context that it is further necessary for the generation of fertile plants to start from a callus culture from which a dividing protoplast culture having the ability to regenerate to plants can be obtained. After an in vitro cultivating period of seven to eight months, Shillito et al. obtain plants which are able to produce viable progeny.

Prioli and Söndahl ((1989) Bio/Technology 7, 589) describe the regeneration and the production of fertile plants from maize protoplasts, the cateto maize inbreeding line Cat 100-1. The authors assume that the regeneration of fertile plants from protoplasts depends on a number of different factors, such as the genotype, the physiological condition of the donor cells and cultivation conditions.

Also the successful transformation of cereal species has already been described, e.g. for barley (Wan and Lexaux, supra; Ritala et al., supra) and for wheat (Nehra et al. (1994) Plant J. 5, 285–297; Altpeter et al. supra).

Once the introduced DNA is integrated into the genome of the plant cell, it remains stable and is also stably inherited to the progeny of the originally transformed cell. Normally, it contains a selection marker conferring resistance against a biocide or antibiotic such as kanamycin G418, bleomycin, hygromycin, methotrexate, glyphosate, streptomycin, sulfonyl urea, gentamycin or phosphinotricin etc. The selection

marker which can be chosen individually should therefore allow selection of transformed cells over cells which are devoid of the introduced DNA.

The transformed cells grow within the plant in the usual manner (see also McCormick et al. (1986) Plant Cell Reports 5, 81–84). The resulting plants can be cultivated in the usual fashion and may be propagated by self-fertilisation or be crossed with plants having the same transformed or other genetic traits. The resulting hybrid individual plants have the respective phenotypic properties. Usually, seeds can be obtained from the plants.

Two or more generations should be grown in order to ensure that the phenotypic trait is stably maintained and inherited. Seeds should also be harvested in order to ensure that the respective phenotype or other characteristics are maintained.

By applying the usual methods, also transgenic lines can be determined, which are homozygous for the new nucleic acid molecules and, their phenotypic behaviour can be examined with respect to altered tocopherol content and compared to that of hemizygous lines.

Expression of the proteins according to the invention having geranylgeranyl reductase activity can be achieved by means of conventional molecular biological and biochemical methods. The skilled person is familiar with these techniques and he is able without any difficulty to choose a suitable detection method, for example, a Northern blot analysis for the detection of geranylgeranyl reductase-specific RNA and for determining the amount of geranylgeranyl reductase-specific RNA accumulation, a Southern blot analysis for the identification of DNA sequences coding for geranylgeranyl reductase or a Western blot analysis for detecting the protein, encoded by the DNA sequences according to the invention, preferably CHL P. Enzymatic activity of geranylgeranyl reductase may, for example, be detected and examined by an enzyme assay, described by Soll and Schultz (1981) in Biochem. Biophys. Res. Commun. 99, 907–912, which is based on the formation of chlorophyll phytyl.

The invention is based on the successful isolation of a cDNA clone coding for geranylgeranyl reductase from a cDNA library from *Nicotiana tabacum* cv. Petit Havana SR1. The sequence of this cDNA clone, which comprises a complete open reading frame, is shown in SEQ ID NO:1. Using the sequence according to SEQ ID NO:1 it was possible to produce transgenic plants which exhibit an altered tocopherol content in comparison to wild-type plants.

The cDNA clone, containing the DNA sequence according to SEQ ID NO:1 was transformed into *Escherichia coli* and the resulting *E. coli* strain was deposited at the Deutsche Sammlung für Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1b, D-38124 Braunschweig, Germany under Deposit Number DSM 11816 on Oct. 16, 1997, in accordance with the Budapest Treaty.

The following examples serve the purpose of illustrating the invention.

EXAMPLES

Example 1

Cloning a tobacco cDNA which codes for a geranylgeranyl reductase (CHL P)

For the identification of a geranylgeranyl reductase cDNA from tobacco a Lambda ZAP II cDNA library (*Nicotiana tabacum* SR1, Stratagene, USA) was screened according to the manufacturer's protocol, by using an EST from *Arabidopsis thaliana* which codes for the locus 4D9T7P. The used EST sequence shows similarity to the known bch P/chl P sequences from *Rhodobacter capsulatus* (Young et al. (1989) Mol. Gen. Genet. 218, 1–12; Bollivar et al. (1994) J. Mol. Biol. 237, 622–640; Bollivar et al. (1994) Biochemistry 33, 12763–12768) und Synechocystis PCC6803 (Addlesee et al. (1996) FEBS Lett. 389, 126–130).

The used hybridisation probe comprises the region of the EST sequence in 4D9T7P (Accession No. T04791) from base 1 to base 364. The probe was isolated from the PRL2 library from *A. thaliana* (vector: λZipLox) (Newman et al. (1994) Plant Physiol. 106:1241–1255) as NotI/SalI restriction fragment, and radioactively labelled with [α-^{32}P]dCTP by nick translation (Life Technologies, Eggenstein, Germany).

Hybridisation was carried out according to the following protocol:

2 hrs prehybridisation at 55° C. with hybridisation solution having the following composition: 5×SSC, 0.1% SDS, 5×Denhardt reagent, 100 μm/ml denatured salmon sperm DNA;

12 hrs main hybridisation at 55° C. with fresh hybridisation solution having the above composition plus radioactively labelled probe;

Washes:

2 times 10 min. at 55° C. with 2×SSC and 0.1% SDS, and

1 time 5 min. at 55° C. with 1×SSC and 0.1% SDS.

The plasmid DNA, isolated by cDNA library screening, was sequenced by conventional methods. The identified chl P cDNA sequence, shown in SEQ ID NO:1, comprises 1510 nucleotides (without polyA-tail); nucleotides 1 to 1392 encode a 52 kDa protein consisting of 464 amino acid residues (including the start codon methionine, and without the stop codon (nucleotides 1393 to 1395)). The deduced amino acid sequence of the CHL P protein is shown in SEQ ID NO:2. The nucleotide sequence shown in SEQ ID NO:1 comprises a 3' untranslated region from nucleotide 1396 to 1510.

For DNA, RNA isolation, sequence analysis, restriction, cloning, gel electrophoresis, radioactive labelling, Southern, Northern and Western Blot analysis, hybridisation and similar procedures conventional methods were employed, as described in standard laboratory manuals, such as Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Example 2

Transformation of tobacco plants and regeneration of intact plants

For the production of transgenic plants which overexpress CHL P and which, thus, exhibit an increased tocopherol content in comparison to non-transformed plants, the DNA sequence according to SEQ ID NO:1 was cut out of the vector with the restriction enzymes BamHI and SalI, located in the multiple cloning site of the pBluescript vector, and ligated in sense orientation behind the CaMV 35S promoter into the binary vector BinAR-TX (Höfgen and Willmitzer (1990) Plant Science 66, 221–230), a pBIB derivative (Becker (1990) Nucleic Acid Res. 18, 203), digested with the same restriction endonucleases. For the purpose of illustration, a restriction map of the vector BinAR-TX is enclosed as FIG. **3**.

Instead of the mentioned binary vector BinAR-TX, any other vector suitable for plant transformation can be used for constructing a chimeric gene, comprising a fusion of the CaMV 35S promoter or any other promoter which provides transcription and translation in plant cells, and DNA sequences which code for CHL P.

COR-AAD1-00001509

The recombinant vector pCHLPbin was then introduced into *Agrobacterium tumefaciens* (strain GV2260; (Horsch et al. (1985) Science 227, 1229–1231) and used for the transformation of tobacco plants (SNN) via the leaf disc transformation technique (Horsch et al., supra).

For this purpose an overnight culture of the resulting *Agrobacterium tumefaciens* clone was centrifuged for 10 minutes at 5000 rpm, and the bacteria were resuspended in 2YT-medium. Young tobacco leaves of a sterile culture (*Nicotiana tabacum* cv. Samsun NN) were cut into small slices of about 1 cm in diameter and incubated for a short period of time in the bacteria suspension. Subsequently, the leaf slices were placed on MS medium (Murashige and Skoog (1962) Physiol. Plant 15, 473; 0.7% Agar), and kept in the dark for two days. Then, the leaf slices were placed on MS medium (0.7% Agar) with 1.6% glucose, 1 mg/l benzylaminopurine, 0.2 mg/l naphthyl acidic acid, 500 mg/l claforan (cefotaxime, Hoechst, Franfurt, Germany) and 50 mg/l kanamycin, for shoot induction. The medium was changed every seven to ten days. After the development of shoots, the leaf slices were transferred into glass vessels, containing the same medium. Developing choots were cut off and placed on MS medium with 2% saccharose and 250 mg/l claforan, and regerated to whole plants.

Example 3

Analysis of transgenic tobacco plants containing the recombinant vector pCHLPbin

Transgenic tobacco plants were transformed, selected and regenerated, as described above. After root growth in sterile culture about 100 independent primary transformants were transferred to soil in the green house. The tobacco plants were kept in the green house at 60% humidity and 20–25° C. for 16 hours at light and 18–20° C. for 8 hours in the dark.

The transformants with normal or increased tocopherol and/or chlorophyll contents showed neither abnormal appearance or habitus nor an altered growth rate in comparison to control plants.

Several of the primary transformants showed an up to 4 times to 6 times increased tocopherol content in comparison to wild-type plants. This increase in tocopherol content could be further raised in the progeny of the T1 and T2 generation and in homozygous progeny plants, obtained by usual self-fertilisation and subsequent determination of the segregation pattern of the seeds on kanamycin-containing medium.

Furthermore, it could be observed that the tocopherol contents in transgenic plants were further increased under stress conditions, such as found during the cultivation at low or increased temperatures or under high power light, as well as in senescent leaves, as compared to control plants.

Tocopherol was measured according to the following method:

Leaf discs were homogenised in liquid nitrogen and extracted three times in methanol. The extracts were collected and eluted on a LiCrospher 100 HPLC RP-18-column (Merck, Darmstadt, Germany) at flow of 1 ml/min with the following gradient: 94% solvent B (100% methanol)/6% solvent A (30% methanol, 10% 0.1 M ammonium acetate, pH 5.1) for 7 min., for further 17 min. 99% solvent B/1% solvent A, then further 26 min. 94% solvent B/6% solvent A.

In an alternative procedure the collected extracts were analysed by means of HPLC in a isocratic gradient (gradient as follows: 2% solvent A [10% methanol and 10% acidic acid] and 98% methanol (solvent B); flow rate 1 ml/min). For detection, a Waters LC-module-device with Shimadzu RF 551 fluorescence detector (295 nmex, 325 nmem) was used.

The results of a tocopherol assay are shown in a bar graph in FIG. **4**. The comparison of leaves 6, 9, 12 (numbering starting from the top of the plant) of the transformants 28 and 30 with the corresponding leaves of the control plant (SNN) demonstrates that the transgenic lines have a tocopherol content which is up to 6 times higher, compared to wild-type plants.

Irrespective of their ability to grow on kanamycin-containing medium, the transgenic tobacco plants were also analysed by Southern blot hybridisation. After hybridisation with a labelled cDNA fragment for CHL P, additional radioactively labelled bands could be detected using the genomic DNA of the transformants, digested with restriction enzymes, compared to genomic DNA of control plants.

A Northern blot analysis revealed an increased amount of specific RNA in the transformants, compared to the CHL P-RNA contents in control plants.

Further, an increased geranylgeranyl reductase expression could be determined in transgenic plants by Western blot analysis. The transformants showed an increased amount of CHL P protein in comparison to control plants.

Moreover, it could be confirmed by plastid import experiments (carried out according to Grimm et al. (1989) Plant Mol. Biol. 13, 583–593) that the CHL P pre-protein, encoded by the sequence according to SEQ ID NO:1, was imported into the plastides after in vitro transcription and translation.

Example 4

Construction of CHL P antisense constructs and transfer to tobacco

While the transgenic plants produced and analysed according to Examples 2 and 3 exhibit an increased tocopherol content due to overexpression of the DNA sequence according to the invention, the following antisense construct was constructed and transferred to tobacco in order to produce transgenic tobacco plants having reduced CHL P activity.

The cDNA sequence according to SEQ ID NO:1 was cut out of the vector using the restriction enzymes KpnI and XbaI, located in the multiple cloning site of the pBluescript vector, and fused with the 35S promoter of cauliflower mosaic virus in antisense orientation in the binary vector BinAR-TX (cf. Example 2) digested with the same restriction enzymes. The resulting recombinant vector pCHLPASbin was transferred to tobacco via *Agrobacterium tumefaciens*-mediated leaf disc transformation, as described in Example 2. Then, transgenic plants were regenerated. Approximately 100 independent transgenic lines were regenerated and proved for the insertion of copies of the transgene by standard methods, such as Southern blotting.

The transformants showed a growth rate slower than that of control plants, lower pigmentation, reduced CHL P-specific RNA and reduced CHL P protein content, high amount of geranylgeranyl chlorophyll (up to 50% of the total chlorophyll content, in comparison to 100% of phytyl chlorophyll in wild-type plants) as well as a reduced chlorophyll and tocopherol content.

Example 5

Over expression of active geranylgeranyl reductase in *Escherichia coli*

For the production of expression clones which overexpress the recombinant CHL P in *E. coli,* the open reading frame for a putatively mature (processed) protein was amplified from the DNA sequence according to SEQ:ID NO. 1 by PCR (1 min 94° C.; 2 min 60° C.; 3 min 72° C.; 25 cycles) using the oligonucleotid primers

COR-AAD1-00001510

CSYN 1 (SEQ ID NO:3) 5'-cgc cat ggg ccg caa tct tcg tgt tgc ggt-3'

and

CSYN 2 (SEQ ID NO:4) 5'-gca gat ctg tcc att tcc ctt ctt agt gca-3'

The amplified PCR fragment was purified and digested with the restriction enzymes NcoI and BglII, and ligated into the expression vector pQE60 (Qiagen, Hilden, Germany), cut with the same enzymes. The initiation codon ATG (forming part of the recognition sequence for restriction enzyme NcoI) was followed by the chl P sequence, starting with nucleotide 148 of the open reading frame of the CHL P cDNA sequence. As a consequence, the methionine is followed by a glycine, corresponding to amino acid residue 50 of the peptide sequence deduced from the cDNA sequence.

For expression of the plant CHL P, *E. coli* strains XL 1 Blue (Stratagene, LaJolla, Calif., USA) or SG 13009 (Gottesmann et al. (1981) J. Bacteriol. 148, 256–273) were transformed with the recombinant vector. After induction of transcription of the recombinant gene by addition of IPTG (a protein having a molecular weight of approximately 47 kDa was expressed in the *E. coli* strains. The protein could be detected in the pellet fraction of the bacterial extract and was purified from the total extract under denaturating conditions using a Ni-affinity column according to manufacturer's instructions (Qiagen, Hilden, Germany).

The purified protein was injected into rabbits for immunisation.

The protein, purified from the total extract, was confirmed to have geranylgeranyl reductase acitivity in a combined enzyme assay with bacterial bacteriochlorophyll synthase by using chlorophyllide and GGPP. The enzyme assay was carried out according to the protocol in Oster et al. (1997) J. Biol. Chem. 272, 9671–9676.

The pigments chlorophyll-GG and chlorophyll-phytol were separated by HPLC on a column (4×250 nm) filled with RP 18 Gromsel 120 ODS5, at 1.2 ml/min flow rate with the following gradient consisting of 60% acetone (solvent A) and 100% acetone (solvent B):t_0 75% solvent A and 25% solvent B, 2 min; within t_{2-4} to 45% solvent A and 55% solvent B; within t_{4-13} to 30% solvent A and 70% solvent B; within t_{13-17} to 100% solvent B; t_{17-21} 100% solvent B isocratic; subsequently within 5 min. to 75% solvent A and 25% solvent B; then further 5 min. 75% solvent A and 25% solvent B isocratic. Tetrapyrroles were detected by use of a fluorescence detector (λ_{ex} 425 nm, λ_{em} 665 nm).

Example 6

Co-expression of the CHL P gene and the HPD gene in *Nicotiana tabacum*

By use of the oligonucleotid primers

hpdoli1 (SEQ ID NO:5) 5'-tta ggt acc atg ggc cac caa acc gcc gcc gtt tca g-3'

and

hpdoli2 (SEQ ID NO:6) 5'-tga gtc gac cac aat cct tta gtt ggt tct tct tct tg-2'.

the sequence of the HPD cDNA (Accession No.: AF 000228) between nucleotide 37 and 1404 was amplified from an *Arabidopsis thaliana* cDNA library, cloned and sequenced. The amplified fragment was digested with restriction endonucleases KpnI and SalI, and ligated into the binary vector Bin-Hyg-TX, also digested with KpnI and SalI. The vector Bin-Hyg-TX is a pBIB derivative (Becker, supra), as is the vector BinAR-TX, used in Example 2), which enables expression of a coding region, ligated into the multiple cloning site, under control of the 35S RNA promoter of cauliflower mosaic virus. In contrast to the vector Bin-Hyg-TX, used for expression of the CHL P gene, which allows selection in plant cells against kanamycine, the binary vector Bin-Hyg-TX carries a hygromycine resistance gene as the selectable marker for plant cells. For the purpose of illustration, a restriction map of the vector Bin-Hyg-TX is provided in FIG. **5**.

Instead of the mentioned binary vector Bin-Hyg-TX any vector suitable for the transformation of plants can be used for constructing a chimeric gene, comprising a fusion of the CaMV 35S promoter or any other promoter, which provides transcription and translation in plant cells, and DNA sequences coding for HPD.

Then, the transfer of the resulting recombinant vector pBinHygHPD onto tobacco SNN was carried out via *Agrobacterium tumefaciens,* as described in the above Example 2, whereby transgenic tobacco shoots were selected on hygromycine-containing medium. Plants obtained after regeneration were used as control plants.

In addition, the transformants 28 and 30, described in Example 2, as well as further transformants, which overexpress the CHL P gene under control of the 35S RNA promoter, were transformed via leaf disc transformation with the recombinant vector pBinHygHPD, and intact plants were regenerated under selection on medium containing both kanamycine and hygromycine.

Successful transfer and expression of the chimeric HPD gene was confirmed in all transgenic plants by suitable Southern and Northern blot hybridisation experiments, using the above-mentioned PCR fragment coding for HPD as the probe.

A comparison of the tocopherol contents (analysis as described in Example 3) in leaves of transgenic plants which only express the CHL P gene (see Example 3, transformants 28 and 30), and plants which co-express the HPD gene and the CHL P gene (transformants 28+HPD and 30+HPD) revealed that the tocopherol content could be further increased by simultaneous expression of the hydroxyphenyl pyruvate dioxygenase gene.

Alternatively transgenic plants which express the CHL P gene as well as the HPD gene can also be obtained by crossing suitable transgenic lines, particularly crossing homozygous "CHL P lines" with homozygous "HPD lines".

An additional increase in tocopherol content can be expected in the double transformants (and double crosses) under stress conditions (e.g. increased temperatures, light stress and the like).

Apart from the above-described double transformants, which express HPD and CHL P, also transgenic HPD lines were analysed with respect to tocopherol content and the influene of stress conditions on their tocopherol content. It could be demonstrated that over-expression of HPD in tobacco leaves results in a 2- to 3-fold increase in tocopherol content, and that the tocopherol content is particularly increased under stress conditions, compared to non-transgenic control plants.

FIG. **6** shows tocopherol contents in the leaves 4, 7 and 10 of 12-week old transgenic tobacco lines (# 2, 6 and 33) which express the Arabidopsis enzyme hydroxyphenyl pyruvate dioxygenase (HPD), in comparison to control plants (SNN). The plants were cultivated either at 38° C. or at 10° C., and a light intensity of approximately 200 μmol photon/m^2/s. Tocopherol is increasingly accumulated in the plants with increasing leaf age. Particularly in older leaves, the tocopherol content increases much more in the plants cultivated at 38° C., and reaches a 2 times higher amount in

COR-AAD1-00001511

comparison to control plants. In comparison, the tocopherol contents in the transformants kept at low temperature was only slightly increased in comparison to wild-type plants.

These results indicate that in the case of an increased need for tocopherol, e.g. particularly under stress conditions, HPD specially favours the generation of these antioxidants in transgenic plants.

Example 7

Increased resistance of the transgenic plants against oxidative stress

Transgenic plants produced according to Examples 2 and 3 were analysed in leaf disc incubation experiments with respect to the influence of inhibitory and oxidative substances. 15 seedlings were incubated in 20 mM potassium phosphate buffer (pH 7.1) for 10 hrs in light. Plants were incubated in control samples (water), in samples with 3.3 μM (low concentration=LC) or 33 μM (high concentration=HC) Acifluorfen (available from BASF, Ludwigshafen, Germany), an inhibitor of protoporphyrinogen oxidase, and in samples with 1.7 μM (low concentration=LC) or 17 μM (high concentration=LC) Rose Bengal (acid red, 4,5,6,7-tetrachloro-2',4',5',7'-tetraiodofluorescein, available from Sigma-Aldrich, Deisenhofen, Germany), which produces reactive oxygen species. The content of tocopherol is about 2 to 3 times higher in the analysed transgenic lines (28, 30) in the buffer control samples as well as under oxidative stress conditions, compared to wild-type plants (SNN). The results are given in FIG. **7**.

The results indicate that due to their increased tocopherol content, as compared to wild-type plants, the plants according to the invention exhibit increased tolerance against oxidative stress. Therefore, an increased antioxidative protection of cellular membranes against reactive oxygen species can be expected in the plants according to the invention.

Example 8

Tocopherol content in seeds of transgenic plants

Tocopherol was extracted according to the above-described procedure from seeds of transgenic plants which exhibit an increased tocopherol content in leaves compared to wild-type plants due to the expression of CHL P under control of the CaMV 35S promoter (see Example 2). The results of the tocopherol quantification by means of HPLC are shown in FIG. **8**, in comparison to tobacco control plants.

Besides α-tocopherol also γ-tocopherol was quantified. The latter is generally present in tobacco seeds in higher amounts; the ratio of γ-tocopherol to α-tocopherol is about 10:1 in tobacco seeds. The contents of both forms of tocopherol, particularly α-tocopherol, are 2- to 3-times higher in transgenic plants having increased geranylgeranyl reductase expression, in comparison to control plants.

These results which reflect the impact of constitutive CHL P expression under control of the 35S promoter on the tocopherol content in seed, clearly indicate that by use of a seed-specific promoter (or a promoter which is specifically inducible in seed tissue) expression of geranylgeranyl reductase in seed tissue and as a consequence the tocopherol content in seeds of transgenic plants can be further increased.

Should a molecular biological procedure in any way not have been adequately described, it was carried out following standard methods, as described, for example, by Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. With respect to the transformation of plants it is referred to generally known review articles as well as the reports and publications mentioned herein.

Description of the Figures

FIG. **1**: SEQ ID NO:1 shows a nucleotide sequence of the chl P cDNA of geranylgeranyl reductase (CHL P) from *Nicotiana tabacum.*

FIG. **2**: SEQ ID NO:2 shows an amino acid sequence of the enzyme CHL P from *N. tabacum,* deduced from the SEQ ID NO:1 shown in FIG. **1**

FIG. **3**: Restriction map of the binary vector BinAR, as used in plant transformation experiments. BinAR (Höfgen and Willmitzer (1990) Plant Science 66, 221) is a Bin19 derivative (Bevan (1984) Nucl. Acids Res. 12, 8711), which contains an expression cassette for the constitutive expression of chimeric genes in plants, the expression cassette being cloned via the EcoRI and HindIII restriction sites of Bin19. The cassette comprises a 770 base pair EcoRI/HindIII fragment, containing the CaMV 35S promoter, part of the pUC18 polylinker as well as the termination signal of the octopine synthase gene (OCS). For insertion of coding sequences the unique restriction sites of the pUC18 polylinker, i.e. KpnI, SmaI, BamHI, XbaI and SalI are particularly useful. As plant selection marker, the binary vector BinAR carries a kanamycine resistence gene.

FIG. **4**: Bar diagram showing the tocopherol content in the leaves 6, 9, 12 (numbering starting from the top of the plant) of transgenic tobacco plants (lines 28 and 30) versus the corresponding leaves of control plants (SNN).

FIG. **5**: Restriction map of the binary vector Bin-Hyg-TX, as used for plant transformation, which is also a pBIB-derivative (Becker, supra; Bevan, supra), containing an expression cassette for the constitutive expression of chimeric genes in plants. For insertion of coding sequences the unique restriction sites of the pUC18 polylinker, i.e. HpaI, KpnI, SmaI, XbaI and SalI are especially useful. As the plant selection marker, the binary vector Bin-Hyg-TX carries a hygromycine resistence gene.

FIG. **6**: Bar diagram of the tocopherol contents in the leaves 4, 7 and 10 in 12-weeks old transgenic lines (# 2, 6, 33), which express the Arabidopsis enzyme HPD, and in control plants (SNN).

FIG. **7**: Tocopherol content in 12-days old seedlings, incubated for 10 hrs at light in 20 mM potassium phosphate buffer (pH 7.1; control=water), with 3.3 μM (LC) or 33 μM (HC) Acifluorfen, and with 1.7 μM (LC) or 17 μM (HC) Rose Bengal (SNN=wild-type control plants, 28 and 39=transgenic lines).

FIG. **8**: Relative α-tocopherol and γ-tocopherol values in seeds of transgenic tobacco plants (# 7, 39, 67, 96) and wild-type tobacco plants (SNN). 100% α-tocopherol corresponds to 6 ng/mg seeds; 100% γ-tocopherol corresponds to 62.8 ng/mg seeds.

COR-AAD1-00001512

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 6

<210> SEQ ID NO 1
<211> LENGTH: 1510
<212> TYPE: DNA
<213> ORGANISM: Nicotiana tabacum

<400> SEQUENCE: 1

atggcttcca ttgctctcaa aactttcacc ggcctccgtc aatcctcgcc ggaaaacaat 60

tccattactc tttctaaatc cctccccttc acccaaaccc accgtaggct ccgaatcaat 120

gcttccaaat ccagcccaag agtcaacggc cgcaatcttc gtgttgcggt ggtgggcggt 180

ggtcctgctg gtggcgccgc cgctgaaaca ctcgccaagg gaggaattga aaccttctta 240

atcgaacgca aaatggacaa ctgcaaaccc tgcggtgggg ccatcccact ttgcatggtg 300

ggagaatttg acctcccttt ggatatcatt gaccggaaag ttacaaagat gaagatgatt 360

tccccatcca acgttgctgt tgatattggt cagactttaa agcctcacga gtacatcggt 420

atggtgcgcc gcgaagtact cgatgcttac ctccgtgacc gcgctgctga agccggagcc 480

tctgttctca acggcttgtt cctcaaaatg gacatgccca aagctcccaa cgcaccttac 540

gtccttcact acacagctta cgactccaaa actaatggcg cgggggagaa gcgtaccctg 600

gaagttgacg ccgttatcgg cgctgacggt gcaaattccc gtgtcgcaaa atccataaac 660

gccggtgact acgagtacgc tattgcattc caagaaagga ttaaaatttc cgatgataaa 720

atgaagtatt acgagaattt agctgaaatg tacgtgggtg atgacgtgtc ccctgatttt 780

tacgggtggg ttttccccaa atgtgaccac gttgccgttg gcactggcac agtcacccac 840

aaagctgaca tcaaaaaatt ccagctagct acaagattga gagctgattc caaaatcacc 900

ggcggaaaaa ttatccgggt cgaggcccac ccgattccag aacacccaag acccagaaga 960

ttacaagaca gagttgcatt ggttggtgat gcggcagggt acgtgaccaa atgttcgggc 1020

gaagggattt acttcgcggc aaagagtgga cgtatgtgtg ctgaagcaat tgttgaaggg 1080

tcagaaatgg gaaaaagaat ggtggacgag agtgatttga ggaagtattt ggagaaatgg 1140

gacaagactt attggccaac gtacaaggtg cttgatatat tgcagaaggt attttacagg 1200

tcgaatccgg cgagggaagc atttgttgaa atgtgcgcag atgagtatgt gcagaagatg 1260

acatttgaca gctatttgta caagaaagta gcaccaggaa acccaattga agacttgaag 1320

cttgctgtga ataccattgg aagtttggtg agagctaatg cactaagaag ggaaatggac 1380

aagctcagtg tataagaaga ttaacagcat taatattttc ttgtaattga aggatttatt 1440

tctcaaatta ctctgtaaac acctttcatc ctgcctttaa tcggatttat gtaacttcat 1500

aatttgagct 1510

<210> SEQ ID NO 2
<211> LENGTH: 464
<212> TYPE: PRT
<213> ORGANISM: Nicotiana tabacum

<400> SEQUENCE: 2

Met Ala Ser Ile Ala Leu Lys Thr Phe Thr Gly Leu Arg Gln Ser Ser
1 5 10 15

Pro Glu Asn Asn Ser Ile Thr Leu Ser Lys Ser Leu Pro Phe Thr Gln
20 25 30

-continued

```
Thr His Arg Arg Leu Arg Ile Asn Ala Ser Lys Ser Ser Pro Arg Val
        35                  40                  45

Asn Gly Arg Asn Leu Arg Val Ala Val Val Gly Gly Gly Pro Ala Gly
    50                  55                  60

Gly Ala Ala Ala Glu Thr Leu Ala Lys Gly Gly Ile Glu Thr Phe Leu
65                  70                  75                  80

Ile Glu Arg Lys Met Asp Asn Cys Lys Pro Cys Gly Gly Ala Ile Pro
                85                  90                  95

Leu Cys Met Val Gly Glu Phe Asp Leu Pro Leu Asp Ile Ile Asp Arg
            100                 105                 110

Lys Val Thr Lys Met Lys Met Ile Ser Pro Ser Asn Val Ala Val Asp
        115                 120                 125

Ile Gly Gln Thr Leu Lys Pro His Glu Tyr Ile Gly Met Val Arg Arg
    130                 135                 140

Glu Val Leu Asp Ala Tyr Leu Arg Asp Arg Ala Ala Glu Ala Gly Ala
145                 150                 155                 160

Ser Val Leu Asn Gly Leu Phe Leu Lys Met Asp Met Pro Lys Ala Pro
                165                 170                 175

Asn Ala Pro Tyr Val Leu His Tyr Thr Ala Tyr Asp Ser Lys Thr Asn
            180                 185                 190

Gly Ala Gly Glu Lys Arg Thr Leu Glu Val Asp Ala Val Ile Gly Ala
        195                 200                 205

Asp Gly Ala Asn Ser Arg Val Ala Lys Ser Ile Asn Ala Gly Asp Tyr
    210                 215                 220

Glu Tyr Ala Ile Ala Phe Gln Glu Arg Ile Lys Ile Ser Asp Asp Lys
225                 230                 235                 240

Met Lys Tyr Tyr Glu Asn Leu Ala Glu Met Tyr Val Gly Asp Asp Val
                245                 250                 255

Ser Pro Asp Phe Tyr Gly Trp Val Phe Pro Lys Cys Asp His Val Ala
            260                 265                 270

Val Gly Thr Gly Thr Val Thr His Lys Ala Asp Ile Lys Lys Phe Gln
        275                 280                 285

Leu Ala Thr Arg Leu Arg Ala Asp Ser Lys Ile Thr Gly Gly Lys Ile
    290                 295                 300

Ile Arg Val Glu Ala His Pro Ile Pro Glu His Pro Arg Pro Arg Arg
305                 310                 315                 320

Leu Gln Asp Arg Val Ala Leu Val Gly Asp Ala Ala Gly Tyr Val Thr
                325                 330                 335

Lys Cys Ser Gly Glu Gly Ile Tyr Phe Ala Ala Lys Ser Gly Arg Met
            340                 345                 350

Cys Ala Glu Ala Ile Val Glu Gly Ser Glu Met Gly Lys Arg Met Val
        355                 360                 365

Asp Glu Ser Asp Leu Arg Lys Tyr Leu Glu Lys Trp Asp Lys Thr Tyr
    370                 375                 380

Trp Pro Thr Tyr Lys Val Leu Asp Ile Leu Gln Lys Val Phe Tyr Arg
385                 390                 395                 400

Ser Asn Pro Ala Arg Glu Ala Phe Val Glu Met Cys Ala Asp Glu Tyr
                405                 410                 415

Val Gln Lys Met Thr Phe Asp Ser Tyr Leu Tyr Lys Lys Val Ala Pro
            420                 425                 430

Gly Asn Pro Ile Glu Asp Leu Lys Leu Ala Val Asn Thr Ile Gly Ser
        435                 440                 445
```

-continued

```
Leu Val Arg Ala Asn Ala Leu Arg Arg Glu Met Asp Lys Leu Ser Val
    450                 455                 460


<210> SEQ ID NO 3
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A primer

<400> SEQUENCE: 3

cgccatgggc cgcaatcttc gtgttgcggt                                    30


<210> SEQ ID NO 4
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A primer

<400> SEQUENCE: 4

gcagatctgt ccatttccct tcttagtgca                                    30


<210> SEQ ID NO 5
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A primer

<400> SEQUENCE: 5

ttaggtacca tgggccacca aaccgccgcc gtttca                             36


<210> SEQ ID NO 6
<211> LENGTH: 38
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A primer

<400> SEQUENCE: 6

tgagtcgacc acaatccttt agttggttct tcttcttg                           38
```

What is claimed is:

1. A transgenic plant transformed with a nucleic acid sequence encoding the protein of SEQ ID NO:2, wherein the protein has the activity of a geranylgeranyl reductase, wherein the nucleic acid sequence is transcribed in said transgenic plant.

2. The plant according to claim **1**, wherein the protein having the activity of a geranylgeranyl reductase is expressed in the transgenic plant.

3. The plant according to claim **1**, which exhibits an increased tocopherol content in comparison to a wild-type plant.

4. The plant according to claim **1**, wherein said plant is a dicot.

5. The plant according to claim **4**, which is a crop plant, food and/or forage plant, selected from the group consisting of rape, soybean, tomato, potato, sugar beet, and clover.

6. The plant according to claim **1**, wherein said plant is a monocot.

7. The plant according to claim **6**, which is selected from the group consisting of a crop plant, a forage plant, a cereal plant a wheat, barley, maize, oats, rye, rice, a sweet grass and pasture grass plant.

8. The plant according to claim **1**, wherein the nucleic acid sequence is integrated into the plant genome.

9. A transgenic plant cell, which contains an isolated nucleic acid sequence coding for the protein of SEQ ID NO:2 and having the activity of a geranylgeranyl reductase, wherein the nucleic acid sequence is transcribed in the transgenic plant cell.

10. A plant cell according to claim **9**, which exhibits an increased tocopherol content in comparison to a non-transformed plant cell.

11. A method for producing a plant or plant cell comprising the following steps:

a) producing a nucleic acid sequence, comprising the following constituents, fused in 5'-3' orientation:
   - a promoter functional in plants, preferably a seed-specific promoter,
   - a nucleic acid sequence which codes for the protein of SEQ ID NO:2, wherein the protein has geranylgeranyl reductase activity, and
   - optionally, a terminal signal for the termination of transcription and addition of a poly-A-tail to the corresponding transcript, as well as, optionally, DNA sequences deduced therefrom;

b) transferring the nucleic acid sequence from step a) to plant cells and, optionally, integration of the nucleic acid sequence into the plant genome

c) if desired, regenerating completely transformed plants, and optionally, propagating the plants.

12. A transgenic plant produced according to the method of claim **11**.

13. A method of obtaining tocopherols comprising obtaining the transgenic plant of claim **1** and isolating tocopherols from the transgenic plant.

14. The method of claim **11**, wherein the plant or plant cell exhibits an increased tocopherol content.

15. The plant according to claim **1** wherein the nucleic acid sequence encodes a protein from tobacco.

16. The plant according to claim **1** wherein the nucleic acid sequence comprises SEQ ID NO:1.

17. The plant according to claim **1** wherein the nucleic acid sequence further encodes a regulatory element.

18. The plant according to claim **1** wherein the nucleic acid sequence further comprises a promoter.

19. The plant according to claim **18** wherein the promoter is a constitutive promoter.

20. The plant according to claim **18** wherein the promoter is an inducible promoter.

21. The plant according to claim **18** wherein the promoter is a tissue-specific promoter.

22. The plant according to claim **18** wherein the promoter is a developmental specific promoter.

23. The plant according to claim **18** wherein the promoter is a seed-specific promoter.

24. The plant according to claim **1** wherein the nucleic acid sequence further encodes enhancer sequences.

25. The plant according to claim **1** wherein the nucleic acid sequence further encodes signal peptides.

26. The plant according to claim **1** wherein the nucleic acid sequence is in anti-sense orientation.

27. A plant cell of the transgenic plant of claim **1**, wherein the plant cell comprises the nucleic acid sequence encoding the protein of SEQ ID NO:2.

28. Progeny of the transgenic plant of claim **1**, wherein each of the progeny comprises the nucleic acid sequence encoding the protein of SEQ ID NO:2.

29. A seed produced by the transgenic plant of claim **1**, wherein the seed comprises the nucleic acid sequence encoding the protein of SEQ ID NO:2.

30. Transgenic propagation material derived from the plant of claim **1**, wherein the material comprises an isolated nucleic acid sequence encoding a protein comprising SEQ ID NO:2.

31. The transgenic propagation material of claim **30**, that is a protoplast, a callus, or a cutting.

* * * * *

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 535 414 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
19.12.2012 Bulletin 2012/51

(21) Application number: **12179273.3**

(22) Date of filing: **29.04.2004**

(51) Int Cl.:
C12N 15/54 *(2006.01)* ***C12N 15/82*** *(2006.01)*
C12N 9/10 *(2006.01)* ***A01H 5/00*** *(2006.01)*

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LI LU MC NL PL PT RO SE SI SK TR

(30) Priority: **29.04.2003 CL 8632003**
30.04.2003 US 427692

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
11153914.4 / 2 322 629
04775913.9 / 1 620 557

(71) Applicants:
- **Pioneer Hi-Bred International Inc.**
  Des Moines, IA 50309 (US)
- **Verdia, Inc.**
  Redwood City, CA 94063 (US)
- **E. I. du Pont de Nemours and Company**
  Wilmington, DE 19898 (US)

(72) Inventors:
- **Castle, Linda A.**
  Moutain View, CA California 94040 (US)
- **Siehl, Dan**
  Menlo Park, CA California 94025 (US)
- **Giver, Lorraine J.**
  Sunnyvale, CA California 94085 (US)
- **Minshull, Jeremy**
  Los Altos, CA California 94022 (US)
- **Ivy, Christina**
  Encinitas, CA California 92024 (US)
- **Chen, Yong Hong**
  Foster City, CA California 94404 (US)
- **Patten, Phillip A.**
  Menlo Park, CA California 94028 (US)
- **Gorton, Rebecca**
  San Francisco, CA California 94102 (US)
- **Duck, Nicholas B.**
  Apex, NC North Carolina 27502 (US)
- **McCutchen, Billy Fred**
  Clive, IA Iowa 50325 (US)
- **Kemble, Roger**
  Wake Forest, NC North Carolina 27587 (US)

(74) Representative: **Bentham, Andrew**
J A Kemp
14 South Square
Gray's Inn
London WC1R 5JJ (GB)

Remarks:
This application was filed on 03-08-2012 as a divisional application to the application mentioned under INID code 62.

(54) Novel glyphosate-n-acetyltransferase (gat) genes

(57) Novel proteins are provided herein, including proteins capable of catalyzing the acetylation of glyphosate and other structurally related proteins. Also provided are novel polynucleotides capable of encoding these proteins, compositions that include one or more of these novel proteins and/or polynucleotides, recombinant cells and transgenic plants comprising these novel compounds, diversification methods involving the novel compounds, and methods of using the compounds. Some of the novel methods and compounds provided herein can be used to render an organism, such as a plant, resistant to glyphosate

EP 2 535 414 A1

Printed by Jouve, 75001 PARIS (FR)

Description

COPYRIGHT NOTIFICATION PURSUANT TO 37 C.F.R. § 1.71(E)

[0001] A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

BACKGROUND OF THE INVENTION

[0002] Crop selectivity to specific herbicides can be conferred by engineering genes into crops which encode appropriate herbicide metabolizing enzymes. In some cases these enzymes, and the nucleic acids that encode them, originate in a plant. In other cases, they are derived from other organisms, such as microbes. *See, e.g.*, Padgette et al. (1996) "New weed control opportunities: Development of soybeans with a Round UP Ready™ gene" and Vasil (1996) "Phosphinothricin-resistant crops", both in Herbicide-Resistant Crops, ed. Duke (CRC Press, Boca Raton, Florida) pp.54-84 and pp. 85-91. Indeed, transgenic plants have been engineered to express a variety of herbicide tolerance/metabolizing genes, from a variety of organisms. For example, acetohydroxy acid synthase, which has been found to make plants that express this enzyme resistant to multiple types of herbicides, has been introduced into a variety of plants (*see, e.g.,* Hattori et al. (1995) Mol. Gen. Genet. 246: 419). Other genes that confer tolerance to herbicides include: a gene encoding a chimeric protein of rat cytochrome P4507A1 and yeast NADPH-cytochrome P450 oxidoreductase (Shiota et al. (1994) Plant Physiol. 106: 17), genes for glutathione reductase and superoxide dismutase (Aono et al. (1995) Plant Cell Physiol. 36: 1687, and genes for various phosphotransferases (Datta et al. (1992) Plant Mol. Biol. 20: 619).

[0003] One herbicide which is the subject of much investigation in this regard is N-phosphonomethylglycine, commonly referred to as glyphosate. Glyphosate is the top selling herbicide in the world, with sales projected to reach $5 billion by 2003. It is a broad spectrum herbicide that kills both broadleaf and grass-type plants. A successful mode of commercial level glyphosate resistance in transgenic plants is by introduction of a modified Agrobacterium CP4 5-enolpyruvylshikimate-3-phosphate synthase (hereinafter referred to as EPSP synthase or EPSPS) gene. The transgene is targeted to the chloroplast where it is capable of continuing to synthesize EPSP from phosphoenolpyruvic acid (PEP) and shikimate-3-phosphate in the presence of glyphosate. In contrast, the native EPSP synthase is inhibited by glyphosate. Without the transgene, plants sprayed with glyphosate quickly die due to inhibition of EPSP synthase which halts the downstream pathway needed for aromatic amino acid, hormone, and vitamin biosynthesis. The CP4 glyphosate-resistant soybean transgenic plants are marketed, e.g., by Monsanto under the name "Round UP Ready™."

[0004] In the environment, the predominant mechanism by which glyphosate is degraded is through soil microflora metabolism. The primary metabolite of glyphosate in soil has been identified as aminomethylphosphonic acid (AMPA), which is ultimately converted into ammonia, phosphate and carbon dioxide. The proposed metabolic scheme that describes the degradation of glyphosate in soil through the AMPA pathway is shown in Fig. 8. An alternative metabolic pathway for the breakdown of glyphosate by certain soil bacteria, the sarcosine pathway, occurs via initial cleavage of the C-P bond to give inorganic phosphate and sarcosine, as depicted in Fig. 9.

[0005] Another successful herbicide/transgenic crop package is glufosinate (phosphinothricin) and the Liberty Link™ trait marketed, *e.g.,* by Aventis. Glufosinate is also a broad spectrum herbicide. Its target is the glutamate synthase enzyme of the chloroplast. Resistant plants carry the bar gene from *Streptomyces hygroscopicus* and achieve resistance by the N-acetylation activity of bar, which modifies and detoxifies glufosinate.

[0006] An enzyme capable of acetylating the primary amine of AMPA is reported in PCT Application No. WO00/29596. The enzyme was not described as being able to acetylate a compound with a secondary amine (e.g., glyphosate).

[0007] While a variety of herbicide resistance strategies are available as noted above, additional approaches would have considerable commercial value. The present invention provides novel polynucleotides and polypeptides for conferring herbicide tolerance, as well as numerous other benefits as will become apparent during review of the disclosure.

SUMMARY OF THE INVENTION

[0008] The present invention provides methods and reagents for rendering an organism, such as a plant, resistant to glyphosate by one or more of the embodiments described below.

[0009] One embodiment of the invention provides novel polypeptides referred to herein as glyphosate-N-acetyltransferase ("GAT") polypeptides. GAT polypeptides are characterized by their structural similarity to one another, e.g., in terms of sequence similarity when the GAT polypeptides are aligned with one another. GAT polypeptides of the present invention possess glyphosate-N-acetyltransferase activity, i.e., the ability to catalyze the acetylation of glyphosate. These GAT polypeptides transfer the acetyl group from acetyl CoA to the N of glyphosate. In addition, some GAT polypeptides

COR-AAD1-00001518

transfer the propionyl group of propionyl CoA to the N of glyphosate. Some GAT polypeptides are also capable of catalyzing the acetylation of glyphosate analogs and/or glyphosate metabolites, *e.g.*, aminomethylphosphonic acid. Exemplary GAT polypeptides correspond to SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972.

[0010] Also provided are novel polynucleotides referred to herein as GAT polynucleotides, e.g., SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952. GAT polynucleotides are characterized by their ability to encode GAT polypeptides. In some embodiments of the invention, a GAT polynucleotide is engineered for better plant expression by replacing one or more parental codons with a synonymous codon that is preferentially used in plants relative to the parental codon. In other embodiments, a GAT polynucleotide is modified by the introduction of a nucleotide sequence encoding an N-terminal chloroplast transit peptide. In other embodiments, a GAT polynucleotide is modified by the insertion of one or more G+C containing codons (such as GCG or GCT) immediately downstream of and adjacent to the initiating Met codon.

[0011] GAT polypeptides, GAT polynucleotides and glyphosate-N-acetyltransferase activity are described in more detail below. The invention further includes certain fragments of the GAT polypeptides and GAT polynucleotides described herein.

[0012] The invention includes non-native variants of the polypeptides and polynucleotides described herein, wherein one or more amino acid of the encoded polypeptide has been mutated.

[0013] In certain preferred embodiments, the GAT polypeptides of the present invention are characterized as follows. When optimally aligned with a reference amino acid sequence selected from the group consisting of SEQ ID NO: 300, 445, and 457 to generate a similarity score of at least 460 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, one or more of the following positions conform to the following restrictions: (i) at positions 18 and 38, a Z5 amino acid residue; (ii) at position 62, a Z1 amino acid residue; (iii) at position 124, a Z6 amino acid residue; and (iv) at position 144, a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P.

[0014] The invention further provides an isolated or recombinant polypeptide comprising an amino acid sequence selected from the groups consisting of: (a) an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) an amino acid sequence that is at least 98% identical to SEQ NO:579; (e) an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) an amino acid sequence that is at least 95% identical to SEQ ID NO:697; (g) an amino acid sequence that is at least 96% identical to SEQ ID NO:712; (h) an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (1) an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) an amino acid sequence that is at least 96% identical to SEQ ID NO:612;and (n) an amino acid sequence that is at least 98% identical to SEQ ID NO:590.

[0015] The invention further provides an isolated or recombinant polypeptide comprising an amino acid sequence selected from the groups consisting of: (a) an amino acid sequence that is at least 96% identical to positions 2-146 of SEQ ID NO:919 (such as, for example, SEQ ID NO:917, 919, 921, 923, 925, 927, 833, 835, 839, 843, 845, 859, 863, 873, 877, 891, 895, 901, 905, 907, 913, 915, or 950); (b) an amino acid sequence that is at least 97% identical to positions 2-146 of SEQ ID NO:929 (such as, for example, SEQ ID NO:929, 931, 835, 843, 849, or 867); (c) an amino acid sequence

COR-AAD1-00001519

that is at least 98% identical to positions 2-146 of SEQ ID NO:847 (such as, for example, SEQ ID NO:845 or 847); (d) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:851; (e) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:853; (f) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:855 (such as, for example, SEQ ID NO:835 or 855); (g) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:857; (h) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:861 (such as, for example, SEQ ID NO:839, 861, or 883); (i) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:871; (j) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:875; (k) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:881; (1) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO: 885 (such as, for example, SEQ ID NO:845 or 885); (m) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:887; (n) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO: 889 (such as, for example, SEQ ID NO: 863, 889, 891, or 903); (o) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:893; (p) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:897; (q) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:899; (r) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:909 (such as, for example, SEQ ID NO:883 or 909); (s) an amino acid sequence that is at least 98% identical to positions 2-146 of SEQ ID NO:911; (t) an amino acid sequence that is at least 99% identical to positions 2-146 of SEQ ID NO:837; (u) an amino acid sequence that is at least 99% identical to positions 2-146 of SEQ ID NO:841; (v) an amino acid sequence that is at least 99% identical to positions 2-146 of SEQ ID NO:865; (w) an amino acid sequence that is at least 99% identical to positions 2-146 of SEQ ID NO:869; and (x) an amino acid sequence that is at least 99% identical to positions 2-146 of SEQ ID NO:879. In some embodiments of the invention, the amino acid sequence of the polypeptide comprises Met, Met-Ala, or Met-Ala-Ala on the N-terminal side of the amino acid corresponding to position 2 of the reference amino acid sequence.

[0016] The invention further provides an isolated or recombinant polypeptide comprising an amino acid sequence that is at least 95% identical to positions 2-146 of SEQ ID NO:929 and which comprises a Gly or an Asn residue at the amino acid position corresponding to position 33 of SEQ ID NO:929 (such as, for example, SEQ ID NO:837, 849, 893, 897, 905, 921, 927, 929 or 931). In some embodiments of the invention, the amino acid sequence of the polypeptide comprises Met, Met-Ala, or Met-Ala-Ala on the N-terminal side of the amino acid corresponding to position 2 of the reference amino acid sequence.

[0017] The invention further provides a nucleic acid construct comprising a polynucleotide of the invention. The construct can be a vector, such as a plant transformation vector. In some aspects a vector of the invention will comprise a T-DNA sequence. The construct can optionally include a regulatory sequence (*e.g.,* a promoter) operably linked to a GAT polynucleotide, where the promoter is heterologous with respect to the polynucleotide and effective to cause sufficient expression of the encoded polypeptide to enhance the glyphosate tolerance of a plant cell transformed with the nucleic acid construct.

[0018] In some aspects of the invention, a GAT polynucleotide functions as a selectable marker, *e.g.,* in a plant, bacteria, actinomycete, yeast, algae or other fungi. For example, an organism that has been transformed with a vector including a GAT polynucleotide selectable marker can be selected based on its ability to grow in the presence of glyphosate. A GAT marker gene can be used for selection or screening for transformed cells expressing the gene.

[0019] The invention further provides vectors with stacked traits, *i.e.,* vectors that encode a GAT polypeptide and that also include a second polynucleotide sequence encoding a second polypeptide that confers a detectable phenotypic trait upon a cell of organism expressing the second polypeptide at an effective level, for example disease resistance or pest resistance. The detectable phenotypic trait can also function as a selectable marker, *e.g.,* by conferring herbicide resistance or by providing some sort of visible marker.

[0020] In one embodiment, the invention provides a composition comprising two or more polynucleotides of the invention. Preferably, the GAT polynucleotides encode GAT polypeptides having different kinetic parameters, i.e., a GAT variant having a lower K_m can be combined with one having a higher k_{cat}. In a further embodiment, the different GAT polynucleotides may be coupled to a chloroplast transit sequence or other signal sequence thereby providing GAT polypeptide expression in different cellular compartments, organelles or secretion of one or more of the GAT polypeptides.

[0021] Accordingly, compositions containing two or more GAT polynucleotides or encoded polypeptides are a feature of the invention. In some cases, these compositions are libraries of nucleic acids containing, e.g., at least 3 or more such nucleic acids. Compositions produced by digesting the nucleic acids of the invention with a restriction endonuclease, a DNAse or an RNAse, or otherwise fragmenting the nucleic acids, e.g., mechanical shearing, chemical cleavage, *etc.*, are also a feature of the invention, as are compositions produced by incubating a nucleic acid of the invention with deoxyribonucleotide triphosphates and a nucleic acid polymerase, such as a thermostable nucleic acid polymerase.

[0022] Cells transduced by a vector of the invention, or which otherwise incorporate a nucleic acid of the invention, are an aspect of the invention. In a preferred embodiment, the cells express a polypeptide encoded by the nucleic acid of the invention.

[0023] In some embodiments, the cells incorporating the nucleic acids of the invention are plant cells. Transgenic

COR-AAD1-00001520

plants, transgenic plant cells, and transgenic plant explants incorporating the nucleic acids of the invention are also a feature of the invention. In some embodiments, the transgenic plants, transgenic plant cells, or transgenic plant explants express an exogenous polypeptide with glyphosate-N-acetyltransferase activity encoded by the nucleic acid of the invention. The invention also provides transgenic seeds produced by the transgenic plants of the invention.

[0024] The invention further provides transgenic plants, transgenic plant cells, transgenic plant explants, or transgenic seeds having enhanced tolerance to glyphosate due to the expression of a polypeptide with glyphosate-N-acetyltransferase activity and a polypeptide that imparts glyphosate tolerance by another mechanism, such as a glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthase and/or a glyphosate-tolerant glyphosate oxido-reductase. In a further embodiment, the invention provides transgenic plants or transgenic plant explants having enhanced tolerance to glyphosate, as well as tolerance to an additional herbicide due to the expression of a polypeptide with glyphosate-N-acetyltransferase activity, a polypeptide that imparts glyphosate tolerance by another mechanism, such as a glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthase and/or a glyphosate-tolerant glyphosate oxido-reductase and a polypeptide imparting tolerance to the additional herbicide, such as a mutated hydroxyphenylpyruvatedioxygenase, a sulfonamide-tolerant acetolactate synthase, a sulfonamide-tolerant acetohydroxy acid synthase, an imidazolinone-tolerant acetolactate synthase, an imidazolinone-tolerant acetohydroxy acid synthase, a phosphinothricin acetyltransferase and a mutated protoporphyrinogen oxidase.

[0025] The invention also provides transgenic plants, transgenic plant cells, transgenic plant explants, or transgenic seeds having enhanced tolerance to glyphosate, as well as tolerance to an additional herbicide due to the expression of a polypeptide with glyphosate-N-acetyltransferase activity and a polypeptide imparting tolerance to an additional herbicide, such as, a mutated hydroxyphenylpyruvatedioxygenase, a sulfonamide-tolerant acetolactate synthase, a sulfonamide-tolerant acetohydroxy acid synthase, an imidazolinone-tolerant acetolactate synthase, an imidazolinone-tolerant acetohydroxy acid synthase, a phosphinothricin acetyltransferase and a mutated protoporphyrinogen oxidase. The invention also provides transgenic plants, transgenic plant cells, transgenic plant explants, or transgenic seeds having enhanced tolerance to glyphosate as well as additional desirable traits which may be conferred by one or more additional transgenes.

[0026] Methods of producing the polypeptides of the invention by introducing the nucleic acids encoding them into cells and then expressing and optionally recovering them from the cells or culture medium are a feature of the invention. In preferred embodiments, the cells expressing the polypeptides of the invention are transgenic plant cells.

[0027] Methods of increasing the expression level of a polypeptide of the invention in a plant or plant cell by inserting into the polypeptide coding sequence one or two G/C-rich codons (such as GCG or GCT) immediately adjacent to and downstream of the initiating methionine ATG codon, and/or substituting in the polypeptide coding sequence one or more codons which are less frequently utilized in plants for codons encoding the same amino acid(s) which are more frequently utilized in plants, and introducing the modified coding sequence into a plant or plant cell and expressing the modified coding sequence, are also a feature of the invention.

[0028] Polypeptides that are specifically bound by a polyclonal antisera that reacts against an antigen derived from SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972 but not to a naturally occurring related sequence, *e.g.,* such as a peptide represented by a subsequence of those of GenBank accession number CAA70664, as well as antibodies which are produced by administering an antigen derived from any one or more of SEQ ID NO:568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972 and/or which bind specifically to such antigens and which do not specifically bind to a naturally occurring polypeptide corresponding to those of GenBank accession number CAA70664, are all features of

COR-AAD1-00001521

the invention.

[0029] Another aspect of the invention relates to methods of polynucleotide diversification to produce novel GAT polynucleotides and polypeptides by recombining or mutating the nucleic acids of the invention *in vitro* or *in vivo*. In an embodiment, the recombination produces at least one library of recombinant GAT polynucleotides. The libraries so produced are embodiments of the invention, as are cells comprising the libraries. Furthermore, methods of producing a modified GAT polynucleotide by mutating a nucleic acid of the invention are embodiments of the invention. Recombinant and mutant GAT polynucleotides and polypeptides produced by the methods of the invention are also embodiments of the invention.

[0030] In some aspects of the invention, diversification is achieved by using recursive recombination, which can be accomplished *in vitro, in vivo, in silico,* or a combination thereof. Some examples of diversification methods described in more detail below are family shuffling methods and synthetic shuffling methods. The invention provides methods for producing a glyphosate-resistant transgenic plant or plant cell that involve transforming a plant or plant cell with a polynucleotide encoding a glyphosate-N-acetyltransferase, and optionally regenerating a transgenic plant from the transformed plant cell. In some aspects the polynucleotide is a GAT polynucleotide, optionally a GAT polynucleotide derived from a bacterial source. In some aspects of the invention, the method can comprise growing the transformed plant or plant cell in a concentration of glyphosate that inhibits the growth of a wild-type plant of the same species without inhibiting the growth of the transformed plant. The method can comprise growing the transformed plant or plant cell or progeny of the plant or plant cell in increasing concentrations of glyphosate and/or in a concentration of glyphosate that is lethal to a wild-type plant or plant cell of the same species. A glyphosate-resistant transgenic plant produced by this method can be propagated, for example by crossing it with a second plant, such that at least some progeny of the cross display glyphosate tolerance.

[0031] The invention further provides methods for selectively controlling weeds in a field containing a crop that involve planting the field with crop seeds or plants which are glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate N-acetyltransferase, and applying to the crop and weeds in the field a sufficient amount of glyphosate to control the weeds without significantly affecting the crop.

[0032] The invention further provides methods for controlling weeds in a field and preventing the emergence of glyphosate-resistant weeds in a field containing a crop which involve planting the field with crop seeds or plants that are glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate-N-acetyltransferase and a gene encoding a polypeptide imparting glyphosate tolerance by another mechanism, such as a glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthase and/or a glyphosate-tolerant glyphosate oxido-reductase and applying to the crop and the weeds in the field a sufficient amount of glyphosate to control the weeds without significantly affecting the crop.

[0033] In a further embodiment the invention provides methods for controlling weeds in a field and preventing the emergence of herbicide resistant weeds in a field containing a crop which involve planting the field with crop seeds or plants that are glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate-N-acetyltransferase, a gene encoding a polypeptide imparting glyphosate tolerance by another mechanism, such as a glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthase and/or a glyphosate-tolerant glyphosate oxido-reductase and a gene encoding a polypeptide imparting tolerance to an additional herbicide, such as a mutated hydroxyphenylpyruvatedioxygenase, a sulfonamide-tolerant acetolactate synthase, a sulfonamide-tolerant acetohydroxy acid synthase, an imidazolinone-tolerant acetolactate synthase, an imidazolinone-tolerant acetohydroxy acid synthase, a phosphinothricin acetyltransferase and a mutated protoporphyrinogen oxidase and applying to the crop and the weeds in the field a sufficient amount of glyphosate and an additional herbicide, such as, a hydroxyphenylpyruvatedioxygenase inhibitor, sulfonamide, imidazolinone, bialaphos, phosphinothricin, azafenidin, butafenacil, sulfosate, glufosinate, and a protox inhibitor to control the weeds without significantly affecting the crop.

[0034] The invention further provides methods for controlling weeds in a field and preventing the emergence of herbicide resistant weeds in a field containing a crop which involve planting the field with crop seeds or plants that are glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate-N-acetyltransferase and a gene encoding a polypeptide imparting tolerance to an additional herbicide, such as a mutated hydroxyphenylpyruvatedioxygenase, a sulfonamide-tolerant acetolactate synthase, a sulfonamide-tolerant acetohydroxy acid synthase, an imidazolinone-tolerant acetolactate synthase, an imidazolinone-tolerant acetohydroxy acid synthase, a phosphinothricin acetyltransferase and a mutated protoporphyrinogen oxidase and applying to the crop and the weeds in the field a sufficient amount of glyphosate and an additional herbicide, such as a hydroxyphenylpyruvatedioxygenase inhibitor, sulfonamide, imidazolinone, bialaphos, phosphinothricin, azafenidin, butafenacil, sulfosate, glufosinate, and a protox inhibitor to control the weeds without significantly affecting the crop.

[0035] The invention further provides methods for producing a genetically transformed plant that is tolerant to glyphosate that involve inserting into the genome of a plant cell a recombinant, double-stranded DNA molecule comprising: (i) a promoter which functions in plant cells to cause the production of an RNA sequence; (ii) a structural DNA sequence that causes the production of an RNA sequence which encodes a GAT; and (iii) a 3' non-translated region which functions in plant cells to cause the addition of a stretch of polyadenyl nucleotides to the 3' end of the RNA sequence; where the

COR-AAD1-00001522

promoter is heterologous with respect to the structural DNA sequence and adapted to cause sufficient expression of the encoded polypeptide to enhance the glyphosate tolerance of a plant cell transformed with the DNA molecule; obtaining a transformed plant cell; and regenerating from the transformed plant cell a genetically transformed plant which has increased tolerance to glyphosate.

[0036] The invention further provides methods for producing a crop that involve growing a crop plant that is glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate N-acetyltransferase, under conditions such that the crop plant produces a crop; and harvesting a crop from the crop plant. These methods often include applying glyphosate to the crop plant at a concentration effective to control weeds. Exemplary crop plants include cotton, corn, and soybean.

[0037] The invention also provides computers, computer readable medium and integrated systems, including databases that are composed of sequence records including character strings corresponding to SEQ ID NO: 1-10, 16, 48, 190, 193, 196,202,205,268,300,442,445,448,454,457,515-830 and 832-972. Such integrated systems optionally include one or more instruction set for selecting, aligning, translating, reverse-translating or viewing any one or more character strings corresponding to SEQ ID NO: 1-10, 16, 48, 190, 193, 196, 202, 205, 268, 300, 442, 445, 448, 454, 457, 515-830 and 832-972, with each other and/or with any additional nucleic acid or amino acid sequence.

BRIEF DESCRIPTION OF THE FIGURES

[0038]

Figure 1 depicts the N-acetylation of glyphosate catalyzed by a glyphosate-N-acetyltransferase ("GAT").
Figure 2 illustrates mass spectroscopic detection ofN-acetylglyphosate produced by an exemplary Bacillus culture expressing a native GAT activity.
Figure 3 is a table illustrating the relative identity between GAT sequences isolated from different strains of bacteria and yitI from *Bacillus subtilis*.
Figure 4 is a map of the plasmid pMAXY2120 for expression and purification of the GAT enzyme from *E. coli* cultures.
Figure 5 is a mass spectrometry output showing increased N-acetylglyphosate production over time in a typical GAT enzyme reaction mix.
Figure 6 is a plot of the kinetic data of a GAT enzyme from which a K_M of 2.9 mM for glyphosate was calculated.
Figure 7 is a plot of the kinetic data taken from the data of Figure 6 from which a K_M of 2 μM was calculated for Acetyl CoA.
Figure 8 is a scheme that describes the degradation of glyphosate in soil through the AMPA pathway.
Figure 9 is a scheme that describes the sarcosine pathway of glyphosate degradation.
Figure 10 is the BLOSUM62 matrix.
Figure 11 is a map of the plasmid pMAXY2190.
Figure 12 depicts a T-DNA construct with *gat* selectable marker.
Figure 13 depicts a yeast expression vector with *gat* selectable marker.
Figure 14 illustrates effect of glyphosate on plant height at tasseling.
Figures 15A and 15B provide a comparison of the kinetic parameters K_m and k_{cat}/K_m, respectively, for various GAT enzymes assayed in either the absence of added KCl (unshaded bars) or in the presence of 20 mM KCl (shaded bars) as described in Example 18. Error bars represent the standard deviation of multiple assays, where available.
Figures 16A, 16B and 16C provide a comparison of the kinetic parameters K_m, k_{cat}, and k_{cat}/K_m, respectively, of various GAT enzymes of the invention (unshaded bars) to the kinetic parameters of some further evolved GAT enzymes of the invention (shaded bars), as described in Example 19. Error bars represent the standard deviation of multiple assays, where available.
Figure 17 depicts remaining GAT activity after incubation at various temperatures as described in Example 16.
Figure 18 depicts the effect of pH on K_{cat} and K_M as described in Example 30.

DETAILED DESCRIPTION OF THE INVENTION

[0039] The present invention relates to a novel class of enzymes exhibiting N-acetyltransferase activity. In one aspect, the invention relates to a novel class of enzymes capable of acetylating glyphosate and glyphosate analogs, *e.g.*, enzymes possessing glyphosate-N-acetyltransferase ("GAT") activity. Such enzymes are characterized by the ability to acetylate the secondary amine of a compound. In some aspects of the invention, this compound is an herbicide, *e.g.*, glyphosate, as illustrated schematically in Figure 1. This compound can also be a glyphosate analog or a metabolic product of glyphosate degradation, *e.g.*, aminomethylphosphonic acid. Although the acetylation of glyphosate is a key catalytic step in one metabolic pathway for catabolism of glyphosate, the enzymatic acetylation of glyphosate by naturally-occurring, isolated, or recombinant enzymes has not been previously described. Thus, the nucleic acids and polypeptides

COR-AAD1-00001523

of the invention provide a new biochemical pathway for engineering herbicide resistance.

[0040] In one aspect, the invention provides novel genes encoding GAT polypeptides. Isolated and recombinant GAT polynucleotides corresponding to naturally occurring polynucleotides, as well as recombinant and engineered, e.g., diversified, GAT polynucleotides are a feature of the invention. GAT polynucleotides are exemplified by SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952. Specific GAT polynucleotide and polypeptide sequences are provided as examples to help illustrate the invention, and are not intended to limit the scope of the genus of GAT polynucleotides and polypeptides described and/or claimed herein.

[0041] The invention also provides methods for generating and selecting diversified libraries to produce additional GAT polynucleotides, including polynucleotides encoding GAT polypeptides with improved and/or enhanced characteristics, *e.g.,* altered K_m for glyphosate, increased rate of catalysis, increased stability, *etc.,* based upon selection of a polynucleotide constituent of the library for the new or improved activities described herein. Such polynucleotides are especially favorably employed in the production of glyphosate-resistant transgenic plants.

[0042] The GAT polypeptides of the invention exhibit a novel enzymatic activity. Specifically, the enzymatic acetylation of the synthetic herbicide glyphosate has not been recognized prior to the present invention. Thus, the polypeptides herein described, e.g., as exemplified by SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931,946, 948, 950, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972 define a novel biochemical pathway for the detoxification of glyphosate that is functional *in vivo,* e.g., in plants.

[0043] Accordingly, the nucleic acids and polypeptides of the invention are of significant utility in the generation of glyphosate-resistant plants by providing new nucleic acids, polypeptides and biochemical pathways for the engineering of herbicide selectivity in transgenic plants.

DEFINITIONS

[0044] Before describing the present invention in detail, it is to be understood that this invention is not limited to particular compositions or biological systems, which can, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting. As used in this specification and the appended claims, the singular forms "a," "an," and "the" include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to "a device" includes a combination of two or more such devices, reference to "a gene fusion construct" includes mixtures of constructs, and the like.

[0045] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. Although any methods and materials similar or equivalent to those described herein can be used in the practice for testing of the present invention, specific examples of appropriate materials and methods are described herein.

[0046] In describing and claiming the present invention, the following terminology will be used in accordance with the definitions set out below.

[0047] Accordingly, for purposes of the present invention, the term "glyphosate" should be considered to include any herbicidally effective form of N-phosphonomethylglycine (including any salt thereof) and other forms which result in the production of the glyphosate anion *in planta.* The term "glyphosate analog" refers to any structural analog of glyphosate that has the ability to inhibit EPSPS at levels such that the glyphosate analog is herbicidally effective.

[0048] As used herein, the term "glyphosate-N-acetyltransferase activity" or "GAT activity" refers to the ability to catalyze the acetylation of the secondary amine group of glyphosate, as illustrated, for example, in Figure 1. A "glyphosate-

COR-AAD1-00001524

N-acetyltransferase" or "GAT" is an enzyme that catalyzes the acetylation of the amine group of glyphosate, a glyphosate analog, and/or a glyphosate primary metabolite *(i.e.,* AMPA or sarcosine). In some preferred embodiments of the invention, a GAT is able to transfer the acetyl group from Acetyl CoA to the secondary amine of glyphosate and the primary amine of AMPA. In addition, some GATs are also able to transfer the propionyl group ofpropionyl CoA to glyphosate, indicating that GAT is also an acyltransferase. The exemplary GATs described herein are active from about pH 5-9, with optimal activity in the range of about pH 6.5-8.0. Activity can be quantified using various kinetic parameters which are well known in the art, *e.g.,* k_{cat}, K_M, and k_{cat}/K_M. These kinetic parameters can be determined as described below in Example 7 or Example 19.

[0049] The terms "polynucleotide," "nucleotide sequence," and "nucleic acid" are used to refer to a polymer of nucleotides (A, C, T, U, G, *etc.* or naturally occurring or artificial nucleotide analogues), *e.g.,* DNA or RNA, or a representation thereof, *e.g.,* a character string, *etc.,* depending on the relevant context. A given polynucleotide or complementary polynucleotide can be determined from any specified nucleotide sequence.

[0050] Similarly, an "amino acid sequence" is a polymer of amino acids (a protein, polypeptide, *etc.*) or a character string representing an amino acid polymer, depending on context. The terms "protein," "polypeptide," and "peptide" are used interchangeably herein.

[0051] A polynucleotide, polypeptide, or other component is "isolated" when it is partially or completely separated from components with which it is normally associated (other proteins, nucleic acids, cells, synthetic reagents, *etc.*). A nucleic acid or polypeptide is "recombinant" when it is artificial or engineered, or derived from an artificial or engineered protein or nucleic acid. For example, a polynucleotide that is inserted into a vector or any other heterologous location, *e.g.,* in a genome of a recombinant organism, such that it is not associated with nucleotide sequences that normally flank the polynucleotide as it is found in nature is a recombinant polynucleotide. A protein expressed *in vitro* or *in vivo* from a recombinant polynucleotide is an example of a recombinant polypeptide. Likewise, a polynucleotide sequence that does not appear in nature, for example a variant of a naturally occurring gene, is recombinant.

[0052] The terms "glyphosate-N-acetyltransferase polypeptide" and "GAT polypeptide" are used interchangeably to refer to any of a family of novel polypeptides provided herein.

[0053] The terms "glyphosate-N-acetyltransferase polynucleotide" and "GAT polynucleotide" are used interchangeably to refer to a polynucleotide that encodes a GAT polypeptide.

[0054] A "subsequence" or "fragment" is any portion of an entire sequence.

[0055] Numbering of an amino acid or nucleotide polymer corresponds to numbering of a selected amino acid polymer or nucleic acid when the position of a given monomer component (amino acid residue, incorporated nucleotide, *etc.*) of the polymer corresponds to the same residue position in a selected reference polypeptide or polynucleotide.

[0056] A vector is a composition for facilitating cell transduction/transformation by a selected nucleic acid, or expression of the nucleic acid in the cell. Vectors include, *e.g.,* plasmids, cosmids, viruses, YACs, bacteria, poly-lysine, chromosome integration vectors, episomal vectors, *etc.*

[0057] "Substantially an entire length of a polynucleotide or amino acid sequence" refers to at least about 70%, generally at least about 80%, or typically about 90% or more of a sequence.

[0058] As used herein, an "antibody" refers to a protein comprising one or more polypeptides substantially or partially encoded by immunoglobulin genes or fragments of immunoglobulin genes. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon and mu constant region genes, as well as myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. A typical immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively. Antibodies exist as intact immunoglobulins or as a number of well characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce F(ab)'2, a dimer of Fab which itself is a light chain joined to VH-CH1 by a disulfide bond. The F(ab)'2 may be reduced under mild conditions to break the disulfide linkage in the hinge region thereby converting the (Fab')2 dimer into an Fab' monomer. The Fab' monomer is essentially a Fab with part of the hinge region (see, Paul, ed. (1998) Fundamental Immunology (4th Edition, Raven Press, NY), for a more detailed description of other antibody fragments). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such Fab' fragments may be synthesized *de novo* either chemically or by utilizing recombinant DNA methodology. Thus, the term antibody as used herein also includes antibody fragments either produced by the modification of whole antibodies or synthesized de novo using recombinant DNA methodologies. Antibodies include single chain antibodies, including single chain Fv (sFv) antibodies in which a variable heavy and a variable light chain are joined together (directly or through a peptide linker) to form a continuous polypeptide.

[0059] A "chloroplast transit peptide" is an amino acid sequence which is translated in conjunction with a protein and

COR-AAD1-00001525

directs the protein to the chloroplast or other plastid types present in the cell in which the protein is made. "Chloroplast transit sequence" refers to a nucleotide sequence that encodes a chloroplast transit peptide.

[0060] A "signal peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the secretory system (Chrispeels (1991) Ann. Rev. Plant Phys. Plant Mol. Biol. 42: 21-53). If the protein is to be directed to a vacuole, a vacuolar targeting signal can further be added, or if to the endoplasmic reticulum, an endoplasmic reticulum retention signal may be added. If the protein is to be directed to the nucleus, any signal peptide present should be removed and instead a nuclear localization signal. included (Raikhel. (1992) Plant Phys. 100: 1627-1632).

[0061] The terms "diversification" and "diversity," as applied to a polynucleotide, refers to generation of a plurality of modified forms of a parental polynucleotide, or plurality of parental polynucleotides. In the case where the polynucleotide encodes a polypeptide, diversity in the nucleotide sequence of the polynucleotide can result in diversity in the corresponding encoded polypeptide, *e.g.* a diverse pool of polynucleotides encoding a plurality of polypeptide variants. In some embodiments of the invention, this sequence diversity is exploited by screening/selecting a library of diversified polynucleotides for variants with desirable functional attributes, *e.g.,* a polynucleotide encoding a GAT polypeptide with enhanced functional characteristics.

[0062] The term "encoding" refers to the ability of a nucleotide sequence to code for one or more amino acids. The term does not require a start or stop codon. An amino acid sequence can be encoded in any one of six different reading frames provided by a polynucleotide sequence and its complement.

[0063] When used herein, the term "artificial variant" refers to a polypeptide having GAT activity, which is encoded by a modified GAT polynucleotide, *e.g.,* a modified form of any one of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952 or of a naturally occurring GAT polynucleotide isolated from an organism. The modified polynucleotide, from which an artificial variant is produced when expressed in a suitable host, is obtained through human intervention by modification of a GAT polynucleotide.

[0064] The term "nucleic acid construct" or "polynucleotide construct" means a nucleic acid molecule, either single-stranded or double-stranded, which is isolated from a naturally occurring gene or which has been modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature. The term nucleic acid construct is synonymous with the term "expression cassette" when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present invention.

[0065] The term "control sequences" is defined herein to include all components, which are necessary or advantageous for the expression of a polypeptide of the present invention. Each control sequence may be native or foreign to the nucleotide sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader sequence, polyadenylation sequence, propeptide sequence, promoter sequence, signal peptide sequence, and transcription terminator sequence. At a minimum, the control sequences include a promoter and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleotide sequence encoding a polypeptide.

[0066] The term "operably linked" is defined herein as a configuration in which a control sequence is appropriately placed at a position relative to the coding sequence of the DNA sequence such that the control sequence directs the expression of a polypeptide.

[0067] When used herein the term "coding sequence" is intended to cover a nucleotide sequence, which directly specifies the amino acid sequence of its protein product. The boundaries of the coding sequence are generally determined by an open reading frame, which usually begins with the ATG start codon. The coding sequence typically includes a DNA, cDNA, and/or recombinant nucleotide sequence.

[0068] In the present context, the term "expression" includes any step involved in the production of the polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

[0069] In the present context, the term "expression vector" covers a DNA molecule, linear or circular, that comprises a segment encoding a polypeptide of the invention, and which is operably linked to additional segments that provide for its transcription.

[0070] The term "host cell", as used herein, includes any cell type which is susceptible to transformation with a nucleic acid construct.

COR-AAD1-00001526

[0071] The term "plant" includes whole plants, shoot vegetative organs/structures (*e.g.,* leaves, stems and tubers), roots, flowers and floral organs/structures *(e.g.,* bracts, sepals, petals, stamens, carpels, anthers and ovules), seed (including embryo, endosperm, and seed coat) and fruit (the mature ovary), plant tissue (*e.g.,* vascular tissue, ground tissue, and the like) and cells *(e.g.,* guard cells, egg cells, trichomes and the like), and progeny of same. The class of plants that can be used in the method of the invention is generally as broad as the class of higher and lower plants amenable to transformation techniques, including angiosperms (monocotyledonous and dicotyledonous plants), gymnosperms, ferns, and multicellular algae. It includes plants of a variety of ploidy levels, including aneuploid, polyploid, diploid, haploid and hemizygous.

[0072] The term "heterologous" as used herein describes a relationship between two or more elements which indicates that the elements are not normally found in proximity to one another in nature. Thus, for example, a polynucleotide sequence is "heterologous to" an organism or a second polynucleotide sequence if it originates from a foreign species, or, if from the same species, is modified from its original form. For example, a promoter operably linked to a heterologous coding sequence refers to a coding sequence from a species different from that from which the promoter was derived, or, if from the same species, a coding sequence which is not naturally associated with the promoter *(e.g.,* a genetically engineered coding sequence or an allele from a different ecotype or variety). An example of a heterologous polypeptide is a polypeptide expressed from a recombinant polynucleotide in a transgenic organism. Heterologous polynucleotides and polypeptides are forms of recombinant molecules.

[0073] A variety of additional terms are defined or otherwise characterized herein.

GLYPHOSATE-N-ACETYLTRANSFERASES

[0074] In one aspect, the invention provides a novel family of isolated or recombinant enzymes referred to herein as "glyphosate-N-acetyltransferases," "GATs," or "GAT enzymes." GATs are enzymes that have GAT activity, preferably sufficient activity to confer some degree of glyphosate tolerance upon a transgenic plant engineered to express the GAT. Some examples of GATs include GAT polypeptides, described in more detail below.

[0075] GAT-mediated glyphosate tolerance is a complex function of GAT activity, GAT expression levels in the transgenic plant, the particular plant, and numerous other factors, including but not limited to the nature and timing of herbicide application. One of skill in the art can determine without undue experimentation the level of GAT activity required to effect glyphosate tolerance in a particular context.

[0076] GAT activity can be characterized using the conventional kinetic parameters k_{cat}, K_M, and k_{cat}/K_M. k_{cat} can be thought of as a measure of the rate of acetylation, particularly at high substrate concentrations, K_M is a measure of the affinity of the GAT for its substrates *(e.g.,* acetyl CoA, propionyl CoA and glyphosate), and k_{cat}/K_M is a measure of catalytic efficiency that takes both substrate affinity and catalytic rate into account. k_{cat}/K_m is particularly important in the situation where the concentration of a substrate is at least partially rate-limiting. In general, a GAT with a higher k_{cat} or k_{cat}/K_M is a more efficient catalyst than another GAT with lower k_{cat} or k_{cat}/K_M. A GAT with a lower K_M is a more efficient catalyst than another GAT with a higher K_M. Thus, to determine whether one GAT is more effective than another, one can compare kinetic parameters for the two enzymes. The relative importance of k_{cat}, k_{cat}/K_M and K_M will vary depending upon the context in which the GAT will be expected to function, e.g., the anticipated effective concentration of glyphosate relative to the K_M for glyphosate. GAT activity can also be characterized in terms of any of a number of functional characteristics, including but not limited to stability, susceptibility to inhibition, or activation by other molecules.

GLYPHOSATE-N-ACETYLTRANSFERASE POLYPEPTIDES

[0077] In one aspect, the invention provides a novel family of isolated or recombinant polypeptides referred to herein as "glyphosate-N-acetyltransferase polypeptides" or "GAT polypeptides." GAT polypeptides are characterized by their structural similarity to a novel family of GATs. Many but not all GAT polypeptides are GATs. The distinction is that GATs are defined in terms of function, whereas GAT polypeptides are defined in terms of structure. A subset of the GAT polypeptides consists of those GAT polypeptides that have GAT activity, preferably at a level that will function to confer glyphosate resistance upon a transgenic plant expressing the protein at an effective level. Some preferred GAT polypeptides for use in conferring glyphosate tolerance have a k_{cat} of at least 1 min^{-1}, or more preferably at least 10 min^{-1}, 100 min^{-1} or 1000 min^{-1}. Other preferred GAT polypeptides for use in conferring glyphosate tolerance have a K_M no greater than 100 mM, or more preferably no greater than 10 mM, 1 mM, or 0.1 mM. Still other preferred GAT polypeptides for use in conferring glyphosate tolerance have a k_{cat}/K_M of at least 1 $mM^{-1}min^{-1}$ or more, preferably at least 10 $mM^{-1}min^{-1}$, 100 $mM^{-1}min^{-1}$, 1000 $mM^{-1}min^{-1}$, or 10,000 $mM^{-1}min^{-1}$.

[0078] Exemplary GAT polypeptides have been isolated and characterized from a variety of bacterial strains. One example of a monomeric GAT polypeptide that has been isolated and characterized has a molecular radius of approximately 17 kD. An exemplary GAT enzyme isolated from a strain of *B. licheniformis,* SEQ ID NO:7, exhibits a K_m for glyphosate of approximately 2.9 mM and a K_m for acetyl CoA of approximately 2 μM, with a k_{cat} equal to 6/minute.

COR-AAD1-00001527

[0079] The term "GAT polypeptide" refers to any polypeptide comprising an amino acid sequence that can be optimally aligned with an amino acid sequence selected from the group consisting of SEQ ID NO:300, 445, and 457 to generate a similarity score of at least 460 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, wherein at least one of the following positions conform to the following restrictions: (i) at positions 18 and 38, there is a Z5 amino acid residue; (ii) at position 62, there is a Z1 amino acid residue; (iii) at position 124, there is a Z6 amino acid residue; and (iv) at position 144, there is a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P. Some aspects of the invention pertain to GAT polypeptides comprising an amino acid sequence that can be optimally aligned with an amino acid sequence selected from the group consisting of SEQ ID NO: 300, 445, and 457 to generate a similarity score of at least 440, 445, 450, 455, 460, 465, 470, 475, 480, 485, 490, 495, 500, 505, 510, 515, 520, 525, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 600, 605, 610, 615, 620, 625, 630, 635, 640, 645, 650, 655, 660, 665, 670, 675, 680, 685, 690, 695, 700, 705, 710, 715, 720, 725, 730, 735, 740, 745, 750, 755, or 760 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, wherein one or more of the following positions conform to the following restrictions: (i) at positions 18 and 38, a Z5 amino acid residue; (ii) at position 62, a Z1 amino acid residue; (iii) at position 124, a Z6 amino acid residue; and (iv) at position 144, a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P.

[0080] Two sequences are "optimally aligned" when they are aligned for similarity scoring using a defined amino acid substitution matrix (e.g., BLOSUM62), gap existence penalty and gap extension penalty so as to arrive at the highest score possible for that pair of sequences. Amino acid substitution matrices and their use in quantifying the similarity between two sequences are well-known in the art and described, *e.g.,* in Dayhoff et al. (1978) "A model of evolutionary change in proteins" in "Atlas of Protein Sequence and Structure," Vol. 5, Suppl. 3 (ed. M.O. Dayhoff), pp. 345-352. Natl. Biomed. Res. Found., Washington, DC and Henikoff et al. (1992) Proc. Nat'l. Acad. Sci. USA 89: 10915-10919. The BLOSUM62 matrix (Fig. 10) is often used as a default scoring substitution matrix in sequence alignment protocols such as Gapped BLAST 2.0. The gap existence penalty is imposed for the introduction of a single amino acid gap in one of the aligned sequences, and the gap extension penalty is imposed for each additional empty amino acid position inserted into an already opened gap. The alignment is defined by the amino acids positions of each sequence at which the alignment begins and ends, and optionally by the insertion of a gap or multiple gaps in one or both sequences so as to arrive at the highest possible score. While optimal alignment and scoring can be accomplished manually, the process is facilitated by the use of a computer-implemented alignment algorithm, *e.g.,* gapped BLAST 2.0, described in Altschul et al. (1997) Nucl. Acids Res. 25: 3389-3402, and made available to the public at the National Center for Biotechnology Information (NCBI) Website (www.ncbi.nlm.nih.gov). Optimal alignments, including multiple alignments, can be prepared using, *e.g.,* PSI-BLAST, available through the NCBI website and described by Altschul et al. (1997) Nucl. Acids Res. 25:3389-3402.

[0081] With respect to an amino acid sequence that is optimally aligned with a reference sequence, an amino acid residue "corresponds to" the position in the reference sequence with which the residue is paired in the alignment. The "position" is denoted by a number that sequentially identifies each amino acid in the reference sequence based on its position relative to the N-terminus. For example, in SEQ ID NO:300, position 1 is M, position 2 is I, position 3 is E, *etc.* When a test sequence is optimally aligned with SEQ ID NO:300, a residue in the test sequence that aligns with the E at position 3 is said to "correspond to position 3" of SEQ ID NO:300. Owing to deletions, insertion, truncations, fusions, *etc.,* that must be taken into account when determining an optimal alignment, in general the amino acid residue number in a test sequence as determined by simply counting from the N-terminal will not necessarily be the same as the number of its corresponding position in the reference sequence. For example, in a case where there is a deletion in an aligned test sequence, there will be no amino acid that corresponds to a position in the reference sequence at the site of deletion. Where there is an insertion in an aligned reference sequence, that insertion will not correspond to any amino acid position in the reference sequence. In the case of truncations or fusions there can be stretches of amino acids in either the reference or aligned sequence that do not correspond to any amino acid in the corresponding sequence.

[0082] The term "GAT polypeptide" further refers to any polypeptide comprising an amino acid sequence selected from the group consisting of: (a) an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) an amino acid sequence that is at least 97% identical to SEQ NO:578; (c) an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) an amino acid sequence that is at least 95% identical to SEQ ID NO:697; (g) an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) an amino acid sequence that is at least 89% identical to SEQ ID NO: 677; (j) an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) an amino acid sequence that is at

COR-AAD1-00001528

least 98% identical to SEQ ID NO:707; (l) an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) an amino acid sequence that is at least 96% identical to SEQ ID NO:612;and (n) an amino acid sequence that is at least 98% identical to SEQ ID NO:590.

[0083] The term "GAT polypeptide" further refers to any polypeptide comprising an amino acid sequence having at least 89% sequence identity with residues 1-96 of the amino acid sequence of SEQ ID NO:677; an amino acid sequence having at least 95% sequence identity with residues 1-96 of the amino acid sequence of SEQ ID NO:697; an amino acid sequence having at least 96% sequence identity with residues 1-96 of the amino acid sequence selected from the group consisting of SEQ ID NO:584, 612, and 721; an amino acid sequence having at least 97% sequence identity with residues 1-96 of the amino acid sequence selected from the group consisting of SEQ ID NO:578, 613, and 621; an amino acid sequence having at least 98% sequence identity with residues 1-96 of the amino acid sequence selected from the group consisting of SEQ ID NO:577, 579, 590, 602, 616, and 707.

[0084] The term "GAT polypeptide" further refers to any polypeptide comprising an amino acid sequence having at least 89% sequence identity with residues 51-146 of the amino acid sequence of SEQ ID NO:677; an amino acid sequence having at least 95% sequence identity with residues 51-146 of the amino acid sequence of SEQ ID NO:697; an amino acid sequence having at least 96% sequence identity with residues 51-146 of the amino acid sequence selected from the group consisting of SEQ ID NO:584, 612, and 721; an amino acid sequence having at least 97% sequence identity with residues 51-146 of the amino acid sequence selected from the group consisting of SEQ ID NO:578, 613, and 621; an amino acid sequence having at least 98% sequence identity with residues 51-146 of the amino acid sequence selected from the group consisting of SEQ ID NO:577, 579, 590, 602, 616, and 707.

[0085] The term "GAT polypeptide" further refers to any polypeptide comprising an amino acid sequence selected from the group consisting of: (a) an amino acid sequence that is at least 96% identical to residues 2-146 of SEQ ID NO: 919; (b) an amino acid sequence that is at least 97% identical to residues 2-146 of SEQ ID NO:929; (c) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:847; (d) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:851; (e) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:853; (f) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:855 (such as, for example, SEQ ID NO:835 or 855); (g) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:857; (h) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO: 861; (i) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:871; (j) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:875; (k) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:881; (1) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:885; (m) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO: 887; (n) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:889; (o) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:893; (p) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:897; (q) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:899; (r) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO: 909; (s) an amino acid sequence that is at least 98% identical to residues 2-146 of SEQ ID NO:911; (t) an amino acid sequence that is at least 99% identical to residues 2-146 of SEQ ID NO:837; (u) an amino acid sequence that is at least 99% identical to residues 2-146 of SEQ ID NO:841; (v) an amino acid sequence that is at least 99% identical to residues 2-146 of SEQ ID NO:865; (w) an amino acid sequence that is at least 99% identical to residues 2-146 of SEQ ID NO: 869; and (x) an amino acid sequence that is at least 99% identical to residues 2-146 of SEQ ID NO:879.

[0086] The term "GAT polypeptide" further refers to any polypeptide comprising an amino acid sequence that is at least 95% identical to residues 2-146 of SEQ ID NO:929 and which comprises a Gly or an Asn residue at the amino acid position corresponding to position 33 of SEQ ID NO:929.

[0087] The term "GAT polypeptide" further refers to any polypeptide comprising an amino acid sequence that shares at least 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity with an exemplary GAT polypeptide disclosed herein. Thus, for example, GAT polypeptides of the invention include polypeptides comprising an amino acid sequence that shares at least 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity with any of SEQ ID NO: 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972.

[0088] As used herein, the term "identity" or "percent identity" when used with respect to a particular pair of aligned amino acid sequences refers to the percent amino acid sequence identity that is obtained by ClustalW analysis (version W 1.8 available from European Bioinformatics Institute, Cambridge, UK), counting the number of identical matches in the alignment and dividing such number of identical matches by the greater of (i) the length of the aligned sequences, and (ii) 96, and using the following default ClustalW parameters to achieve slow/accurate pairwise alignments — Gap Open Penalty:10; Gap Extension Penalty:0.10; Protein weight matrix: Gonnet series; DNA weight matrix: IUB; Toggle Slow/Fast pairwise alignments = SLOW or FULL Alignment.

[0089] In another aspect, the invention provides an isolated or recombinant polypeptide that comprises at least 20, or

COR-AAD1-00001529

alternatively, at least 50, at least 75, at least 100, at least 125, at least 130, at least 135, at least 140, at least 141, at least 142, at least 143, at least 144 or at least 145 contiguous amino acids of an amino acid sequence selected from the group consisting of: (a) an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) an amino acid sequence that is at least 95% identical to SEQ ID NO: 697; (g) an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (1) an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) an amino acid sequence that is at least 96% identical to SEQ ID NO:612;and (n) an amino acid sequence that is at least 98% identical to SEQ ID NO:590.

[0090] In another aspect, the invention provides a polypeptide comprising residues 2-146 of an amino acid sequence selected from the group consisting of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, and 825. In some embodiments of the invention, the amino acid sequence of the polypeptide comprises Met, Met-Ala, or Met-Ala-Ala on the N-terminal side of the amino acid corresponding to position 2 of the reference amino acid sequence.

[0091] Some preferred GAT polypeptides of the invention can be optimally aligned with a reference amino acid sequence selected from the group consisting of SEQ ID NO:300, 445, and 457 to generate a similarity score of at least 460 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, wherein at least one of the following positions conforms to the following restrictions: (i) at positions 18 and 38, there is a Z5 amino acid residue; (ii) at position 62, there is a Z1 amino acid residue; (iii) at position 124, there is a Z6 amino acid residue; and (iv) at position 144, there is a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P, and further wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions: (a) at positions 2, 4, 15, 19, 26, 28, 31, 45, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 123, 129, 139, and/or 145 the amino acid residue is B1; and (b) at positions 3, 5, 8,10,11,14,17,24,27,32,37,47,48, 49, 52, 57, 58, 61, 63, 68, 69, 79, 80, 82, 83, 89, 92, 100, 101, 104, 119, 120, 125, 126, 128, 131, and/or 143 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T. When used to specify an amino acid or amino acid residue, the single letter designations A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y have their standard meaning as used in the art and as provided in Table 1 herein.

[0092] Some preferred GAT polypeptides of the invention can be optimally aligned with a reference amino acid sequence selected from the group consisting of SEQ ID NO: 300, 445, and 457 to generate a similarity score of at least 460 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, wherein at least one of the following positions conforms to the following restrictions: (i) at positions 18 and 38, there is a Z5 amino acid residue; (ii) at position 62, there is a Z1 amino acid residue; (iii) at position 124, there is a Z6 amino acid residue; and (iv) at position 144, there is a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P, and further wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions: (a) at positions 2, 4, 15, 19, 26, 28, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 129, 139, and/or 145 the amino acid residue is Z1; (b) at positions 31 and/or 45 the amino acid residue is Z2; (c) at position 8 the amino acid residue is Z3; (d) at position 89 the amino acid residue is Z3 or Z6; (e) at positions 82, 92, 101 and/or 120 the amino acid residue is Z4; (f) at positions 3, 11, 27 and/or 79 the amino acid residue is Z5; (g) at position 18 the amino acid residue is Z4 or Z5; (h) at position 123 the amino acid residue is Z1 or Z2; (i) at positions 12, 33, 35, 39, 53, 59, 112, 132, 135, 140, and/or 146 the amino acid residue is Z1 or Z3; (j) at position 30 the amino acid residue is Z1; (k) at position 6 the amino acid residue is Z6; (1) at position 81 the amino acid residue is Z2 or Z4; (m) at position 113 the amino acid residue is Z3; (n) at position 138 the amino acid residue is Z4; (o) at position 142 the amino acid residue is Z2; (p) at positions 57 and/or 126 the amino acid residue is Z3 or Z4; (q) at position 5, 17, and 61 the amino acid residue is Z4; (r) at position 24 the amino acid residue is Z3; (s) at position 104 the amino acid residue is Z5; (t) at positions 52, and/or 69 the amino acid residue is Z3; (u) at positions 14 and/or 119 the amino acid

COR-AAD1-00001530

residue is Z5; (v) at positions 10, 32, 63, and/or 83 the amino acid residue is Z5; (w) at positions 48 and/or 80 the amino acid residue is Z6; (x) at position 40 the amino acid residue is Z1 or Z2; (y) at position 96 the amino acid residue is Z3 or Z5; (z) at position 65 the amino acid residue is Z3, Z4, or Z6; (aa) at positions 84 and/or 115 the amino acid residue is Z3; (ab) at position 93 the amino acid residue is Z4; (ac) at position 130 the amino acid residue is Z2; (ad) at position 58 the amino acid residue is Z3, Z4 or Z6; (ae) at position 47 the amino acid residue is Z4 or Z6; (af) at positions 49 and/or 100 the amino acid residue is Z3 or Z4; (ag) at position 68 the amino acid residue is Z4 or Z5; (ah) at position 143 the amino acid residue is Z4; (ai) at position 131 the amino acid residue is Z5; (aj) at positions 125 and/or 128 the amino acid residue is Z5; (ak) at position 67 the amino acid residue is Z3 or Z4; (al) at position 60 the amino acid residue is Z5; and (am) at position 37 the amino acid residue is Z4 or Z6; wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting of N, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

[0093] Some preferred GAT polypeptides of the invention further comprise the amino acid residues in the amino acid sequence that correspond to the positions specified in (a)-(am), wherein at least 90% conform to the amino acid residue restrictions specified in (a)-(am).

[0094] Some preferred GAT polypeptides of the invention additionally comprise amino acid residues in the amino acid sequence that correspond to the following positions, wherein at least 90% conform to the following restrictions: (a) at positions 1, 7, 9, 13, 20, 36, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and (b) at positions 16, 21, 22, 23, 25, 29, 34, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99, 102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

[0095] Some preferred GAT polypeptides of the invention additionally comprise amino acid residues in the amino acid sequence that correspond to the following positions, wherein at least 90% conform to the following restrictions: (a) at positions 1, 7, 9, 13, 20, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and (b) at positions 16, 21, 22, 23, 25, 29, 34, 36, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99,102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

[0096] Some preferred GAT polypeptides of the invention additionally comprise amino acid residues in the amino acid sequence that correspond to the following positions, wherein at least 90% conform to the following restrictions: (a) at positions 1, 7, 9, 20, 42, 50, 72, 75, 76, 78, 94, 98, 110, 121, and/or 141 the amino acid residue is Z1; (b) at positions 13, 46, 56, 70, 107, 117, and/or 118 the amino acid residue is Z2; (c) at positions 23, 55, 71, 77, 88, and/or 109 the amino acid residue is Z3; (d) at positions 16, 21, 41, 73, 85, 99, and/or 111 the amino acid residue is Z4; (e) at positions 34 and/or 95 the amino acid residue is Z5; (f) at position 22, 25, 29, 43, 44, 66, 74, 87, 102, 108, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is Z6; wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting of N, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting ofD and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

[0097] Some preferred GAT polypeptides of the invention further comprise an amino acid residue at position 36 which is selected from the group consisting of Z1 and Z3. Some preferred GAT polypeptides of the invention further comprise an amino acid residue at position 64 which is selected from the group consisting of Z1 and Z2.

[0098] Some preferred GAT polypeptides of the invention further comprise amino acid residues in the amino acid sequence that correspond to the following positions, wherein at least 80% conform to the following restrictions: (a) at position 2 the amino acid residue is I or L; (b) at position 3 the amino acid residue is E; (c) at position 4 the amino acid residue is V or I; (d) at position 5 the amino acid residue is K; (e) at position 6 the amino acid residue is P; (f) at position 8 the amino acid residue is N; (g) at position 10 the amino acid residue is E; (h) at position 11 the amino acid residue is D or E; (i) at position 12 the amino acid residue is T; (j) at position 14 the amino acid residue is E or D; (k) at position 15 the amino acid residue is L; (l) at position 17 the amino acid residue is H; (m) at position 18 the amino acid residue is R, E or K; (n) at position 19 the amino acid residue is I or V; (o) at position 24 the amino acid residue is Q; (p) at position 26 the amino acid residue is M, L, V or I; (q) at position 27 the amino acid residue is E; (r) at position 28 the amino acid residue is A or V; (s) at position 30 the amino acid residue is M; (t) at position 31 the amino acid residue is Y or F; (u) at position 32 the amino acid residue is E or D; (v) at position 33 the amino acid residue is T or S; (w) at position 35 the amino acid residue is L; (x) at position 37 the amino acid residue is R, G, E or Q; (y) at position 39 the amino acid residue is A or S; (z) at position 40 the amino acid residue is F or L; (aa) at position 45 the amino acid residue is Y or F; (ab) at position 47 the amino acid residue is R or G; (ac) at position 48 the amino acid residue is G; (ad) at

COR-AAD1-00001531

position 49 the amino acid residue is K, R, or Q; (ae) at position 51 the amino acid residue is I or V; (af) at position 52 the amino acid residue is S; (ag) at position 53 the amino acid residue is I or V; (ah) at position 54 the amino acid residue is A; (ai) at position 57 the amino acid residue is H or N; (aj) at position 58 the amino acid residue is Q, K, R or P; (ak) at position 59 the amino acid residue is A; (al) at position 60 the amino acid residue is E; (am) at position 61 the amino acid residue is H or R; (an) at position 63 the amino acid residue is E or D; (ao) at position 65 the amino acid residue is E, P or Q; (ap) at position 67 the amino acid residue is Q or R; (aq) at position 68 the amino acid residue is K or E; (ar) at position 69 the amino acid residue is Q; (as) at position 79 the amino acid residue is E; (at) at position 80 the amino acid residue is G; (au) at position 81 the amino acid residue is Y, H or F; (av) at position 82 the amino acid residue is R; (aw) at position 83 the amino acid residue is E or D; (ax) at position 84 the amino acid residue is Q; (ay) at position 86 the amino acid residue is A; (az) at position 89 the amino acid residue is G, T or S; (ba) at position 90 the amino acid residue is L; (bb) at position 91 the amino acid residue is L, I or V; (bc) at position 92 the amino acid residue is R or K; (bd) at position 93 the amino acid residue is H; (be) at position 96 the amino acid residue is E or Q; (bf) at position 97 the amino acid residue is I; (bg) at position 100 the amino acid residue is K or N ; (bh) at position 101 the amino acid residue is K or R; (bi) at position 103 the amino acid residue is A or V; (bj) at position 104 the amino acid residue is D; (bk) at position 105 the amino acid residue is M, L or I; (bl) at position 106 the amino acid residue is L; (bm) at position 112 the amino acid residue is T or A; (bn) at position 113 the amino acid residue is S or T; (bo) at position 114 the amino acid residue is A; (bp) at position 115 the amino acid residue is S; (bq) at position 119 the amino acid residue is K or R; (br) at position 120 the amino acid residue is K or R; (bs) at position 123 the amino acid residue is F or L; (bt) at position 125 the amino acid residue is E; (bu) at position 126 the amino acid residue is Q or H; (bv) at position 128 the amino acid residue is E or D; (bw) at position 129 the amino acid residue is V or I; (bx) at position 130 the amino acid residue is F; (by) at position 131 the amino acid residue is D or E; (bx) at position 132 the amino acid residue is T; (ca) at position 135 the amino acid residue is V; (cb) at position 138 the amino acid residue is H; (cc) at position 139 the amino acid residue is I; (cd) at position 140 the amino acid residue is L or M; (ce) at position 142 the amino acid residue is Y; (cf) at position 143 the amino acid residue is K or R; (cg) at position 145 the amino acid residue is L or I; and (ch) at position 146 the amino acid residue is T.

[0099] Some preferred GAT polypeptides of the invention further comprise amino acid residues in the amino acid sequence that correspond to the positions specified in (a) — (ch) above, wherein at least 90% conform to the amino acid residue restrictions specified in (a) — (ch).

[0100] Some preferred GAT polypeptides of the invention can be optimally aligned with a reference amino acid sequence selected from the group consisting of SEQ ID NO: 300,445, and 457 to generate a similarity score of at least 460 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, wherein at least one of the following positions conforms to the following restrictions: (i) at positions 18 and 38, there is a Z5 amino acid residue; (ii) at position 62, there is a Z1 amino acid residue; (iii) at position 124, there is a Z6 amino acid residue; and (iv) at position 144, there is a Z2 amino acid residue; wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P, and further wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions: (a) at positions 9, 76, 94 and 110 the amino acid residue is A; (b) at positions 29 and 108 the amino acid residue is C; (c) at position 34 the amino acid residue is D; (d) at position 95 the amino acid residue is E; (e) at position 56 the amino acid residue is F; (f) at positions 43, 44, 66, 74, 87, 102, 116, 122, 127 and 136 the amino acid residue is G; (g) at position 41 the amino acid residue is H; (h) at position 7 the amino acid residue is I; (i) at position 85 the amino acid residue is K; (j) at positions 20, 42, 50, 78 and 121 the amino acid residue is L; (k) at positions 1 and 141 the amino acid residue is M; (l) at positions 23 and 109 the amino acid residue is N; (m) at positions 22, 25, 133, 134 and 137 the amino acid residue is P; (n) at position 71 the amino acid residue is Q; (o) at positions 16, 21, 73, 99 and 111 the amino acid residue is R; (p) at position 55 the amino acid residue is S; (q) at position 77 the amino acid residue is T; (r) at position 107 the amino acid residue is W; and (s) at position 13, 46, 70 and 118 the amino acid residue is Y.

[0101] Some preferred GAT polypeptides of the invention further comprise amino acid sequences wherein the amino acid residues meet at least one of the following restrictions: (a) at position 36 the amino acid residue is M, L, or T; (b) at position 72 the amino acid residue is L or I; (c) at position 75 the amino acid residue is M or V; (d) at position 64 the amino acid residue is L, I, or F; (e) at position 88 the amino acid residue is T or S; and (f) at position 117 the amino acid residue is Y or F.

[0102] Some preferred GAT polypeptides of the invention comprise an amino acid sequence wherein the amino acid residues meet at least one of the following additional restrictions: (a) at position 14 the amino acid residue is D; (b) at position 18 the amino acid residue is E; (c) at position 26 the amino acid residue is M or V; (e) at position 30 the amino acid residue is I; (f) at position 32 the amino acid residue is D; (g) at position 36 the amino acid residue is M or T; (i) at position 37 the amino acid residue is C; (j) at position 38 the amino acid residue is D; (j) at position 53 the amino acid residue is V; (k) at position 58 the amino acid residue is R; (l) at position 61 the amino acid residue is R; (m) at position

COR-AAD1-00001532

62 the amino acid residue is L; (n) at position 64 the amino acid residue is I or F; (o) at position 65 the amino acid residue is P; (p) at position 72 the amino acid residue is I; (q) at position 75 the amino acid residue is V; (r) at position 88 the amino acid residue is T; (s) at position 89 the amino acid residue is G; (t) at position 91 the amino acid residue is L; (u) at position 98 the amino acid residue is I; (v) at position 105 the amino acid residue I; (w) at position 112 the amino acid residue is A; (x) at position 124 the amino acid residue is G or C; (y) at position 128 the amino acid residue is D; (z) at position 140 the amino acid residue is M; (aa) at position 143 the amino acid residue is R; and (ab) at position 144 the amino acid residue is W.

[0103] Some preferred GAT polypeptides of the invention comprise an amino acid sequence wherein of the amino acid residues that correspond to the positions specified in (a) through (ab) as described above, at least 80% conform to the amino acid residue restrictions specified in (a) through (ab).

[0104] Some preferred GAT polypeptides of the invention have an amino acid sequence that comprises amino acid residues at least one of which meets the following additional restrictions: (a) at position 41 the amino acid residue is H; (b) at position 138 the amino acid residue is H; (c) at position 34 the amino acid residue is N; and (d) at position 55 the amino acid residue is S.

[0105] Some preferred GAT polypeptides of the invention comprise an amino acid sequence selected from the group consisting of: (a) an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) an amino acid sequence that is at least 95% identical to SEQ ID NO: 697; (g) an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) an amino acid sequence that is at least 96% identical to SEQ NO:584; (k) an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (1) an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) an amino acid sequence that is at least 96% identical to SEQ ID NO:612;and (n) an amino acid sequence that is at least 98% identical to SEQ ID NO:590.

[0106] Some preferred GAT polypeptides of the invention comprise an amino acid sequence selected from the group consisting of: (a) an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) an amino acid sequence that is at least 95% identical to SEQ ID NO: 697; (g) an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (l) an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) an amino acid sequence that is at least 96% identical to SEQ ID NO:612;and (n) an amino acid sequence that is at least 98% identical to SEQ ID NO:590, wherein at least one of the following positions further conforms to the following restrictions: (i) at positions 18 and 38, there is a Z5 amino acid residue; (ii) at position 62, there is a Z1 amino acid residue; (iii) at position 124, there is a Z6 amino acid residue; and (iv) at position 144, there is a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P.

[0107] Some preferred GAT polypeptides of the invention comprise an amino acid sequence selected from the group consisting of: (a) an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) an amino acid sequence that is at least 95% identical to SEQ ID NO: 697; (g) an amino acid sequence that is at least 96% identical to SEQ NO:721; (h) an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (l) an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) an amino acid sequence that is at least 96% identical to SEQ ID NO:612;and (n) an amino acid sequence that is at least 98% identical to SEQ ID NO:590, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following additional restrictions: (a) at positions 2, 4, 15, 19, 26, 28, 31, 45, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 123, 129, 139, and/or 145 the amino acid residue is B1; and (b) at positions 3, 5, 8, 10, 11, 14, 17, 24, 27, 32, 37, 47, 48, 49, 52, 57, 58, 61, 63, 68, 69, 79, 80, 82, 83, 89, 92, 100, 101, 104, 119, 120, 125, 126, 128, 131, and/or 143 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

COR-AAD1-00001533

[0108] Some preferred GAT polypeptides of the invention comprise an amino acid sequence selected from the group consisting of: (a) an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) an amino acid sequence that is at least 95% identical to SEQ ID NO: 697; (g) an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (1) an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) an amino acid sequence that is at least 96% identical to SEQ ID NO:612;and (n) an amino acid sequence that is at least 98% identical to SEQ ID NO:590, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following additional restrictions: (a) at positions 2, 4, 15, 19, 26, 28, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 129, 139, and/or 145 the amino acid residue is Z1; (b) at positions 31 and/or 45 the amino acid residue is Z2; (c) at position 8 the amino acid residue is Z3; (d) at position 89 the amino acid residue is Z3 or Z6; (e) at positions 82, 92, 101 and/or 120 the amino acid residue is Z4; (f) at positions 3, 11, 27 and/or 79 the amino acid residue is Z5; (g) at position 18 the amino acid residue is Z4 or Z5; (h) at position 123 the amino acid residue is Z1 or Z2; (i) at positions 12, 33, 35, 39, 53, 59, 112, 132, 135, 140, and/or 146 the amino acid residue is Z1 or Z3; (j) at position 30 the amino acid residue is Z1; (k) at position 6 the amino acid residue is Z6; (1) at position 81 the amino acid residue is Z2 or Z4; (m) at position 113 the amino acid residue is Z3; (n) at position 138 the amino acid residue is Z4; (o) at position 142 the amino acid residue is Z2; (p) at positions 57 and/or 126 the amino acid residue is Z3 or Z4; (q) at position 5, 17, and 61 the amino acid residue is Z4; (r) at position 24 the amino acid residue is Z3; (s) at position 104 the amino acid residue is Z5; (t) at positions 52, and/or 69 the amino acid residue is Z3; (u) at positions 14 and/or 119 the amino acid residue is Z5; (v) at positions 10, 32, 63, and/or 83 the amino acid residue is Z5; (w) at positions 48 and/or 80 the amino acid residue is Z6; (x) at position 40 the amino acid residue is Z1 or Z2; (y) at position 96 the amino acid residue is Z3 or Z5; (z) at position 65 the amino acid residue is Z3, Z4, or Z6; (aa) at positions 84 and/or 115 the amino acid residue is Z3; (ab) at position 93 the amino acid residue is Z4; (ac) at position 130 the amino acid residue is Z2; (ad) at position 58 the amino acid residue is Z3, Z4 or Z6; (ae) at position 47 the amino acid residue is Z4 or Z6; (af) at positions 49 and/or 100 the amino acid residue is Z3 or Z4; (ag) at position 68 the amino acid residue is Z4 or Z5; (ah) at position 143 the amino acid residue is Z4; (ai) at position 131 the amino acid residue is Z5; (aj) at positions 125 and/or 128 the amino acid residue is Z5; (ak) at position 67 the amino acid residue is Z3 or Z4; (al) at position 60 the amino acid residue is Z5; and (am) at position 37 the amino acid residue is Z4 or Z6; wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting ofN, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

[0109] Some preferred GAT polypeptides of the invention further comprise amino acid residues in the amino acid sequence that correspond to the positions specified in (a)-(am), wherein at least 90% conform to the amino acid residue restrictions specified in (a)-(am).

[0110] Some preferred GAT polypeptides of the invention comprise amino acid residues in the amino acid sequence that correspond to the following positions wherein at least 90% conform to the following additional restrictions: (a) at positions 1, 7, 9, 13, 20, 36, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and (b) at positions 16, 21, 22, 23, 25, 29, 34, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99, 102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

[0111] Some preferred GAT polypeptides of the invention comprise amino acid residues in the amino acid sequence that correspond to the following positions wherein at least 90% conform to the following additional restrictions: (a) at positions 1, 7, 9, 13, 20, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and (b) at positions 16, 21, 22, 23, 25, 29, 34, 36, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99,102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting ofR, N, D, C, Q, E, G, H, K, P, S, and T.

[0112] Some preferred GAT polypeptides of the invention comprise amino acid residues in the amino acid sequence that correspond to the following positions wherein at least 90% conform to the following additional restrictions: (a) at positions 1, 7, 9, 20, 42, 50, 72, 75, 76, 78, 94, 98, 110, 121, and/or 141 the amino acid residue is Z1; (b) at positions 13, 46, 56, 70, 107, 117, and/or 118 the amino acid residue is Z2; (c) at positions 23, 55, 71, 77, 88, and/or 109 the amino acid residue is Z3; (d) at positions 16, 21, 41, 73, 85, 99, and/or 111 the amino acid residue is Z4; (e) at positions 34 and/or 95 the amino acid residue is Z5; (f) at position 22, 25, 29, 43, 44, 66, 74, 87, 102, 108, 116, 122, 127, 133,

COR-AAD1-00001534

134, 136, and/or 137 the amino acid residue is Z6; wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting of N, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

[0113] Some preferred GAT polypeptides of the invention further comprise an amino acid sequence wherein the amino acid residue at position 36 is selected from the group consisting of Z1 and Z3. Some preferred GAT polypeptides of the invention further comprise an amino acid sequence wherein the amino acid residue at position 64 is selected from the group consisting of Z1 and Z2.

[0114] Some preferred GAT polypeptides of the invention comprise an amino acid sequence wherein of the amino acid residues that correspond to the following positions, at least 80% conform to the following additional restrictions: (a) at position 2 the amino acid residue is I or L; (b) at position 3 the amino acid residue is E; (c) at position 4 the amino acid residue is V or I; (d) at position 5 the amino acid residue is K; (e) at position 6 the amino acid residue is P; (f) at position 8 the amino acid residue is N; (g) at position 10 the amino acid residue is E; (h) at position 11 the amino acid residue is D or E; (i) at position 12 the amino acid residue is T; (j) at position 14 the amino acid residue is E or D; (k) at position 15 the amino acid residue is L; (1) at position 17 the amino acid residue is H; (m) at position 18 the amino acid residue is R, E or K; (n) at position 19 the amino acid residue is I or V; (o) at position 24 the amino acid residue is Q; (p) at position 26 the amino acid residue is M, L, V or I; (q) at position 27 the amino acid residue is E; (r) at position 28 the amino acid residue is A or V; (s) at position 30 the amino acid residue is M; (t) at position 31 the amino acid residue is Y or F; (u) at position 32 the amino acid residue is E or D; (v) at position 33 the amino acid residue is T or S; (w) at position 35 the amino acid residue is L; (x) at position 37 the amino acid residue is R, G, E or Q; (y) at position 39 the amino acid residue is A or S; (z) at position 40 the amino acid residue is F or L; (aa) at position 45 the amino acid residue is Y or F; (ab) at position 47 the amino acid residue is R or G; (ac) at position 48 the amino acid residue is G; (ad) at position 49 the amino acid residue is K, R, or Q; (ae) at position 51 the amino acid residue is I or V; (af) at position 52 the amino acid residue is S; (ag) at position 53 the amino acid residue is I or V; (ah) at position 54 the amino acid residue is A; (ai) at position 57 the amino acid residue is H or N; (aj) at position 58 the amino acid residue is Q, K, R or P; (ak) at position 59 the amino acid residue is A; (al) at position 60 the amino acid residue is E; (am) at position 61 the amino acid residue is H or R; (an) at position 63 the amino acid residue is E or D; (ao) at position 65 the amino acid residue is E, P or Q; (ap) at position 67 the amino acid residue is Q or R; (aq) at position 68 the amino acid residue is K or E; (ar) at position 69 the amino acid residue is Q; (as) at position 79 the amino acid residue is E; (at) at position 80 the amino acid residue is G; (au) at position 81 the amino acid residue is Y, H or F; (av) at position 82 the amino acid residue is R; (aw) at position 83 the amino acid residue is E or D; (ax) at position 84 the amino acid residue is Q; (ay) at position 86 the amino acid residue is A; (az) at position 89 the amino acid residue is G, T or S; (ba) at position 90 the amino acid residue is L; (bb) at position 91 the amino acid residue is L, I or V; (bc) at position 92 the amino acid residue is R or K; (bd) at position 93 the amino acid residue is H; (be) at position 96 the amino acid residue is E or Q; (bf) at position 97 the amino acid residue is I; (bg) at position 100 the amino acid residue is K or N ; (bh) at position 101 the amino acid residue is K or R; (bi) at position 103 the amino acid residue is A or V; (bj) at position 104 the amino acid residue is D; (bk) at position 105 the amino acid residue is M, L or I; (bl) at position 106 the amino acid residue is L; (bm) at position 112 the amino acid residue is T or A; (bn) at position 113 the amino acid residue is S or T; (bo) at position 114 the amino acid residue is A; (bp) at position 115 the amino acid residue is S; (bq) at position 119 the amino acid residue is K or R; (br) at position 120 the amino acid residue is K or R; (bs) at position 123 the amino acid residue is F or L; (bt) at position 125 the amino acid residue is E; (bu) at position 126 the amino acid residue is Q or H; (bv) at position 128 the amino acid residue is E or D; (bw) at position 129 the amino acid residue is V or I; (bx) at position 130 the amino acid residue is F; (by) at position 131 the amino acid residue is D or E; (bx) at position 132 the amino acid residue is T; (ca) at position 135 the amino acid residue is V; (cb) at position 138 the amino acid residue is H; (cc) at position 139 the amino acid residue is I; (cd) at position 140 the amino acid residue is L or M; (ce) at position 142 the amino acid residue is Y; (cf) at position 143 the amino acid residue is K or R; (cg) at position 145 the amino acid residue is L or I; and (ch) at position 146 the amino acid residue is T.

[0115] Some preferred GAT polypeptides of the invention comprise an amino acid sequence in which of the residues that correspond to the positions specified in (a) - (ch) above, at least 90% conform to the amino acid residue restrictions specified in (a) - (ch).

[0116] Some preferred GAT polypeptides of the invention comprise an amino acid sequence selected from the group consisting of: (a) an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) an amino acid sequence that is at least 95% identical to SEQ ID NO: 697; (g) an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) an amino acid sequence that is at least 89% identical to SEQ ID NO:677;

COR-AAD1-00001535

(j) an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (1) an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) an amino acid sequence that is at least 96% identical to SEQ ID NO:612;and (n) an amino acid sequence that is at least 98% identical to SEQ ID NO:590, further wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions: (a) at positions 9, 76, 94 and 110 the amino acid residue is A; (b) at positions 29 and 108 the amino acid residue is C; (c) at position 34 the amino acid residue is D; (d) at position 95 the amino acid residue is E; (e) at position 56 the amino acid residue is F; (f) at positions 43, 44, 66, 74, 87, 102, 116, 122, 127 and 136 the amino acid residue is G; (g) at position 41 the amino acid residue is H; (h) at position 7 the amino acid residue is I; (i) at position 85 the amino acid residue is K; (j) at positions 20, 42, 50, 78 and 121 the amino acid residue is L; (k) at positions 1 and 141 the amino acid residue is M; (1) at positions 23 and 109 the amino acid residue is N; (m) at positions 22, 25, 133, 134 and 137 the amino acid residue is P; (n) at position 71 the amino acid residue is Q; (o) at positions 16, 21, 73, 99 and 111 the amino acid residue is R; (p) at position 55 the amino acid residue is S; (q) at position 77 the amino acid residue is T; (r) at position 107 the amino acid residue is W; and (s) at position 13, 46, 70 and 118 the amino acid residue is Y.

[0117] Some preferred GAT polypeptides of the invention further comprise an amino acid sequence in which at least one of the following criteria is met: (a) at position 14 the amino acid residue is D; (b) at position 18 the amino acid residue is E; (c) at position 26 the amino acid residue is M or V; (e) at position 30 the amino acid residue is I; (f) at position 32 the amino acid residue is D; (g) at position 36 the amino acid residue is M or T; (i) at position 37 the amino acid residue is C; (j) at position 38 the amino acid residue is D; (j) at position 53 the amino acid residue is V; (k) at position 58 the amino acid residue is R; (1) at position 61 the amino acid residue is R; (m) at position 62 the amino acid residue is L; (n) at position 64 the amino acid residue is I or F; (o) at position 65 the amino acid residue is P; (p) at position 72 the amino acid residue is I; (q) at position 75 the amino acid residue is V; (r) at position 88 the amino acid residue is T; (s) at position 89 the amino acid residue is G; (t) at position 91 the amino acid residue is L; (u) at position 98 the amino acid residue is I; (v) at position 105 the amino acid residue I; (w) at position 112 the amino acid residue is A; (x) at position 124 the amino acid residue is G or C; (y) at position 128 the amino acid residue is D; (z) at position 140 the amino acid residue is M; (aa) at position 143 the amino acid residue is R; and (ab) at position 144 the amino acid residue is W.

[0118] Some preferred GAT polypeptides of the invention further comprise an amino acid sequence wherein of the amino acid residues that correspond to the positions specified in (a) through (ab) as described above, at least 80% conform to the amino acid residue restrictions specified in (a) through (ab).

[0119] Some preferred GAT polypeptides of the invention further comprise an amino acid sequence wherein the following conditions are also met: (a) at position 41 the amino acid residue is H; (b) at position 138 the amino acid residue is H; (c) at position 34 the amino acid residue is N; and (d) at position 55 the amino acid residue is S.

[0120] Some preferred GAT polypeptides of the invention when optimally aligned with a reference amino acid sequence selected from the group consisting of SEQ ID NO: 300, 445, and 457 to generate a similarity score of at least 460 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, have amino acid sequences such that one or more of the following positions conform to the following restrictions: (i) at positions 18 and 38, there is a Z5 amino acid residue; (ii) at position 62, there is a Z1 amino acid residue; (iii) at position 124, there is a Z6 amino acid residue; and (iv) at position 144, there is a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P. In certain of the aforementioned GAT polypeptides, the amino acid residue in the polypeptide corresponding to position 28 is V, I or A. Valine or isoleucine at position 28 generally correlates with reduced K_M, while alanine at that position generally correlates with increased k_{cat}. Threonine at position 89 and arginine at position 58 generally correlates with reduced K_M. Other preferred GAT polypeptides are characterized by having 127 (*i.e.*, an I at position 27), M30, D34, S35, R37, S39, H41, G48, K49, N57, Q58, P62, T62, Q65, Q67, K68, V75, E83, S89, A96, E96, R101, T112, A114, K119, K120, E128, V129, D131, T131, V132, V134, V135, H138, R144, I145, or T146, or any combination thereof.

[0121] Some preferred GAT polypeptides of the invention comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699, 701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823 and 825.

[0122] In another aspect, the invention provides an isolated or recombinant polypeptide that comprises at least 20, or alternatively, at least 50, at least 75, at least 100, at least 125, at least 130, at least 135, at least 140, at least 141, at least 142, at least 143, at least 144 or at least 145 contiguous amino acids of an amino acid sequence selected from

COR-AAD1-00001536

the groups consisting of: (a) an amino acid sequence that is at least 96% identical to SEQ ID NO:919 (such as, for example, SEQ ID NO:917, 919, 921, 923, 925, 927, 833, 835, 839, 843, 845, 859, 863, 873, 877, 891, 895, 901, 905, 907, 913, 915, or 950); (b) an amino acid sequence that is at least 97% identical to SEQ ID NO:929 (such as, for example, SEQ ID NO:929, 931, 835, 843, 849, or 867); (c) an amino acid sequence that is at least 98% identical to SEQ ID NO: 847 (such as, for example, SEQ ID NO:845 or 847); (d) an amino acid sequence that is at least 98% identical to SEQ ID NO:851; (e) an amino acid sequence that is at least 98% identical to SEQ ID NO:853; (f) an amino acid sequence that is at least 98% identical to SEQ ID NO:855 (such as, for example, SEQ ID NO:835 or 855); (g) an amino acid sequence that is at least 98% identical to SEQ ID NO:857; (h) an amino acid sequence that is at least 98% identical to SEQ ID NO:861 (such as, for example, SEQ ID NO:839, 861, or 883); (i) an amino acid sequence that is at least 98% identical to SEQ ID NO:871; (j) an amino acid sequence that is at least 98% identical to SEQ ID NO:875; (k) an amino acid sequence that is at least 98% identical to SEQ ID NO:881; (1) an amino acid sequence that is at least 98% identical to SEQ ID NO:885 (such as, for example, SEQ ID NO:845 or 885); (m) an amino acid sequence that is at least 98% identical to SEQ ID NO:887; (n) an amino acid sequence that is at least 98% identical to SEQ ID NO:889 (such as, for example, SEQ ID NO: 863, 889, 891, or 903); (o) an amino acid sequence that is at least 98% identical to SEQ ID NO: 893; (p) an amino acid sequence that is at least 98% identical to SEQ ID NO:897; (q) an amino acid sequence that is at least 98% identical to SEQ ID NO:899; (r) an amino acid sequence that is at least 98% identical to SEQ ID NO:909 (such as, for example, SEQ ID NO:883 or 909); (s) an amino acid sequence that is at least 98% identical to SEQ ID NO: 911; (t) an amino acid sequence that is at least 99% identical to SEQ ID NO:837; (u) an amino acid sequence that is at least 99% identical to SEQ ID NO:841; (v) an amino acid sequence that is at least 99% identical to SEQ ID NO:865; (w) an amino acid sequence that is at least 99% identical to SEQ ID NO:869; and (x) an amino acid sequence that is at least 99% identical to SEQ ID NO:879.

[0123] In another aspect, the invention provides an isolated or recombinant polypeptide that comprises at least 20, or alternatively, at least 50, at least 75, at least 100, at least 125, at least 130, at least 135, at least 140, at least 141, at least 142, at least 143, at least 144 or at least 145 contiguous amino acids of an amino acid sequence that is at least 95% identical to SEQ ID NO:929 and which comprises a Gly or an Asn residue at the amino acid position corresponding to position 33 of SEQ ID NO:929 (such as, for example, SEQ ID NO:837, 849, 893, 897, 905, 921, 927, 929 or 931).

[0124] In another aspect, the invention provides a polypeptide comprising residues 2-146 of an amino acid sequence selected from the group consisting of SEQ ID NO: 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972. In some embodiments of the invention, the amino acid sequence of the polypeptide comprises Met, Met-Ala, or Met-Ala-Ala on the N-terminal side of the amino acid corresponding to position 2 of the reference amino acid sequence.

[0125] Some preferred GAT polypeptides of the invention comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 946, 948, and 950.

[0126] The invention further provides preferred GAT polypeptides that are characterized by a combination of the foregoing amino acid residue position restrictions.

[0127] In addition, the invention provides GAT polynucleotides encoding the preferred GAT polypeptides described above, and complementary nucleotide sequences thereof.

[0128] Some aspects of the invention pertain particularly to the subset of any of the above-described categories of GAT polypeptides having GAT activity, as described herein. These GAT polypeptides are preferred, for example, for use as agents for conferring glyphosate resistance upon a plant. Examples of desired levels of GAT activity are described herein.

[0129] In one aspect, the GAT polypeptides comprise an amino acid sequence encoded by a recombinant or isolated form of naturally occurring nucleic acids isolated from a natural source, e.g., a bacterial strain. Wild-type polynucleotides encoding such GAT polypeptides may be specifically screened for by standard techniques known in the art. The polypeptides defined by SEQ ID NO:6 - 10, for example, were discovered by expression cloning of sequences from *Bacillus* strains exhibiting GAT activity, as described in more detail below.

[0130] The invention also includes isolated or recombinant polypeptides which are encoded by an isolated or recombinant polynucleotide comprising a nucleotide sequence which hybridizes under stringent conditions over substantially the entire length of a nucleotide sequence selected from the group consisting of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528; 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764,

COR-AAD1-00001537

768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, and 824, their complements, and nucleotide sequences encoding an amino acid sequence selected from the group consisting of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, and 825, including their complements.

[0131] The invention also includes isolated or recombinant polypeptides which are encoded by an isolated or recombinant polynucleotide comprising a nucleotide sequence which hybridizes under stringent conditions over substantially the entire length of a nucleotide sequence selected from the group consisting of SEQ ID NO: 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, and 930, their complements, and nucleotide sequences encoding an amino acid sequence selected from the group consisting of SEQ ID NO: 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972.

[0132] The invention further includes any polypeptide having GAT activity that is encoded by a fragment of any of the GAT-encoding polynucleotides described herein.

[0133] The invention also provides fragments of GAT polypeptides that can be spliced together to form a functional GAT polypeptide. Splicing can be accomplished *in vitro* or *in vivo,* and can involve cis- or trans-splicing (i.e., intramolecular or intermolecular splicing). The fragments themselves can, but need not, have GAT activity. For example, two or more segments of a GAT polypeptide can be separated by inteins; removal of the intein sequence by cis-splicing results in a functional GAT polypeptide. In another example, an encrypted GAT polypeptide can be expressed as two or more separate fragments; trans-splicing of these segments results in recovery of a functional GAT polypeptide. Various aspects of cis- and trans-splicing, gene encryption, and introduction of intervening sequences are described in more detail in U.S. Patent Application Serial Nos. 09/517,933 and 09/710,686, both of which are incorporated by reference herein in their entirety.

[0134] In general, the invention includes any polypeptide encoded by a modified GAT polynucleotide derived by mutation, recursive sequence recombination, and/or diversification of the polynucleotide sequences described herein. In some aspects of the invention, a GAT polypeptide is modified by single or multiple amino acid substitutions, a deletion, an insertion, or a combination of one or more of these types of modifications. Substitutions can be conservative or non-conservative, can alter function or not, and can add new function. Insertions and deletions can be substantial, such as the case of a truncation of a substantial fragment of the sequence, or in the fusion of additional sequence, either internally or at N or C terminal. In some embodiments of the invention, a GAT polypeptide is part of a fusion protein comprising a functional addition such as, for example, a secretion signal, a chloroplast transit peptide, a purification tag, or any of the numerous other functional groups that will be apparent to the skilled artisan and which are described in more detail elsewhere in this specification.

[0135] Polypeptides of the invention may contain one or more modified amino acid. The presence of modified amino acids may be advantageous in, for example, (a) increasing polypeptide *in vivo* half-life, (b) reducing or increasing polypeptide antigenicity, and (c) increasing polypeptide storage stability. Amino acid(s) are modified, for example, co-translationally or post-translationally during recombinant production (e.g., N-linked glycosylation at N-X-S/T motifs during expression in mammalian cells) or modified by synthetic means.

[0136] Non-limiting examples of a modified amino acid include a glycosylated amino acid, a sulfated amino acid, a prenlyated (e.g., farnesylated, geranylgeranylated) amino acid, an acetylated amino acid, an acylated amino acid, a PEG-ylated amino acid, a biotinylated amino acid, a carboxylated amino acid, a phosphorylated amino acid, and the like. References adequate to guide one of skill in the modification of amino acids are replete throughout the literature. Example protocols are found in Walker (1998) Protein Protocols on CD-ROM (Humana Press, Towata, NJ).

[0137] Recombinant methods for producing and isolating GAT polypeptides of the invention are described herein. In addition to recombinant production, the polypeptides may be produced by direct peptide synthesis using solid-phase techniques (e.g., Stewart et al. (1969) Solid-Phase Peptide Synthesis (WH Freeman Co, San Francisco); and Merrifield (1963) J. Am. Chem. Soc. 85: 2149-2154). Peptide synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer, Foster City, CA) in accordance with the instructions provided by the manufacturer. For example, subsequences may be chemically synthesized separately and combined using chemical methods to provide full-length GAT polypeptides. Peptides can also be ordered from a variety of sources.

COR-AAD1-00001538

[0138] In another aspect of the invention, a GAT polypeptide of the invention is used to produce antibodies which have, e.g., diagnostic uses, for example, related to the activity, distribution, and expression of GAT polypeptides, for example, in various tissues of a transgenic plant.

[0139] GAT homologue polypeptides for antibody induction do not require biological activity; however, the polypeptide or oligopeptide must be antigenic. Peptides used to induce specific antibodies may have an amino acid sequence consisting of at least 10 amino acids, preferably at least 15 or 20 amino acids. Short stretches of a GAT polypeptide may be fused with another protein, such as keyhole limpet hemocyanin, and an antibody produced against the chimeric molecule.

[0140] Methods of producing polyclonal and monoclonal antibodies are known to those of skill in the art, and many antibodies are available. *See, e.g.,* Coligan (1991) Current Protocols in Immunology (Wiley/Greene, NY); Harlow and Lane (1989) Antibodies: A Laboratory Manual (Cold Spring Harbor Press, NY); Stites et al. (eds.) Basic and Clinical Immunology, 4th ed. (Lange Medical Publications, Los Altos, CA), and references cited therein; Goding (1986) Monoclonal Antibodies: Principles and Practice, 2d ed. (Academic Press, New York, NY); and Kohler and Milstein (1975) Nature 256: 495-497. Other suitable techniques for antibody preparation include selection of libraries of recombinant antibodies in phage or similar vectors. *See,* Huse et al. (1989) Science 246: 1275-1281; and Ward et al. (1989) Nature 341: 544-546. Specific monoclonal and polyclonal antibodies and antisera will usually bind with a K_D of at least about 0.1 μM, preferably at least about 0.01 μM or better, and most typically and preferably, 0.001 μM or better.

[0141] Additional details antibody of production and engineering techniques can be found in Borrebaeck, ed. (1995) Antibody Engineering, 2nd ed. (Freeman and Company, NY); McCafferty et al. (1996) Antibody Engineering, A Practical Approach (IRL at Oxford Press, Oxford, England); and Paul (1995) Antibody Engineering Protocols (Humana Press, Towata, NJ).

Sequence Variations

[0142] GAT polypeptides of the present invention include conservatively modified variations of the sequences disclosed herein as SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972. Such conservatively modified variations comprise substitutions, additions or deletions which alter, add or delete a single amino acid or a small percentage of amino acids (typically less than about 5%, more typically less than about 4%, 2%, or 1 %) in any of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972.

[0143] For example, a conservatively modified variation (e.g., deletion) of the 146 amino acid polypeptide identified herein as SEQ ID NO:6 will have a length of at least 140 amino acids, preferably at least 141 amino acids, more preferably at least 144 amino acids, and still more preferably at least 145 amino acids, corresponding to a deletion of less than about 5%, 4%, 2% or about 1%, or less of the polypeptide sequence.

[0144] Another example of a conservatively modified variation (e.g., a "conservatively substituted variation") of the polypeptide identified herein as SEQ ID NO:6 will contain "conservative substitutions," according to the six substitution groups set forth in Table 2, in up to about 7 residues (*i.e.,* less than about 5%) of the 146 amino acid polypeptide.

[0145] The GAT polypeptide sequence homologues of the invention, including conservatively substituted sequences, can be present as part of larger polypeptide sequences such as occur in a GAT polypeptide, in a GAT fusion with a signal sequence, e.g., a chloroplast targeting sequence, or upon the addition of one or more domains for purification of the protein (e.g., poly his segments, FLAG tag segments, *etc.*). In the latter case, the additional functional domains have

COR-AAD1-00001539

little or no effect on the activity of the GAT portion of the protein, or where the additional domains can be removed by post synthesis processing steps such as by treatment with a protease.

Defining Polypeptides by Immunoreactivity

[0146] Because the polypeptides of the invention provide a new class of enzymes with a defined activity, *i.e.,* the acetylation and acylation of glyphosate, the polypeptides also provide new structural features which can be recognized, e.g., in immunological assays. The generation of antisera which specifically binds the polypeptides of the invention, as well as the polypeptides which are bound by such antisera, are a feature of the invention.

[0147] The invention includes GAT polypeptides that specifically bind to or that are specifically immunoreactive with an antibody or antisera generated against an immunogen comprising an amino acid sequence selected from one or more of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972. To eliminate cross-reactivity with other GAT homologues, the antibody or antisera is subtracted with available related proteins, such as those represented by the proteins or peptides corresponding to GenBank accession numbers available as of the filing date of this application, and exemplified by CAA70664, Z99109 and Y09476. Where the accession number corresponds to a nucleic acid, a polypeptide encoded by the nucleic acid is generated and used for antibody/antisera subtraction purposes. Figure 3 tabulates the relative identity between exemplary GAT sequences and the most closely related sequence available in Genbank, YitI. The function of native YitI has yet to be elucidated, but the enzyme has been shown to possess detectable GAT activity.

[0148] In one typical format, the immunoassay uses a polyclonal antiserum which was raised against one or more polypeptides comprising one or more of the sequences corresponding to one or more of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699, 701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972, or a substantial subsequence thereof (*i.e.*, at least about 30% of the full length sequence provided). The full set of potential polypeptide immunogens derived from SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972 are collectively referred to below as "the immunogenic polypeptide(s)." The resulting antisera is optionally selected to have low cross-reactivity against other related sequences and any such cross-reactivity is removed by immunoabsorbtion with one or more of the related sequences, prior to use of the polyclonal antiserum in the immunoassay.

[0149] In order to produce antisera for use in an immunoassay, one or more of the immunogenic polypeptide(s) is produced and purified as described herein. For example, recombinant protein may be produced in a bacterial cell line. An inbred strain of mice (used in this assay because results are more reproducible due to the virtual genetic identity of the mice) is immunized with the immunogenic polypeptide(s) in combination with a standard adjuvant, such as Freund's

COR-AAD1-00001540

adjuvant, using a standard mouse immunization protocol (*see,* Harlow and Lane (1988) Antibodies, A Laboratory Manual (Cold Spring Harbor Publications, New York), for a standard description of antibody generation, immunoassay formats and conditions that can be used to determine specific immunoreactivity). Alternatively, one or more synthetic or recombinant polypeptides derived from the sequences disclosed herein is conjugated to a carrier protein and used as an immunogen.

[0150] Polyclonal sera are collected and titered against the immunogenic polypeptide(s) in an immunoassay, for example, a solid phase immunoassay with one or more of the immunogenic proteins immobilized on a solid support. Polyclonal antisera with a titer of 10^6 or greater are selected, pooled and subtracted with related polypeptides, e.g., those identified from GENBANK as noted, to produce subtracted, pooled, titered polyclonal antisera.

[0151] The subtracted, pooled, titered polyclonal antisera are tested for cross reactivity against the related polypeptides. Preferably at least two of the immunogenic GATs are used in this determination, preferably in conjunction with at least two related polypeptides, to identify antibodies which are specifically bound by the immunogenic polypeptide(s).

[0152] In this comparative assay, discriminatory binding conditions are determined for the subtracted, titered polyclonal antisera which result in at least about a 5-10 fold higher signal to noise ratio for binding of the titered polyclonal antisera to the immunogenic GAT polypeptides as compared to binding to the related polypeptides. That is, the stringency of the binding reaction is adjusted by the addition of non-specific competitors such as albumin or non-fat dry milk, or by adjusting salt conditions, temperature, or the like. These binding conditions are used in subsequent assays for determining whether a test polypeptide is specifically bound by the pooled, subtracted polyclonal antisera. In particular, a test polypeptide which shows at least a 2-5 fold higher signal to noise ratio than the control polypeptide under discriminatory binding conditions, and at least about a ½ signal to noise ratio as compared to the immunogenic polypeptide(s), shares substantial structural similarity with the immunogenic polypeptide(s) as compared to known GAT, and is, therefore a polypeptide of the invention.

[0153] In another example, immunoassays in the competitive binding format are used for the detection of a test polypeptide. For example, as noted, cross-reacting antibodies are removed from the pooled antisera mixture by immunoabsorption with the control GAT polypeptides. The immunogenic polypeptide(s) are then immobilized to a solid support which is exposed to the subtracted pooled antisera. Test proteins are added to the assay to compete for binding to the pooled, subtracted antisera. The ability of the test protein(s) to compete for binding to the pooled, subtracted antisera as compared to the immobilized protein(s) is compared to the ability of the immunogenic polypeptide(s) added to the assay to compete for binding (the immunogenic polypeptide(s) compete effectively with the immobilized immunogenic polypeptide(s) for binding to the pooled antisera). The percent cross-reactivity for the test proteins is calculated, using standard calculations.

[0154] In a parallel assay, the ability of the control proteins to compete for binding to the pooled, subtracted antisera is optionally determined as compared to the ability of the immunogenic polypeptide(s) to compete for binding to the antisera. Again, the percent cross-reactivity for the control polypeptides is calculated, using standard calculations. Where the percent cross-reactivity is at least 5-10x higher for the test polypeptides, the test polypeptides are said to specifically bind the pooled, subtracted antisera.

[0155] In general, the immunoabsorbed and pooled antisera can be used in a competitive binding immunoassay as described herein to compare any test polypeptide to the immunogenic polypeptide(s). In order to make this comparison, the two polypeptides are each assayed at a wide range of concentrations and the amount of each polypeptide required to inhibit 50% of the binding of the subtracted antisera to the immobilized protein is determined using standard techniques. If the amount of the test polypeptide required is less than twice the amount of the immunogenic polypeptide(s) that is required, then the test polypeptide is said to specifically bind to an antibody generated to the immunogenic polypeptide (s), provided the amount is at least about 5-10x higher for a control polypeptide.

[0156] As a final determination of specificity, the pooled antisera is optionally fully immunosorbed with the immunogenic polypeptide(s) (rather than the control polypeptides) until little or no binding of the subtracted, pooled antisera to the immunogenic polypeptide(s) is detectable. This fully immunosorbed antisera is then tested for reactivity with the test polypeptide. If little or no reactivity is observed (i.e., no more than 2x the signal to noise ratio observed for binding of the fully immunosorbed antisera to the immunogenic polypeptide(s)), then the test polypeptide is specifically bound by the antisera elicited by the immunogenic polypeptide(s).

GLYPHOSATE-N-ACETYLTRANSFERASE POLYNUCLEOTIDES

[0157] In one aspect, the invention provides a novel family of isolated or recombinant polynucleotides referred to herein as "glyphosate-N-acetyltransferase polynucleotides" or "GAT polynucleotides." GAT polynucleotide sequences are characterized by the ability to encode a GAT polypeptide. In general, the invention includes any nucleotide sequence that encodes any of the novel GAT polypeptides described herein. In some aspects of the invention, a GAT polynucleotide that encodes a GAT polypeptide with GAT activity is preferred.

[0158] In one aspect, the GAT polynucleotides comprise recombinant or isolated forms of naturally occurring nucleic

COR-AAD1-00001541

acids isolated from an organism, e.g., a bacterial strain. Exemplary GAT polynucleotides, e.g., SEQ ID NO:1 - 5, were discovered by expression cloning of sequences from *Bacillus* strains exhibiting GAT activity. Briefly, a collection of approximately 500 *Bacillus* and *Pseudomonas* strains were screened for native ability to N-acetylate glyphosate. Strains were grown in LB overnight, harvested by centrifugation, permeabilized in dilute toluene, and then washed and resuspended in a reaction mix containing buffer, 5 mM glyphosate, and 200 μM acetyl-CoA. The cells were incubated in the reaction mix for between 1 and 48 hours, at which time an equal volume of methanol was added to the reaction. The cells were then pelleted by centrifugation and the supernatant was filtered before analysis by parent ion mode mass spectrometry. The product of the reaction was positively identified as N-acetylglyphosate by comparing the mass spectrometry profile of the reaction mix to an N-acetylglyphosate standard as shown in Figure 2. Product detection was dependent on inclusion of both substrates (acetyl CoA and glyphosate) and was abolished by heat denaturing the bacterial cells.

[0159] Individual GAT polynucleotides were then cloned from the identified strains by functional screening. Genomic DNA was prepared and partially digested with Sau3A1 enzyme. Fragments of approximately 4 kb were cloned into an *E. coli* expression vector and transformed into electrocompetent *E. coli.* Individual clones exhibiting GAT activity were identified by mass spectrometry following a reaction as described previously except that the toluene wash was replaced by permeabilization with PMBS. Genomic fragments were sequenced and the putative GAT polypeptide-encoding open reading frame was identified. Identity of the GAT gene was confirmed by expression of the open reading frame in *E. coli* and detection of high levels of N-acetylglyphosate produced from reaction mixtures.

[0160] In another aspect of the invention, GAT polynucleotides are produced by diversifying, e.g., recombining and/or mutating one or more naturally occurring, isolated, or recombinant GAT polynucleotides. As described in more detail elsewhere herein, it is often possible to generate diversified GAT polynucleotides encoding GAT polypeptides with superior functional attributes, e.g., increased catalytic function, increased stability, or higher expression level, than a GAT polynucleotide used as a substrate or parent in the diversification process.

[0161] The polynucleotides of the invention have a variety of uses in, for example: recombinant production (i.e., expression) of the GAT polypeptides of the invention; as transgenes (e.g., to confer herbicide resistance in transgenic plants); as selectable markers for transformation and plasmid maintenance; as immunogens; as diagnostic probes for the presence of complementary or partially complementary nucleic acids (including for detection of natural GAT coding nucleic acids); as substrates for further diversity generation, e.g., recombination reactions or mutation reactions to produce new and/or improved GAT homologues, and the like.

[0162] It is important to note that certain specific, substantial and credible utilities of GAT polynucleotides do not require that the polynucleotide encode a polypeptide with substantial GAT activity. For example, GAT polynucleotides that do not encode active enzymes can be valuable sources of parental polynucleotides for use in diversification procedures to arrive at GAT polynucleotide variants, or non-GAT polynucleotides, with desirable functional properties (e.g., high k_{cat} or k_{cat}/K_m, low K_m, high stability towards heat or other environmental factors, high transcription or translation rates, resistance to proteolytic cleavage, reducing antigenicity, etc.). For example, nucleotide sequences encoding protease variants with little or no detectable activity have been used as parent polynucleotides in DNA shuffling experiments to produce progeny encoding highly active proteases (Ness et al. (1999) Nature Biotech. 17:893-96).

[0163] Polynucleotide sequences produced by diversity generation methods or recursive sequence recombination ("RSR") methods (e.g., DNA shuffling) are a feature of the invention. Mutation and recombination methods using the nucleic acids described herein are a feature of the invention. For example, one method of the invention includes recursively recombining one or more nucleotide sequences of the invention as described above and below with one or more additional nucleotides. The recombining steps are optionally performed *in vivo, ex vivo, in silico* or *in vitro.* This diversity generation or recursive sequence recombination produces at least one library of recombinant modified GAT polynucleotides. Polypeptides encoded by members of this library are included in the invention.

[0164] Also contemplated are uses of polynucleotides, also referred to herein as oligonucleotides, typically having at least 12 bases, preferably at least 15, more preferably at least 20, 30, or 50 or more bases, which hybridize under stringent or highly stringent conditions to a GAT polynucleotide sequence. The polynucleotides may be used as probes, primers, sense and antisense agents, and the like, according to methods as noted herein.

[0165] In accordance with the present invention, GAT polynucleotides, including nucleotide sequences that encode GAT polypeptides, fragments of GAT polypeptides, related fusion proteins, or functional equivalents thereof, are used in recombinant DNA molecules that direct the expression of the GAT polypeptides in appropriate host cells, such as bacterial or plant cells. Due to the inherent degeneracy of the genetic code, other nucleic acid sequences which encode substantially the same or a functionally equivalent amino acid sequence can also be used to clone and express the GAT polynucleotides.

[0166] The invention provides GAT polynucleotides that encode transcription and/or translation products that are subsequently spliced to ultimately produce functional GAT polypeptides. Splicing can be accomplished *in vitro* or *in vivo,* and can involve cis- or trans-splicing. The substrate for splicing can be polynucleotides (e.g., RNA transcripts) or polypeptides. An example of cis-splicing of a polynucleotide is where an intron inserted into a coding sequence is removed and

COR-AAD1-00001542

the two flanking exon regions are spliced to generate a GAT polypeptide encoding sequence. An example of trans-splicing would be where a GAT polynucleotide is encrypted by separating the coding sequence into two or more fragments that can be separately transcribed and then spliced to form the full-length GAT encoding sequence. The use of a splicing enhancer sequence (which can be introduced into a construct of the invention) can facilitate splicing either in cis or trans. Cis- and trans-splicing of polypeptides are described in more detail elsewhere herein and in U.S. Patent Application Serial Nos. 09/517,933 and 09/710,686.

[0167] Thus, some GAT polynucleotides do not directly encode a full-length GAT polypeptide, but rather encode a fragment or fragments of a GAT polypeptide. These GAT polynucleotides can be used to express a functional GAT polypeptide through a mechanism involving splicing, where splicing can occur at the level of polynucleotide (e.g., intron/exon) and/or polypeptide (e.g., intern/extern). This can be useful, for example, in controlling expression of GAT activity, since functional GAT polypeptide will only be expressed if all required fragments are expressed in an environment that permits splicing processes to generate functional product. In another example, introduction of one or more insertion sequences into a GAT polynucleotide can facilitate recombination with a low homology polynucleotide; use of an intron or intein for the insertion sequence facilitates the removal of the intervening sequence, thereby restoring function of the encoded variant.

[0168] As will be understood by those of skill in the art, it can be advantageous to modify a coding sequence to enhance its expression in a particular host. The genetic code is redundant with 64 possible codons, but most organisms preferentially use a subset of these codons. The codons that are utilized most often in a species are called optimal codons, and those not utilized very often are classified as rare or low-usage codons (*see, e.g.,* Zhang et al. (1991) Gene 105: 61-72). Codons can be substituted to reflect the preferred codon usage of the host, a process sometimes called "codon optimization" or "controlling for species codon bias."

[0169] Optimized coding sequences containing codons preferred by a particular prokaryotic or eukaryotic host (*see also,* Murray et al. (1989) Nucl. Acids Res. 17:477-508) can be prepared, for example, to increase the rate of translation or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, as compared with transcripts produced from a non-optimized sequence. Translation stop codons can also be modified to reflect host preference. For example, preferred stop codons for *S. cerevisiae* and mammals are UAA and UGA, respectively. The preferred stop codon for monocotyledonous plants is UGA, whereas insects and *E. coli* prefer to use UAA as the stop codon (Dalphin et al. (1996) Nucl. Acids Res. 24: 216-218). Methodology for optimizing a nucleotide sequence for expression in a plant is provided, for example, in U.S. Patent No. 6,015,891, and the references cited therein.

[0170] One embodiment of the invention includes a GAT polynucleotide having optimal codons for expression in a relevant host, e.g., a transgenic plant host. This is particularly desirable when a GAT polynucleotide of bacterial origin is introduced into a transgenic plant, e.g., to confer glyphosate resistance to the plant.

[0171] The polynucleotide sequences of the present invention can be engineered in order to alter a GAT polynucleotide for a variety of reasons, including but not limited to alterations which modify the cloning, processing and/or expression of the gene product. For example, alterations may be introduced using techniques that are well known in the art, e.g., site-directed mutagenesis, to insert new restriction sites, alter glycosylation patterns, change codon preference, introduce splice sites, *etc.*

[0172] As described in more detail herein, the polynucleotides of the invention include sequences which encode novel GAT polypeptides and sequences complementary to the coding sequences, and novel fragments of coding sequences and complements thereof. The polynucleotides can be in the form of RNA or in the form of DNA, and include mRNA, cRNA, synthetic RNA and DNA, genomic DNA and cDNA. The polynucleotides can be double-stranded or single-stranded, and if single-stranded, can be the coding strand or the non-coding (anti-sense, complementary) strand. The polynucleotides optionally include the coding sequence of a GAT polypeptide (i) in isolation, (ii) in combination with an additional coding sequence, so as to encode, e.g., a fusion protein, a pre-protein, a prepro-protein, or the like, (iii) in combination with non-coding sequences, such as introns or inteins, control elements such as a promoter, an enhancer, a terminator element, or 5' and/or 3' untranslated regions effective for expression of the coding sequence in a suitable host, and/or (iv) in a vector or host environment in which the GAT polynucleotide is a heterologous gene. Sequences can also be found in combination with typical compositional formulations of nucleic acids, including in the presence of carriers, buffers, adjuvants, excipients and the like.

[0173] Polynucleotides and oligonucleotides of the invention can be prepared by standard solid-phase methods, according to known synthetic methods. Typically, fragments of up to about 100 bases are individually synthesized, then joined (*e.g.,* by enzymatic or chemical ligation methods, or polymerase mediated methods) to form essentially any desired continuous sequence. For example, polynucleotides and oligonucleotides of the invention can be prepared by chemical synthesis using, *e.g.,* the classical phosphoramidite method described by Beaucage et al. (1981) Tetrahedron Letters 22:1859-69, or the method described by Matthes et al. (1984) EMBO J. 3: 801-05, *e.g.,* as is typically practiced in automated synthetic methods. According to the phosphoramidite method, oligonucleotides are synthesized, e.g., in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors.

[0174] In addition, essentially any nucleic acid can be custom ordered from any of a variety of commercial sources,

COR-AAD1-00001543

such as The Midland Certified Reagent Company (mcrc@oligos.com), The Great American Gene Company (www.genco.com), ExpressGen Inc. (www.expressgen.com), Operon Technologies Inc. (Alameda, CA) and many others. Similarly, peptides and antibodies can be custom ordered from any of a variety of sources, such as PeptidoGenic (pkim@ccnet.com), HTI Bio-products, Inc. (www.htibio.com), BMA Biomedicals Ltd (U.K.), Bio.Synthesis, Inc., and many others.

[0175] Polynucleotides may also be synthesized by well-known techniques as described in the technical literature. See, *e.g.,* Carruthers et al., Cold Spring Harbor Symp. Quant. Biol. 47: 411-418 (1982), and Adams et al. (1983) J. Am. Chem. Soc. 105: 661. Double stranded DNA fragments may then be obtained either by synthesizing the complementary strand and annealing the strands together under appropriate conditions, or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

[0176] General texts which describe molecular biological techniques useful herein, including mutagenesis, include Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, Volume 152 (Academic Press, Inc., San Diego, CA); Sambrook et al. (1989) Molecular Cloning - A Laboratory Manual, 2nd ed., Volumes 1-3 (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York); and Ausubel et al., eds. (2000) Current Protocols in Molecular Biology (Greene Publishing Associates, Inc. and John Wiley & Sons, Inc.). Examples of techniques sufficient to direct persons of skill through *in vitro* amplification methods, including the polymerase chain reaction (PCR), the ligase chain reaction (LCR), Qβ-replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA) are found in Berger, Sambrook, and Ausubel, as well as in Mullis et al. (1987) U.S. Patent No. 4,683,202; Innis et al., eds. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press Inc. San Diego, CA); Arnheim & Levinson (October 1, 1990) Chemical and Engineering News 36-47; The Journal Of NIH Research (1991) 3: 81-94; Kwoh et al. (1989) Proc. Nat'l. Acad. Sci. USA 86: 1173; Guatelli et al. (1990) Proc. Nat'l. Acad. Sci. USA 87: 1874; Lomell et al. (1989) J. Clin. Chem. 35: 1826; Landegren et al. (1988) Science 241: 1077-1080; Van Brunt (1990) Biotechnology 8: 291-294; Wu and Wallace (1989) Gene 4: 560; Barringer et al. (1990) Gene 89: 117, and Sooknanan and Malek (1995) Biotechnology 13: 563-564. Improved methods of cloning *in vitro* amplified nucleic acids are described in Wallace et al. U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids by PCR are summarized in Cheng et al. (1994) Nature 369: 684-685 and the references cited therein, in which PCR amplicons of up to 40kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. *See,* Ausbel, Sambrook and Berger, *all supra.*

[0177] One aspect of the invention provides an isolated or recombinant polynucleotide selected from the group consisting of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952.

[0178] Preferred polynucleotides of the present invention include an isolated or recombinant polynucleotide sequence encoding and amino acid sequence that can be optimally aligned with a reference amino acid sequence selected from the group consisting of SEQ ID NO: 3 00, 445, and 457 to generate a similarity score of at least 460 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, wherein one or more of the following positions conform to the following restrictions: (i) at positions 18 and 38, there is a Z5 amino acid residue; (ii) at position 62, there is a Z1 amino acid residue; (iii) at position 124, there is a Z6 amino acid residue; and (iv) at position 144, there is a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P, and further wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions: (a) at positions 2, 4, 15, 19, 26, 28, 31, 45, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 123, 129, 139, and/or 145 the amino acid residue is B1; and (b) at positions 3, 5, 8, 10, 11, 14, 17, 24, 27, 32, 37, 47, 48, 49, 52, 57, 58, 61, 63, 68, 69, 79, 80, 82, 83, 89, 92, 100, 101, 104, 119, 120, 125, 126, 128, 131, and/or 143 the amino acid residue is B2; wherein B1 is an amino, acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T. When used to specify an amino acid or amino acid residue, the single letter designations A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y have their standard meaning as used in the art and as provided in Table 1 herein.

[0179] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with a reference amino acid sequence selected from the group consisting

COR-AAD1-00001544

of SEQ ID NO: 300, 445, and 457 to generate a similarity score of at least 460 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, one or more of the following positions conform to the following restrictions: (i) at positions 18 and 38, there is a Z5 amino acid residue; (ii) at position 62, there is a Z1 amino acid residue; (iii) at position 124, there is a Z6 amino acid residue; and (iv) at position 144, there is a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P, and further wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions: (a) at positions 2, 4, 15, 19, 26, 28, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 129, 139, and/or 145 the amino acid residue is Z1; (b) at positions 31 and/or 45 the amino acid residue is Z2; (c) at position 8 the amino acid residue is Z3; (d) at position 89 the amino acid residue is Z3 or Z6; (e) at positions 82, 92, 101 and/or 120 the amino acid residue is Z4; (f) at positions 3, 11, 27 and/or 79 the amino acid residue is Z5; (g) at position 18 the amino acid residue is Z4 or Z5; (h) at position 123 the amino acid residue is Z1 or Z2; (i) at positions 12, 33, 35, 39, 53, 59, 112, 132, 135, 140, and/or 146 the amino acid residue is Z1 or Z3; (j) at position 30 the amino acid residue is Z1; (k) at position 6 the amino acid residue is Z6; (1) at position 81 the amino acid residue is Z2 or Z4; (m) at position 113 the amino acid residue is Z3; (n) at position 138 the amino acid residue is Z4; (o) at position 142 the amino acid residue is Z2; (p) at positions 57 and/or 126 the amino acid residue is Z3 or Z4; (q) at position 5, 17, and 61 the amino acid residue is Z4; (r) at position 24 the amino acid residue is Z3; (s) at position 104 the amino acid residue is Z5; (t) at positions 52 and/or 69 the amino acid residue is Z3; (u) at positions 14 and/or 119 the amino acid residue is Z5; (v) at positions 10, 32, 63, and/or 83 the amino acid residue is Z5; (w) at positions 48 and/or 80 the amino acid residue is Z6; (x) at position 40 the amino acid residue is Z1 or Z2; (y) at position 96 the amino acid residue is Z3 or Z5; (z) at position 65 the amino acid residue is Z3, Z4, or Z6; (aa) at positions 84 and/or 115 the amino acid residue is Z3; (ab) at position 93 the amino acid residue is Z4; (ac) at position 130 the amino acid residue is Z2; (ad) at position 58 the amino acid residue is Z3, Z4, or Z6; (ae) at position 47 the amino acid residue is Z4 or Z6; (af) at positions 49 and/or 100 the amino acid residue is Z3 or Z4; (ag) at position 68 the amino acid residue is Z4 or Z5; (ah) at position 143 the amino acid residue is Z4; (ai) at position 131 the amino acid residue is Z5; (aj) at positions 125 and/or 128 the amino acid residue is Z5; (ak) at position 67 the amino acid residue is Z3 or Z4; (a1) at position 60 the amino acid residue is Z5; and (am) at position 37 the amino acid residue is Z4 or Z6; wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting of N, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting ofD and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

[0180] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence further comprising the amino acid residues in the amino acid sequence that correspond to the positions specified in (a)-(am), at least 90% conform to the amino acid residue restrictions specified in(a)-(am).

[0181] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with SEQ ID NO: 300, 445, or 457, at least 90% of the amino acid residues in the amino acid sequence conform to the following restrictions: (a) at positions 1, 7, 9, 13, 20, 36, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and (b) at positions 16, 21, 22, 23, 25, 29, 34, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99, 102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting ofR, N, D, C, Q, E, G, H, K, P, S, and T.

[0182] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with SEQ ID NO: 300, 445, or 457, at least 90% of the amino acid residues in the amino acid sequence conform to the following restrictions: (a) at positions 1, 7, 9, 13, 20, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and (b) at positions 16, 21, 22, 23, 25, 29, 34, 36, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99,102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

[0183] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with SEQ ID NO: 300,445, or 457, at least 90% of the amino acid residues in the amino acid sequence conform to the following restrictions: (a) at positions 1, 7, 9, 20, 42, 50, 72, 75, 76, 78, 94, 98, 110, 121, and/or 141 the amino acid residue is Z1; (b) at positions 13, 46, 56, 70, 107, 117, and/or 118 the amino acid residue is Z2; (c) at positions 23, 55, 71, 77, 88, and/or 109 the amino acid residue is Z3; (d) at positions 16, 21, 41, 73, 85, 99, and/or 111 the amino acid residue is Z4; (e) at positions 34 and/or 95 the amino acid residue is Z5; (f) at position 22, 25, 29, 43, 44, 66, 74, 87, 102, 108, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is Z6; wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting ofN, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting

COR-AAD1-00001545

of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

[0184] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence further comprising at position 36 an amino acid residue selected from the group consisting of Z1 and Z3. Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence further comprising at position 64 an amino acid residue selected from the group consisting of Z1 and Z2.

[0185] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with SEQ ID NO: 300, 445, or 457, at least 80% of the amino acid residues in the amino acid sequence conform to the following restrictions: (a) at position 2 the amino acid residue is I or L; (b) at position 3 the amino acid residue is E; (c) at position 4 the amino acid residue is V or I; (d) at position 5 the amino acid residue is K; (e) at position 6 the amino acid residue is P; (f) at position 8 the amino acid residue is N; (g) at position 10 the amino acid residue is E; (h) at position 11 the amino acid residue is D or E; (i) at position 12 the amino acid residue is T; (j) at position 14 the amino acid residue is E or D; (k) at position 15 the amino acid residue is L; (1) at position 17 the amino acid residue is H; (m) at position 18 the amino acid residue is R, E or K; (n) at position 19 the amino acid residue is I or V; (o) at position 24 the amino acid residue is Q; (p) at position 26 the amino acid residue is M, L, V or I; (q) at position 27 the amino acid residue is E; (r) at position 28 the amino acid residue is A or V; (s) at position 30 the amino acid residue is M; (t) at position 31 the amino acid residue is Y or F; (u) at position 32 the amino acid residue is E or D; (v) at position 33 the amino acid residue is T or S; (w) at position 35 the amino acid residue is L; (x) at position 37 the amino acid residue is R, G, E or Q; (y) at position 39 the amino acid residue is A or S; (z) at position 40 the amino acid residue is F or L ; (aa) at position 45 the amino acid residue is Y or F; (ab) at position 47 the amino acid residue is R or G; (ac) at position 48 the amino acid residue is G; (ad) at position 49 the amino acid residue is K, R, or Q; (ae) at position 51 the amino acid residue is I or V; (af) at position 52 the amino acid residue is S; (ag) at position 53 the amino acid residue is I or V; (ah) at position 54 the amino acid residue is A; (ai) at position 57 the amino acid residue is H or N; (aj) at position 58 the amino acid residue is Q, K, R or P; (ak) at position 59 the amino acid residue is A; (al) at position 60 the amino acid residue is E; (am) at position 61 the amino acid residue is H or R; (an) at position 63 the amino acid residue is E or D; (ao) at position 65 the amino acid residue is E, P or Q; (ap) at position 67 the amino acid residue is Q or R; (aq) at position 68 the amino acid residue is K or E; (ar) at position 69 the amino acid residue is Q; (as) at position 79 the amino acid residue is E; (at) at position 80 the amino acid residue is G; (au) at position 81 the amino acid residue is Y, H or F; (av) at position 82 the amino acid residue is R; (aw) at position 83 the amino acid residue is E or D; (ax) at position 84 the amino acid residue is Q; (ay) at position 86 the amino acid residue is A; (az) at position 89 the amino acid residue is G, T or S; (ba) at position 90 the amino acid residue is L; (bb) at position 91 the amino acid residue is L, I or V; (bc) at position 92 the amino acid residue is R or K; (bd) at position 93 the amino acid residue is H; (be) at position 96 the amino acid residue is E or Q; (bf) at position 97 the amino acid residue is I; (bg) at position 100 the amino acid residue is K or N ; (bh) at position 101 the amino acid residue is K or R; (bi) at position 103 the amino acid residue is A or V; (bj) at position 104 the amino acid residue is D; (bk) at position 105 the amino acid residue is M, L or I; (bl) at position 106 the amino acid residue is L; (bm) at position 112 the amino acid residue is T or A; (bn) at position 113 the amino acid residue is S or T; (bo) at position 114 the amino acid residue is A; (bp) at position 115 the amino acid residue is S; (bq) at position 119 the amino acid residue is K or R; (br) at position 120 the amino acid residue is K or R; (bs) at position 123 the amino acid residue is F or L; (bt) at position 125 the amino acid residue is E; (bu) at position 126 the amino acid residue is Q or H; (bv) at position 128 the amino acid residue is E or D; (bw) at position 129 the amino acid residue is V or I; (bx) at position 130 the amino acid residue is F; (by) at position 131 the amino acid residue is D or E; (bx) at position 132 the amino acid residue is T; (ca) at position 135 the amino acid residue is V; (cb) at position 138 the amino acid residue is H; (cc) at position 139 the amino acid residue is I; (cd) at position 140 the amino acid residue is L or M; (ce) at position 142 the amino acid residue is Y; (cf) at position 143 the amino acid residue is K or R; (cg) at position 145 the amino acid residue is L or I; and (ch) at position 146 the amino acid residue is T.

[0186] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with SEQ ID NO: 300, 445, or 457, at least 90% of the amino acid residues in the amino acid sequence conform to the amino acid residue restrictions specified in (a) - (ch) above.

[0187] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence that when optimally aligned with a reference amino acid sequence selected from the group consisting of SEQ ID NO: 300, 445, and 457 to generate a similarity score of at least 460 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, one or more of the following positions conform to the following restrictions: (i) at positions 18 and 38, there is a Z5 amino acid residue; (ii) at position 62, there is a Z1 amino acid residue; (iii) at position 124, there is a Z6 amino acid residue; and (iv) at position 144, there is a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P, further wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions: (a) at positions 9, 76, 94 and 110 the amino acid residue is A; (b) at positions 29 and 108 the amino acid residue is C; (c) at position

COR-AAD1-00001546

34 the amino acid residue is D; (d) at position 95 the amino acid residue is E; (e) at position 56 the amino acid residue is F; (f) at positions 43, 44, 66, 74, 87, 102, 116, 122, 127 and 136 the amino acid residue is G; (g) at position 41 the amino acid residue is H; (h) at position 7 the amino acid residue is I; (i) at position 85 the amino acid residue is K; (j) at positions 20, 42, 50, 78 and 121 the amino acid residue is L; (k) at positions 1 and 141 the amino acid residue is M; (l) at positions 23 and 109 the amino acid residue is N; (m) at positions 22, 25, 133, 134 and 137 the amino acid residue is P; (n) at position 71 the amino acid residue is Q; (o) at positions 16, 21, 73, 99 and 111 the amino acid residue is R; (p) at position 55 the amino acid residue is S; (q) at position 77 the amino acid residue is T; (r) at position 107 the amino acid residue is W; and (s) at position 13, 46, 70 and 118 the amino acid residue is Y.

[0188] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence which conforms to at least one of the following additional restrictions: (a) at position 36 the amino acid residue is M, L, or T; (b) at position 72 the amino acid residue is L or I; (c) at position 75 the amino acid residue is M or V; (d) at position 64 the amino acid residue is L, I, or F; (e) at position 88 the amino acid residue is T or S; (f) at position 117 the amino acid residue is Y or F.

[0189] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence in which at least one of the following additional conditions is met: (a) at position 14 the amino acid residue is D; (b) at position 18 the amino acid residue is E; (c) at position 26 the amino acid residue is M or V; (e) at position 30 the amino acid residue is I; (f) at position 32 the amino acid residue is D; (g) at position 36 the amino acid residue is M or T; (i) at position 37 the amino acid residue is C; (j) at position 38 the amino acid residue is D; (j) at position 53 the amino acid residue is V; (k) at position 58 the amino acid residue is R; (1) at position 61 the amino acid residue is R; (m) at position 62 the amino acid residue is L; (n) at position 64 the amino acid residue is I or F; (o) at position 65 the amino acid residue is P; (p) at position 72 the amino acid residue is I; (q) at position 75 the amino acid residue is V; (r) at position 88 the amino acid residue is T; (s) at position 89 the amino acid residue is G; (t) at position 91 the amino acid residue is L; (u) at position 98 the amino acid residue is I; (v) at position 105 the amino acid residue I; (w) at position 112 the amino acid residue is A; (x) at position 124 the amino acid residue is G or C; (y) at position 128 the amino acid residue is D; (z) at position 140 the amino acid residue is M; (aa) at position 143 the amino acid residue is R; and (ab) at position 144 the amino acid residue is W.

[0190] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence wherein, of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) through (ab) as described above, at least 80% conform to the amino acid residue restrictions specified in (a) through (ab).

[0191] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence which conforms to at least one of the following additional restrictions: (a) at position 41 the amino acid residue is H; (b) at position 138 the amino acid residue is H; (c) at position 34 the amino acid residue is N; and (d) at position 55 the amino acid residue is S.

[0192] Some preferred isolated or recombinant polynucleotides of the invention are selected from the group consisting of: (a) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) a nucleotide sequence encoding an amino acid sequence that is at least 95% identical to SEQ ID NO:697; (g) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) a nucleotide sequence encoding an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (1) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO: 612;and (n) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:590.

[0193] Some preferred isolated or recombinant polynucleotides of the invention are selected from the group consisting of: (a) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) a nucleotide sequence encoding an amino acid sequence that is at least 95% identical to SEQ ID NO:697; (g) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) a nucleotide sequence encoding an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical

COR-AAD1-00001547

to SEQ ID NO:707; (l) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO: 612; and (n) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:590, wherein the following positions conform to the following restrictions: (i) at positions 18 and 38, there is a Z5 amino acid residue; (ii) at position 62, there is a Z1 amino acid residue; (iii) at position 124, there is a Z6 amino acid residue; and (iv) at position 144, there is a Z2 amino acid residue, wherein: Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid residue selected from the group consisting of F, W, and Y; Z5 is an amino acid residue selected from the group consisting of D and E; and Z6 is an amino acid residue selected from the group consisting of C, G, and P.

[0194] Some preferred isolated or recombinant polynucleotides of the invention are selected from the group consisting of: (a) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) a nucleotide sequence encoding an amino acid sequence that is at least 95% identical to SEQ ID NO:697; (g) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) a nucleotide sequence encoding an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (l) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO: 612; and (n) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:590, further wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions: (a) at positions 2, 4, 15, 19, 26, 28, 31, 45, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 123, 129, 139, and/or 145 the amino acid residue is B1; and (b) at positions 3, 5, 8, 10, 11, 14, 17, 24, 27, 32, 37, 47, 48, 49, 52, 57, 58, 61, 63, 68, 69, 79, 80, 82, 83, 89, 92, 100, 101, 104, 119, 120, 125, 126, 128, 131, and/or 143 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

[0195] Some preferred isolated or recombinant polynucleotides of the invention are selected from the group consisting of: (a) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) a nucleotide sequence encoding an amino acid sequence that is at least 95% identical to SEQ ID NO:697; (g) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) a nucleotide sequence encoding an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (l) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO: 612;and (n) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:590, and further wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions: (a) at positions 2, 4, 15, 19, 26, 28, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 129, 139, and/or 145 the amino acid residue is Z1; (b) at positions 31 and/or 45 the amino acid residue is Z2; (c) at position 8 the amino acid residue is Z3; (d) at position 89 the amino acid residue is Z3 or Z6; (e) at positions 82, 92, 101 and/or 120 the amino acid residue is Z4; (f) at positions 3, 11, 27 and/or 79 the amino acid residue is Z5; (g) at position 18 the amino acid residue is Z4 or Z5; (h) at position 123 the amino acid residue is Z1 or Z2; (i) at positions 12, 33, 35, 39, 53, 59, 112, 132, 135, 140, and/or 146 the amino acid residue is Z1 or Z3; (j) at position 30 the amino acid residue is Z1; (k) at position 6 the amino acid residue is Z6; (l) at position 81 the amino acid residue is Z2 or Z4; (m) at position 113 the amino acid residue is Z3; (n) at position 138 the amino acid residue is Z4; (o) at position 142 the amino acid residue is Z2; (p) at positions 57 and/or 126 the amino acid residue is Z3 or Z4; (q) at position 5, 17, and 61 the amino acid residue is Z4; (r) at position 24 the amino acid residue is Z3; (s) at position 104 the amino acid residue is Z5; (t) at positions 52, and/or 69 the amino acid residue is Z3; (u) at positions 14 and/or 119 the amino acid residue is Z5; (v) at positions 10, 32, 63, and/or 83 the amino acid residue is Z5 ; (w) at positions 48 and/or 80 the amino acid residue is Z6; (x) at position 40 the amino acid residue is Z1 or Z2; (y) at position 96 the amino acid residue is Z3 or Z5; (z) at position 65 the amino acid residue is Z3, Z4, or Z6; (aa) at positions 84 and/or 115 the amino acid residue is Z3;

COR-AAD1-00001548

(ab) at position 93 the amino acid residue is Z4; (ac) at position 130 the amino acid residue is Z2; (ad) at position 58 the amino acid residue is Z3, Z4 or Z6; (ae) at position 47 the amino acid residue is Z4 or Z6; (af) at positions 49 and/or 100 the amino acid residue is Z3 or Z4; (ag) at position 68 the amino acid residue is Z4 or Z5; (ah) at position 143 the amino acid residue is Z4; (ai) at position 131 the amino acid residue is Z5; (aj) at positions 125 and/or 128 the amino acid residue is Z5; (ak) at position 67 the amino acid residue is Z3 or Z4; (al) at position 60 the amino acid residue is Z5; and (am) at position 37 the amino acid residue is Z4 or Z6; wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting ofN, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

[0196] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence further wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a)-(am), at least 90% conform to the amino acid residue restrictions specified in(a)-(am).

[0197] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence in which of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following additional restrictions: (a) at positions 1, 7, 9, 13, 20, 36, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and (b) at positions 16, 21, 22, 23, 25, 29, 34, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99, 102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

[0198] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with SEQ ID NO: 300, 445, or 457, at least 90% of the amino acid residues in the amino acid sequence conform to the following restrictions: (a) at positions 1, 7, 9, 13, 20, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and (b) at positions 16, 21, 22, 23, 25, 29, 34, 36, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99,102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2; wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

[0199] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with SEQ ID NO: 300, 445, or 457, at least 90% of the amino acid residues in the amino acid sequence conform to the following restrictions: (a) at positions 1, 7, 9, 20, 42, 50, 72, 75, 76, 78, 94, 98, 110, 121, and/or 141 the amino acid residue is Z1; (b) at positions 13, 46, 56, 70, 107, 117, and/or 118 the amino acid residue is Z2; (c) at positions 23, 55, 71, 77, 88, and/or 109 the amino acid residue is Z3; (d) at positions 16, 21, 41, 73, 85, 99, and/or 111 the amino acid residue is Z4; (e) at positions 34 and/or 95 the amino acid residue is Z5; (f) at position 22, 25, 29, 43, 44, 66, 74, 87, 102, 108, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is Z6; wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting ofN, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

[0200] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence further comprising at position 36 an amino acid residue selected from the group consisting of Z1 and Z3. Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence further comprising at position 64 an amino acid residue selected from the group consisting of Z1 and Z2.

[0201] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with SEQ ID NO: 300, 445, or 457, at least 80% of the amino acid residues in the amino acid sequence conform to the following restrictions: (a) at position 2 the amino acid residue is I or L; (b) at position 3 the amino acid residue is E; (c) at position 4 the amino acid residue is V or I; (d) at position 5 the amino acid residue is K; (e) at position 6 the amino acid residue is P; (f) at position 8 the amino acid residue is N; (g) at position 10 the amino acid residue is E; (h) at position 11 the amino acid residue is D or E; (i) at position 12 the amino acid residue is T; (j) at position 14 the amino acid residue is E or D; (k) at position 15 the amino acid residue is L; (1) at position 17 the amino acid residue is H; (m) at position 18 the amino acid residue is R, E or K; (n) at position 19 the amino acid residue is I or V; (o) at position 24 the amino acid residue is Q; (p) at position 26 the amino acid residue is M, L, V or I; (q) at position 27 the amino acid residue is E; (r) at position 28 the amino acid residue is A or V; (s) at position 30 the amino acid residue is M; (t) at position 31 the amino acid residue is Y or F; (u) at position 32 the amino acid residue is E or D; (v) at position 33 the amino acid residue is T or S; (w) at position 35 the amino acid residue is L; (x) at position 37 the amino acid residue is R, G, E or Q; (y) at position 39 the amino acid residue is A or S; (z), at position 40 the amino acid residue is F or L ; (aa) at position 45 the amino acid residue is Y or F; (ab) at position 47 the amino acid residue is R or G; (ac) at position 48 the amino acid residue is G; (ad) at position 49 the amino acid residue is K, R, or Q; (ae) at position 51 the amino acid residue is I or V; (af) at position 52 the amino acid residue is S; (ag) at position 53 the amino

COR-AAD1-00001549

acid residue is I or V; (ah) at position 54 the amino acid residue is A; (ai) at position 57 the amino acid residue is H or N; (aj) at position 58 the amino acid residue is Q, K, R or P; (ak) at position 59 the amino acid residue is A; (al) at position 60 the amino acid residue is E; (am) at position 61 the amino acid residue is H or R; (an) at position 63 the amino acid residue is E or D; (ao) at position 65 the amino acid residue is E, P or Q; (ap) at position 67 the amino acid residue is Q or R; (aq) at position 68 the amino acid residue is K or E; (ar) at position 69 the amino acid residue is Q; (as) at position 79 the amino acid residue is E; (at) at position 80 the amino acid residue is G; (au) at position 81 the amino acid residue is Y, H or F; (av) at position 82 the amino acid residue is R; (aw) at position 83 the amino acid residue is E or D; (ax) at position 84 the amino acid residue is Q; (ay) at position 86 the amino acid residue is A; (az) at position 89 the amino acid residue is G, T or S; (ba) at position 90 the amino acid residue is L; (bb) at position 91 the amino acid residue is L, I or V; (bc) at position 92 the amino acid residue is R or K; (bd) at position 93 the amino acid residue is H; (be) at position 96 the amino acid residue is E or Q; (bf) at position 97 the amino acid residue is I; (bg) at position 100 the amino acid residue is K or N ; (bh) at position 101 the amino acid residue is K or R; (bi) at position 103 the amino acid residue is A or V; (bj) at position 104 the amino acid residue is D; (bk) at position 105 the amino acid residue is M, L or I; (bl) at position 106 the amino acid residue is L; (bm) at position 112 the amino acid residue is T or A; (bn) at position 113 the amino acid residue is S or T; (bo) at position 114 the amino acid residue is A; (bp) at position 115 the amino acid residue is S; (bq) at position 119 the amino acid residue is K or R; (br) at position 120 the amino acid residue is K or R; (bs) at position 123 the amino acid residue is F or L; (bt) at position 125 the amino acid residue is E; (bu) at position 126 the amino acid residue is Q or H; (bv) at position 128 the amino acid residue is E or D; (bw) at position 129 the amino acid residue is V or I; (bx) at position 130 the amino acid residue is F; (by) at position 131 the amino acid residue is D or E; (bx) at position 132 the amino acid residue is T; (ca) at position 135 the amino acid residue is V; (cb) at position 138 the amino acid residue is H; (cc) at position 139 the amino acid residue is I; (cd) at position 140 the amino acid residue is L or M; (ce) at position 142 the amino acid residue is Y; (cf) at position 143 the amino acid residue is K or R; (cg) at position 145 the amino acid residue is L or I; and (ch) at position 146 the amino acid residue is T.

[0202] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with SEQ ID NO: 3 00, 445, or 457, at least 90% of the amino acid residues in the amino acid sequence conform to the amino acid residue restrictions specified in (a) - (ch) above.

[0203] Some preferred isolated or recombinant polynucleotides of the invention are selected from the group consisting of: (a) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:577; (b) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:578; (c) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:621; (d) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:579; (e) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:602; (f) a nucleotide sequence encoding an amino acid sequence that is at least 95% identical to SEQ ID NO:697; (g) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:721; (h) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:613; (i) a nucleotide sequence encoding an amino acid sequence that is at least 89% identical to SEQ ID NO:677; (j) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:584; (k) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:707; (1) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:616; (m) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO: 612; and (n) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:590, and further wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions: (a) at positions 9, 76, 94 and 110 the amino acid residue is A; (b) at positions 29 and 108 the amino acid residue is C; (c) at position 34 the amino acid residue is D; (d) at position 95 the amino acid residue is E; (e) at position 56 the amino acid residue is F; (f) at positions 43, 44, 66, 74, 87, 102, 116, 122, 127 and 136 the amino acid residue is G; (g) at position 41 the amino acid residue is H; (h) at position 7 the amino acid residue is I; (i) at position 85 the amino acid residue is K; (j) at positions 20, 42, 50, 78 and 121 the amino acid residue is L; (k) at positions 1 and 141 the amino acid residue is M; (1) at positions 23 and 109 the amino acid residue is N; (m) at positions 22, 25, 133, 134 and 137 the amino acid residue is P; (n) at position 71 the amino acid residue is Q; (o) at positions 16, 21, 73, 99 and 111 the amino acid residue is R; (p) at position 55 the amino acid residue is S; (q) at position 77 the amino acid residue is T; (r) at position 107 the amino acid residue is W; and (s) at position 13, 46, 70 and 118 the amino acid residue is Y.

[0204] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence further comprising at least one amino acid residue that meets the following criteria: (a) at position 14 the amino acid residue is D; (b) at position 18 the amino acid residue is E; (c) at position 26 the amino acid residue is M or V; (e) at position 30 the amino acid residue is I; (f) at position 32 the amino acid residue is D; (g) at position 36 the amino acid residue is M or T; (i) at position 37 the amino acid residue is C; (j) at position 38 the amino acid residue is D; (j) at position 53 the amino acid residue is V; (k) at position 58 the amino acid residue is R; (1) at position 61 the amino acid residue is R; (m) at position 62 the amino acid residue is L; (n) at position 64 the amino acid residue is I or F; (o) at position 65

COR-AAD1-00001550

the amino acid residue is P; (p) at position 72 the amino acid residue is I; (q) at position 75 the amino acid residue is V; (r) at position 88 the amino acid residue is T; (s) at position 89 the amino acid residue is G; (t) at position 91 the amino acid residue is L; (u) at position 98 the amino acid residue is I; (v) at position 105 the amino acid residue I; (w) at position 112 the amino acid residue is A; (x) at position 124 the amino acid residue is G or C; (y) at position 128 the amino acid residue is D; (z) at position 140 the amino acid residue is M; (aa) at position 143 the amino acid residue is R; and (ab) at position 144 the amino acid residue is W.

[0205] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence such that when the sequence is optimally aligned with SEQ ID NO: 300, 445, or 457, at least 80% of the amino acid residues in the amino acid sequence conform to the amino acid residue restrictions specified in (a) through (ab) above.

[0206] Some preferred isolated or recombinant polynucleotides of the invention comprise a nucleotide sequence which encodes an amino acid sequence selected from the group consisting of: (a) an amino acid sequence that is at least 96% identical to SEQ ID NO:919 (such as, for example, a nucleotide sequence which encodes SEQ ID NO:917, 919, 921, 923, 925, 927, 833, 835, 839, 843, 845, 859, 863, 873, 877, 891, 895, 901, 905, 907, 913, 915, or 950); (b) an amino acid sequence that is at least 97% identical to SEQ ID NO:929 (such as, for example, a nucleotide sequence which encodes SEQ ID NO:929, 931, 835, 843, 849, or 867); (c) an amino acid sequence that is at least 98% identical to SEQ ID NO:847 (such as, for example, a nucleotide sequence which encodes SEQ ID NO:845 or 847); (d) an amino acid sequence that is at least 98% identical to SEQ ID NO:851; (e) an amino acid sequence that is at least 98% identical to SEQ ID NO:853; (f) an amino acid sequence that is at least 98% identical to SEQ ID NO:855 (such as, for example, a nucleotide sequence which encodes SEQ ID NO:835 or 855); (g) an amino acid sequence that is at least 98% identical to SEQ ID NO:857; (h) an amino acid sequence that is at least 98% identical to SEQ ID NO:861 (such as, for example, a nucleotide sequence which encodes SEQ ID NO:839, 861, or 883); (i) an amino acid sequence that is at least 98% identical to SEQ ID NO:871; (j) an amino acid sequence that is at least 98% identical to SEQ ID NO:875; (k) an amino acid sequence that is at least 98% identical to SEQ ID NO:881; (1) an amino acid sequence that is at least 98% identical to SEQ ID NO:885 (such as, for example, a nucleotide sequence which encodes SEQ ID NO:845 or 885); (m) an amino acid sequence that is at least 98% identical to SEQ ID NO:887; (n) an amino acid sequence that is at least 98% identical to SEQ ID NO:889 (such as, for example, a nucleotide sequence which encodes SEQ ID NO: 863, 889, 891, or 903); (o) an amino acid sequence that is at least 98% identical to SEQ ID NO:893; (p) an amino acid sequence that is at least 98% identical to SEQ ID NO:897; (q) an amino acid sequence that is at least 98% identical to SEQ ID NO:899; (r) an amino acid sequence that is at least 98% identical to SEQ ID NO:909 (such as, for example, a nucleotide sequence which encodes SEQ ID NO:883 or 909); (s) an amino acid sequence that is at least 98% identical to SEQ ID NO:911; (t) an amino acid sequence that is at least 99% identical to SEQ ID NO:837; (u) an amino acid sequence that is at least 99% identical to SEQ ID NO:841; (v) an amino acid sequence that is at least 99% identical to SEQ ID NO:865; (w) an amino acid sequence that is at least 99% identical to SEQ ID NO:869; and (x) an amino acid sequence that is at least 99% identical to SEQ ID NO:879.

[0207] Some preferred isolated or recombinant polynucleotides of the invention are selected from the group consisting of: (a) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:919 (for example, a nucleotide sequence such as SEQ ID NO:916, 918, 920, 922, 924, 926, 832, 834, 838, 842, 844, 858, 862, 872, 876, 890, 894, 900, 904, 906, 912, 914, 939, 940, 941, 942, 943, 944, 949, 951 or 952); (b) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:929 (for example, a nucleotide sequence such as SEQ ID NO:928, 930, 834, 842, 848, 866, 936 or 937); (c) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:847 (for example, a nucleotide sequence such as SEQ ID NO: 844 or 846); (d) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO: 851 (for example, a nucleotide sequence such as SEQ ID NO:850); (e) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:853 (for example, a nucleotide sequence such as SEQ ID NO: 852); (f) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:855 (for example, a nucleotide sequence such as SEQ ID NO:834 or 854); (g) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:857 (for example, a nucleotide sequence such as SEQ ID NO: 856); (h) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:861 (for example, a nucleotide sequence such as SEQ ID NO:838, 860, or 882); (i) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:871 (for example, a nucleotide sequence such as SEQ ID NO:870); (j) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:875 (for example, a nucleotide sequence such as SEQ ID NO:874); (k) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:881 (for example, a nucleotide sequence such as SEQ ID NO: 880); (1) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:885 (for example, a nucleotide sequence such as SEQ ID NO:844 or 884); (m) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:887 (for example, a nucleotide sequence such as SEQ ID NO: 886); (n) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:889 (for example, a nucleotide sequence such as SEQ ID NO: 862, 888, 890, or 902); (o) a nucleotide sequence encoding an

COR-AAD1-00001551

amino acid sequence that is at least 98% identical to SEQ ID NO:893 (for example, a nucleotide sequence such as SEQ ID NO:892); (p) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO: 897 (for example, a nucleotide sequence such as SEQ ID NO:896); (q) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:899 (for example, a nucleotide sequence such as SEQ ID NO: 898); (r) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:909 (for example, a nucleotide sequence such as SEQ ID NO:882 or 908); (s) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:911 (for example, a nucleotide sequence such as SEQ ID NO: 910); (t) a nucleotide sequence encoding an amino acid sequence that is at least 99% identical to SEQ ID NO:837 (for example, a nucleotide sequence such as SEQ ID NO:836); (u) a nucleotide sequence encoding an amino acid sequence that is at least 99% identical to SEQ ID NO:841 (for example, a nucleotide sequence such as SEQ ID NO:840); (v) a nucleotide sequence encoding an amino acid sequence that is at least 99% identical to SEQ ID NO:865 (for example, a nucleotide sequence such as SEQ ID NO:864); (w) a nucleotide sequence encoding an amino acid sequence that is at least 99% identical to SEQ ID NO:869 (for example, a nucleotide sequence such as SEQ ED NO:868); and (x) a nucleotide sequence encoding an amino acid sequence that is at least 99% identical to SEQ ID NO:879 (for example, a nucleotide sequence such as SEQ ID NO:878).

[0208] Some preferred isolated or recombinant polynucleotides of the invention comprise a nucleotide sequence encoding an amino acid sequence that is at least 95% identical to SEQ ID NO:929 and which comprises a Gly or an Asn residue at the amino acid position corresponding to position 33 of SEQ ID NO:929 (such as, for example, a nucleotide sequence which encodes SEQ ID NO:837, 849, 893, 897, 905, 921, 927, 929 or 931). Some preferred isolated or recombinant polynucleotides of the invention comprise a nucleotide sequence encoding an amino acid sequence that is at least 95% identical to SEQ ID NO:929 and which comprises a Gly or an Asn residue at the amino acid position corresponding to position 33 of SEQ ID NO:929 (for example, a nucleotide sequence such as SEQ ID NO:836, 848, 892, 896, 904, 920, 926, 928, 930, 938).

[0209] Some preferred isolated or recombinant polynucleotides of the invention encode an amino acid sequence which further comprises one or more amino acid residues meeting the following criteria: (a) at position 41 the amino acid residue is H; (b) at position 138 the amino acid residue is H; (c) at position 34 the amino acid residue is N; and (d) at position 55 the amino acid residue is S.

[0210] While description of the polypeptides of the invention is sometimes expressed herein as a list of possible restrictions on what amino acid residues are found at particular positions, in some embodiments, a polypeptide of the invention meets all of a particular set of possible restrictions. That is, in some instances herein, a list of possible restrictions is expressed as a list of options joined by the conjunction "and/or," and in some embodiments, each such conjunction operates as an "and" rather than an "or." In some embodiments, possible restrictions which are expressed as alternate possibilities are all found in the polypeptide of the invention; this is only true where the alternate possibilities are not mutually exclusive.

Sequence Variations

[0211] It will be appreciated by those skilled in the art that due to the degeneracy of the genetic code, a multitude of nucleotide sequences encoding GAT polypeptides of the invention may be produced, some of which bear substantial identity to the nucleic acid sequences explicitly disclosed herein.

Table 1

Codon Table						
Amino acids			Codon			
Alanine	Ala	A	GCA	GCC	GCG	GCU
Cysteine	Cys	C	UGC	UGU		
Aspartic acid	Asp	D	GAC	GAU		
Glutamic acid	Glu	E	GAA	GAG		
Phenylalanine	Phe	F	UUC	UUU		
Glycine	Gly	G	GGA	GGC	GGG	GGU
Histidine	His	H	CAC	CAU		
Isoleucine	Ile	I	AUA	AUC	AUU	
Lysine	Lys	K	AAA	AAG		

COR-AAD1-00001552

(continued)

Amino acids			Codon					
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

[0212] For instance, inspection of the codon table (Table 1) shows that codons AGA, AGG, CGA, CGC, CGG, and CGU all encode the amino acid arginine. Thus, at every position in the nucleic acids of the invention where an arginine is specified by a codon, the codon can be altered to any of the corresponding codons described above without altering the encoded polypeptide. It is understood that U in an RNA sequence corresponds to T in a DNA sequence.

[0213] Using as an example the nucleic acid sequence corresponding to nucleotides 1-15 of SEQ ID NO:1 (ATG ATT GAA GTC AAA (SEQ ID NO:826)), a silent variation of this sequence includes AGT ATC GAG GTG AAG (SEQ ID NO: 827); both sequences encode the amino acid sequence MIEVK (SEQ ID NO:828), which corresponds to amino acids 1-5 of SEQ ID NO:6.

[0214] Such "silent variations" are one species of "conservatively modified variations," as discussed below. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine) can be modified by standard techniques to encode a functionally identical polypeptide. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in any described sequence. The invention provides each and every possible variation of nucleic acid sequence encoding a polypeptide of the invention that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code (e.g., as set forth in Table 1) as applied to the nucleic acid sequence encoding a GAT homologue polypeptide of the invention. All such variations of every nucleic acid herein are specifically provided and described by consideration of the sequence in combination with the genetic code. Any variant can be produced as noted herein.

[0215] A group of two or more different codons that, when translated in the same context, all encode the same amino acid, are referred to herein as "synonymous codons." As described herein, in some aspects of the invention a GAT polynucleotide is engineered for optimized codon usage in a desired host organism, for example a plant host. The term "optimized" or "optimal" are not meant to be restricted to the very best possible combination of codons, but simply indicates that the coding sequence as a whole possesses an improved usage of codons relative to a precursor polynucleotide from which it was derived. Thus, in one aspect the invention provides a method for producing a GAT polynucleotide variant by replacing at least one parental codon in a nucleotide sequence with a synonymous codon that is preferentially used in a desired host organism, e.g., a plant, relative to the parental codon.

[0216] "Conservatively modified variations" or, simply, "conservative variations" of a particular nucleic acid sequence refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or, where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. One of skill will recognize that individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids (typically less than 5%, more typically less than 4%, 2% or 1%, or less) in an encoded sequence are "conservatively modified variations" where the alterations result in the deletion of an amino acid, addition of an amino acid, or substitution of an amino acid with a chemically similar amino acid.

[0217] Conservative substitution tables providing functionally similar amino acids are well known in the art. Table 2 sets forth six groups which contain amino acids that are "conservative substitutions" for one another.

Table 2

Conservative Substitution Groups			
1	Alanine (A)	Serine (S)	Threonine (T)
2	Aspartic acid (D)	Glutamic acid (E)	

COR-AAD1-00001553

(continued)

Conservative Substitution Groups				
3	Asparagine (N)	Glutamine (Q)		
4	Arginine (R)	Lysine (K)		
5	Isoleucine (I)	Leucine (L)	Methionine (M)	Valine (V)
6	Phenylalanine (F)	Tyrosine (Y)	Tryptophan (W)	

[0218] Thus, "conservatively substituted variations" of a listed polypeptide sequence of the present invention include substitutions of a small percentage, typically less than 5%, more typically less than 2% and often less than 1%, of the amino acids of the polypeptide sequence, with a conservatively selected amino acid of the same conservative substitution group. Thus, a conservatively substituted variation of a polypeptide of the invention can contain 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 substitutions with a conservatively substituted variation of the same conservative substitution group.

[0219] For example, a conservatively substituted variation of the polypeptide identified herein as SEQ ID NO:6 will contain "conservative substitutions" according to the six groups defined above, in up to 7 residues *(i.e.,* 5% of the amino acids) in the 146 amino acid polypeptide.

[0220] In a further example, if four conservative substitutions were localized in the region corresponding to amino acids 21 to 30 of SEQ ID NO:6, examples of conservatively substituted variations of this region,
RPN QPL EAC M (SEQ ID NO:829), include:

KPQ QPV ESC M (SEQ ID NO:830) and
KPN NPL DAC V (SEQ ID NO:831) and the like, in accordance with the conservative substitutions listed in Table 2 (in the above example, conservative substitutions are underlined). The listing of a protein sequence herein, in conjunction with the above substitution table, provides an express listing of all conservatively substituted proteins.

[0221] Finally, the addition of sequences which do not alter the encoded activity of a nucleic acid molecule, such as the addition of a non-functional or non-coding sequence, is a conservative variation of the basic nucleic acid.

[0222] One of skill will appreciate that many conservative variations of the nucleic acid constructs which are disclosed yield a functionally identical construct. For example, as discussed above, owing to the degeneracy of the genetic code, "silent substitutions" (*i.e.,* substitutions in a nucleic acid sequence which do not result in an alteration in an encoded polypeptide) are an implied feature of every nucleic acid sequence which encodes an amino acid. Similarly, "conservative amino acid substitutions," in one or a few amino acids in an amino acid sequence are substituted with different amino acids with highly similar properties, are also readily identified as being highly similar to a disclosed construct. Such conservative variations of each disclosed sequence are a feature of the present invention.

[0223] Non-conservative modifications of a particular nucleic acid are those which substitute any amino acid not characterized as a conservative substitution. For example, any substitution which crosses the bounds of the six groups set forth in Table 2. These include substitutions of basic or acidic amino acids for neutral amino acids, (e.g., Asp, Glu, Asn, or Gln for Val, Ile, Leu or Met), aromatic amino acid for basic or acidic amino acids (e.g., Phe, Tyr or Trp for Asp, Asn, Glu or Gln) or any other substitution not replacing an amino acid with a like amino acid.

Nucleic Acid Hybridization

[0224] Nucleic acids "hybridize" when they associate, typically in solution. Nucleic acids hybridize due to a variety of well-characterized physico-chemical forces, such as hydrogen bonding, solvent exclusion, base stacking and the like. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes, Part I, Chapter 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays," (Elsevier, New York ("Tijssen")), as well as in Ausubel, *supra,* Hames and Higgins (1995) Gene Probes 1, IRL Press at Oxford University Press, Oxford, England ("Hames and Higgins 1") and Hames and Higgins (1995) Gene Probes 2, IRL Press at Oxford University Press, Oxford, England ("Hames and Higgins 2") and provide details on the synthesis, labeling, detection and quantification of DNA and RNA, including oligonucleotides.

[0225] "Stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments, such as Southern and northern hybridizations, are sequence dependent, and are different under different environmental parameters. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993), supra, and in Hames and Higgins 1 and Hames and Higgins 2, supra.

[0226] For purposes of the present invention, generally, "highly stringent" hybridization and wash conditions are se-

COR-AAD1-00001554

lected to be about 5°C or less lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH (as noted below, highly stringent conditions can also be referred to in comparative terms). The T_m is the temperature (under defined ionic strength and pH) at which 50% of the test sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected to be equal to the T_m for a particular probe.

[0227] The T_m of a nucleic acid duplex indicates the temperature at which the duplex is 50% denatured under the given conditions and its represents a direct measure of the stability of the nucleic acid hybrid. Thus, the T_m corresponds to the temperature corresponding to the midpoint in transition from helix to random coil and it depends on length, nucleotide composition, and ionic strength for long stretches of nucleotides.

[0228] After hybridization, unhybridized nucleic acid material can be removed by a series of washes, the stringency of which can be adjusted depending upon the desired results. Low stringency washing conditions (e.g., using higher salt and lower temperature) increase sensitivity, but can produce nonspecific hybridization signals and high background signals. Higher stringency conditions (e.g., using lower salt and higher temperature that is closer to the hybridization temperature) lowers the background signal, typically with only the specific signal remaining. See Rapley, R. and Walker, J.M. eds., Molecular Biomethods Handbook (Humana Press, Inc. 1998) (hereinafter "Rapley and Walker"), which is incorporated herein by reference in its entirety for all purposes.

[0229] The T_m of a DNA-DNA duplex can be estimated using Equation 1 as follows:

$$T_m\,(°C) = 81.5°C + 16.6\,(\log_{10}M) + 0.41\,(\%G + C) - 0.72\,(\%f) - 500/n,$$

where M is the molarity of the monovalent cations (usually Na+), (%G + C) is the percentage of guanosine (G) and cytosine (C) nucleotides, (%f) is the percentage of formalize and n is the number of nucleotide bases (*i.e.*, length) of the hybrid. *See* Rapley and Walker, *supra*.

[0230] The T_m of an RNA-DNA duplex can be estimated by using Equation 2 as follows:

$T_m\,(°C) = 79.8°C + 18.5\,(\log_{10}M) + 0.58\,(\%G + C) - 11.8(\%G + C)^2 - 0.56\,(\%f) - 820/n$, where M is the molarity of the monovalent cations (usually Na+), (%G + C) is the percentage of guanosine (G) and cytosine (C) nucleotides, (%f) is the percentage of formamide and n is the number of nucleotide bases *(i.e.,* length) of the hybrid. *Id.*

[0231] Equations 1 and 2 are typically accurate only for hybrid duplexes longer than about 100-200 nucleotides. Id.

[0232] The T_m of nucleic acid sequences shorter than 50 nucleotides can be calculated as follows:

$$T_m\,(°C) = 4(G + C) + 2(A + T),$$

where A (adenine), C, T (thymine), and G are the numbers of the corresponding nucleotides.

[0233] An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is 50% formalin with 1 mg of heparin at 42°C, with the hybridization being carried out overnight. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (*see* Sambrook, *supra* for a description of SSC buffer). Often the high stringency wash is preceded by a low stringency wash to remove background probe signal. An example low stringency wash is 2x SSC at 40°C for 15 minutes.

[0234] In general, a signal to noise ratio of 2.Sx-Sx (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization. Detection of at least stringent hybridization between two sequences in the context of the present invention indicates relatively strong structural similarity or homology to, e.g., the nucleic acids of the present invention provided in the sequence listings herein.

[0235] As noted, "highly stringent" conditions are selected to be about 5° C or less lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. Target sequences that are closely related or identical to the nucleotide sequence of interest (e.g., "probes") can be identified under highly stringent conditions. Lower stringency conditions are appropriate for sequences that are less complementary. See, e.g., Rapley and Walker, supra.

[0236] Comparative hybridization can be used to identify nucleic acids of the invention, and this comparative hybridization method is a preferred method of distinguishing nucleic acids of the invention. Detection of highly stringent hybridization between two nucleotide sequences in the context of the present invention indicates relatively strong structural similarity/homology to, e.g., the nucleic acids provided in the sequence listing herein. Highly stringent hybridization between two nucleotide sequences demonstrates a degree of similarity or homology of structure, nucleotide base composition, arrangement or order that is greater than that detected by stringent hybridization conditions. In particular, detection of highly stringent hybridization in the context of the present invention indicates strong structural similarity or structural homology (e.g., nucleotide structure, base composition, arrangement or order) to, e.g., the nucleic acids

COR-AAD1-00001555

provided in the sequence listings herein. For example, it is desirable to identify test nucleic acids that hybridize to the exemplar nucleic acids herein under stringent conditions.

[0237] Thus, one measure of stringent hybridization is the ability to hybridize to one of the listed nucleic acids (*e.g.*, nucleic acid sequences SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952, and complementary polynucleotide sequences thereof), under highly stringent conditions (or very stringent conditions, or ultra-high stringency hybridization conditions, or ultra-ultra high stringency hybridization conditions). Stringent hybridization (as well as highly stringent, ultra-high stringency, or ultra-ultra high stringency hybridization conditions) and wash conditions can easily be determined empirically for any test nucleic acid. For example, in determining highly stringent hybridization and wash conditions, the hybridization and wash conditions are gradually increased (e.g., by increasing temperature, decreasing salt concentration, increasing detergent concentration and/or increasing the concentration of organic solvents, such as formalin, in the hybridization or wash), until a selected set of criteria are met. For example, the hybridization and wash conditions are gradually increased until a probe comprising one or more nucleic acid sequences selected from SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952, and complementary polynucleotide sequences thereof, binds to a perfectly matched complementary target (again, a nucleic acid comprising one or more nucleic acid sequences selected from SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952, and complementary polynucleotide sequences thereof), with a signal to noise ratio that is at least about 2.5x, and optionally about 5x or more as high as that observed for hybridization of the probe to an unmatched target. In this case, the unmatched target is a nucleic acid corresponding to a nucleic acid (other than those in the accompanying sequence listing) that is present in a public database such as GenBank™ at the time of filing of the subject application. Such sequences can be identified in GenBank by one of skill. Examples include Accession Nos. Z99109 and Y09476. Additional such sequences can be identified in e.g., GenBank, by one of ordinary skill in the art.

[0238] A test nucleic acid is said to specifically hybridize to a probe nucleic acid when it hybridizes at least ½ as well to the probe as to the perfectly matched complementary target, *i.e.*, with a signal to noise ratio at least ½ as high as hybridization of the probe to the target under conditions in which the perfectly matched probe binds to the perfectly matched complementary target with a signal to noise ratio that is at least about 2x-10x, and occasionally 20x, 50x or greater than that observed for hybridization to any of the unmatched polynucleotides of Accession Nos. Z99109 and Y09476.

[0239] Ultra high-stringency hybridization and wash conditions are those in which the stringency of hybridization and wash conditions are increased until the signal to noise ratio for binding of the probe to the perfectly matched complementary target nucleic acid is at least 10x as high as that observed for hybridization to any of the unmatched target nucleic acids of Genbank Accession numbers Z99109 and Y09476. A target nucleic acid which hybridizes to a probe under such conditions, with a signal to noise ratio of at least ½ that of the perfectly matched complementary target

COR-AAD1-00001556

nucleic acid is said to bind to the probe under ultra-high stringency conditions.

[0240] Similarly, even higher levels of stringency can be determined by gradually increasing the hybridization and/or wash conditions of the relevant hybridization assay. For example, those in which the stringency of hybridization and wash conditions are increased until the signal to noise ratio for binding of the probe to the perfectly matched complementary target nucleic acid is at least 10x 20x, 50x, 100x, or 500x or more as high as that observed for hybridization to any of the unmatched target nucleic acids of Genbank Accession numbers Z99109 and Y09476. A target nucleic acid which hybridizes to a probe under such conditions, with a signal to noise ratio of at least ½ that of the perfectly matched complementary target nucleic acid is said to bind to the probe under ultra-ultra-high stringency conditions.

[0241] Target nucleic acids which hybridize to the nucleic acids represented by SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952 under high, ultra-high and ultra-ultra high stringency conditions are a feature of the invention. Examples of such nucleic acids include those with one or a few silent or conservative nucleic acid substitutions as compared to a given nucleic acid sequence.

[0242] Nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code, or when antisera or antiserum generated against one or more of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 946, 948, 950, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972, which has been subtracted using the polypeptides encoded by known nucleotide sequences, including those of Genbank Accession number CAA70664. Further details on immunological identification of polypeptides of the invention are found below. Additionally, for distinguishing between duplexes with sequences of less than about 100 nucleotides, a TMAC1 hybridization procedure known to those of ordinary skill in the art can be used. *See, e.g.,* Sorg, U. et al. Nucleic Acids Res. (Sept. 11, 1991) 19 (17), incorporated herein by reference in its entirety for all purposes.

[0243] In one aspect, the invention provides a nucleic acid which comprises a unique subsequence in a nucleic acid selected from SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952. The unique subsequence is unique as compared to a nucleic acid corresponding to any of Genbank Accession numbers Z99109 and Y09476. Such unique subsequences can be determined by aligning any of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836,

COR-AAD1-00001557

838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952 against the complete set of nucleic acids represented by GenBank accession numbers Z99109 and Y09476 or other related sequences available in public databases as of the filing date of the subject application. Alignment can be performed using the BLAST algorithm set to default parameters. Any unique subsequence is useful, e.g., as a probe to identify the nucleic acids of the invention.

[0244] Similarly, the invention includes a polypeptide which comprises a unique subsequence in a polypeptide selected from: SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 946, 948, 950, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972. Here, the unique subsequence is unique as compared to a polypeptide corresponding to that of GenBank accession number CAA70664. Here again, the polypeptide is aligned against the sequences represented by accession number CAA70664. Note that if the sequence corresponds to a non-translated sequence such as a pseudo gene, the corresponding polypeptide is generated simply by in *silico* translation of the nucleic acid sequence into an amino acid sequence, where the reading frame is selected to correspond to the reading frame of homologous GAT polynucleotides.

[0245] The invention also provides for target nucleic acids which hybridize under stringent conditions to a unique coding oligonucleotide which encodes a unique subsequence in a polypeptide selected from SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 946, 948, 950, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972 wherein the unique subsequence is unique as compared to a polypeptide corresponding to any of the control polypeptides. Unique sequences are determined as noted above.

[0246] In one example, the stringent conditions are selected such that a perfectly complementary oligonucleotide to the coding oligonucleotide hybridizes to the coding oligonucleotide with at least about a 2.5x-10x higher, preferably at least about a 5-10x higher signal to noise ratio than for hybridization of the perfectly complementary oligonucleotide to a control nucleic acid corresponding to any of the control polypeptides. Conditions can be selected such that higher ratios of signal to noise are observed in the particular assay which is used, e.g., about 15x, 20x, 30x, 50x or more. In this example, the target nucleic acid hybridizes to the unique coding oligonucleotide with at least a 2x higher signal to noise ratio as compared to hybridization of the control nucleic acid to the coding oligonucleotide. Again, higher signal to noise ratios can be selected, e.g., about 2.5x, 5x, 10x, 20x, 30x, 50x or more. The particular signal will depend on the label used in the relevant assay, e.g., a fluorescent label, a colorimetric label, a radioactive label, or the like.

Vectors, Promoters and Expression Systems,

[0247] The present invention also includes recombinant constructs comprising one or more of the nucleic acid sequences as broadly described above. The constructs comprise a vector, such as, a plasmid, a cosmid, a phage, a virus, a bacterial artificial chromosome (BAC), a yeast artificial chromosome (YAC), or the like, into which a nucleic acid sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available.

[0248] As previously discussed, general texts which describe molecular biological techniques useful herein, including the use of vectors, promoters and many other relevant topics, include Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology Volume 152, (Academic Press, Inc., San Diego, CA) ("Berger"); Sambrook et al., Molecular Cloning - A Laboratory Manual, 2d ed., Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New

COR-AAD1-00001558

York, 1989 ("Sambrook") and Current Protocols in Molecular Biology, F.M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 1999) ("Ausubel"). Examples of protocols sufficient to direct persons of skill through *in vitro* amplification methods, including the polymerase chain reaction (PCR), the ligase chain reaction (LCR), Qβ-replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA), *e.g.*, for the production of the homologous nucleic acids of the invention are found in Berger, Sambrook, and Ausubel, as well as in Mullis et al. (1987) U.S. Patent No. 4,683,202; Innis et al., eds. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press Inc. San Diego, CA) ("Innis"); Arnheim & Levinson (October 1, 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3: 81-94; Kwoh et al. (1989) Proc. Natl. Acad. Sci. USA 86: 1173; Guatelli et al. (1990) Proc. Nat'l. Acad. Sci. USA 87: 1874; Lomell et al. (1989) J. Clin. Chem 35: 1826; Landegren et al. (1988) Science 241: 1077-1080; Van Brunt (1990) Biotechnology 8: 291-294; Wu and Wallace (1989) Gene 4:560; Barringer et at. (1990) Gene 89:117; and Sooknanan and Malek (1995) Biotechnology 13: 563-564. Improved methods for cloning *in vitro* amplified nucleic acids are described in Wallace et al., U.S. Pat. No. 5,426,039. Improved methods for amplifying large nucleic acids by PCR are summarized in Cheng et al. (1994) Nature 369: 684-685 and the references cited therein, in which PCR amplicons of up to 40kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. *See, e.g.,* Ausubel, Sambrook and Berger, all supra.

[0249] The present invention also relates to engineered host cells that are transduced (transformed or transfected) with a vector of the invention (e.g., an invention cloning vector or an invention expression vector), as well as the production of polypeptides of the invention by recombinant techniques. The vector may be, for example, a plasmid, a viral particle, a phage, *etc.* The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants, or amplifying the GAT homologue gene. Culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to those skilled in the art and in the references cited herein, including, e.g., Sambrook, Ausubel and Berger, as well as e.g., Freshney (1994) Culture of Animal Cells: A Manual of Basic Technique, 3rd ed. (Wiley- Liss, New York) and the references cited therein.

[0250] GAT polypeptides of the invention can be produced in non-animal cells such as plants, yeast, fungi, bacteria and the like. In addition to Sambrook, Berger and Ausubel, details regarding non-animal cell culture can be found in Payne et al. (1992) Plant Cell and Tissue Culture in Liquid Systems (John Wiley & Sons, Inc. New York, NY); Gamborg and Phillips, eds. (1995) Plant Cell, Tissue and Organ Culture: Fundamental Methods/ Springer Lab Manual (Springer-Verlag, Berlin); and Atlas and Parks, eds., The Handbook of Microbiological Media (1993) CRC Press, Boca Raton, FL.

[0251] Polynucleotides of the present invention can be incorporated into any one of a variety of expression vectors suitable for expressing a polypeptide. Suitable vectors include chromosomal, nonchromosomal and synthetic DNA sequences, *e.g.*, derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, pseudorabies, adenovirus, adeno-associated viruses, retroviruses and many others. Any vector that transduces genetic material into a cell, and, if replication is desired, which is replicable and viable in the relevant host can be used.

[0252] When incorporated into an expression vector, a polynucleotide of the invention is operatively linked to an appropriate transcription control sequence (promoter) to direct mRNA synthesis. Examples of such transcription control sequences particularly suited for use in transgenic plants include the cauliflower mosaic virus (CaMV), figwort mosaic virus (FMV) and strawberry vein banding virus (SVBV) promoters, described in U.S. Provisional Application No. 60/245,354. Other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses and which can be used in some embodiments of the invention include SV40 promoter, *E. coli* lac or trp promoter, and the phage lambda P_L promoter. An expression vector optionally contains a ribosome binding site for translation initiation, and a transcription terminator, such as PinII. The vector also optionally includes appropriate sequences for amplifying expression, e.g., an enhancer.

[0253] In addition, the expression vectors of the present invention optionally contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells. Usually, the selectable marker gene will encode antibiotic or herbicide resistance. Suitable genes include those coding for resistance to the antibiotic spectinomycin or streptomycin (e.g., the aada gene), the streptomycin phosphotransferase (SPT) gene coding for streptomycin resistance, the neomycin phosphotransferase (NPTII) gene encoding kanamycin or geneticin resistance, the hygromycin phosphotransferase (HPT) gene coding for hygromycin resistance. Additional selectable marker genes include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, and tetracycline or ampicillin resistance in *E. coli.*

[0254] Suitable genes coding for resistance to herbicides include those which act to inhibit the action of acetolactate synthase (ALS), in particular the sulfonylurea-type herbicides (e.g., the acetolactate synthase (ALS) gene containing mutations leading to such resistance in particular the S4 and/or Hra mutations), those which act to inhibit the action of glutamine synthase, such as phosphinothricin or basta (e.g., the *bar* gene), or other such genes known in the art. The *bar* gene encodes resistance to the herbicide basta and the ALS gene encodes resistance to the herbicide chlorsulfuron.

COR-AAD1-00001559

In some instances, the modified GAT genes are used as selectable markers.

[0255] Vectors of the present invention can be employed to transform an appropriate host to permit the host to express an inventive protein or polypeptide. Examples of appropriate expression hosts include: bacterial cells, such as *E. coli, B. subtilis, Streptomyces,* and *Salmonella typhimurium*; fungal cells, such as *Saccharomyces cerevisiae, Pichia pastoris,* and *Neurospora crassa*; insect cells such as *Drosophila* and *Spodoptera frugiperda*; mammalian cells such as CHO, COS, BHK, HEK 293 or Bowes melanoma; or plant cells or explants, *etc.* It is understood that not all cells or cell lines need to be capable of producing fully functional GAT polypeptides; for example, antigenic fragments of a GAT polypeptide may be produced. The present invention is not limited by the host cells employed.

[0256] In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the GAT polypeptide. For example, when large quantities of GAT polypeptide or fragments thereof are needed for commercial production or for induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be desirable. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the GAT polypeptide coding sequence may be ligated into the vector in-frame with sequences for the amino-terminal Met and the subsequent 7 residues of beta-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke & Schuster (1989) J. Biol. Chem. 264: 5503-5509); pET vectors (Novagen, Madison WI); and the like.

[0257] Similarly, in the yeast *Saccharomyces cerevisiae* a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase and PGH may be used for production of the GAT polypeptides of the invention. For reviews, see Ausubel (supra) and Grant et al. (1987) Methods in Enzymology 153:516-544.

[0258] In mammalian host cells, a variety of expression systems, including viral-based systems, may be utilized. In cases where an adenovirus is used as an expression vector, a coding sequence, e.g., of a GAT polypeptide, is optionally ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion of a GAT polypeptide coding region into a nonessential E1 or E3 region of the viral genome will result in a viable virus capable of expressing a GAT in infected host cells (Logan and Shenk (1984) Proc. Nat'l. Acad. Sci. USA 81:3655-3659). In addition, transcription enhancers, such as the rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

[0259] Similarly, in plant cells, expression can be driven from a transgene integrated into a plant chromosome, or cytoplasmically from an episomal or viral nucleic acid. In the case of stably integrated transgenes, it is often desirable to provide sequences capable of driving constitutive or inducible expression of the GAT polynucleotides of the invention, for example, using viral, e.g., CaMV, or plant derived regulatory sequences. Numerous plant derived regulatory sequences have been described, including sequences which direct expression in a tissue specific manner, e.g., TobRB7, patatin B33, GRP gene promoters, the rbcS-3A promoter, and the like. Alternatively, high level expression can be achieved by transiently expressing exogenous sequences of a plant viral vector, e.g., TMV, BMV, *etc.* Typically, transgenic plants constitutively expressing a GAT polynucleotide of the invention will be preferred, and the regulatory sequences are selected to insure constitutive stable expression of the GAT polypeptide.

[0260] Typical vectors useful for expression of nucleic acids in higher plants are well known in the art and include vectors derived from the tumor-inducing (Ti) plasmid of *Agrobacterium tumefaciens* described by Rogers et al. (1987) Meth. Enzymol. 153: 253-277. Exemplary *A. tumefaciens* vectors useful herein are plasmids pKYLX6 and pKYLX7 of Schardl et al. (1987) Gene 61:1-11 and Berger et al. (1989) Proc. Nat'l. Acad. Sci. U.S.A. 86: 8402-8406. Another useful vector herein is plasmid pBI101.2 that is available from Clontech Laboratories, Inc. (Palo Alto, CA). A variety of plant viruses that can be employed as vectors are known in the art and include cauliflower mosaic virus (CaMV), geminivirus, brome mosaic virus, and tobacco mosaic virus.

[0261] In some embodiments of the present invention, a GAT polynucleotide construct suitable for transformation of plant cells is prepared. For example, a desired GAT polynucleotide can be incorporated into a recombinant expression cassette to facilitate introduction of the gene into a plant and subsequent expression of the encoded polypeptide. An expression cassette will typically comprise a GAT polynucleotide, or functional fragment thereof, operably linked to a promoter sequence and other transcriptional and translational initiation regulatory sequences which will direct expression of the sequence in the intended tissues (e.g., entire plant, leaves, seeds) of the transformed plant.

[0262] For example, a strongly or weakly constitutive plant promoter can be employed which will direct expression of the GAT polypeptide in all tissues of a plant. Such promoters are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens,* the ubiquitin 1 promoter, the Smas promoter, the cinnamyl alcohol dehydrogenase promoter (U.S. Patent No. 5,683,439), the *Nos* promoter, the pEmu promoter, the rubisco promoter, the GRP1-8 promoter and other transcription initiation regions from various plant genes known to those of skill. In situations in which over expression of a GAT polynucleotide is detrimental to the plant or otherwise undesirable, one of skill, upon review of this disclosure, will recognize that weak constitutive promoters can be used for low-levels of expression. In those cases where high levels of expression is not harmful to the plant, a strong promoter, *e.g.*, a t-RNA or other pol III promoter, or a strong pol II promoter, such as the cauliflower

COR-AAD1-00001560

mosaic virus promoter, can be used.

[0263] Alternatively, a plant promoter may be under environmental control. Such promoters are referred to here as "inducible" promoters. Examples of environmental conditions that may effect transcription by inducible promoters include pathogen attack, anaerobic conditions, or the presence of light. In particular, examples of inducible promoters are the Adh1 promoter which is inducible by hypoxia or cold stress, the Hsp70 promoter which is inducible by heat stress, and the PPDK promoter which is inducible by light. Also useful are promoters which are chemically inducible.

[0264] The promoters used in the present invention can be "tissue-specific" and, as such, under developmental control in that the polynucleotide is expressed only in certain tissues, such as leaves, roots, fruit, flowers and/or seeds. An exemplary promoter is the anther specific promoter 5126 (U.S. Patent Nos. 5,689,049 and 5,689,051). Examples of seed-preferred promoters include, but are not limited to, 27 kD gamma zein promoter and waxy promoter, Boronat et al. 1986) Plant Sci. 47, 95-102; Reina et al. (1990) Nucleic Acids Res. 18 (21): 6426; and Kloesgen et al. (1986) Mol. Gen. Genet. 203: 237-244. Promoters that express in the embryo, pericarp, and endosperm are disclosed in U.S. Patent Application Ser. Nos. 60/097,233 filed August 20, 1998 and 60/098,230 filed August 28, 1998. The disclosures each of these are incorporated herein by reference in their entirety. In embodiments in which one or more nucleic acid sequences endogenous to the plant system are incorporated into the construct, the endogenous promoters (or variants thereof) from these genes can be employed for directing expression of the genes in the transfected plant. Tissue-specific promoters can also be used to direct expression of heterologous polynucleotides.

[0265] In general, the particular promoter used in the expression cassette in plants depends on the intended application. Either heterologous or non-heterologous (*i.e.*, endogenous) promoters can be employed to direct expression of the nucleic acids of the present invention. These promoters can also be used, for example, in expression cassettes to drive expression of antisense nucleic acids to reduce, increase, or alter the concentration and/or composition of the proteins of the present invention in a desired tissue. Any of a number of promoters which direct transcription in plant cells are suitable. The promoter can be either constitutive or inducible. In addition to the promoters noted above, promoters of bacterial origin which operate in plants include the octopine synthase promoter, the nopaline synthase promoter and other promoters derived from native Ti plasmids (*see,* Herrara-Estrella et al. (1983) Nature 303: 209-213). Viral promoters include the 35S and 19S RNA promoters of cauliflower mosaic virus (Odell et al. (1985) Nature 313: 810-812). Other plant promoters include the ribulose-1,3-bisphosphate carboxylase small subunit promoter and the phaseolin promoter. The promoter sequence from the E8 gene and other genes may also be used. The isolation and sequence of the E8 promoter is described in detail in Deikman and Fischer (1988) EMBO J. 7: 3315-3327.

[0266] To identify candidate promoters, the 5' portions of a genomic clone is analyzed for sequences characteristic of promoter sequences. For instance, promoter sequence elements include the TATA box consensus sequence (TATAAT), which is usually 20 to 30 base pairs upstream of the transcription start site. In plants, further upstream from the TATA box, at positions -80 to -100, there is typically a promoter element with a series of adenines surrounding the trinucleotide G (or T) as described by Messing et al. (1983) Genetic Engineering in Plants, eds. Kosage, et al., pp. 221-227.

[0267] In preparing polynucleotide constructs, e.g., vectors, of the invention, sequences other than the promoter and the cojoined polynucleotide can also be employed. If normal polypeptide expression is desired, a polyadenylation region at the 3'-end of a GAT-encoding region can be included. The polyadenylation region can be derived, for example, from a variety of plant genes, or from T-DNA. The 3' end sequence to be added can be derived from, for example, the nopaline synthase or octopine synthase genes, or alternatively from another plant gene, or less preferably from any other eukaryotic gene.

[0268] An intron sequence can be added to the 5' untranslated region of the coding sequence or the partial coding sequence to increase the amount of the mature message that accumulates. See for example Buchman and Berg (1988) Mol. Cell Biol. 8: 4395-4405 and Callis et al. (1987) Genes Dev. 1: 1183-1200. Use of maize introns Adhl, intron 1, 2, and 6, and the Bronze-1 intron are known in the art. See generally, Freeling and Walbot, eds. (1994) The Maize Handbook (Springer, New York), chapter 116.

[0269] The construct can also include a marker gene which confers a selectable phenotype on plant cells. For example, the marker may encode biocide tolerance, particularly antibiotic tolerance, such as tolerance to kanamycin, G418, bleomycin, hygromycin, or herbicide tolerance, such as tolerance to chlorsulfuron, or phosphinothricin (the active ingredient in the herbicides bialaphos and Basta).

[0270] Specific initiation signals can aid in efficient translation of a GAT polynucleotide-encoding sequence of the present invention. These signals can include, e.g., the ATG initiation codon and adjacent sequences. In cases where a GAT polypeptide-encoding sequence, its initiation codon and upstream sequences are inserted into an appropriate expression vector, no additional translational control signals may be needed. However, in cases where only the coding sequence (e.g., a mature protein coding sequence), or a portion thereof, is inserted, exogenous transcriptional control signals including the initiation codon must be provided. Furthermore, the initiation codon must be in the correct reading frame to ensure transcription of the entire insert. Exogenous transcriptional elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate to the cell system in use (Scharf et al. (1994) Results Probl. Cell Differ. 20: 125-62 and Bittner et al. (1987)

COR-AAD1-00001561

Methods in Enzymol 153: 516-544).

Secretion/Localization Sequences

[0271] Polynucleotides of the invention can also be fused, for example, in-frame to nucleic acids encoding a secretion/localization sequence, to target polypeptide expression to a desired cellular compartment, membrane, or organelle of a host cell, or to direct polypeptide secretion to the periplasmic space or into the cell culture media. Such sequences are known to those of skill, and include secretion leader peptides, organelle targeting sequences (e.g., nuclear localization sequences, ER retention signals, mitochondrial transit sequences, and chloroplast transit sequences), membrane localization/anchor sequences (e.g., stop transfer sequences, GPI anchor sequences), and the like.

[0272] In a preferred embodiment, a polynucleotide of the invention is fused in frame with an N-terminal chloroplast transit sequence (or chloroplast transit peptide sequence) derived from a gene encoding a polypeptide that is normally targeted to the chloroplast. Such sequences are typically rich in serine and threonine; are deficient in aspartate, glutamate, and tyrosine; and generally have a central domain rich in positively charged amino acids.

Expression Hosts

[0273] In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a eukaryotic cell, such as a mammalian cell, a yeast cell, or a plant cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, electroporation, or other common techniques (Davis et al., Basic Methods in Molecular Biology).

[0274] A host cell is optionally chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the protein include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation. Post-translational processing that cleaves a "pre" or a "prepro" form of the protein may also be important for correct insertion, folding and/or function. Different host cells such as *E. coli, Bacillus sp.,* yeast or mammalian cells such as CHO, HeLa, BHK, MDCK, 293, WI38, *etc.* have specific cellular machinery and characteristic mechanisms, e.g., for post-translational activities and may be chosen to ensure the desired modification and processing of the introduced, foreign protein.

[0275] For long-term, high-yield production of recombinant proteins, stable expression systems can be used. For example, plant cells, explants or tissues, e.g. shoots, or leaf discs, which stably express a polypeptide of the invention are transduced using expression vectors which contain viral origins of replication or endogenous expression elements and a selectable marker gene. Following the introduction of the vector, cells may be allowed to grow for a period determined to be appropriate for the cell type, e.g., 1 or more hours for bacterial cells, 1-4 days for plant cells, 2-4 weeks for some plant explants, in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. For example, transgenic plants expressing the polypeptides of the invention can be selected directly for resistance to the herbicide, glyphosate. Resistant embryos derived from stably transformed explants can be proliferated, e.g., using tissue culture techniques appropriate to the cell type.

[0276] Host cells transformed with a nucleotide sequence encoding a polypeptide of the invention are optionally cultured under conditions suitable for the expression and recovery of the encoded protein from cell culture. The protein or fragment thereof produced by a recombinant cell may be secreted, membrane-bound, or contained intracellularly, depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing GAT polynucleotides of the invention can be designed with signal sequences which direct secretion of the mature polypeptides through a prokaryotic or eukaryotic cell membrane.

Additional Polypeptide Sequences

[0277] Polynucleotides of the present invention may also comprise a coding sequence fused in-frame to a marker sequence that, e.g., facilitates purification of the encoded polypeptide. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, a sequence which binds glutathione (e.g., GST), a hemagglutinin (HA) tag (corresponding to an epitope derived from the influenza hemagglutinin protein; Wilson et al. (1984) Cell 37: 767), maltose binding protein sequences, the FLAG epitope utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle, WA), and the like. The inclusion of a protease-cleavable polypeptide linker sequence between the purification domain and the GAT homologue sequence is useful to facilitate purification. One expression vector contemplated for use in the compositions and methods described herein provides for expression of a fusion protein comprising a polypeptide of the invention fused to a polyhistidine region separated by an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC

COR-AAD1-00001562

(immobilized metal ion affinity chromatography, as described in Porath et al. (1992) Protein Expression and Purification 3: 263-281) while the enterokinase cleavage site provides a means for separating the GAT homologue polypeptide from the fusion protein. pGEX vectors (Promega; Madison, WI) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to ligand-agarose beads (*e.g.*, glutathione-agarose in the case of GST-fusions) followed by elution in the presence of free ligand.

Polypeptide Production and Recovery

[0278] Following transduction of a suitable host and growth of the host cells to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Microbial cells employed in the expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, or other methods, which are well known to those skilled in the art.

[0279] As noted, many references are available for the culture and production of many cells, including cells of bacterial, plant, animal (especially mammalian) and archebacterial origin. *See e.g.*, Sambrook, Ausubel, and Berger (*all supra*), as well as Freshney (1994) Culture of Animal Cells: A Manual of Basic Technique, 3rd ed. (Wiley-Liss, New York) and the references cited therein; Doyle and Griffiths (1997) Mammalian Cell Culture: Essential Techniques (John Wiley and Sons, NY); Humason (1979) Animal Tissue Techniques, 4th ed. (W.H. Freeman and Company); and Ricciardelli, et al. (1989) In vitro Cell Dev. Biol. 25: 1016-1024. For plant cell culture and regeneration see, Payne et al. (1992) Plant Cell and Tissue Culture in Liquid Systems (John Wiley & Sons, Inc., New York, NY); Gamborg and Phillips, eds. (1995) Plant Cell, Tissue and Organ Culture: Fundamental Methods/ Springer Lab Manual (Springer-Verlag, Berlin); Jones, ed. (1984) Plant Gene Transfer and Expression Protocols (Humana Press, Totowa, New Jersey); and Croy, ed. (1993) Plant Molecular Biology (Bios Scientific Publishers, Oxford, U.K.), ISBN 0 12 198370 6. Cell culture media in general are set forth in Atlas and Parks, eds. (1993) The Handbook of Microbiological Media (CRC Press, Boca Raton, FL). Additional information for cell culture is found in available commercial literature such as the Life Science Research Cell Culture Catalogue (1998) from Sigma-Aldrich, Inc. (St Louis, MO) ("Sigma-LSRCCC") and, *e.g.*, The Plant Culture Catalogue and supplement (1997) also from Sigma-Aldrich, Inc. (St Louis, MO) ("Sigma-PCCS"). Further details regarding plant cell transformation and transgenic plant production are found below.

[0280] Polypeptides of the invention can be recovered and purified from recombinant cell cultures by any of a number of methods well known in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography (e.g., using any of the tagging systems noted herein), hydroxylapatite chromatography, and lectin chromatography. Protein refolding steps can be used, as desired, in completing the configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed in the final purification steps. In addition to the references noted supra, a variety of purification methods are well known in the art, including, e.g., those set forth in Sandana (1997) Bioseparation of Proteins (Academic Press, Inc.; Bollag et al. (1996) Protein Methods, 2nd ed. (Wiley-Liss, NY); Walker (1996) The Protein Protocols Handbook (Humana Press, NJ), Harris and Angal (1990) Protein Purification Applications: A Practical Approach (IRL Press at Oxford, Oxford, England); Harris and Angal Protein Purification Methods: A Practical Approach (IRL Press at Oxford, Oxford, England); Scopes (1993) Protein Purification: Principles and Practice, 3rd ed. (Springer Verlag, NY); Janson and Ryden (1998) Protein Purification: Principles, High Resolution Methods and Applications, 2nd ed. (Wiley-VCH, NY); and Walker (1998) Protein Protocols on CD-ROM (Humana Press, NJ).

[0281] In some cases, it is desirable to produce the GAT polypeptide of the invention in a large scale suitable for industrial and/or commercial applications. In such cases bulk fermentation procedures are employed. Briefly, a GAT polynucleotide, e.g., a polynucleotide comprising any one of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952, or other nucleic acids encoding GAT polypeptides of the invention can be cloned into an expression vector. For example, U.S. Patent No. 5,955,310 to Widner et al. "METHODS FOR PRODUCING A POLYPEPTIDE IN A BACILLUS CELL," de-

COR-AAD1-00001563

scribes a vector with tandem promoters, and stabilizing sequences operably linked to a polypeptide encoding sequence. After inserting the polynucleotide of interest into a vector, the vector is transformed into a bacterial, *e.g.,* a *Bacillus subtilis* strain PL1801IIE (amyE, apr, npr, spoIIE::Tn917) host. The introduction of an expression vector into a *Bacillus* cell may, for instance, be effected by protoplast transformation (see, e.g., Chang and Cohen (1979) Mol. Gen. Genet. 168:111), by using competent cells (see, e.g., Young and Spizizin (1961) J. Bacteriol. 81:823, or Dubnau and Davidoff-Abelson (1971) J. Mol. Biol. 56: 209), by electroporation (see, e.g., Shigekawa and Dower (1988) Biotechniques 6: 742), or by conjugation (see, e.g., Koehler and Thorne (1987) J. Bacteriol 169: 5271), see also, Ausubel, Sambrook and Berger, all supra.

[0282] The transformed cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods that are known in the art. For example, the cell may be cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection). The secreted polypeptide can be recovered directly from the medium.

[0283] The resulting polypeptide may be isolated by methods known in the art. For example, the polypeptide may be isolated from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation. The isolated polypeptide may then be further purified by a variety of procedures known in the art including, but not limited to, chromatography (e.g., ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing), differential solubility (e.g., ammonium sulfate precipitation), or extraction (see, *e.g.,* Bollag et al. (1996) Protein Methods, 2nd ed. (Wiley-Liss, NY) and Walker (1996) The Protein Protocols Handbook (Humana Press, NJ).

[0284] Cell-free transcription/translation systems can also be employed to produce polypeptides using DNAs or RNAs of the present invention. Several such systems are commercially available. A general guide to *in vitro* transcription and translation protocols is found in Tymms (1995) In vitro Transcription and Translation Protocols: Methods in Molecular Biology (Garland Publishing, NY), vol. 37.

SUBSTRATES AND FORMATS FOR SEQUENCE RECOMBINATION

[0285] The polynucleotides of the invention are optionally used as substrates for a variety of diversity generating procedures, e.g., mutation, recombination and recursive recombination reactions, in addition to their use in standard cloning methods as set forth in, e.g., Ausubel, Berger and Sambrook, to produce additional GAT polynucleotides and polypeptides with desired properties. A variety of diversity generating protocols are available and described in the art. The procedures can be used separately, and/or in combination to produce one or more variants of a polynucleotide or set of polynucleotides, as well variants of encoded proteins. Individually and collectively, these procedures provide robust, widely applicable ways of generating diversified polynucleotides and sets of polynucleotides (including, e.g., polynucleotide libraries) useful, *e.g.,* for the engineering or rapid evolution of polynucleotides, proteins, pathways, cells and/or organisms with new and/or improved characteristics. The process of altering the sequence can result in, for example, single nucleotide substitutions, multiple nucleotide substitutions, and insertion or deletion of regions of the nucleic acid sequence.

[0286] While distinctions and classifications are made in the course of the ensuing discussion for clarity, it will be appreciated that the techniques are often not mutually exclusive. Indeed, the various methods can be used singly or in combination, in parallel or in series, to access diverse sequence variants.

[0287] The result of any of the diversity generating procedures described herein can be the generation of one or more polynucleotides, which can be selected or screened for polynucleotides that encode proteins with or which confer desirable properties. Following diversification by one or more of the methods described herein, or otherwise available to one of skill, any polynucleotides that are produced can be selected for a desired activity or property, e.g. altered K_m for glyphosate, altered K_m for acetyl CoA, use of alternative cofactors (e.g., propionyl CoA) increased k_{cat}, *etc.* This can include identifying any activity that can be detected, for example, in an automated or automatable format, by any of the assays in the art. For example, GAT homologs with increased specific activity can be detected by assaying the conversion of glyphosate to N-acetylglyphosate, e.g., by mass spectrometry. Alternatively, improved ability to confer resistance to glyphosate can be assayed by growing bacteria transformed with a nucleic acid of the invention on agar containing increasing concentrations of glyphosate or by spraying transgenic plants incorporating a nucleic acid of the invention with glyphosate. A variety of related (or even unrelated) properties can be evaluated, in serial or in parallel, at the discretion of the practitioner. Additional details regarding recombination and selection for herbicide tolerance can be found, e.g., in "DNA SHUFFLING TO PRODUCE HERBICIDE RESISTANT CROPS" (U.S. Pub. No. 2002/0058249) filed August 12,1999.

[0288] Descriptions of a variety of diversity generating procedures, including multigene shuffling and methods for

COR-AAD1-00001564

generating modified nucleic acid sequences encoding multiple enzymatic domains, are found the following publications and the references cited therein: Soong, N. et al. (2000) Nat. Genet. 25(4): 436-39; Stemmer, et al. (1999) Tumor Targeting 4: 1-4; Ness et al. (1999) Nature Biotech. 17:893-896; Chang et al. (1999) Nature Biotech. 17: 793-797; Minshull and Stemmer (1999) Current Opinion in Chemical Biology 3: 284-290; Christians et al. (1999) Nature Biotech. 17: 259-264; Crameri et al. (1998) Nature 391: 288-291; Crameri et al. (1997) Nature Biotech. 15: 436-438; Zhang et al. (1997) Proc. Nat'l. Acad. Sci. USA 94: 4504-4509; Patten et al. (1997) Current Opinion in Biotech. 8: 724-733; Crameri et al. (1996) Nature Med. 2:100-103; Crameri et al. (1996) Nature Biotech. 14:315-319; Gates et al. (1996) J. Mol. Biol. 255: 373-386; Stemmer (1996) "Sexual PCR and Assembly PCR" in The Encyclopedia of Molecular Biology (VCH Publishers, New York) pp.447-457; Crameri and Stemmer (1995) BioTechniques 18: 194-195; Stemmer et al., (1995) Gene 164: 49-53; Stemmer (1995) Science 270: 1510; Stemmer (1995) Bio/Technology 13: 549-553; Stemmer (1994) Nature 370: 389-391; and Stemmer (1994) Proc. Nat'l. Acad. Sci. USA 91:10747-10751.

[0289] Mutational methods of generating diversity include, for example, site-directed mutagenesis (Ling et al. (1997) "Approaches to DNA mutagenesis: an overview" Anal Biochem. 254(2): 157-178; Dale et al. (1996) "Oligonucleotide-directed random mutagenesis using the phosphorothioate method" Methods Mol. Biol. 57:369-374; Smith (1985) "In vitro mutagenesis" Ann. Rev. Genet. 19:423-462; Botstein & Shortle (1985) "Strategies and applications of in vitro mutagenesis" Science 229:1193-1201; Carter (1986) "Site-directed mutagenesis" Biochem. J. 237:1-7; and Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis" in Nucleic Acids & Molecular Biology (Eckstein, F. and Lilley, D.M.J. eds., Springer Verlag, Berlin)); mutagenesis using uracil containing templates (Kunkel (1985) "Rapid and efficient site-specific mutagenesis without phenotypic selection" Proc. Natl. Acad. Sci. USA 82:488-492; Kunkel et al. (1987) "Rapid and efficient site-specific mutagenesis without phenotypic selection" Methods in Enzymol. 154, 367-382; and Bass et al. (1988) "Mutant Trp repressors with new DNA-binding specificities" Science 242:240-245); oligonucleotide-directed mutagenesis (Methods in Enzymol. 100: 468-500 (1983); Methods in Enzymol. 154: 329-350 (1987); Zoller & Smith (1982) "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment" Nucleic Acids Res. 10:6487-6500; Zoller & Smith (1983) "Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors" Methods in Enzymol. 100:468-500; and Zoller & Smith (1987) "Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template" Methods in Enzymol. 154:329-350); phosphorothioate-modified DNA mutagenesis (Taylor et al. (1985) "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA" Nucl. Acids Res. 13: 8749-8764; Taylor et al. (1985) "The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA" Nucl. Acids Res. 13: 8765-8787; Nakamaye & Eckstein (1986) "Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis" Nucl. Acids Res. 14: 9679-9698; Sayers et al. (1988) "Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis" Nucl. Acids Res. 16:791-802; and Sayers et al. (1988) "Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide" Nucl. Acids Res. 16: 803-814); mutagenesis using gapped duplex DNA (Kramer et al. (1984) "The gapped duplex DNA approach to oligonucleotide-directed mutation construction" Nucl. Acids Res. 12: 9441-9456; Kramer & Fritz (1987) Methods in Enzymol. "Oligonucleotide-directed construction of mutations via gapped duplex DNA" 154: 350-367; Kramer et al. (1988) "Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations" Nucl. Acids Res. 16: 7207; and Fritz et al. (1988) "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro" Nucl. Acids Res. 16: 6987-6999).

[0290] Additional suitable methods include point mismatch repair (Kramer et al. (1984) "Point Mismatch Repair" Cell 38:879-887), mutagenesis using repair-deficient host strains (Carter et al. (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors" Nucl. Acids Res. 13: 4431-4443; and Carter (1987) "Improved oligonucleotide-directed mutagenesis using M13 vectors" Methods in Enzymol. 154: 382-403), deletion mutagenesis (Eghtedarzadeh & Henikoff (1986) "Use of oligonucleotides to generate large deletions" Nucl. Acids Res. 14: 5115), restriction-selection and restriction-purification (Wells et al. (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" Phil. Trans. R. Soc. Lond. A 317: 415-423), mutagenesis by total gene synthesis (Nambiar et al. (1984) "Total synthesis and cloning of a gene coding for the ribonuclease S protein" Science 223: 1299-1301; Sakamar and Khorana (1988) "Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)" Nucl. Acids Res. 14: 6361-6372; Wells et al. (1985) "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites" Gene 34:315-323; and Grundström et al. (1985) "Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis" Nucl. Acids Res. 13: 3305-3316); double-strand break repair (Mandecki (1986); Arnold (1993) "Protein engineering for unusual environments" Current Opinion in Biotechnology 4:450-455; and "Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis" Proc. Natl. Acad. Sci. USA, 83:7177-7181). Additional details on many of the above methods can be found in Methods in Enzymology Volume 154, which also describes useful controls for troubleshooting problems with various mutagenesis methods.

COR-AAD1-00001565

[0291] Additional details regarding various diversity generating methods can be found in the following U.S. patents, PCT publications, and EPO publications: U.S. Pat. No. 5,605,793 to Stemmer (February 25, 1997), "Methods for *In vitro* Recombination;" U.S. Pat. No. 5,811,238 to Stemmer et al. (September 22, 1998) "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" U.S. Pat. No. 5,830,721 to Stemmer et al. (November 3, 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly;" U.S. Pat. No. 5,834,252 to Stemmer, et al. (November 10, 1998) "End-Complementary Polymerase Reaction;" U.S. Pat. No. 5,837,458 to Minshull, et al. (November 17, 1998), "Methods and Compositions for Cellular and Metabolic Engineering;" WO 95/22625, Stemmer and Crameri, "Mutagenesis by Random Fragmentation and Reassembly;" WO 96/33207 by Stemmer and Lipschutz "End Complementary Polymerase Chain Reaction;" WO 97/20078 by Stemmer and Crameri "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" WO 97/35966 by Minshull and Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering;" WO 99/41402 by Punnonen et al. "Targeting of Genetic Vaccine Vectors;" WO 99/41383 by Punnonen et al. "Antigen Library Immunization;" WO 99/41369 by Punnonen et al. "Genetic Vaccine Vector Engineering;" WO 99/41368 by Punnonen et al. "Optimization of Immunomodulatory Properties of Genetic Vaccines;" EP 752008 by Stemmer and Crameri, "DNA Mutagenesis by Random Fragmentation and Reassembly;" EP 0932670 by Stemmer "Evolving Cellular DNA Uptake by Recursive Sequence Recombination;" WO 99/23107 by Stemmer et al., "Modification of Virus Tropism and Host Range by Viral Genome Shuffling;" WO 99/21979 by Apt et al., "Human Papillomavirus Vectors;" WO 98/31837 by del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination;" WO 98/27230 by Patten and Stemmer, "Methods and Compositions for Polypeptide Engineering;" WO 98/13487 by Stemmer et al., "Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection;" WO 00/00632, "Methods for Generating Highly Diverse Libraries;" WO 00/09679, "Methods for Obtaining *in vitro* Recombined Polynucleotide Sequence Banks and Resulting Sequences;" WO 98/42832 by Arnold et al., "Recombination of Polynucleotide Sequences Using Random or Defined Primers;" WO 99/29902 by Arnold et al., "Method for Creating Polynucleotide and Polypeptide Sequences;" WO 98/41653 by Vind, "An *in vitro* Method for Construction of a DNA Library;" WO 98/41622 by Borchert et al., "Method for Constructing a Library Using DNA Shuffling;" WO 98/42727 by Pati and Zarling, "Sequence Alterations using Homologous Recombination;" WO 00/18906 by Patten et al., "Shuffling of Codon-Altered Genes;" WO 00/04190 by del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Recombination;" WO 00/42561 by Crameri et al., "Oligonucleotide Mediated Nucleic Acid Recombination;" WO 00/42559 by Selifonov and Stemmer "Methods of Populating Data Structures for Use in Evolutionary Simulations;" WO 00/42560 by Selifonov et al., "Methods for Making Character Strings, Polynucleotides & Polypeptides Having Desired Characteristics;" WO 01/23401 by Welch et al., "Use of Codon-Varied Oligonucleotide Synthesis for Synthetic Shuffling;" and WO 01/64864 "Single-Stranded Nucleic Acid Template-Mediated Recombination and Nucleic Acid Fragment Isolation" by Affholter.

[0292] Certain U.S. applications provide additional details regarding various diversity generating methods, including "SHUFFLING OF CODON ALTERED GENES" by Patten *et al.* filed September 28, 1999, (USSN 09/407,800); "EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE SEQUENCE RECOMBINATION", by del *Cardayre et al.* filed July 15, 1998 (USSN 09/166,188), and July 15, 1999 (U.S. Patent No. 6,379,964); "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" by Crameri *et al.*, filed September 28, 1999 (U.S. Patent No. 6,376,246); "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" by *Crameri et al.*, filed January 18, 2000 (WO 00/42561); "USE OF CODON-BASED OLIGONUCLEOTIDE SYNTHESIS FOR SYNTHETIC SHUFFLING" by Welch *et al.*, filed September 28, 1999 (U.S. Patent No. 6,436,675); "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov *et al.*, filed January 18, 2000, (WO 00/42560); "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov *et al.*, filed July 18, 2000 (USSN 09/618,579); "METHODS OF POPULATING DATA STRUCTURES FOR USE IN EVOLUTIONARY SIMULATIONS" by Selifonov and Stemmer (WO 00/42559), filed January 18, 2000; and "SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED RECOMBINATION AND NUCLEIC ACID FRAGMENT ISOLATION" by Affholter (USSN 60/186,482, filed March 2, 2000).

[0293] In brief, several different general classes of sequence modification methods, such as mutation, recombination, *etc.* are applicable to the present invention and set forth in the references above. That is, alterations to the component nucleic acid sequences to produced modified gene fusion constructs can be performed by any number of the protocols described, either before cojoining of the sequences, or after the cojoining step. The following exemplify some of the different types of preferred formats for diversity generation in the context of the present invention, including, e.g., certain recombination based diversity generation formats.

[0294] Nucleic acids can be recombined *in vitro* by any of a variety of techniques discussed in the references above, including e.g., DNAse digestion of nucleic acids to be recombined followed by ligation and/or PCR reassembly of the nucleic acids. For example, sexual PCR mutagenesis can be used in which random (or pseudo random, or even non-random) fragmentation of the DNA molecule is followed by recombination, based on sequence similarity, between DNA molecules with different but related DNA sequences, *in vitro,* followed by fixation of the crossover by extension in a polymerase chain reaction. This process and many process variants is described in several of the references above,

COR-AAD1-00001566

e.g., in Stemmer (1994) Proc. Nat'l. Acad. Sci. USA 91:10747-10751.

[0295] Similarly, nucleic acids can be recursively recombined *in vivo, e.g.,* by allowing recombination to occur between nucleic acids in cells. Many such *in* vivo recombination formats are set forth in the references noted above. Such formats optionally provide direct recombination between nucleic acids of interest, or provide recombination between vectors, viruses, plasmids, etc., comprising the nucleic acids of interest, as well as other formats. Details regarding such procedures are found in the references noted above.

[0296] Whole genome recombination methods can also be used in which whole genomes of cells or other organisms are recombined, optionally including spiking of the genomic recombination mixtures with desired library components (e.g., genes corresponding to the pathways of the present invention). These methods have many applications, including those in which the identity of a target gene is not known. Details on such methods are found, e.g., in WO 98/31837 by del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination;" and in, *e.g.,* WO 00/04190 by del Cardayre et al., also entitled "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination." Thus, any of these processes and techniques for recombination, recursive recombination, and whole genome recombination, alone or in combination, can be used to generate the modified nucleic acid sequences and/or modified gene fusion constructs of the present invention.

[0297] Synthetic recombination methods can also be used, in which oligonucleotides corresponding to targets of interest are synthesized and reassembled in PCR or ligation reactions which include oligonucleotides which correspond to more than one parental nucleic acid, thereby generating new recombined nucleic acids. Oligonucleotides can be made by standard nucleotide addition methods, or can be made, e.g., by tri-nucleotide synthetic approaches. Details regarding such approaches are found in the references noted above, including, *e.g.,* WO 00/42561 by Crameri et al., "Oligonucleotide Mediated Nucleic Acid Recombination;" WO 01/23401 by Welch et al., "Use of Codon-Varied Oligonucleotide Synthesis for Synthetic Shuffling;" WO 00/42560 by Selifonov et al., "Methods for Making Character Strings, Polynucleotides and Polypeptides Having Desired Characteristics;" and WO 00/42559 by Selifonov and Stemmer "Methods of Populating Data Structures for Use in Evolutionary Simulations."

[0298] *In silico* methods of recombination can be effected in which genetic algorithms are used in a computer to recombine sequence strings which correspond to homologous (or even non-homologous) nucleic acids. The resulting recombined sequence strings are optionally converted into nucleic acids by synthesis of nucleic acids which correspond to the recombined sequences, *e.g.,* in concert with oligonucleotide synthesis gene reassembly techniques. This approach can generate random, partially random or designed variants. Many details regarding *in silico* recombination, including the use of genetic algorithms, genetic operators and the like in computer systems, combined with generation of corresponding nucleic acids (and/or proteins), as well as combinations of designed nucleic acids and/or proteins (*e.g.*, based on cross-over site selection) as well as designed, pseudo-random or random recombination methods are described in WO 00/42560 by Selifonov et al., "Methods for Making Character Strings, Polynucleotides and Polypeptides Having Desired Characteristics" and WO 00/42559 by Selifonov and Stemmer "Methods of Populating Data Structures for Use in Evolutionary Simulations." Extensive details regarding *in silico* recombination methods are found in these applications. This methodology is generally applicable to the present invention in providing for recombination of nucleic acid sequences and/or gene fusion constructs encoding proteins involved in various metabolic pathways (such as, for example, carotenoid biosynthetic pathways, ectoine biosynthetic pathways, polyhydroxyalkanoate biosynthetic pathways, aromatic polyketide biosynthetic pathways, and the like) *in silico* and/or the generation of corresponding nucleic acids or proteins.

[0299] Many methods of accessing natural diversity, *e.g.,* by hybridization of diverse nucleic acids or nucleic acid fragments to single-stranded templates, followed by polymerization and/or ligation to regenerate full-length sequences, optionally followed by degradation of the templates and recovery of the resulting modified nucleic acids can be similarly used. In one method employing a single-stranded template, the fragment population derived from the genomic library (ies) is annealed with partial, or, often approximately full length ssDNA or RNA corresponding to the opposite strand. Assembly of complex chimeric genes from this population is then mediated by nuclease-base removal of non-hybridizing fragment ends, polymerization to fill gaps between such fragments and subsequent single stranded ligation. The parental polynucleotide strand can be removed by digestion (*e.g.,* if RNA or uracil-containing), magnetic separation under denaturing conditions (if labeled in a manner conducive to such separation) and other available separation/purification methods. Alternatively, the parental strand is optionally co-purified with the chimeric strands and removed during subsequent screening and processing steps. Additional details regarding this approach are found, *e.g.,* in "Single-Stranded Nucleic Acid Template-Mediated Recombination and Nucleic Acid Fragment Isolation" by Affholter, WO 01/64864.

[0300] In another approach, single-stranded molecules are converted to double-stranded DNA (dsDNA) and the dsDNA molecules are bound to a solid support by ligand-mediated binding. After separation of unbound DNA, the selected DNA molecules are released from the support and introduced into a suitable host cell to generate a library of enriched sequences which hybridize to the probe. A library produced in this manner provides a desirable substrate for further diversification using any of the procedures described herein.

[0301] Any of the preceding general recombination formats can be practiced in a reiterative fashion (e.g., one or more cycles of mutation/recombination or other diversity generation methods, optionally followed by one or more selection

COR-AAD1-00001567

methods) to generate a more diverse set of recombinant nucleic acids.

[0302] Mutagenesis employing polynucleotide chain termination methods have also been proposed *(see e.g.,* U.S. Patent No. 5,965,408, "Method of DNA reassembly by interrupting synthesis" to Short, and the references above), and can be applied to the present invention. In this approach, double stranded DNAs corresponding to one or more genes sharing regions of sequence similarity are combined and denatured, in the presence or absence of primers specific for the gene. The single stranded polynucleotides are then annealed and incubated in the presence of a polymerase and a chain terminating reagent (*e.g.*, ultraviolet, gamma or X-ray irradiation; ethidium bromide or other intercalators; DNA binding proteins, such as single strand binding proteins, transcription activating factors, or histones; polycyclic aromatic hydrocarbons; trivalent chromium or a trivalent chromium salt; or abbreviated polymerization mediated by rapid thermocycling; and the like), resulting in the production of partial duplex molecules. The partial duplex molecules, *e.g.,* containing partially extended chains, are then denatured and reannealed in subsequent rounds of replication or partial replication resulting in polynucleotides which share varying degrees of sequence similarity and which are diversified with respect to the starting population of DNA molecules. Optionally, the products, or partial pools of the products, can be amplified at one or more stages in the process. Polynucleotides produced by a chain termination method, such as described above, are suitable substrates for any other described recombination format.

[0303] Diversity also can be generated in nucleic acids or populations of nucleic acids using a recombinational procedure termed "incremental truncation for the creation of hybrid enzymes" ("ITCHY") described in Ostermeier et al. (1999) "A combinatorial approach to hybrid enzymes independent of DNA homology" Nature Biotech 17:1205. This approach can be used to generate an initial library of variants which can optionally serve as a substrate for one or more *in vitro* or *in vivo* recombination methods. See, also, Ostermeier et al. (1999) "Combinatorial Protein Engineering by Incremental Truncation," Proc. Natl. Acad. Sci. USA, 96: 3562-67; and Ostermeier et al. (1999), "Incremental Truncation as a Strategy in the Engineering of Novel Biocatalysts," Biological and Medicinal Chemistry, 7: 2139-44.

[0304] Mutational methods which result in the alteration of individual nucleotides or groups of contiguous or non-contiguous nucleotides can be favorably employed to introduce nucleotide diversity into the nucleic acid sequences and/or gene fusion constructs of the present invention. Many mutagenesis methods are found in the above-cited references; additional details regarding mutagenesis methods can be found in following, which can also be applied to the present invention.

[0305] For example, error-prone PCR can be used to generate nucleic acid variants. Using this technique, PCR is performed under conditions where the copying fidelity of the DNA polymerase is low, such that a high rate of point mutations is obtained along the entire length of the PCR product. Examples of such techniques are found in the references above and, *e.g.,* in Leung et al. (1989) Technique 1:11-15 and Caldwell et al. (1992) PCR Methods Applic. 2: 28-33. Similarly, assembly PCR can be used, in a process which involves the assembly of a PCR product from a mixture of small DNA fragments. A large number of different PCR reactions can occur in parallel in the same reaction mixture, with the products of one reaction priming the products of another reaction.

[0306] Oligonucleotide directed mutagenesis can be used to introduce site-specific mutations in a nucleic acid sequence of interest. Examples of such techniques are found in the references above and, e.g., in Reidhaar-Olson et al. (1988) Science 241:53-57. Similarly, cassette mutagenesis can be used in a process that replaces a small region of a double stranded DNA molecule with a synthetic oligonucleotide cassette that differs from the native sequence. The oligonucleotide can contain, *e.g.,* completely and/or partially randomized native sequence(s).

[0307] Recursive ensemble mutagenesis is a process in which an algorithm for protein mutagenesis is used to produce diverse populations of phenotypically related mutants, members of which differ in amino acid sequence. This method uses a feedback mechanism to monitor successive rounds of combinatorial cassette mutagenesis. Examples of this approach are found in Arkin & Youvan (1992) Proc. Nat'l. Acad. Sci. USA 89:7811-7815.

[0308] Exponential ensemble mutagenesis can be used for generating combinatorial libraries with a high percentage of unique and functional mutants. Small groups of residues in a sequence of interest are randomized in parallel to identify, at each altered position, amino acids which lead to functional proteins. Examples of such procedures are found in Delegrave & Youvan (1993) Biotech. Res. 11:1548-1552.

[0309] *In vivo* mutagenesis can be used to generate random mutations in any cloned DNA of interest by propagating the DNA, *e.g.,* in a strain of *E. coli* that carries mutations in one or more of the DNA repair pathways. These "mutator" strains have a higher random mutation rate than that of a wild-type parent. Propagating the DNA in one of these strains will eventually generate random mutations within the DNA. Such procedures are described in the references noted above.

[0310] Other procedures for introducing diversity into a genome, *e.g.* a bacterial; fungal, animal or plant genome can be used in conjunction with the above described and/or referenced methods. For example, in addition to the methods above, techniques have been proposed which produce nucleic acid multimers suitable for transformation into a variety of species (see, *e.g.,* Schellenberger U.S. Patent No. 5,756,316 and the references above). Transformation of a suitable host with such multimers, consisting of genes that are divergent with respect to one another, (*e.g.*, derived from natural diversity or through application of site directed mutagenesis, error prone PCR, passage through mutagenic bacterial strains, and the like), provides a source of nucleic acid diversity for DNA diversification, *e.g.,* by *an in vivo* recombination

COR-AAD1-00001568

process as indicated above.

[0311] Alternatively, a multiplicity of monomeric polynucleotides sharing regions of partial sequence similarity can be transformed into a host species and recombined *in vivo* by the host cell. Subsequent rounds of cell division can be used to generate libraries, members of which, include a single, homogenous population, or pool of monomeric polynucleotides. Alternatively, the monomeric nucleic acids can be recovered by standard techniques, *e.g.,* PCR and/or cloning, and recombined in any of the recombination formats, including recursive recombination formats, described above.

[0312] Methods for generating multispecies expression libraries have been described (in addition to the references noted above, see, *e.g.,* Peterson et al. (1998) U.S. Pat. No. 5,783,431 "METHODS FOR GENERATING AND SCREENING NOVEL METABOLIC PATHWAYS;" and Thompson, et al. (1998) U.S. Pat. No. 5,824,485 METHODS FOR GENERATING AND SCREENING NOVEL METABOLIC PATHWAYS) and their use to identify protein activities of interest has been proposed (in addition to the references noted above, see, Short (1999) U.S. Pat. No. 5,958,672 "PROTEIN ACTIVITY SCREENING OF CLONES HAVING DNA FROM UNCULTIVATED MICROORGANISMS"). Multispecies expression libraries include, in general, libraries comprising cDNA or genomic sequences from a plurality of species or strains, operably linked to appropriate regulatory sequences, in an expression cassette. The cDNA and/or genomic sequences are optionally randomly ligated to further enhance diversity. The vector can be a shuttle vector suitable for transformation and expression in more than one species of host organism, *e.g.,* bacterial species or eukaryotic cells: In some cases, the library is biased by preselecting sequences which encode a protein of interest, or which hybridize to a nucleic acid of interest. Any such libraries can be provided as substrates for any of the methods herein described.

[0313] The above described procedures have been largely directed to increasing nucleic acid and/or encoded protein diversity. However, in many cases, not all of the diversity is useful, *e.g.,* functional, and contributes merely to increasing the background of variants that must be screened or selected to identify the few favorable variants. In some applications, it is desirable to preselect or prescreen libraries (*e.g.*, an amplified library, a genomic library, a cDNA library, a normalized library, *etc.)* or other substrate nucleic acids prior to diversification, e.g., by recombination-based mutagenesis procedures, or to otherwise bias the substrates towards nucleic acids that encode functional products. For example, in the case of antibody engineering, it is possible to bias the diversity generating process toward antibodies with functional antigen binding sites by taking advantage of *in vivo* recombination events prior to manipulation by any of the described methods. For example, recombined CDRs derived from B cell cDNA libraries can be amplified and assembled into framework regions (*e.g.*, Jirholt et al. (1998) "Exploiting sequence space: shuffling in vivo formed complementarity determining regions into a master framework" Gene 215: 471) prior to diversifying according to any of the methods described herein.

[0314] Libraries can be biased towards nucleic acids which encode proteins with desirable enzyme activities. For example, after identifying a clone from a library which exhibits a specified activity, the clone can be mutagenized using any known method for introducing DNA alterations. A library comprising the mutagenized homologues is then screened for a desired activity, which can be the same as or different from the initially specified activity. An example of such a procedure is proposed in Short (1999) U.S. Patent No. 5,939,250 for "PRODUCTION OF ENZYMES HAVING DESIRED ACTIVITIES BY MUTAGENESIS." Desired activities can be identified by any method known in the art. For example, WO 99/10539 proposes that gene libraries can be screened by combining extracts from the gene library with components obtained from metabolically rich cells and identifying combinations which exhibit the desired activity. It has also been proposed (*e.g.*, WO 98/58085) that clones with desired activities can be identified by inserting bioactive substrates into samples of the library, and detecting bioactive fluorescence corresponding to the product of a desired activity using a fluorescent analyzer, *e.g.,* a flow cytometry device, a CCD, a fluorometer, or a spectrophotometer.

[0315] Libraries can also be biased towards nucleic acids which have specified characteristics, *e.g.,* hybridization to a selected nucleic acid probe. For example, WO 99/10539 proposes that polynucleotides encoding a desired activity (*e.g.*, an enzymatic activity, for example: a lipase, an esterase, a protease, a glycosidase, a glycosyl transferase, a phosphatase, a kinase, an oxygenase, a peroxidase, a hydrolase, a hydratase, a nitrilase, a transaminase, an amidase or an acylase) can be identified from among genomic DNA sequences. In particular, single stranded DNA molecules from a population of genomic DNA are hybridized to a ligand-conjugated probe. The genomic DNA can be derived from either a cultivated or uncultivated microorganism, or from an environmental sample. Alternatively, the genomic DNA can be derived from a multicellular organism, or a tissue derived therefrom. Second strand synthesis can be conducted directly from the hybridization probe used in the capture, with or without prior release from the capture medium or by a wide variety of other strategies known in the art. Alternatively, the isolated single-stranded genomic DNA population can be fragmented without further cloning and used directly in, *e.g.,* a recombination-based approach, that employs a single-stranded template, as described above.

[0316] "Non-stochastic" methods of generating nucleic acids and polypeptides are described in Short "Non-Stochastic Generation of Genetic Vaccines and Enzymes" WO 00/46344. These methods, including proposed non-stochastic polynucleotide reassembly and site-saturation mutagenesis methods can be applied to the present invention as well. Random or semi-random mutagenesis using doped or degenerate oligonucleotides is also described in, *e.g.,* Arkin and Youvan (1992) "Optimizing nucleotide mixtures to encode specific subsets of amino acids for semi-random mutagenesis"

COR-AAD1-00001569

Biotechnology 10:297-300; Reidhaar-Olson et al. (1991) "Random mutagenesis of protein sequences using oligonucleotide cassettes" Methods Enzymol. 208:564-86; Lim and Sauer (1991) "The role of internal packing interactions in determining the structure and stability of a protein" J. Mol. Biol. 219:359-76; Breyer and Sauer (1989) "Mutational analysis of the fine specificity of binding of monoclonal antibody 51F to lambda repressor" J. Biol. Chem. 264:13355-60); "Walk-Through Mutagenesis" (Crea, R; U.S. Patent Nos. 5,830,650 and 5,798,208, and EP Patent 0527809 B1.

[0317] It will be readily appreciated that any of the above described techniques suitable for enriching a library prior to diversification can also be used to screen the products, or libraries of products, produced by the diversity generating methods. Any of the above described methods can be practiced recursively or in combination to alter nucleic acids, *e.g.,* GAT encoding polynucleotides.

[0318] Kits for mutagenesis, library construction and other diversity generation methods are also commercially available. For example, kits are available from, *e.g.,* Stratagene (*e.g.*, QuickChange™ site-directed mutagenesis kit; and Chameleon™ double-stranded, site-directed mutagenesis kit); Bio/Can Scientific, Bio-Rad (*e.g.*, using the Kunkel method described above); Boehringer Mannheim Corp.; Clonetech Laboratories; DNA Technologies; Epicentre Technologies (*e.g.*, 5 prime 3 prime kit); Genpak Inc.; Lemargo Inc.; Life Technologies (Gibco BRL); New England Biolabs; Pharmacia Biotech; Promega Corp.; Quantum Biotechnologies; Amersham International plc (*e.g.*, using the Eckstein method above); and Anglian Biotechnology Ltd (*e.g.*, using the Carter/Winter method above).

[0319] The above references provide many mutational formats, including recombination, recursive recombination, recursive mutation and combinations of recombination with other forms of mutagenesis, as well as many modifications of these formats. Regardless of the diversity generation format that is used, the nucleic acids of the present invention can be recombined (with each other, or with related (or even unrelated) sequences) to produce a diverse set of recombinant nucleic acids for use in the gene fusion constructs and modified gene fusion constructs of the present invention, including, *e.g.,* sets of homologous nucleic acids, as well as corresponding polypeptides.

[0320] Many of the above-described methodologies for generating modified polynucleotides generate a large number of diverse variants of a parental sequence or sequences. In some preferred embodiments of the invention the modification technique (*e.g.*, some form of shuffling) is used to generate a library of variants that is then screened for a modified polynucleotide or pool of modified polynucleotides encoding some desired functional attribute, *e.g.,* improved GAT activity. Exemplary enzymatic activities that can be screened for include catalytic rates (conventionally characterized in terms of kinetic constants such as k_{cat} and K_M), substrate specificity, and susceptibility to activation or inhibition by substrate, product or other molecules (*e.g.*, inhibitors or activators).

[0321] One example of selection for a desired enzymatic activity entails growing host cells under conditions that inhibit the growth and/or survival of cells that do not sufficiently express an enzymatic activity of interest, *e.g.* the GAT activity. Using such a selection process can eliminate from consideration all modified polynucleotides except those encoding a desired enzymatic activity. For example, in some embodiments of the invention host cells are maintained under conditions that inhibit cell growth or survival in the absence of sufficient levels of GAT, *e.g.,* a concentration of glyphosate that is lethal or inhibits the growth of a wild-type plant of the same variety that either lacks or does not express a GAT polynucleotide. Under these conditions, only a host cell harboring a modified nucleic acid that encodes enzymatic activity or activities able to catalyze production of sufficient levels of the product will survive and grow. Some embodiments of the invention employ multiple rounds of screening at increasing concentrations of glyphosate or a glyphosate analog.

[0322] In some embodiments of the invention, mass spectrometry is used to detect the acetylation of glyphosate, or a glyphosate analog or metabolite. The use of mass spectrometry is described in more detail in the Examples below.

[0323] For convenience and high throughput it will often be desirable to screen/select for desired modified nucleic acids in a microorganism, e.g., a bacteria such as *E. coli.* On the other hand, screening in plant cells or plants can in some cases be preferable where the ultimate aim is to generate a modified nucleic acid for expression in a plant system.

[0324] In some preferred embodiments of the invention throughput is increased by screening pools of host cells expressing different modified nucleic acids, either alone or as part of a gene fusion construct. Any pools showing significant activity can be deconvoluted to identify single clones expressing the desirable activity.

[0325] The skilled artisan will recognize that the relevant assay, screening or selection method will vary depending upon the desired host organism and other parameters known in the art. It is normally advantageous to employ an assay that can be practiced in a high-throughput format.

[0326] In high-throughput assays, it is possible to screen up to several thousand different variants in a single day. For example, each well of a microtiter plate can be used to run a separate assay, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single variant.

[0327] In addition to fluidic approaches, it is possible, as mentioned above, simply to grow cells on media plates that select for the desired enzymatic or metabolic function. This approach offers a simple and high-throughput screening method.

[0328] A number of well known robotic systems have also been developed for solution phase chemistries useful in assay systems. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark

COR-AAD1-00001570

Corporation, Hopkinton, MA; and Orca, Hewlett-Packard, Palo Alto, CA) which mimic the manual synthetic operations performed by a scientist. Any of the above devices are suitable for application to the present invention. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein with reference to the integrated system will be apparent to persons skilled in the relevant art.

[0329] High-throughput screening systems are commercially available *(see, e.g.,* Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, *etc.).* These systems typically automate entire procedures including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the particular assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization.

[0330] The manufacturers of such systems provide detailed protocols for the various high throughput devices. Thus, for example, Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like. Microfluidic approaches to reagent manipulation have also been developed, *e.g.,* by Caliper Technologies (Mountain View, CA).

[0331] Optical images viewed (and, optionally, recorded) by a camera or other recording device *(e.g.,* a photodiode and data storage device) are optionally further processed in any of the embodiments herein, *e.g.,* by digitizing the image and/or storing and analyzing the image on a computer. A variety of commercially available peripheral equipment and software is available for digitizing, storing and analyzing a digitized video or digitized optical image, e.g., using PC (Intel x86 or Pentium chip compatible DOS™, OS™ WINDOWS™, WINDOWS NT™ or WINDOWS 95™ based machines), MACINTOSH™, or UNIX based *(e.g.,* SUN™ work station) computers.

[0332] One conventional system carries light from the assay device to a cooled charge-coupled device (CCD) camera, a common use in the art. A CCD camera includes an array of picture elements (pixels). The light from the specimen is imaged on the CCD. Particular pixels corresponding to regions of the specimen (e.g., individual hybridization sites on an array of biological polymers) are sampled to obtain light intensity readings for each position. Multiple pixels are processed in parallel to increase speed. The apparatus and methods of the invention are easily used for viewing any sample, *e.g.* by fluorescent or dark field microscopic techniques.

OTHER POLYNUCLEOTIDE COMPOSITIONS

[0333] The invention also includes compositions comprising two or more polynucleotides of the invention (*e.g.*, as substrates for recombination). The composition can comprise a library of recombinant nucleic acids, where the library contains at least 2, 3, 5, 10, 20, or 50 or more polynucleotides. The polynucleotides are optionally cloned into expression vectors, providing expression libraries.

[0334] The invention also includes compositions produced by digesting one or more polynucleotide of the invention with a restriction endonuclease, an RNAse, or a DNAse (*e.g.*, as is performed in certain of the recombination formats noted above); and compositions produced by fragmenting or shearing one or more polynucleotide of the invention by mechanical means (*e.g.*, sonication, vortexing, and the like), which can also be used to provide substrates for recombination in the methods above. Similarly, compositions comprising sets of oligonucleotides corresponding to more than one nucleic acid of the invention are useful as recombination substrates and are a feature of the invention. For convenience, these fragmented, sheared, or oligonucleotide synthesized mixtures are referred to as fragmented nucleic acid sets.

[0335] Also included in the invention are compositions produced by incubating one or more of the fragmented nucleic acid sets in the presence of ribonucleotide- or deoxyribonucelotide triphosphates and a nucleic acid polymerase. This resulting composition forms a recombination mixture for many of the recombination formats noted above. The nucleic acid polymerase may be an RNA polymerase, a DNA polymerase, or an RNA-directed DNA polymerase (*e.g.*, a "reverse transcriptase"); the polymerase can be, *e.g.,* a thermostable DNA polymerase (such as, VENT, TAQ, or the like).

INTEGRATED SYSTEMS

[0336] The present invention provides computers, computer readable media and integrated systems comprising character strings corresponding to the sequence information herein for the polypeptides and nucleic acids herein, including, e.g., those sequences listed herein and the various silent substitutions and conservative substitutions thereof.

[0337] For example, various methods and genetic algorithms (GAs) known in the art can be used to detect homology or similarity between different character strings, or can be used to perform other desirable functions such as to control output files, provide the basis for making presentations of information including the sequences and the like. Examples include BLAST, discussed *supra.*

[0338] Thus, different types of homology and similarity of various stringency and length can be detected and recognized in the integrated systems described herein. For example, many homology determination methods have been designed for comparative analysis of sequences ofbiopolymers, for spell-checking in word processing, and for data retrieval from various databases. With an understanding of double-helix pair-wise complement interactions among 4 principal nucle-

COR-AAD1-00001571

obases in natural polynucleotides, models that simulate annealing of complementary homologous polynucleotide strings can also be used as a foundation of sequence alignment or other operations typically performed on the character strings corresponding to the sequences herein (*e.g.*, word-processing manipulations, construction of figures comprising sequence or subsequence character strings, output tables, *etc.*). An example of a software package with GAs for calculating sequence similarity is BLAST, which can be adapted to the present invention by inputting character strings corresponding to the sequences herein.

[0339] Similarly, standard desktop applications such as word processing software (*e.g.*, Microsoft Word™ or Corel WordPerfect™) and database software (*e.g.*, spreadsheet software such as Microsoft Excel™, Corel Quattro Pro™, or database programs such as Microsoft Access™ or Paradox™) can be adapted to the present invention by inputting a character string corresponding to the GAT homologues of the invention (either nucleic acids or proteins, or both). For example, the integrated systems can include the foregoing software having the appropriate character string information, *e.g.*, used in conjunction with a user interface (*e.g.*, a GUI in a standard operating system such as a Windows, Macintosh or LINUX system) to manipulate strings of characters. As noted, specialized alignment programs such as BLAST can also be incorporated into the systems of the invention for alignment of nucleic acids or proteins (or corresponding character strings).

[0340] Integrated systems for analysis in the present invention typically include a digital computer with GA software for aligning sequences, as well as data sets entered into the software system comprising any of the sequences herein. The computer can be, *e.g.*, a PC (Intel x86 or Pentium chip compatible DOS™, OS2™ WINDOWS™ WINDOWS NT™, WINDOWS95™, WINDOWS98™ LINUX based machine, a MACINTOSH™, Power PC, or a UNIX based (*e.g.*, SUN™ work station) machine) or other commercially common computer which is known to one of skill. Software for aligning or otherwise manipulating sequences is available, or can easily be constructed by one of skill using a standard programming language such as Visualbasic, Fortran, Basic, Java, or the like.

[0341] Any controller or computer optionally includes a monitor which is often a cathode ray tube ("CRT") display, a flat panel display (*e.g.*, active matrix liquid crystal display, liquid crystal display), or others. Computer circuitry is often placed in a box which includes numerous integrated circuit chips, such as a microprocessor, memory, interface circuits, and others. The box also optionally includes a hard disk drive, a floppy disk drive, a high capacity removable drive such as a writeable CD-ROM, and other common peripheral elements. Inputting devices such as a keyboard or mouse optionally provide for input from a user and for user selection of sequences to be compared or otherwise manipulated in the relevant computer system.

[0342] The computer typically includes appropriate software for receiving user instructions, either in the form of user input into set parameter fields, *e.g.*, in a GUI, or in the form of preprogrammed instructions, *e.g.*, preprogrammed for a variety of different specific operations. The software then converts these instructions to appropriate language for instructing the operation of the fluid direction and transport controller to carry out the desired operation.

[0343] The software can also include output elements for controlling nucleic acid synthesis (*e.g.*, based upon a sequence or an alignment of a sequences herein) or other operations which occur downstream from an alignment or other operation performed using a character string corresponding to a sequence herein. Nucleic acid synthesis equipment can, accordingly, be a component in one or more integrated systems herein.

[0344] In an additional aspect, the present invention provides kits embodying the methods, composition, systems and apparatus herein. Kits of the invention optionally comprise one or more of the following: (1) an apparatus, system, system component or apparatus component as described herein; (2) instructions for practicing the methods described herein, and/or for operating the apparatus or apparatus components herein and/or for using the compositions herein; (3) one or more GAT compositions or components; (4) a container for holding components or compositions, and, (5) packaging materials.

[0345] In a further aspect, the present invention provides for the use of any apparatus, apparatus component, composition or kit herein, for the practice of any method or assay herein, and/or for the use of any apparatus or kit to practice any assay or method herein.

HOST CELLS AND ORGANISMS

[0346] The host cell can be eukaryotic, for example, a eukaryotic cell, a plant cell, an animal cell, a protoplast, or a tissue culture cell. The host cell optionally comprises a plurality of cells, for example, an organism. Alternatively, the host cell can be prokaryotic including, but not limited to, bacteria *(i.e.*, gram positive bacteria, purple bacteria, green sulfur bacteria, green non-sulfur bacteria, cyanobacteria, spirochetes, thermatogales, flavobacteria, and bacteroides) and archaebacteria (*i.e.*, Korarchaeota, Thermoproteus, Pyrodictium, Thermococcales, Methanogens, Archaeoglobus, and extreme Halophiles).

[0347] Transgenic plants, or plant cells, incorporating the GAT nucleic acids, and/or expressing the GAT polypeptides of the invention are a feature of the invention. The transformation of plant cells and protoplasts can be carried out in essentially any of the various ways known to those skilled in the art of plant molecular biology, including, but not limited

COR-AAD1-00001572

to, the methods described herein. *See,* in general, Methods in Enzymology, Vol. 153 (Recombinant DNA Part D) Wu and Grossman (eds.) 1987, Academic Press; and Weising et al., Ann. Rev. Genet. 22: 421-477 (1988), incorporated herein by reference. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation, PEG-mediated transfection, particle bombardment, silicon fiber delivery, or microinjection of plant cell protoplasts or embryogenic callus. See, *e.g.,* Tomes, et al. (1995) "Direct DNA Transfer into Intact Plant Cells Via Microprojectile Bombardment," in Plant Cell, Tissue and Organ Culture, Fundamental Methods, eds. Gamborg and Phillips (Springer-Verlag, Berlin), pp.197-213. Further methods for transforming various host cells are disclosed in Klein et al. (1992) "Transformation of microbes, plants and animals by particle bombardment" BiolTechnol. 10 (3): 286-291.

[0348] The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski et al. (1984) EMBO J. 3:2717-2722. Electroporation techniques are described in Fromm et al. (1985) Proc. Natl. Acad. Sci. 82:5824. Ballistic transformation techniques are described in Klein et al. (1987) Nature 327: 70-73.

[0349] Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria. See, U.S. Patent No. 5,591,616.

[0350] *Agrobacterium tumefaciens*-meditated transformation techniques are well described in the scientific literature. See, for example Horsch et al. (1984) Science 233: 496-498, and Fraley et al. (1983) Proc. Natl. Acad. Sci. 80:4803. For instance, *Agrobacterium* transformation of maize is described in U.S. Patent Nos. 5,550,318 and 5,981,840.

[0351] Other methods of transformation include *(1) Agrobacterium rhizogenes*-mediated transformation (see, *e.g.,* Lichtenstein and Fuller In: Genetic Engineering, Vol. 6, PWJ Rigby, ed., London, Academic Press, 1987; Lichtenstein, C. P., and Draper, J,. In: DNA Cloning, Vol. II, D. M. Glover, Ed., Oxford, IRI Press, 1985;WO 88/02405 describes the use *of A. rhizogenes* strain A4 and its Ri plasmid along with *A. tumefaciens* vectors pARC8 or pARC16); (2) liposome-mediated DNA uptake (see, *e.g.,* Freeman et al. (1984) Plant Cell Physiol. 25:1353; (3) the vortexing method (see, *e.g.,* Kindle (1990) Proc. Natl. Acad. Sci. USA 87:1228.

[0352] DNA can also be introduced into plants by direct DNA transfer into pollen as described by Zhou et al. (1983) Methods in Enzymology 101:433; D. Hess (1987) Intern Rev. Cytol. 107:367; and Luo et al. (1988) Plant Mol. Biol. Reporter 6:165. Expression of polypeptide coding nucleic acids can be obtained by injection of the DNA into reproductive organs of a plant as described by Pena et al. (1987) Nature 325:274. DNA can also be injected directly into the cells of immature embryos and the rehydration of desiccated embryos as described by Neuhaus et al. (1987) Theor. Appl. Genet. 75: 30; and Benbrook et al. (1986) in Proceedings Bio Expo 1986, Butterworth, Stoneham, Mass., pp. 27-54.

[0353] Animal and lower eukaryotic (*e.g.,* yeast) host cells are competent or rendered competent for transfection by various means. There are several well-known methods of introducing DNA into animal cells. These methods include: calcium phosphate precipitation; fusion of the recipient cells with bacterial protoplasts containing the DNA; treatment of the recipient cells with liposomes containing the DNA; DEAE dextran; electroporation; biolistics; and microinjection of the DNA directly into the cells. The transfected cells are cultured by means well known in the art. See, Kuchler, R.J. (1977) Biochemical Methods in Cell Culture and Virology (Dowden, Hutchinson and Ross, Inc.). As used herein, the term "transformation" means alteration of the genotype of a host plant by the introduction of a nucleic acid sequence, *e.g.,* a "heterologous" or "foreign" nucleic acid sequence. The heterologous nucleic acid sequence need not necessarily originate from a different source but it will, at some point, have been external to the cell into which is introduced.

[0354] In addition to Berger, Ausubel and Sambrook, useful general references for plant cell cloning, culture and regeneration include Jones, ed. (1995) Plant Gene Transfer and Expression Protocols– Methods in Molecular Biology, volume 49 (Humana Press, Towata, NJ); Payne et al. (1992) Plant Cell and Tissue Culture in Liquid Systems (John Wiley & Sons, Inc. New York, NY) ("Payne"); and Gamborg and Phillips, eds. (1995) Plant Cell, Tissue and Ozgan Culture; Fundamental Methods/Springer Lab Manual, (Springer-Verlag, Berlin) ("Gamborg"). A variety of cell culture media are described in Atlas and Parks, eds. The Handbook of Microbiological Media_(CRC Press, Boca Raton, FL) ("Atlas"). Additional information for plant cell culture is found in available commercial literature such as the Life Science Research Cell Culture Catalogue (1998) from Sigma-Aldrich, Inc. (St Louis, MO) (Sigma-LSRCCC) and, *e.g.,* the Plant Culture Catalogue and supplement (1997) also from Sigma-Aldrich, Inc. (St Louis, MO) (Sigma-PCCS). Additional details regarding plant cell culture are found in Croy, ed. (1993) Plant Molecular Biology (Bios Scientific Publishers, Oxford, UK).

[0355] In an embodiment of this invention, recombinant vectors including one or more GAT polynucleotides, suitable for the transformation of plant cells are prepared. A DNA sequence encoding for the desired GAT polypeptide, *e.g.,* selected from among SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790,

COR-AAD1-00001573

792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952, is conveniently used to construct a recombinant expression cassette which can be introduced into the desired plant. In the context of the present invention, an expression cassette will typically comprise a selected GAT polynucleotide operably linked to a promoter sequence and other transcriptional and translational initiation regulatory sequences which are sufficient to direct the transcription of the GAT sequence in the intended tissues (*e.g.*, entire plant, leaves, roots, *etc.)* of the transformed plant.

[0356] A number of promoters can be used in the practice of the present invention. The promoters can be selected based on the desired outcome. That is, the nucleic acids can be combined with constitutive, tissue-preferred, or other promoters for expression in plants.

[0357] Constitutive promoters include, for example, the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO 99/43838 and U.S. Patent No. 6,072,050; the core CaMV 35S promoter (Odell et al. (1985) Nature 313:810-812); rice actin (McElroy et al. (1990) Plant Cell 2:163-171); ubiquitin (Christensen et al. (1989) Plant Mol. Biol. 12:619-632 and Christensen et al. (1992) Plant Mol. Biol. 18:675-689); pEMU (Last et al. (1991) Theor. Appl. Genet. 81:581-588); MAS (Velten et al. (1984) EMBO J. 3:2723-2730); ALS promoter (U.S. Patent No. 5,659,026), and the like. Other constitutive promoters include, for example, those disclosed in U.S. Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611.

[0358] Chemical-regulated promoters can be used to modulate the expression of a gene in a plant through the application of an exogenous chemical regulator. Depending upon the objective, the promoter may be a chemical-inducible promoter, where application of the chemical induces gene expression, or a chemical-repressible promoter, where application of the chemical represses gene expression. Chemical-inducible promoters are known in the art and include, but are not limited to, the maize In2-2 promoter, which is activated by benzene sulfonamide herbicide safeners; the maize GST promoter, which is activated by hydrophobic electrophilic compounds that are used as pre-emergent herbicides; and the tobacco PR-1a promoter, which is activated by salicylic acid. Other chemical-regulated promoters of interest include steroid-responsive promoters. See, for example, the glucocorticoid-inducible promoter in Schena et al. (1991) Proc. Natl. Acad. Sci. USA 88:10421-10425 and McNellis et al. (1998) Plant J. 14(2):247-257 and the tetracycline-inducible and tetracycline-repressible promoters for example, Gatz et al. (1991) Mol. Gen. Genet. 227:229-237, and U.S. Patent Nos. 5,814,618 and 5,789,156, herein incorporated by reference.

[0359] Tissue-preferred promoters can also be utilized to target GAT expression within a particular plant tissue. Tissue-preferred promoters include those disclosed in Yamamoto et al. (1997) Plant J. 12(2):255-265; Kawamata et al. (1997) Plant Cell Physiol. 38(7):792-803; Hansen et al. (1997) Mol. Gen Genet. 254(3):337-343; Russell et al. (1997) Transgenic Res. 6(2):157-168; Rinehart et al. (1996) Plant Physiol. 112(3):1331-1341; Van Camp et al. (1996) Plant Physiol. 112 (2):525-535; Canevascini et al. (1996) Plant Physiol. 112(2):513-524; Yamamoto et al. (1994) Plant Cell Physiol. 35(5): 773-778; Lam (1994) Results Probl. Cell Differ. 20:181-196; Orozco et al. (1993) Plant Mol Biol. 23(6):1129-1138; Matsuoka et al. (1993) Proc Natl. Acad. Sci. USA 90(20):9586-9590; and Guevara-Garcia et al. (1993) Plant J. 4(3): 495-505. Such promoters can be modified, if necessary, for weak expression.

[0360] Leaf-specific promoters are known in the art. See, for example, Yamamoto et al. (1997) Plant J. 12(2):255-265; Kwon et al. (1994) Plant Physiol. 105:357-67; Yamamoto et al. (1994) Plant Cell Physiol. 35(5):773-778; Gotor et al. (1993) Plant J. 3:509-18; Orozco et al. (1993) Plant Mol. Biol. 23(6):1129-1138; and Matsuoka et al. (1993) Proc. Natl. Acad. Sci. USA 90(20):9586-9590.

[0361] Root-preferred promoters are known and can be selected from the many available from the literature or isolated *de novo* from various compatible species. See, for example, Hire et al. (1992) Plant Mol. Biol. 20(2):207-218 (soybean root-specific glutamine synthetase gene); Keller et al. (1991) Plant Cell 3(10):1051-1061 (root-specific control element in the GRP 1.8 gene of French bean); Sanger et al. (1990) Plant Mol. Biol. 14(3):433-443 (root-specific promoter of the mannopine synthase (MAS) gene of *Agrobacterium tumefaciens);* and Miao et al. (1991) Plant Cell 3(1):11-22 (full-length cDNA clone encoding cytosolic glutamine synthetase (GS), which is expressed in roots and root nodules of soybean). See also Bogusz et al. (1990) Plant Cell 2(7):633-641, which discloses two root-specific promoters isolated from hemoglobin genes from the nitrogen-fixing nonlegume *Parasponia andersonii* and the related non-nitrogen-fixing nonlegume *Trema tomentosa.* The promoters of these genes were linked to a ß-glucuronidase reporter gene and introduced into both the nonlegume *Nicotiana tabacum* and the legume *Lotus corniculatus,* and in both instances root-specific promoter activity was preserved. Leach *et al.* (1991) describe their analysis of the promoters of the highly expressed *rolC* and *rolD* root-inducing genes of *Agrobacterium rhizogenes* (see Plant Science (Limerick) 79(1):69-76). They concluded that enhancer and tissue-preferred DNA determinants are dissociated in those promoters. Teeri et al. (1989) EMBO J. 8(2):343-350 used gene fusion to *lacZ* to show that the *Agrobacterium* T-DNA gene encoding octopine synthase is especially active in the epidermis of the root tip and that the TR2' gene is root specific in the intact plant and stimulated by wounding in leaf tissue, which is an especially desirable combination of characteristics for use with an insecticidal or larvicidal gene. The TR1' gene, fused to *nptII* (neomycin phosphotransferase II), showed similar characteristics. Additional

COR-AAD1-00001574

root-preferred promoters include the VfENOD-GRP3 gene promoter (Kuster et al. (1995) Plant Mol. Biol. 29(4):759-772); the ZRP2 promoter (U.S. Patent No. 5,633,636); the IFS1 promoter (U.S. Patent Application Serial No. 10/104,706) and the *rolB* promoter (Capana et al. (1994) Plant Mol. Biol. 25(4):681-691). See also U.S. Patent Nos. 5,837,876; 5,750,386; 5,459,252; 5,401,836; 5,110,732; and 5,023,179.

[0362] "Seed-preferred" promoters include both "seed-specific" promoters (those promoters active during seed development such as promoters of seed storage proteins) as well as "seed-germinating" promoters (those promoters active during seed germination). See Thompson et al. (1989) BioEssays 10:108, herein incorporated by reference. Such seed-preferred promoters include, but are not limited to, *Cim1* (cytokinin-induced message); *cZ19B1* (maize 19 kDa zein); *milps* (myo-inositol-1-phosphate synthase); and *celA* (cellulose synthase) (see U.S. Patent No. 6,225,529, herein incorporated by reference). Gamma-zein is an endosperm-specific promoter. Glob-1 is an embryo-specific promoter. For dicots, seed-specific promoters include, but are not limited to, bean β-phaseolin, napkin, β-conglycinin, soybean lectin, cruciferin, and the like. For monocots, seed-specific promoters include, but are not limited to, maize 15 kDa zein, 22 kDa zein, 27 kDa zein, g-zein, waxy; shrunken 1, shrunken 2, globulin 1, *etc.* See also WO 00/12733, which discloses seed-preferred promoters from *end1* and *end2* genes; herein incorporated by reference.

[0363] In particular, a strongly or weakly constitutive plant promoter that directs expression of a GAT nucleic acid in all tissues of a plant can be favorably employed. Such promoters are active under most environmental conditions and states of development or cell differentiation. In addition to the promoters mentioned above examples of constitutive promoters include the 1'- or 2'-promoter *of Agrobacterium tumefaciens,* and other transcription initiation regions from various plant genes known to those of skill. Where over expression of a GAT polypeptide of the invention is detrimental to the plant, one of skill will recognize that weak constitutive promoters can be used for low-levels of expression. Generally, by "weak promoter" a promoter that drives expression of a coding sequence at a low level is intended. By "low level" levels from about 1/1000 transcripts to about 1/100,000 transcripts, to about as low as 1/500,000 transcripts per cell are intended. Alternatively, it is recognized that weak promoters also include promoters that are expressed in only a few cells and not in others to give a total low level of expression. Where a promoter is expressed at unacceptably high levels, portions of the promoter sequence can be deleted or modified to decrease expression levels. In those cases where high levels of expression is not harmful to the plant, a strong promoter, *e.g.,* a t-RNA, or other pol III promoter, or a strong pol II promoter, (*e.g.*, the cauliflower mosaic virus promoter, CaMV, 35S promoter) can be used.

[0364] Alternatively, a plant promoter can be under environmental control. Such promoters are referred to as "inducible" promoters. Examples of environmental conditions that may alter transcription by inducible promoters include pathogen attack, anaerobic conditions, or the presence of light. In some cases, it is desirable to use promoters that are "tissue-specific" and/or are under developmental control such that the GAT polynucleotide is expressed only in certain tissues or stages of development, e.g., leaves, roots, shoots, *etc.* Endogenous promoters of genes related to herbicide tolerance and related phenotypes are particularly useful for driving expression of GAT nucleic acids, e.g., P450 monooxygenases, glutathione-S-transferases, homoglutathione-S-transferases, glyphosate oxidases and 5-enolpyruvylshikimate-2-phosphate synthases.

[0365] Tissue specific promoters can also be used to direct expression of heterologous structural genes, including the GAT polynucleotides described herein. Thus the promoters can be used in recombinant expression cassettes to drive expression of any gene whose expression is desirable in the transgenic plants of the invention, *e.g.,* GAT and/or other genes conferring herbicide resistance or tolerance, genes which influence other useful characteristics, *e.g.,* heterosis. Similarly, enhancer elements, *e.g.,* derived from the 5' regulatory sequences or intron of a heterologous gene, can also be used to improve expression of a heterologous structural gene, such as a GAT polynucleotide.

[0366] In general, the particular promoter used in the expression cassette in plants depends on the intended application. Any of a number of promoters which direct transcription in plant cells can be suitable. The promoter can be either constitutive or inducible. In addition to the promoters noted above, promoters of bacterial origin which operate in plants include the octopine synthase promoter, the nopaline synthase promoter and other promoters derived from Ti plasmids. *See,* Herrera-Estrella et al. (1983) Nature 303:209. Viral promoters include the 35S and 19S RNA promoters of CaMV. *See,* Odell et al. (1985) Nature 313:810. Other plant promoters include the ribulose-1,3-bisphosphate carboxylase small subunit promoter and the phaseolin promoter. The promoter sequence from the E8 gene *(see,* Deikman and Fischer (1988) EMBO J 7:3315*)* and other genes are also favorably used. Promoters specific for monocotyledonous species are also considered (McElroy and Brettell (1994) "Foreign gene expression in transgenic cereals" Trends Biotech. 12: 62-68.) Alternatively, novel promoters with useful characteristics can be identified from any viral, bacterial, or plant source by methods, including sequence analysis, enhancer or promoter trapping, and the like, known in the art.

[0367] In preparing expression vectors of the invention, sequences other than the promoter and the GAT encoding gene are also favorably used. If proper polypeptide expression is desired, a polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA. Signal/localization peptides, which, *e.g.,* facilitate translocation of the expressed polypeptide to internal organelles (*e.g.*, chloroplasts) or extracellular secretion, can also be employed.

[0368] The vector comprising the GAT polynucleotide also can include a marker gene which confers a selectable

COR-AAD1-00001575

phenotype on plant cells. For example, the marker may encode biocide tolerance, particularly antibiotic tolerance, such as tolerance to kanamycin, G418, bleomycin, hygromycin, or herbicide tolerance, such as tolerance to chlorosulfuron, or phophinothricin. Reporter genes, which are used to monitor gene expression and protein localization via visualizable reaction products (*e.g.*, beta-glucuronidase, beta-galactosidase, and chloramphenicol acetyltransferase) or by direct visualization of the gene product itself (*e.g.*, green fluorescent protein, GFP; Sheen et al. (1995) The Plant Journal 8: 777) can be used for, *e.g.,* monitoring transient gene expression in plant cells. Transient expression systems can be employed in plant cells, for example, in screening plant cell cultures for herbicide tolerance activities.

PLANT TRANSFORMATION

Protoplasts

[0369] Numerous protocols for establishment of transformable protoplasts from a variety of plant types and subsequent transformation of the cultured protoplasts are available in the art and are incorporated herein by reference. For examples, *see,* Hashimoto et al. (1990) Plant Physiol. 93: 857; Fowke and Constabel, eds.(1994) Plant Protoplasts; Saunders et al. (1993) Applications of Plant In vitro Technology Symposium, UPM 16-18; and Lyznik et al. (1991) BioTechniques 10:295, each of which is incorporated herein by reference.

Chloroplasts

[0370] Chloroplasts are a site of action of some herbicide tolerance activities, and, in some instances, the GAT polynucleotide is fused to a chloroplast transit sequence peptide to facilitate translocation of the gene products into the chloroplasts. In these cases, it can be advantageous to transform the GAT polynucleotide into the chloroplasts of the plant host cells. Numerous methods are available in the art to accomplish chloroplast transformation and expression (e.g., Daniell et al. (1998) Nature Biotech. 16:346; O'Neill et al. (1993) The Plant Journal 3:729; and Maliga (1993) TIBTECH 11:1). The expression construct comprises a transcriptional regulatory sequence functional in plants operably linked to a polynucleotide encoding the GAT polypeptide. Expression cassettes that are designed to function in chloroplasts (such as an expression cassette including a GAT polynucleotide) include the sequences necessary to ensure expression in chloroplasts. Typically, the coding sequence is flanked by two regions of homology to the chloroplastid genome to effect a homologous recombination with the chloroplast genome; often a selectable marker gene is also present within the flanking plastid DNA sequences to facilitate selection of genetically stable transformed chloroplasts in the resultant transplastonic plant cells *(see, e.g.,* Maliga (1993) and Daniell (1998) supra, and references cited therein).

General transformation methods

[0371] DNA constructs of the invention can be introduced into the genome of the desired plant host by a variety of conventional techniques. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. *See, e.g.,* Payne, Gamborg, Croy, Jones, *etc.* all *supra,* as well as, *e.g.,* Weising et al. (1988) Ann. Rev. Genet. 22:421 and U.S. Patent Nos. 5,889,191, 5,889,190, 5,866,785, 5,589,367 and 5,316,931, herein incorporated by reference.

[0372] A variety of other transformation protocols are contemplated in the present invention. Transformation protocols as well as protocols for introducing nucleotide sequences into plants may vary depending on the type of plant or plant cell, i.e., monocot or dicot, targeted for transformation. Suitable methods of introducing nucleotide sequences into plant cells and subsequent insertion into the plant genome include microinjection (Crossway et al. (1986) Biotechniques 4: 320-334), electroporation (Riggs et al. (1986) Proc. Natl. Acad. Sci. USA 83:5602-5606), *Agrobacterium*-mediated transformation (U.S. Patent Nos. 5,563,055 and 5,981,840), direct gene transfer (Paszkowski et al. (1984) EMBO J. 3: 2717-2722), and ballistic particle acceleration (see, for example, U.S. Patent Nos. 4,945,050; U.S. Patent No. 5,879,918; 5,886,244; 5,932,782; Tomes et al. (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment, in Plant Cell, Tissue, and Organ Culture: Fundamental Methods, Eds., Gamborg and Phillips (Springer-Verlag, Berlin); McCabe et al. (1988) Biotechnology 6:923-926); and *Lec1* transformation (WO 00/28058). See also, Weissinger et al. (1988) Ann. Rev. Genet. 22:421-477; Sanford et al. (1987) Particulate Science and Technology 5:27-37 (onion); Christou et al. (1988) Plant Physiol. 87:671-674 (soybean); McCabe et al. (1988) Bio/Technology 6:923-926 (soybean); Finer and McMullen (1991) In vitro Cell Dev. Biol. 27P:175-182 (soybean); Singh et al. (1998) Theor. Appl. Genet. 96:319-324 (soybean); Datta et al. (1990) Biotechnology 8:736-740 (rice); Klein et al. (1988) Proc. Natl. Acad. Sci. USA 85:4305-4309 (maize); Klein et al. (1988) Biotechnology 6:559-563 (maize); U.S. Patent Nos. 5,240,855; 5,322,783 and 5,324,646; Klein et al. (1988) Plant Physiol. 91:440-444 (maize); Fromm et al. (1990) Biotechnology 8:833-839 (maize); Hooykaas-Van Slogteren et al. (1984) Nature (London) 311:763-764; U.S. Patent No. 5,736,369 (cereals); Bytebier et al. (1987) Proc. Natl. Acad. Sci. USA 84:5345-5349 (Liliaceae); De Wet et al. (1985) in The Experimental Manipulation of Ovule

COR-AAD1-00001576

Tissues, Eds., Chapman et al. (Longman, New York), pp. 197-209 (pollen); Kaeppler et al. (1990) Plant Cell Reports 9: 415-418 and Kaeppler et al. (1992) Theor. Appl. Genet. 84:560-566 (whisker-mediated transformation); D'Halluin et al. (1992) Plant Cell 4:1495-1505 (electroporation); Li et al. (1993) Plant Cell Reports 12:250-255 and Christou and Ford (1995) Annals of Botany 75:407-413 (rice); Osjoda et al. (1996) Nature Biotechnology 14:745-750 (maize via *Agrobacterium tumefaciens);* all of which are herein incorporated by reference.

[0373] For example, DNAs can be introduced directly into the genomic DNA of a plant cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the DNA constructs can be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the Agrobacterium host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the plant cell is infected by the bacteria.

[0374] Microinjection techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski et al (1984) EMBO J 3:2717. Electroporation techniques are described in Fromm et al. (1985) Proc Nat'l Acad Sci USA 82:5824. Ballistic transformation techniques are described in Klein et al. (1987) Nature 327:70; and Weeks et al. Plant Physiol 102:1077.

[0375] In some embodiments, Agrobacterium mediated transformation techniques are used to transfer the GAT sequences of the invention to transgenic plants. Agrobacterium-mediated transformation is widely used for the transformation of dicots, however, certain monocots can also be transformed by Agrobacterium. For example, Agrobacterium transformation of rice is described by Hiei et al. (1994) Plant J. 6:271; US Patent No. 5,187,073; US Patent No. 5,591,616; Li et al. (1991) Science in China 34:54; and Raineri et al. (1990) BiolTechnology 8: 33. Transformed maize, barley, triticale and asparagus by Agrobacterium mediated transformation have also been described (Xu et al. (1990) Chinese J Bot 2:81).

[0376] Agrobacterium mediated transformation techniques take advantage of the ability of the tumor-inducing (Ti) plasmid of *A. tumefaciens* to integrate into a plant cell genome, to co-transfer a nucleic acid of interest into a plant cell. Typically, an expression vector is produced wherein the nucleic acid of interest, such as a GAT polynucleotide of the invention, is ligated into an autonomously replicating plasmid which also contains T-DNA sequences. T-DNA sequences typically flank the expression cassette nucleic acid of interest and comprise the integration sequences of the plasmid. In addition to the expression cassette, T-DNA also typically includes a marker sequence, *e.g.,* antibiotic resistance genes. The plasmid with the T-DNA and the expression cassette are then transfected into Agrobacterium cells. Typically, for effective transformation of plant cells, the *A. tumefaciens* bacterium also possesses the necessary *vir* regions on a plasmid, or integrated into its chromosome. For a discussion of Agrobacterium mediated transformation, *see,* Firoozabady and Kuehnle, (1995) in Plant Cell Tissue and Organ Culture Fundamental Methods, eds. Gamborg and Phillips.

[0377] In certain embodiments the polynucleotides of the present invention can be stacked with any combination of polynucleotide sequences of interest in order to create plants with a desired phenotype. For example, the polynucleotides of the present invention may be stacked with any other polynucleotides encoding polypeptides having pesticidal and/or insecticidal activity, such as *Bacillus thuringiensis* toxic proteins (described in U.S. Patent Nos. 5,366,892; 5,747,450; 5,737,514; 5,723,756; 5,593,881; and Geiser et al. (1986) Gene 48:109), lectins (Van Damme et al. (1994) Plant Mol. Biol. 24:825, pentin (described in U.S. Patent No. 5,981,722), and the like. The combinations generated can also include multiple copies of any one of the polynucleotides of interest. The polynucleotides of the present invention can also be stacked with any other gene or combination of genes to produce plants with a variety of desired trait combinations including, but not limited to, traits desirable for animal feed such as high oil genes (e.g., U.S. Patent No. 6,232,529); balanced amino acids (e.g., hordothionins (U.S. Patent Nos. 5,990,389; 5,885,801; 5,885,802; and 5,703,409); barley high lysine (Williamson et al. (1987) Eur. J. Biochem. 165:99-106; and WO 98/20122) and high methionine proteins (Pedersen et al. (1986) J. Biol. Chem. 261:6279; Kirihara et al. (1988) Gene 71:359; and Musumura et al. (1989) Plant Mol. Biol. 12:123)); increased digestibility (e.g., modified storage proteins (U.S. Application Serial No. 10/053,410, filed November 7, 2001); and thioredoxins (U.S. Application Serial No. 10/005,429, filed December 3, 2001)); the disclosures of which are herein incorporated by reference.

[0378] The polynucleotides of the present invention can also be stacked with traits desirable for disease or herbicide resistance (*e.g.,* fumonisin detoxification genes (U.S. Patent No. 5,792,931); avirulence and disease resistance genes (Jones et al. (1994) Science 266:789; Martin et al. (1993) Science 262:1432; Mindrinos et al. (1994) Cell 78:1089); acetolactate synthase (ALS) mutants that lead to herbicide resistance such as the S4 and/or Hra mutations; inhibitors of glutamine synthase such as phosphinothricin or basta (e.g., bar gene); and glyphosate resistance (EPSPS gene)); and traits desirable for processing or process products such as high oil (e.g., U.S. Patent No. 6,232,529); modified oils (e.g., fatty acid desaturase genes (U.S. Patent No. 5,952,544; WO 94/11516)); modified starches (e.g., ADPG pyrophosphorylases (AGPase), starch synthases (SS), starch branching enzymes (SBE), and starch debranching enzymes (SDBE)); and polymers or bioplastics (e.g., U.S. Patent No. 5.602,321; beta-ketothiolase, polyhydroxybutyrate synthase, and acetoacetyl-CoA reductase (Schubert et al. (1988) J. Bacteriol. 170:5837-5847) facilitate expression ofpolyhydroxyalkanoates (PHAs)); the disclosures of which are herein incorporated by reference. One could also combine the poly-

COR-AAD1-00001577

nucleotides of the present invention with polynucleotides providing agronomic traits such as male sterility (e.g., see U.S. Patent No. 5.583,210), stalk strength, flowering time, or transformation technology traits such as cell cycle regulation or gene targeting (e.g., WO 99/61619, WO 00/17364, and WO 99/25821); the disclosures of which are herein incorporated by reference.

[0379] These stacked combinations can be created by any method including, but not limited to, cross-breeding plants by any conventional or TopCross methodology, or genetic transformation. If the traits are stacked by genetically transforming the plants, the polynucleotide sequences of interest can be combined at any time and in any order. For example, a transgenic plant comprising one or more desired traits can be used as the target to introduce further traits by subsequent transformation. The traits can be introduced simultaneously in a co-transformation protocol with the polynucleotides of interest provided by any combination of transformation cassettes. For example, if two sequences will be introduced, the two sequences can be contained in separate transformation cassettes (trans) or contained on the same transformation cassette (cis). Expression of the sequences can be driven by the same promoter or by different promoters. In certain cases, it may be desirable to introduce a transformation cassette that will suppress the expression of the polynucleotide of interest. This may be combined with any combination of other suppression cassettes or overexpression cassettes to generate the desired combination of traits in the plant. It is further recognized that polynucleotide sequences can be stacked at a desired genomic location using a site-specific recombination system. See, for example, WO99/25821, WO99/25854, WO99/25840, WO99/25855, and WO99/25853, all of which are herein incorporated by reference.

Regeneration of Transgenic Plants

[0380] Transformed plant cells which are derived by plant transformation techniques, including those discussed above, can be cultured to regenerate a whole plant which possesses the transformed genotype (*i.e.*, a GAT polynucleotide), and thus the desired phenotype, such as acquired resistance (*i.e.*, tolerance) to glyphosate or a glyphosate analog. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with the desired nucleotide sequences. For transformation and regeneration of maize see, Gordon-Kamm et al., The Plant Cell, 2:603-618 (1990). Alternatively, selection for glyphosate resistance conferred by the GAT polynucleotide of the invention can be performed. Plant regeneration from cultured protoplasts is described in Evans et al. (1983) Protoplasts Isolation and Culture, Handbook of Plant Cell Culture, pp 124-176, Macmillan Publishing Company, New York; and Binding (1985) Regeneration of Plants, Plant Protoplasts pp 21-73, CRC Press, Boca Raton. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee et al. (1987) Ann Rev of Plant Phys 38:467. *See also, e.g.,* Payne and Gamborg.

[0381] Transformed plant cells, calli or explant can be cultured on regeneration medium in the dark for several weeks, generally about 1 to 3 weeks to allow the somatic embryos to mature. Preferred regeneration media include media containing MS salts. The plant cells, calli or explant are then typically cultured on rooting medium in a light/dark cycle until shoots and roots develop. Methods for plant regeneration are known in the art and preferred methods are provided by Kamo et al., (Bot. Gaz. 146(3):324-334, 1985); West et al., (The Plant Cell 5:1361-1369, 1993); and Duncan et al. (Planta 165:322-332, 1985).

[0382] Small plantlets can then be transferred to tubes containing rooting medium and allowed to grow and develop more roots for approximately another week. The plants can then be transplanted to soil mixture in pots in the greenhouse.

[0383] The regeneration of plants containing the foreign gene introduced by *Agrobacterium* can be achieved as described by Horsch et al., Science, 227:1229-1231 (1985) and Fraley et al., Proc. Natl. Acad. Sci. U.S.A., 80:4803 (1983). This procedure typically produces shoots within two to four weeks and these transformant shoots are then transferred to an appropriate root-inducing medium containing the selective agent and an antibiotic to prevent bacterial growth. Transgenic plants of the present invention may be fertile or sterile.

[0384] Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee et al., Ann. Rev. of Plant Phys. 38:467-486 (1987). The regeneration of plants from either single plant protoplasts or various explants is well known in the art. See, for example, Methods for Plant Molecular Biology, A. Weissbach and H. Weissbach, eds., Academic Press, Inc., San Diego, Calif. (1988). For maize cell culture and regeneration see generally, The Maize Handbook, Freeling and Walbot, eds., Springer, New York (1994); Corn and Corn Improvement, 3rd Ed., Sprague and Dudley eds., American Society of Agronomy, Madison, Wisconsin (1988).

[0385] After transformation with Agrobacterium, the explants typically are transferred to selection medium. One of skill will realize that the selection medium depends on the selectable marker that was co-transfected into the explants. After a suitable length of time, transformants will begin to form shoots. After the shoots are about 1-2 cm in length, the shoots should be transferred to a suitable root and shoot medium. Selection pressure should be maintained in the root and shoot medium.

[0386] Typically, the transformants will develop roots in about 1-2 weeks and form plantlets. After the plantlets are

COR-AAD1-00001578

about 3-5 cm in height, they are placed in sterile soil in fiber pots. Those of skill in the art will realize that different acclimation procedures are used to obtain transformed plants of different species. For example, after developing a root and shoot, cuttings, as well as somatic embryos of transformed plants, are transferred to medium for establishment of plantlets. For a description of selection and regeneration of transformed plants, *see, e.g.,* Dodds and Roberts (1995) Experiments in Plant Tissue Culture, 3rd Ed., Cambridge University Press.

[0387] There are also methods for Agrobacterium transformation of Arabidopsis using vacuum infiltration (Bechtold N., Ellis J. and Pelletier G" 1993, In planta Agrobacterium mediated gene transfer by infiltration of adult Arabidopsis thaliana plants. CR Acad Sci Paris Life Sci 316:1194-1199) and simple dipping of flowering plants (Desfeux, C., Clough S.J., and Bent A.F., 2000, Female reproductive tissues are the primary target of Agrobacterium-mediated transformation by the Arabidopsis floral-dip method. Plant Physiol. 123:895-904). Using these methods, transgenic seed are produced without the need for tissue culture.

[0388] There are plant varieties for which effective Agrobacterium-mediated transformation protocols have yet to be developed. For example, successful tissue transformation coupled with regeneration of the transformed tissue to produce a transgenic plant has not been reported for some of the most commercially relevant cotton cultivars. Nevertheless, an approach that can be used with these plants involves stably introducing the polynucleotide into a related plant variety via Agrobacterium-mediated transformation, confirming operability, and then transferring the transgene to the desired commercial strain using standard sexual crossing or back-crossing techniques. For example, in the case of cotton, Agrobacterium can be used to transform a Coker line of *Gossypium hirustum (e.g.,* Coker lines 310, 312, 5110 Deltapine 61 or Stoneville 213), and then the transgene can be introduced into another more commercially relevant *G. hirustum* cultivar by back-crossing.

[0389] The transgenic plants of this invention can be characterized either genotypically or phenotypically to determine the presence of the GAT polynucleotide of the invention. Genotypic analysis can be performed by any of a number of well-known techniques, including PCR amplification of genomic DNA and hybridization of genomic DNA with specific labeled probes. Phenotypic analysis includes, *e.g.,* survival of plants or plant tissues exposed to a selected herbicide such as glyphosate.

[0390] One of skill will recognize that after the expression cassette containing the GAT gene is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

[0391] In vegetatively propagated crops, mature transgenic plants can be propagated by the taking of cuttings or by tissue culture techniques to produce multiple identical plants. Selection of desirable transgenics is made and new varieties are obtained and propagated vegetatively for commercial use. In seed propagated crops, mature transgenic plants can be self crossed to produce a homozygous inbred plant. The inbred plant produces seed containing the newly introduced heterologous nucleic acid. These seeds can be grown to produce plants that would produce the selected phenotype.

[0392] Parts obtained from the regenerated plant, such as flowers, seeds, leaves, branches, fruit, and the like are included in the invention, provided that these parts comprise cells comprising the isolated GAT nucleic acid. Progeny and variants, and mutants of the regenerated plants are also included within the scope of the invention, provided that these parts comprise the introduced nucleic acid sequences.

[0393] Transgenic plants expressing a selectable marker can be screened for transmission of the GAT nucleic acid, for example, by standard immunoblot and DNA detection techniques. Transgenic lines are also typically evaluated on levels of expression of the heterologous nucleic acid. Expression at the RNA level can be determined initially to identify and quantitate expression-positive plants. Standard techniques for RNA analysis can be employed and include PCR amplification assays using oligonucleotide primers designed to amplify only the heterologous RNA templates and solution hybridization assays using heterologous nucleic acid-specific probes. The RNA-positive plants can then be analyzed for protein expression by Western immunoblot analysis using the specifically reactive antibodies of the present invention. In addition, *in situ* hybridization and immunocytochemistry according to standard protocols can be done using heterologous nucleic acid specific polynucleotide probes and antibodies, respectively, to localize sites of expression within transgenic tissue. Generally, a number of transgenic lines are usually screened for the incorporated nucleic acid to identify and select plants with the most appropriate expression profiles.

[0394] A preferred embodiment is a transgenic plant that is homozygous for the added heterologous nucleic acid; *i.e.,* a transgenic plant that contains two added nucleic acid sequences, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually mating (selfing) a heterozygous transgenic plant that contains a single added heterologous nucleic acid, germinating some of the seed produced and analyzing the resulting plants produced for altered cell division relative to a control plant (*i.e.*, native, non-transgenic). Back-crossing to a parental plant and out-crossing with a non-transgenic plant are also contemplated.

[0395] Essentially any plant can be transformed with the GAT polynucleotides of the invention. Suitable plants for the transformation and expression of the novel GAT polynucleotides of this invention include agronomically and horticulturally important species. Such species include, but are not restricted to members of the families: Graminae (including corn, rye, triticale, barley, millet, rice, wheat, oats, *etc.);* Leguminosae (including pea, beans, lentil, peanut, yam bean, cowpeas,

COR-AAD1-00001579

velvet beans, soybean, clover, alfalfa, lupine, *vetch,* lotus, sweet clover, wisteria, and sweetpea); Compositae (the largest family of vascular plants, including at least 1,000 genera, including important commercial crops such as sunflower); and Rosaciae (including raspberry, apricot, almond, peach, rose, *etc.);* as well as nut plants (including, walnut, pecan, hazelnut, *etc.);* and forest trees (including *Pinus, Quercus, Pseutotsuga, Sequoia, Populus,etc.)*

[0396] Additional targets for modification by the GAT polynucleotides of the invention, as well as those specified above, include plants from the genera: *Agrostis, Allium, Antirhinum, Apium, Arachis, Asparagus, Atropa, Avena (e.g.,* oats), *Bambusa, Brassica, Bromus, Browaalia, Camellia, Cannabis, Capsicum, Cicer, Chenopodium, Chichorium, Citrus, Coffea, Coix, Cucumis, Curcubita, Cynodon, Dactylis, Datura, Daucus, Digitalis, Dioscorea, Elaeis, Eleusine, Festuca, Fragaria, Geranium, Gossypium, Glycine, Helianthus, Heterocallis, Hevea, Hordeum (e.g.,* barley), *Hyoscyamus, Ipomoea, Lactuca, Lens, Lilium, Linum, Lolium, Lotus, Lycopersicon, Majorana, Malus, Mangifena, Manihot, Medicago, Nemesia, Nicotiana, Onoblychis, Oryza (e.g.,* rice), *Panicum, Pelargonium, Pennisetum (e.g.,* millet), *Petunia, Pisum, Phaseolus, Phleum, Poa, Prunus, Ranunculus, Raphanus, Ribes, Ricinus, Rubus, Saccharum, Salpiglossis, Secale (e.g.,* rye), *Senecio, Setaria, Sinapis, Solanum, Sorghum, Stenotaphrum, Theobroma, Trifolium, Trigonella, Triticum (e.g.,* wheat), *Vicia, Vigna, Vitis, Zea* (*e.g.,* corn), and the *Olyreae,* the *Pharoideae* and many others. As noted, plants in the family *Graminae* are particularly desirable target plants for the methods of the invention.

[0397] Common crop plants which are targets of the present invention include corn, rice, triticale, rye, cotton, soybean, sorghum, wheat, oats, barley, millet, sunflower, canola, peas, beans, lentils, peanuts, yam beans, cowpeas, velvet beans, clover, alfalfa, lupine, *vetch*, lotus, sweet clover, wisteria, sweetpea and nut plants (*e.g.*, walnut, pecan, *etc).*

[0398] In one aspect, the invention provides a method for producing a crop by growing a crop plant that is glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate N-acetyltransferase, under conditions such that the crop plant produces a crop, and harvesting the crop. Preferably, glyphosate is applied to the plant, or in the vicinity of the plant, at a concentration effective to control weeds without preventing the transgenic crop plant from growing and producing the crop. The application of glyphosate can be before planting, or at any time after planting up to and including the time of harvest. Glyphosate can be applied once or multiple times. The timing of glyphosate application, amount applied, mode of application, and other parameters will vary based upon the specific nature of the crop plant and the growing environment, and can be readily determined by one of skill in the art. The invention further provides a crop produced by this method.

[0399] The invention provides for the propagation of a plant containing a GAT polynucleotide transgene. The plant can be, for example, a monocot or a dicot. In one aspect, propagation entails crossing a plant containing a GAT polynucleotide transgene with a second plant, such that at least some progeny of the cross display glyphosate tolerance.

[0400] In one aspect, the invention provides a method for selectively controlling weeds in a field where a crop is being grown. The method involves planting crop seeds or plants that are glyphosate-tolerant as a result of being transformed with a gene encoding a GAT, *e.g.,* a GAT polynucleotide, and applying to the crop and any weeds a sufficient amount of glyphosate to control the weeds without a significant adverse impact on the crop. It is important to note that it is not necessary for the crop to be totally insensitive to the herbicide, so long as the benefit derived from the inhibition of weeds outweighs any negative impact of the glyphosate or glyphosate analog on the crop or crop plant.

[0401] In another aspect, the invention provides for use of a GAT polynucleotide as a selectable marker gene. In this embodiment of the invention, the presence of the GAT polynucleotide in a cell or organism confers upon the cell or organism the detectable phenotypic trait of glyphosate resistance, thereby allowing one to select for cells or organisms that have been transformed with a gene of interest linked to the GAT polynucleotide. Thus, for example, the GAT polynucleotide can be introduced into a nucleic acid construct, *e.g.,* a vector, thereby allowing for the identification of a host (*e.g.*, a cell or transgenic plant) containing the nucleic acid construct by growing the host in the presence of glyphosate and selecting for the ability to survive and/or grow at a rate that is discernibly greater than a host lacking the nucleic acid construct would survive or grow. A GAT polynucleotide can be used as a selectable marker in a wide variety of hosts that are sensitive to glyphosate, including plants, most bacteria (including *E. coli),* actinomycete, yeasts, algae and fungi. One benefit of using herbicide resistance as a marker in plants, as opposed to conventional antibiotic resistance, is that it obviates the concern of some members of the public that antibiotic resistance might escape into the environment. Some experimental data from experiments demonstrating the use of a GAT polynucleotide as a selectable marker in diverse host systems are described in the Examples section of this specification.

Selection of GAT polynucleotides conferring enhanced glyphosate resistance in Transgenic plants.

[0402] Libraries of GAT encoding nucleic acids diversified according to the methods described herein can be selected for the ability to confer resistance to glyphosate in transgenic plants. Following one or more cycles of diversification and selection, the modified GAT genes can be used as a selection marker to facilitate the production and evaluation of transgenic plants and as a means of conferring herbicide resistance in experimental or agricultural plants. For example, after diversification of any one or more of, *e.g.*, SEQ ID NO:1 - 5 to produce a library of diversified GAT polynucleotides, an initial functional evaluation can be performed by expressing the library of GAT encoding sequences in *E. coli.* The

COR-AAD1-00001580

expressed GAT polypeptides can be purified, or partially purified as described above, and screened for improved kinetics by mass spectrometry. Following one or more preliminary rounds of diversification and selection, the polynucleotides encoding improved GAT polypeptides are cloned into a plant expression vector, operably linked to, *e.g.,* a strong constitutive promoter, such as the CaMV 35S promoter. The expression vectors comprising the modified GAT nucleic acids are transformed, typically by Agrobacterium mediated transformation, into *Arabidopsis thaliana* host plants. For example, *Arabidopsis* hosts are readily transformed by dipping inflorescences into solutions of Agrobacterium and allowing them to grow and set seed. Thousands of seeds are recovered in approximately 6 weeks. The seeds are then collected in bulk from the dipped plants and germinated in soil. In this manner it is possible to generate several thousand independently transformed plants for evaluation, constituting a high throughput (HTP) plant transformation format. Bulk grown seedlings are sprayed with glyphosate and surviving seedlings exhibiting glyphosate resistance survive the selection process, whereas non-transgenic plants and plants incorporating less favorably modified GAT nucleic acids are damaged or killed by the herbicide treatment. Optionally, the GAT encoding nucleic acids conferring improved resistance to glyphosate are recovered, *e.g.,* by PCR amplification using T-DNA primers flanking the library inserts, and used in further diversification procedures or to produce additional transgenic plants of the same or different species. If desired, additional rounds of diversification and selection can be performed using increasing concentrations of glyphosate in each subsequent selection. In this manner, GAT polynucleotides and polypeptides conferring resistance to concentrations of glyphosate useful in field conditions can be obtained.

Herbicide Resistance

[0403] The present invention provides a composition comprising two or more polynucleotides of the invention. Preferably, the GAT polynucleotides encode GAT polypeptides having different kinetic parameters, *i.e.,* a GAT variant having a lower K_m can be combined with one having a higher k_{cat}. In a further embodiment, the different GAT polynucleotides may be coupled to a chloroplast transit sequence or other signal sequence thereby providing GAT polypeptide expression in different cellular compartments, organelles or secretion of one or more of the GAT polypeptides.

[0404] The mechanism of glyphosate resistance of the present invention can be combined with other modes of glyphosate resistance known in the art to produce plants and plant explants with superior glyphosate resistance. For example, glyphosate-tolerant plants can be produced by inserting into the genome of the plant the capacity to produce a higher level of 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) as more fully described in U.S. Patent Nos. 6,248,876 B1; 5,627,061; 5,804,425; 5,633,435; 5,145,783; 4,971,908; 5,312,910; 5,188,642; 4,940,835; 5,866,775; 6,225,114 B1; 6,130,366; 5,310,667; 4,535,060; 4,769,061; 5,633,448; 5,510,471; Re. 36,449; RE 37,287 E; and 5,491,288; and international publications WO 97/04103; WO 00/66746; WO 01/66704; and WO 00/66747, which are incorporated herein by reference in their entireties for all purposes. Glyphosate resistance is also imparted to plants that express a gene that encodes a glyphosate oxido-reductase enzyme as described more fully in U.S. Patent Nos. 5,776,760 and 5,463,175, which are incorporated herein by reference in their entireties for all purposes.

[0405] Further, the mechanism of glyphosate resistance of the present invention may be combined with other modes of herbicide resistance to provide plants and plant explants that are resistant to glyphosate and one or more other herbicides. For example, the hydroxyphenylpyruvatedioxygenases are enzymes that catalyze the reaction in which para-hydroxyphenylpyruvate (HPP) is transformed into homogentisate. Molecules which inhibit this enzyme, and which bind to the enzyme in order to inhibit transformation of the HPP into homogentisate are useful as herbicides. Plants more resistant to certain herbicides are described in U.S. Patent Nos. 6,245,968 B1; 6,268,549; and 6,069,115; and international publication WO 99/23886, which are incorporated herein by reference in their entireties for all purposes.

[0406] Sulfonylurea and imidazolinone herbicides also inhibit growth of higher plants by blocking acetolactate synthase (ALS) or acetohydroxy acid synthase (AHAS). The production of sulfonylurea and imidazolinone tolerant plants is described more fully in U.S. Patent Nos. 5,605,011; 5,013,659; 5,141,870; 5,767,361; 5,731,180; 5,304,732; 4,761,373; 5,331,107; 5,928,937; and 5,378,824; and international publication WO 96/33270, which are incorporated herein by reference in their entireties for all purposes.

[0407] Glutamine synthetase (GS) appears to be an essential enzyme necessary for the development and life of most plant cells. Inhibitors of GS are toxic to plant cells. Glufosinate herbicides have been developed based on the toxic effect due to the inhibition of GS in plants. These herbicides are non-selective. They inhibit growth of all the different species of plants present, causing their total destruction. The development of plants containing an exogenous phosphinothricin acetyltransferase is described in U.S. Patent Nos. 5,969,213; 5,489,520; 5,550,318; 5,874,265; 5,919,675; 5,561,236; 5,648,477; 5,646,024; 6,177,616 B1; and 5,879,903, which are incorporated herein by reference in their entireties for all purposes.

[0408] Protoporphyrinogen oxidase (protox) is necessary for the production of chlorophyll, which is necessary for all plant survival. The protox enzyme serves as the target for a variety of herbicidal compounds. These herbicides also inhibit growth of all the different species of plants present, causing their total destruction. The development of plants containing altered protox activity which are resistant to these herbicides are described in U.S. Patent Nos. 6,288,306

COR-AAD1-00001581

B1; 6,282,837 B1; and 5,767,373; and international publication WO 01/12825, which are incorporated herein by reference in their entireties for all purposes.

[0409] Accordingly, the invention provides methods for selectively controlling weeds in a field containing a crop that involve planting the field with crop seeds or plants which are glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate N-acetyltransferase, and applying to the crop and weeds in the field a sufficient amount of glyphosate to control the weeds without significantly affecting the crop.

[0410] The invention further provides methods for controlling weeds in a field and preventing the emergence of glyphosate-resistant weeds in a field containing a crop which involve planting the field with crop seeds or plants that are glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate-N-acetyltransferase and a gene encoding a polypeptide imparting glyphosate tolerance by another mechanism, such as, a glyphosate-tolerant 5-enolpyruvylshikiniate-3-phosphate synthase and/or a glyphosate-tolerant glyphosate oxido-reductase and applying to the crop and the weeds in the field a sufficient amount of glyphosate to control the weeds without significantly affecting the crop.

[0411] In a further embodiment the invention provides methods for controlling weeds in a field and preventing the emergence of herbicide resistant weeds in a field containing a crop which involve planting the field with crop seeds or plants that are glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate-N-acetyltransferase, a gene encoding a polypeptide imparting glyphosate tolerance by another mechanism, such as, a glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthase and/or a glyphosate-tolerant glyphosate oxido-reductase and a gene encoding a polypeptide imparting tolerance to an additional herbicide, such as, a mutated hydroxyphenylpyruvatedioxygenase, a sulfonamide-tolerant acetolactate synthase, a sulfonamide-tolerant acetohydroxy acid synthase, an imidazolinone-tolerant acetolactate synthase, an imidazolinone-tolerant acetohydroxy acid synthase, a phosphinothricin acetyltransferase and a mutated protoporphyrinogen oxidase and applying to the crop and the weeds in the field a sufficient amount of glyphosate and an additional herbicide, such as, a hydroxyphenylpyruvatedioxygenase inhibitor, sulfonamide, imidazolinone, bialaphos, phosphinothricin, azafenidin, butafenacil, sulfosate, glufosinate, and a protox inhibitor to control the weeds without significantly affecting the crop.

[0412] The invention further provides methods for controlling weeds in a field and preventing the emergence of herbicide resistant weeds in a field containing a crop which involve planting the field with crop seeds or plants that are glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate-N-acetyltransferase and a gene encoding a polypeptide imparting tolerance to an additional herbicide, such as, a mutated hydroxyphenylpyruvatedioxygenase, a sulfonamide-tolerant acetolactate synthase, a sulfonamide-tolerant acetohydroxy acid synthase, an imidazolinone-tolerant acetolactate synthase, an imidazolinone-tolerant acetohydroxy acid synthase, a phosphinothricin acetyltransferase and a mutated protoporphyrinogen oxidase and applying to the crop and the weeds in the field a sufficient amount of glyphosate and an additional herbicide, such as, a hydroxyphenylpyruvatedioxygenase inhibitor, sulfonamide, imidazolinone, bialaphos, phosphinothricin, azafenidin, butafenacil, sulfosate, glufosinate, and a protox inhibitor to control the weeds without significantly affecting the crop.

EXAMPLES

[0413] The following examples are illustrative and not limiting. One of skill will recognize a variety of non-critical parameters that can be altered to achieve essentially similar results.

EXAMPLE 1: ISOLATING NOVEL NATIVE GAT POLYNUCLEOTIDES

[0414] Five native GAT polynucleotides (*i.e.*, GAT polynucleotides that occur naturally in a non-genetically modified organism) were discovered by expression cloning of sequences from *Bacillus* strains exhibiting GAT activity. Their nucleotide sequences were determined and are provided herein as SEQ ID NO:1 - 5. Briefly, a collection of approximately 500 *Bacillus* and *Pseudomonas* strains were screened for native ability to N-acetylate glyphosate. Strains were grown in LB overnight, harvested by centrifugation, permeabilized in dilute toluene, and then washed and resuspended in a reaction mix containing buffer, 5 mM glyphosate, and 200 μM acetyl-CoA. The cells were incubated in the reaction mix for between 1 and 48 hours, at which time an equal volume of methanol was added to the reaction. The cells were then pelleted by centrifugation and the supernatant was filtered before analysis by parent ion mode mass spectrometry. The product of the reaction was positively identified as N-acetylglyphosate by comparing the mass spectrometry profile of the reaction mix to an N-acetylglyphosate standard as shown in Figure 2. Product detection was dependent on inclusion of both substrates (acetyl CoA and glyphosate) and was abolished by heat denaturing the bacterial cells.

[0415] Individual GAT polynucleotides were then cloned from the identified strains by functional screening. Genomic DNA was prepared and partially digested with Sau3A1 enzyme. Fragments of approximately 4 Kb were cloned into an *E. coli* expression vector and transformed into electrocompetent E. *coli*. Individual clones exhibiting GAT activity were identified by mass spectrometry following a reaction as described previously except that the toluene wash was replaced by permeabilization with PMBS. Genomic fragments were sequenced and the putative GAT polypeptide-encoding open

COR-AAD1-00001582

reading frame identified. Identity of the GAT gene was confirmed by expression of the open reading frame in *E. coli* and detection of high levels ofN-acetylglyphosate produced from reaction mixtures.

EXAMPLE 2: CHARACTERIZATION OF A GAT POLYPEPTIDE ISOLATED FROM *B.LICHENIFORMIS* STRAIN B6.

[0416] Genomic DNA from *B. licheniformis* strain B6 was purified, partially digested with Sau3A1 and fragments of 1-10 Kb were cloned into an E. *coli* expression vector. A clone with a 2.5 kb insert conferred the glyphosate-N-acetyltransferase (GAT) activity on the *E. coli* host as determined with mass spectrometry analysis. Sequencing of the insert revealed a single complete open reading frame of 441 base pairs. Subsequent cloning of this open reading frame confirmed that it encoded the GAT enzyme. A plasmid, pMAXY2120, is shown in Figure 4. The gene encoding the GAT enzyme of B6 was transformed into *E. coli* strain XL1 Blue. A 10% innoculum of a saturated culture was added to Luria broth, and the culture was incubated at 37° C for 1 hr. Expression of GAT was induced by the addition of IPTG at a concentration of 1 mM. The culture was incubated a further 4 hrs, following which, cells were harvested by centrifugation and the cell pellet stored at -80° C.

[0417] Lysis of the cells was effected by the addition of 1 ml of the following buffer to 0.2 g of cells: 25 mM HEPES, pH 7.3, 100 mM KCl and 10% methanol (HKM) plus 0.1 mM EDTA, 1 mM DTT, 1 mg/ml chicken egg lysozyme, and a protease inhibitor cocktail obtained from Sigma and used according to the manufacturer's recommendations. After 20 minutes incubation at room temperature (*e.g.*, 22-25° C), lysis was completed with brief sonication. The lysate was centrifuged and the supernatant was desalted by passage through Sephadex G25 equilibrated with HKM. Partial purification was obtained by affinity chromatography on CoA Agarose (Sigma). The column was equilibrated with HKM and the clarified extract was allowed to pass through under hydrostatic pressure. Non-binding proteins were removed by washing the column with HKM, and GAT was eluted with HKM containing 1 mM Coenzyme A. This procedure provided 4-fold purification. At this stage, approximately 65% of the protein staining observed on an SDS polyacrylamide gel loaded with crude lysate was due to GAT, with another 20% due to chloramphenicol acetyltransferase encoded by the vector.

[0418] Purification to homogeneity was obtained by gel filtration of the partially purified protein through Superdex 75 (Pharmacia). The mobile phase was HKM, in which GAT activity eluted at a volume corresponding to a molecular radius of 17 kD. This material was homogeneous as judged by Coomassie staining of a 3 μg sample of GAT subjected to SDS polyacrylamide gel electrophoresis on a 12% acrylamide gel, 1 mm thickness. Purification was achieved with a 6-fold increase in specific activity.

[0419] The apparent K_M for glyphosate was determined on reaction mixtures containing saturating (200 μM) Acetyl CoA, varying concentrations of glyphosate, and 1 μM purified GAT in buffer containing 5 mM morpholine adjusted to pH 7.7 with acetic acid and 20 % ethylene glycol. Initial reaction rates were determined by continuous monitoring of the hydrolysis of the thioester bond of Acetyl CoA at 235 nm (E = 3.4 OD/mM/cm). Hyperbolic saturation kinetics were observed (Figure 5), from which an apparent K_M of 2.9 ± 0.2 (SD) mM was obtained.

[0420] The apparent K_M for Acetyl CoA was determined on reaction mixtures containing 5 mM glyphosate, varying concentrations of Acetyl CoA, and 0.19 μM GAT in buffer containing 5 mM morpholine adjusted to pH 7.7 with acetic acid and 50% methanol. Initial reaction rates were determined using mass spectrometric detection of N-acetyl glyphosate. Five μl were repeatedly injected into the instrument and reaction rates were obtained by plotting reaction time vs area of the integrated peak (Figure 6). Hyperbolic saturation kinetics were observed (Figure 7), from which an apparent K_M of 2 μM was derived. From values for V_{max} obtained at a known concentration of enzyme, a k_{cat} of 6/min was calculated.

EXAMPLE 3: MASS SPECTROMETRY (MS) SCREENING PROCESS

[0421] Sample (5 μl) was drawn from a 96-well microtiter plate at a speed of one sample every 26 seconds and injected into the mass spectrometer (Micromass Quattro LC, triple quadrupole mass spectrometer) without any separation. The sample was carried into the mass spectrometer by a mobile phase of water/methanol (50:50) at a flow rate of 500 Ul/min. Each injected sample was ionized by a negative electrospray ionization process (needle voltage, -3.5 KV; cone voltage, 20 V; source temperature, 120° C; desolvation temperature, 250° C; cone gas flow, 90 L/Hr; and desolvation gas flow, 600 L/Hr). The molecular ions (m/z 210) formed during this process were selected by the first quadrupole for performing collision induced dissociation (CID) in the second quadrupole, where the pressure was set at 5 x 10^{-4} mBar and the collision energy was adjusted to 20 Ev. The third quadrupole was set for only allowing one of the daughter ions (m/z 124) produced from the parent ions (m/z 210) to get into the detector for signal recording. The first and third quadrupoles were set at unit resolution, while the photomultiplier was operated at 650 V. Pure N-acetylglyphosate standards were used for comparison and peak integration was used to estimate concentrations. It was possible to detect less than 200 Nm N-acetylglyphosate by this method.

COR-AAD1-00001583

EXAMPLE 4: DETECTION OF NATIVE OR LOW ACTIVITY GAT ENZYMES

[0422] Native or low activity GAT enzymes typically have a k_{cat} of approximately 1 min^{-1} and a K_M for glyphosate of 1.5-10 Mm. K_M for acetyl CoA was typically less than 25 μM.

[0423] Bacterial cultures were grown in rich medium in deep 96-well plates and 0.5 ml stationary phase cells were harvested by centrifugation, washed with 5 mM morpholine acetate pH 8, and resuspended in 0.1 ml reaction mix containing 200 μM ammonium acetyl CoA, 5 mM ammonium glyphosate, and 5μ g/ml PMBS (Sigma) in 5 mM morpholine acetate, pH 8. The PMBS permeabilizes the cell membrane allowing the substrates and products to move from the cells to the buffer without releasing the entire cellular contents. Reactions were carried out at 25-37°C for 1-48 hours. The reactions were quenched with an equal volume of 100% ethanol and the entire mixture was filtered on a 0.45 μm MAHV Multiscreen filter plate (Millipore). Samples were analyzed using a mass spectrometer as described above and compared to synthetic N-acetylglyphosate standards.

EXAMPLE 5: DETECTION OF HIGH ACTIVITY GAT ENZYMES

[0424] High activity GAT enzymes typically have a k_{cat} up to 400 min^{-1} and a K_M below 0.1 mM glyphosate.

[0425] Genes coding for GAT enzymes were cloned into *E. coli* expression vector pQE80 (Qiagen) and introduced into *E. coli* strain XL1 Blue (Stratagene). Cultures were grown in 150 ul rich medium (LB with 50 ug/ml carbenicllin) in shallow U-bottom 96-well polystyrene plates to late-log phase and diluted 1:9 with fresh medium containing 1 mM IPTG (USB). After 4-8 hours induction, cells were harvested, washed with 5mM morpholine acetate pH 6.8 and resuspended in an equal volume of the same morpholine buffer. Reactions were carried out with up to 10 ul of washed cells. At higher activity levels, the cells were first diluted up to 1:200 and 5 ul was added to 100 ul reaction mix. To measure GAT activity, the same reaction mix as described for low activity was used. However, for detecting highly active GAT enzymes the glyphosate concentration was reduced to 0.15 - 0.5 mM, the pH was reduced to 6.8, and reactions were carried out for 1 hour at 37°C. Reaction workup and MS detection were as described herein.

EXAMPLE 6: PURIFICATION OF GAT ENZYMES

[0426] Enzyme purification was achieved by affinity chromatography of cell lysates on CoA-agarose and gel-filtration on Superdex-75. Quantities of purified GAT enzyme up to 10 mg were obtained as follows: A 100-ml culture of *E. coli* carrying a GAT polynucleotide on a pQE80 vector and grown overnight in LB containing 50 ug/ml carbenicillin was used to inoculate 1L of LB plus 50 ug/ml carbenicillin. After 1 hr, IPTG was added to 1 mM, and the culture was grown a further 6 hr. Cells were harvested by centrifugation. Lysis was effected by suspending the cells in 25 mM HEPES (pH 7.2), 100 mM KCl, 10% methanol (HKM), 0.1 mM EDTA, 1 mM DTT, protease inhibitor cocktail supplied by Sigma-Aldrich and 1 mg/ml of chicken egg lysozyme. After 30 minutes at room temperature, the cells were briefly sonicated. Particulate material was removed by centrifugation, and the lysate was passed through a bed of coenzyme A-Agarose. The column was washed with several bed volumes of HKM and GAT was eluted in 1.5 bed volumes of HKM containing 1 mM acetyl CoA. GAT in the eluate was concentrated by its retention above a Centricon YM 50 ultrafiltration membrane. Further purification was obtained by passing the protein through a Superdex 75 column through a series of 0.6-ml injections. The peak of GAT activity eluted at a volume corresponding to a molecular weight of 17 kD. This method resulted in purification of GAT enzyme to homogeneity with >85% recovery. A similar procedure was used to obtain 0.1 to 0.4 mg quantities of up to 96 shuffled variants at a time. The volume of induced culture was reduced to 1 to 10 ml, coenzyme A-Agarose affinity chromatography was performed in 0.15-ml columns packed in an MAHV filter plate (Millipore) and Superdex 75 chromatography was omitted.

EXAMPLE 7: STANDARD PROTOCOL FOR DETERMINATION OF K_{CAT} AND K_M

[0427] k_{cat} and K_M for glyphosate of purified protein were determined using a continuous spectrophotometric assay, in which hydrolysis of the sulfoester bond of Acetyl CoA was monitored at 235 mn. Reactions were performed at ambient temperature (about 23°C) in the wells of a 96-well assay plate, with the following components present in a final volume of 0.3 ml: 20 mM HEPES, pH 6.8, 10% ethylene glycol, 0.2 mM acetyl CoA, and various concentrations of ammonium glyphosate. In comparing the kinetics of two GAT enzymes, both enzymes were assayed under the same conditions, e.g., both at 23°C. k_{cat} was calculated from V_{max} and the enzyme concentration, determined by Bradford assay. K_M was calculated from the initial reaction rates obtained from concentrations of glyphosate ranging from 0.125 to 10 mM, using the Lineweaver-Burke transformation of the Michaelis-Menten equation. k_{cat}/K_M was determined by dividing the value determined for k_{cat} by the value determined for K_M.

[0428] Using this methodology, kinetic parameters for a number of GAT polypeptides exemplified herein were determined. For example, the k_{cat}, K_M and k_{cat}/K_M for the GAT polypeptide corresponding to SEQ ID NO:445 have been

COR-AAD1-00001584

determined to be 322 min^{-1}, 0.5 mM and 660 $mM^{-1}min^{-1}$, respectively, using the assay conditions described above. The k_{cat}, K_M and k_{cat}/K_M for the GAT polypeptide corresponding to SEQ ID NO:457 have been determined to be 118 min^{-1}, 0.1 mM and 1184 $mM^{-1}min^{-1}$, respectively, using the assay conditions described above. The k_{cat}, K_M and k_{cat}/K_M for the GAT polypeptide corresponding to SEQ ID NO:300 have been determined to be 296 min^{-1}, 0.65 mM and 456 $mM^{-1}min^{-1}$, respectively, using the assay conditions described above. One of skill in the art can use these numbers to confirm that a GAT activity assay is generating kinetic parameters for a GAT suitable for comparison with the values given herein. For example, the conditions used to compare the activity of GATs should yield the same kinetic constants for SEQ ID NO: 300, 445, and 457 (within normal experimental variance) as those reported herein, when the conditions are used to compare a test GAT with the GAT polypeptides exemplified herein.

[0429] K_M for Acetyl CoA was measured using the mass spectrometry method with repeated sampling during the reaction. AcetylCoA and glyphosate (ammonium salts) were placed as 50-fold-concentrated stock solutions into a well of a mass spectrometry sample plate. Reactions were initiated with the addition of enzyme appropriately diluted in a volatile buffer such as morpholine acetate or ammonium carbonate, pH 6.8 or 7.7. The sample was repeatedly injected into the instrument and initial rates were calculated from plots of retention time and peak area. K_M was calculated as for glyphosate.

EXAMPLE 8: SELECTION OF TRANSFORMED *E. COLI*

[0430] An evolved GAT gene (a chimera with a native *B. licheniformis* ribosome binding site (AACTGAAGGAG-GAATCTC; SEQ ID NO:515) attached directly to the 5' end of the GAT coding sequence) was cloned into the expression vector pQE80 (Qiagen) between the EcoRI and HindIII sites, resulting in the plasmid pMAXY2190 (Figure 11). This eliminated the His tag domain from the plasmid and retained the B-lactamase gene conferring resistance to the antibiotics ampicillin and carbenicillin. pMAXY2190 was electroporated (BioRad Gene Pulser) into XL1 Blue (Stratagene) *E. coli* cells. The cells were suspended in SOC rich medium and allowed to recover for one hour. The cells were then gently pelleted, washed one time with M9 minimal media lacking aromatic amino acids (12.8 g/L Na2HPO4.7 H2O, 3.0 g/L KH2PO4, 0.5 g/L NaCl, 1.0 g/L NH4Cl, 0.4% glucose, 2 mM MgSO4, 0.1 mM CaCl2, 10 mg/L thiamine, 10 mg/L proline, 30 mg/L carbenicillin), and resuspended in 20 ml of the same M9 medium. After overnight growth at 37°C at 250 rpm, equal volumes of cells were plated on either M9 medium or M9 plus 1 mM glyphosate medium. pQE80 vector with no GAT gene was similarly introduced into *E. coli* cells and plated for single colonies for comparison. Table 3 presents a summary of the results, demonstrating that GAT activity allows selection and growth of transformed E. *coli* cells with less than 1% background. Note that no IPTG induction was necessary for sufficient GAT activity to allow growth of transformed cells. Transformation was verified by re-isolation of pMAXY2190 from the *E. coli* cells grown in the presence of glyphosate.

Table 3. Glyphosate selection of pMAXY2190 in *E. coli*

Number of colonies		
Plasmid	M9 - glyphosate	M9 + 1 mM glyphosate
pMAXY2190	568	512
pQE80	324	3

EXAMPLE 9: SELECTION OF TRANSFORMED PLANT CELLS

[0431] Agrobacterium-mediated transformation of plant cells occurs at low efficiencies. To allow propagation of transformed cells while inhibiting proliferation of non-transformed cells, a selectable marker is needed. Antibiotic markers for kanamycin and hygromycin and the herbicide modifying gene *bar*, which detoxifies the herbicidal compound phosphinothricin, are examples of selectable markers used in plants (Methods in Molecular Biology, 1995, 49:9-18). Here we demonstrate that GAT activity serves as an efficient selectable marker for plant transformation. An evolved GAT gene (0_5B8), SEQ ID NO:190, was cloned between a plant promoter (enhanced strawberry vein banded virus) and a ubiquinone terminator and introduced into the T-DNA region of the binary vector pMAXY3793 suitable for transformation of plant cells via *Agrobacterium tumefaciens* EHA105 as shown in Figure 12. A screenable GUS marker was present in the T-DNA to allow confirmation of transformation. Transgenic tobacco shoots were generated using glyphosate as the only selecting agent.

[0432] Axillary buds *of Nicotiana tabacum* L. Xanthi were subcultured on half-strength MS medium with sucrose (1.5 %) and Gelrite (0.3 %) under 16-h light (35-42 μEinsteins m^{-2} s^{-1}, cool white fluorescent lamps) at 24 °C every 2-3 weeks. Young leaves were excised from plants after 2-3 weeks subculture and were cut into 3 x 3 mm segments. *A.*

COR-AAD1-00001585

tumefaciens EHA105 was inoculated into LB medium and grown overnight to a density of A600=1.0. Cells were pelleted at 4,000 rpm for 5 minutes and resuspended in 3 volumes of liquid co-cultivation medium composed of Murashige and Skoog (MS) medium (pH 5.2) with 2 mg/L N6-benzyladenine (BA), 1% glucose and 400 uM acetysyringone. The leaf pieces were then fully submerged in 20 ml *of A. tumefaciens* in 100 x 25 mm Petri dishes for 30 min, blotted with autoclaved filter paper, then placed on solid co-cultivation medium (0.3% Gelrite) and incubated as described above. After 3 days of co-cultivation, 20-30 segments were transferred to basal shoot induction (BSI) medium composed of MS solid medium (pH 5.7) with 2 mg/L BA, 3% sucrose, 0.3% Gelrite, 0-200 uM glyphosate, and 400 ug/ml Timentin.

[0433] After 3 weeks, shoots were clearly evident on the explants placed on media with no glyphosate regardless of the presence or absence of the GAT gene. T-DNA transfer from both constructs was confirmed by GUS histochemical staining of leaves from regenerated shoots. Glyphosate concentrations greater than 20 uM completely inhibited any shoot formation from the explants lacking a GAT gene. Explants infected with *A. tumefaciens* with the GAT construct regenerated shoots at glyphosate concentrations up to 200 uM (the highest level tested). Transformation was confirmed by GUS histochemical staining and by PCR fragment amplification of the GAT gene using primers annealing to the promoter and 3' regions. The results are summarized in Table 4.

Table 4. Tobacco shoot regeneration with glyphosate selection.

	Glyphosate concentration % Shoot Regeneration				
Transferred genes	0 uM	20 uM	40 uM	80 uM	200 uM
GUS	100	0	0	0	0
gat and GUS	100	60	30	5	3

EXAMPLE 10: GLYPHOSATE SELECTION OF TRANSFORMED YEAST CELLS

[0434] Selection markers for yeast transformation are usually auxotrophic genes that allow growth of transformed cells on a medium lacking the specific amino acid or nucleotide. Because *Saccharomyces cerevisiae* is sensitive to glyphosate, GAT can also be used as a selectable marker. To demonstrate this, an evolved GAT gene (0_6D10), SEQ ID NO:196, is cloned from the T-DNA vector pMAXY3793 (as shown in Example 9) as a PstI-ClaI fragment containing the entire coding region and ligated into PstI-ClaI digested p424TEF (Gene, 1995, 156:119-122) as shown in Figure 13. This plasmid contains an *E. coli* origin of replication and a gene conferring carbenicillin resistance as well as a TRP1, tryptophan auxotroph selectable marker for yeast transformation.

[0435] The GAT containing construct is transformed into *E. coli* XL1 Blue (Statagene) and plated on LB carbenicillin (50 ug/ml) agar medium. Plasmid DNA is prepared and used to transform yeast strain YPH499 (Stratagene) using a transformation kit (Bio101). Equal amounts of transformed cells are plated on CSM-YNB-glucose medium (Bio101) lacking all aromatic amino acids (tryptophan, tyrosine, and phenylalanine) with added glyphosate. For comparison, p424TEF lacking the GAT gene is also introduced into YPH499 and plated as described. The results demonstrate that GAT activity function will as an efficient selectable marker. The presence of the GAT containing vector in glyphosate selected colonies can be confirmed by re-isolation of the plasmid and restriction digest analysis.

EXAMPLE 11. HERBICIDE SPRAY TESTS OF GAT EXPRESSING TOBACCO PLANTS

[0436] Tobacco shoots generated as described in EXAMPLE 9 were excised from the explants and transferred to basal root induction (BRI) medium composed of half-strength Murashige and Skoog (MS) medium, pH 5.7, with 1.5% sucrose, 0.3% Gelrite, 0-200 uM glyphosate and 400 ug/ml Timentin. Rooted plants and axillary shoots were clonally propagated by cutting the stem and transferring it to fresh BRI medium until the desired number of clones was obtained. Rooted plants were carefully removed from the solid medium. Prior to placing the plants into small pots of soil, the roots were washed to remove any remaining Gelrite. A protective plastic cover was kept over the plants for at least one week until the plants were well established.

[0437] To determine if GAT expressing tobacco plants could tolerate simulated field rate sprays of glyphosate, clonal lines of several events per GAT variant were tested. A typical test was set up as follows: One clone from each event was sprayed with 1 ml of solution containing the isopropylamine salt of glyphosate (Sigma P5671) and 0.125% Triton X-100, pH 6.8 such that the amount of active ingredient sprayed was equivalent to that present in commercial glyphosate products. For example, to achieve 32 oz/acre (1X) of herbicide containing 40% active ingredient ("ai"), 2.4 ul of 40% ai formulation was diluted into 1 ml water and sprayed on a plant in a 4-inch square pot (16 in^2). A mock application (0X) with surfactant only was also included. In some cases a second spray was applied 1-4 weeks later. Plants were kept in

COR-AAD1-00001586

controlled growth rooms at 25 °C and 70% humidity with 16 hr light.

[0438] In this example, 10 events confirmed positive for GATO_6D10 (SEQ ID NO:196), ten for GATO_5D3 (SEQ ID NO:193), 8 events for GATO_5B8 (SEQ ID NO:190), and plants transformed with the vector only (no GAT) were clonally propagated, transferred to soil and sprayed when plants had an average of 5 leaves. Seed-grown wild type plants were also sprayed. After two weeks, the vector only and seed grown plants sprayed with 0.5, 2 or 4X glyphosate stopped growing, wilted, and turned brown. Each of the transgenic GAT plants survived the spraying procedure without signs of glyphosate damage such as chlorosis, leaf elongation, stunting, or browning. All 0X plants were healthy, including the non-GAT control plants. Three weeks later all of the surviving plants were sprayed with an 8X dose. The 0X control plants died within two weeks. Again, all GAT plants survived.

[0439] Tobacco plants transformed with GAT and selected on glyphosate were fertile. Flowering and seed set were not detectably different from wild type plants.

EXAMPLE 12. MENDELIAN INHERITANCE OF GAT GENE AND GLYPHOSATE-TOLERANT PHENOTYPE

[0440] Mendelian inheritance of the GAT gene and glyphosate-tolerant phenotype was demonstrated with transformed Arabidopsis. Columbia type Arabidopsis plants were grown and transformed by the dipping method (Clough, SJ and Bent, AF, (1998) Plant J. 16(6):735-43) with a construct containing the GAT variant called chimera (SEQ ID NO:16). Bulk seed was collected and GAT plants were confirmed by PCR with primers specific to the insert within the T-DNA. T1 seed from individual events were sown on soil with 10-30 seeds per 2-inch square pot. When the first set of true leaves was emerging, pots were sprayed with glyphosate equivalent to 0.5 and 1X commercial product (as calculated in EXAMPLE 11). After two weeks, segregation of the transgene and tolerant phenotype was evident as shown in Table 5.

Table 5. Summary of segregation data for 0.5 and 1X glyphosate-tolerant T1 Arabidopsis

Chimera event (SEQ ID NO:16)	#Survivors	#Dead	Segregation ratio
1	8	11	1:1.4
3	6	22	1:3.7
5	26	2	13:1
13	10	9	1:1
65	46	19	2.4:1
Vector only	0	22	-
Wild-type	0	29	-

[0441] Ratios near 3:1 indicate a single segregating dominant event. Ratios greater than 3:1 1 indicate several segregating inserts. Ratios less than 3:1 1 can be due to small sample size effects, incomplete dominance, or position effects that render expression too low to confer herbicide tolerance. Compared to the controls, it was clear that the GAT gene was transmitted to the T1 generation and conferred glyphosate tolerance.

EXAMPLE 13: PRODUCTION OF GLYPHOSATE-RESISTANT MAIZE EXPRESSING GAT TRANSGENES

[0442] Maize plants expressing GAT variant transgenes were produced using the methods described in U.S. Patent No. 5,981,849, which is incorporated herein by reference. Specifically, *Agrobacterium tumefaciens* vectors were constructed according to methods known in the art. Each vector contained an insert having an ubiquitin promoter and intron, a GAT variant and a PinII terminator. Maize immature embryos were excised and infected with an *Agrobacterium tumefaciens* vector containing the GAT variant of interest. After infection, embryos were transferred and cultured in co-cultivation medium. After co-cultivation, the infected immature embryos were transferred onto media containing 1.0 mM glyphosate (Roundup ULTRA MAX™). This selection lasted until actively growing putative transgenic calli were identified. The putative transgenic callus tissues were sampled for PCR and Western assay (data not shown) to confirm the presence of the GAT gene. The putative transgenic callus tissues were maintained on 1.0 mM glyphosate selection media for further growth and selection before plant regeneration. At regeneration, callus tissue confirmed to be transgenic were transferred onto maturation medium containing 0.1 mM glyphosate and cultured for somatic embryo maturation. Mature embryos were then transferred onto regeneration medium containing 0.1 mM glyphosate for shoot and root formation. After shoots and roots emerged, individual plantlets were transferred into tubes with rooting medium containing 0.1 mM glyphosate. Plantlets with established shoots and roots were transplanted into pots in the greenhouse for further

COR-AAD1-00001587

growth, the generation of T0 spray data and the production of T1 seed.

[0443] In order to evaluate the level of glyphosate resistance of the transgenic maize plants expressing the GAT variant transgenes, T0 plants were sprayed with glyphosate (Roundup ULTRA MAX™) in the greenhouse. Plant resistance levels were evaluated by plant discoloration scores and plant height measurements. Plant discoloration and plant height were evaluated according to the following scales:

Discoloration score at 1, 2, 3 and 4 weeks after spray with glyphosate

9 = no leaf/stem discoloration
7 = minor leaf/stem discoloration
5 = worse leaf/stem discoloration
3 = severely discolored plant or dying plant
1 = dead plant

Plant height measurements
before spraying with glyphosate
after spraying with glyphosate at 1, 2, 3 and 4 weeks
mature plants (at tasseling)

[0444] Two plants were sent to the greenhouse from each event (independent transgenic callus) listed in Table 6. Plant 1 was kept for seed production and was not sprayed with glyphosate. Plant 2 was sprayed at 4x glyphosate (1x glyphosate = 26 ounces/acre) at 14 days after transplanting. The T0 plant discoloration scores with 4x spray at 7 and 14 days after the spray are shown in Tables 6 and 7. Height data at tasseling is shown in Figure 14. An additional experiment was performed in which T0 plants were sprayed with 6x glyphosate. The T0 plant discoloration scores with 6x spray at 10 days after spray are shown in Table 8.

Table 6. Resistance Scores at 7 days after treatment with 4x glyphosate

constructs	# events tested with 4x	% events @ 9	% events @ 7	% events @ <7
18534 (SEQ ID NO:196)	169	30% (50)	59% (101)	11% (18)
18537 (SEQ ID NO:193)	72	40% (29)	54% (39)	6% (4)
18540 (SEQ ID NO:190)	111	32% (36)	61% (67)	7% (8)
total	352	33% (115)	59% (207)	8% (30)

Table 7. Resistance Scores at 14 days after treatment with 4x glyphosate

constructs	# events tested with 4x	% events @ 9
18534 (SEQ ID NO:196)	169	29%(49)
18537 (SEQ ID NO:193)	72	50%(36)
18540 (SEQ ID NO:190)	111	29%(32)
total	352	33%(117)

Table 8. Resistance Scores at 10 days after treatment with 6X glyphosate

constructs	# events tested with 6X	% events with no damage after glyphosate treatment (score= 9)
19286 (SEQ ID NO:814)	312	51%(160)
19288 (SEQ ID NO:549)	310	52%(163)

COR-AAD1-00001588

(continued)

constructs	# events tested with 6X	% events with no damage after glyphosate treatment (score= 9)
19900 (SEQ ID NO:738)	231	56%(129)
19902 (SEQ ID NO:638)	230	42%(96)
21895 (SEQ ID NO:848)	55	30%(17)
21896 (SEQ ID NO:912)	61	61%(37)
21905 (SEQ ID NO:906)	32	70%(25)
total	1231	51 %(627)

EXAMPLE 14: GAT IS ALSO AN ACYLTRANSFERASE

[0445] The ability of GAT variants (B6 (SEQ ID NO:7), 0_6D10 (SEQ ID NO:448), 17-15H3 (SEQ ID NO:601), and 20-8H12c (SEQ ID NO:817)) to transfer the propionyl group from propionyl CoA to glyphosate was tested in reaction mixtures containing 5mM glyphosate or no glyphosate. Propionyl CoA was present at 1 mM. After 30 minutes the reactions were terminated and the presence of free propionyl CoA was determined by the addition of DTNB. All variants showed glyphosate-dependent hydrolysis of propionyl CoA. These results indicate that GAT also functions as an acyl-transferase.

EXAMPLE 15: T1 STUDIES OF GLYPHOSATE-RESISTANT MAIZE EXPRESSING GAT TRANSGENES

[0446] Maize plants expressing GAT variant transgenes 18-28D9b (SEQ ID NO:814) and 17-15H3 (SEQ ID NO:549) were produced using the methods described in Example 13. T1 plants were used for the generation of glyphosate field tolerance data. The T1 plants were treated in the field with four different glyphosate spray treatments (0X, 4X, 8X, and 4X + 4X) for each event. The plants were sprayed at V3 and V8. Plants were scored 10 days after treatment for leaf discoloration and plant height comparisons as described in Example 13. The T1 field spray data correlated well with the results previously obtained in the greenhouse as reported in Example 13. T2 seeds were collected for further studies.

EXAMPLE 16: THERMOSTABILITY OF GAT POLYPEPTIDES

A. EFFECT OF TEMPERATURE VARIATION ON GLYPHOSATE TOLERANCE OF GLYPHOSATE RESISTANT MAIZE EXPRESSING GAT TRANSGENES

[0447] Maize plants expressing GAT variant transgenes 10_4F2 (SEQ ID NO:203), 17-15H3 (SEQ ID NO:549), and 18-28D9b (SEQ ID NO:814) were produced using the methods described in Example 13. The effect of temperature on glyphosate tolerance was evaluated in T1 plants. The T1 plants were grown in cool/cold (day 14° C, night 8° C), warm (day 28° C, night 20° C), and hot (day 37° C, night 20° C) conditions. T1 plants were sprayed at V2 with four different glyphosate spray treatments (0X, 4X, 6X, and 8X). Plants were scored at 5 and 14 days after treatment for leaf discoloration and plant height comparisons as described in Example 13. Visual observations indicated that glyphosate tolerance is not adversely effected by the range of temperatures tested.

B. EFFECT OF TEMPERATURE VARIATION ON GAT ACTIVITY IN VITRO

[0448] *In vitro* thermostability of several GAT polypeptides (DS3 (a native GAT polypeptide corresponding to SEQ ID NO: 8), 6_6D5 (SEQ ID NO: 410), 17-15H3 (SEQ ID NO: 601), 20-8H12 (SEQ ID NO: 739), 22-13B12 (SEQ ID NO: 781) and 401 (a native GAT polypeptide corresponding to SEQ ID NO: 6)) was evaluated in accordance with the following method. The enzymes were distributed to 200 μl strip PCR tubes (VWR, San Francisco, CA) and incubated in a gradient thermocycler (ML Research, Watertown, MA) for 15 minutes at various temperatures between 30°C and 60°C as indicated in Figure 17. Precipitated protein was removed by centrifugation, and surviving enzymatic activity of the remaining soluble protein was measured at 22°C by the continuous spectrophotometric assay, as described in Example 7. Saturating concentrations of glyphosate (10 mM for DS3 (SEQ ID NO: 8), 401 (SEQ ID NO: 6) and 6_6D5 (SEQ ID NO: 410); 5 mM for 17-15H3 (SEQ ID NO: 601), 20-8H12 (SEQ ID NO: 739), and 22-13B12 (SEQ ID NO: 781) and AcCoA (167

COR-AAD1-00001589

μM) were used.

[0449] The data is depicted in Figure 17. Native *(i.e.,* wild type) GAT polypeptides DS3 (SEQ ID NO: 8) and 401 (SEQ ID NO: 6) appeared stable with respect to activity at temperatures up to about 42 to about 44°C. GAT polypeptides that are not native to any organism (i.e., not wild type) appeared stable at temperatures in the range of about 47°C to about 54°C.

[0450] The half lives of several GAT polypeptides were also measured at 37.5°C according to the following procedure. GAT polypeptides 401 (SEQ ID NO: 6), 17-15H3 (SEQ ID NO: 601), 20-8H12 (SEQ ID NO: 739), 22-13B12 (SEQ ID NO: 781), 22-15B4 (SEQ ID NO: 946) and 22-18C5 (SEQ ID NO: 795) were incubated in a matrix of 25 mM Hepes, pH 7.2, 10 mM KCl and 10% methanol ("HKM"). At various timepoints, aliquots were withdrawn and assayed in triplicate at 22°C using the continuous spectrophotometric assay described in Example 7 using saturating concentrations of glyphosate (20 mM for 401, 5 mM for the rest) and AcCoA (167 uM). The standard error at each time point averaged about 2.9%. GAT activity was plotted as a function of incubation time and the data was fitted to a curve for exponential decay ($y=e^{-x}$), where y is enzyme activity and x is time in hours, from which half life was calculated. The data is shown below in Table 9.

Table 9. Half-lives of GAT polypeptides at 37.5 C

Enzyme	Half-life, hrs
401 (SEQ ID NO: 6)	14
17-15H3 (SEQ ID NO: 601)	45
20-8H12 (SEQ ID NO: 739)	54
22-13B12 (SEQ ID NO: 781)	67
22-15B4 (SEQ ID NO: 946)	26
22-18C5 (SEQ ID NO: 795)	43

EXAMPLE 17: PRODUCTION OF GLYPHOSATE-RESISTANT SOYBEAN EXPRESSING GAT TRANSGENES

[0451] Soybean plants expressing GAT variant transgenes were produced using the method of particle gun bombardment (see Klein et al. (1987) Nature 327:70-73) using a DuPont Biolistic PDS1000/He instrument. The selection agent used during the transformation process was hygromycin. Either the hygromycin selectable marker gene remained in the transgenic events or the hygromycin gene was excised by methods known in the art. DNA fragments were prepared with a synthetic constitutive promoter, a GAT variant and PinII terminator. The selectable marker gene, comprising the 35S CaMV promoter, HPT gene and NOS terminator, was cobombarded with the GAT gene variant as described above. Bombarded soybean embryogenic suspension tissue was cultured for one week in the absence of selection agent. Embryogenic suspension tissue was placed in liquid selection medium for 6 weeks. Putative transgenic suspension tissue was sampled for PCR analysis to determine the presence of the GAT gene. Putative transgenic suspension culture tissue was maintained in selection medium for 3 weeks to obtain enough tissue for plant regeneration. Suspension tissue was matured for 4 weeks using standard procedures; matured somatic embryos were desiccated for 4-7 days and then placed on germination induction medium for 2-4 weeks. Germinated plantlets were transferred to soil in cell pack trays for 3 weeks for acclimatization. Plantlets were potted to 10-inch pots in the greenhouse for evaluation of glyphosate resistance.

[0452] To determine the level of glyphosate resistance of transgenic soybeans expressing the GAT variant transgenes, T0 plants were sprayed with glyphosate (Roundup ULTRA MAX™) in the greenhouse. Plant resistance levels were evaluated by plant discoloration scores and plant height measurements.

Discoloration score at 2 weeks after spray with glyphosate

[0453]

COR-AAD1-00001590

9 = no leaf/stem discoloration
7 = minor leaf/stem discoloration
5 = worse leaf/stem discoloration
3 = severely discolored plant or dying plant
1 = dead plant

[0454] One to four plants were sent to the greenhouse from each independent transgenic event. An additional 1-2 plants per event were grown in controlled environment growth chambers for seed production and were not sprayed with glyphosate. The greenhouse plants were sprayed at 1X, 2X or 4X glyphosate (1X glyphosate = 26 ounces/acre of RoundUp ULTRA MAX™) 3-4 weeks after transfer to soil. The T0 plant discoloration scores with 2X and 4X spray rates are shown in Table 10 and Table 11, respectively.

[0455] These results show that soybeans are effectively transformed with GAT gene variants as confirmed by PCR analysis. Transgenic soybeans expressing GAT gene variants are resistant to glyphosate at 2X and 4X spray rates. Events surviving the 4X glyphosate spray rate do show some minor leaf discoloration however within 2 weeks of the spray test, plants recover and demonstrate normal leaf morphology.

Table 10. Resistance Scores at 10 days after treatment with 2X glyphosate.

	#EVENTS TESTED WITH 2X	% EVENTS @ 7-8	% EVENTS @ 3-6
SEQ ID NO:193	27	15%(4)	11%(3)
SEQ ID NO:824	38	8%(3)	74%(23)

Table 11. Resistance Scores at 10 days after treatment with 4X glyphosate.

	#EVENTS TESTED WITH 4X	% EVENTS @ 7-8	% EVENTS @ 3-6
SEQ ID NO:824	23	8%(2)	43%(10)

EXAMPLE 18: EFFECT OF SALT ON GAT KINETICS

[0456] To better approximate the physiological conditions under which the GAT enzymes of the invention are intended to be used (*e.g.*, plant cells), the activities of some GAT enzymes of the invention were re-evaluated in the presence of added salt. Figures 15A and 15B provide a comparison of the kinetic parameters K_m and k_{cat}/K_m, respectively, for native GAT enzymes GAT401 (SEQ ID NO:6), B6 (SEQ ID NO:7), and DS3 (SEQ ID NO:8), and evolved GAT enzymes 0_6D10 (SEQ ID NO:448), 10_4F2 (SEQ ID NO:454), 18-28D9 (SEQ ID NO:618), 17-15H3 (SEQ ID NO:601), 17-10B3 (SEQ ID NO:592), 20-8H12 (SEQ ID NO:739), 20-16A3 (SEQ ID NO:639), and 20-30C6 (SEQ ID NO:683), assayed in either the absence of added KCl (unshaded bars) or in the presence of 20 mM KCl (shaded bars). Protein concentrations were determined using the Bradford assay as described in Example 7. Owing to their extremely low K_ms for glyphosate in the absence of KCl, the kinetic parameters for evolved GAT enzymes 0_6D10, 18-28D9 and 20-8H12 were determined in the absence of KCl using the mass spectrometry assay as described in Example 3, while all other kinetic parameters (either in the absence or presence of KCl) were determined using the continuous spectrophotometric assay as described in Example 7. Error bars represent the standard deviation of multiple assays, where available. Figure 15A shows that addition of salt (20 mM KCl) to the assay buffer significantly increases the K_m value for glyphosate. The k_{cat} value remains relatively unchanged or increases slightly, the net result being a lower observed k_{cat}/K_m value for GAT enzymes assayed in the presence of 20 mM KCl than in the absence of added KCl (Figure 15B).

EXAMPLE 19: FURTHER EVOLVED GAT GENES ENCODING GAT ENZYMES WITH EXTREMELY HIGH ACTIVITY

[0457] Additional iterations of directed molecular evolution yielded further evolved gat genes encoding GAT enzymes exhibiting extremely high GAT activity, *e.g.* exhibiting one or more improved property such as reduced K_m for glyphosate, increased k_{cat}, or increased k_{cat}/K_m compared to previously-described GAT enzymes.

[0458] The further evolved *gat* genes were first selected for growth in *E. coli* in minimal M9 medium as described in Example 8, except that 5 mM rather than 1 mM glyphosate was used in the selection. Proteins were purified as described in Example 6 above.

[0459] Protein concentrations were determined by UV absorbance at 205 nm. The extinction coefficient was determined by the method described by Scopes (1994; Protein Purification, Principles and Practice, Springer, NewYork) according to the formula E (mg ml^{-1} cm^{-1}) = 27 + 120(A_{280}/A_{205}) = 30.5. Prior to quantitation by UV absorbance the protein solution was buffer-exchanged into 50 mM Na_2SO_4 using a NAP-5 column (Amersham-Pharmacia Biotech).

COR-AAD1-00001591

[0460] Exemplary further evolved *gat* coding sequences comprise nucleic acids sequences identified herein as SEQ ID NOs: 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, and 930, which encode further evolved GAT enzymes comprising amino acid sequences identified herein as SEQ ID NOs: 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, and 931, respectively. Some such further evolved GAT enzymes exhibit extremely high GAT activity, in that they exhibit one or more improved property such as reduced K_m for glyphosate, increased k_{cat}, or increased k_{cat}/K_m, compared to previously-described GAT enzymes assayed under the same conditions.

[0461] Figures 16A, 16B and 16C provide a comparison of the kinetic parameters K_m, k_{cat}, and k_{cat}/K_m, respectively, of several previously-described GAT enzymes (unshaded bars) to the kinetic parameters of some further evolved GAT enzymes of the invention (shaded bars), assayed using the continuous spectrophotometric assay in the presence of 20 mM KCl with protein quantified via UV absorbance as described above. Error bars represent the standard deviation of multiple assays, where available. Under these assay conditions, native GAT enzyme GAT401 (SEQ ID NO:6) exhibited a K_m for glyphosate of about 4 mM, a k_{cat} of about 5.4 min^{-1}, and a k_{cat}/K_m of about 1.35 mM^{-1} min^{-1}. When assayed under these conditions, some further evolved GAT enzymes of the invention (shaded bars) exhibit a range of K_m values for glyphosate of less than about 0.4 mM (such as, between about 0.4 mM and 0.1 mM), k_{cat} values of at least about 1000 min^{-1} (such as, between about 1000 min^{-1} and about 2500 min^{-1}), and k_{cat}/K_m values of at least about 4800 mM^{-1} min^{-1} (such as, between about 4800 mM^{-1} min^{-1} and about 8000 mm^{-1} min^{-1}). For example, some further evolved GAT enzymes of the invention exhibit at least about a 7000-fold increase in k_{cat}/K_m over native GAT enzyme GAT401 under these assay conditions.

[0462] Some further evolved GAT enzymes of the invention comprise one or more amino acid residue positions not observed in previously described GAT polypeptides and GAT enzymes, such as, at position 27, a B1, Z1 or A amino acid residue; at position 33, an N or G amino acid residue; at position 46, a B2, Z4, or H amino acid residue; and at position 93, an R amino acid residue; where B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y and V; B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T; Z1 is an amino acid selected from the group consisting of A, I, L, M and V; and Z4 is an amino acid selected from the group consisting of R, H and K. For example, some further evolved GAT enzymes of the invention comprise one or more of: an Ala at position 27 (*i.e.*, Ala27); an Asn or a Gly at position 33 *(i.e.,* Asn33 or Gly33); a His at position 46 *(i.e.,* His46); and an Arg at position 93 *(i.e.,* Arg93), with sequence numbering corresponding to that of, *e.g.,* SEQ ID NO: 907.

[0463] Sequence/activity analyses were performed to identify amino acid residues which correlate positively with a high k_{cat}/K_m (as manifested by a high k_{cat}, a low K_m, or both). Amino acid residues which appear to correlate positively with a high k_{cat}/K_m include Glu14, Asp32, Asn33, Gly38, and Thr62 (sequence numbering corresponding to that of SEQ ID NO:907). Additional GAT enzymes may be constructed by substituting codons for one or more of these residues into the appropriate position(s) of a coding sequence of a template GAT polypeptide. For example, additional GAT enzymes were generated by substituting one or more of codons encoding Glu at codon position 14, Asp at position 32, Asn at position 33, Gly at position 38, and Thr at position 62, into a nucleic acid sequence encoding a template polypeptide, such as GAT 24-5H5 (SEQ ID NO:845) or GAT 25-8H7 (SEQ ID NO:907), two of the further evolved GAT enzymes exhibiting extremely high activity as described above. Exemplary further evolved GAT enzymes generated in this manner, identified herein as R12G1 (SEQ ID NOs917), R12G2 (SEQ ID NO:919), R12G3 (SEQ ID NO:921), R12G4 (SEQ ID NO:923), R12G5 (SEQ ID NO:925), R12G6 (SEQ ID NO:927), R12G7 (SEQ ID NO:929), and R12G8 (SEQ ID NO:931), encoded by nucleic acids identified as SEQ ID NOs: 916, 918, 920, 922, 924, 926, 928, and 930, respectively, exhibited extremely high GAT activities comparable to those of the template polypeptides.

EXAMPLE 20: AMINO ACIDS THAT CORRELATE WITH HIGH GAT ACTIVITY

[0464] The amino acids aspartic acid (Asp, D), histidine (His, H) and cysteine (Cys, C) are known to be associated with the active sites of various acetyltransferase enzymes. To determine if any such residues play a role in GAT activity, all D, C, and H residues of GAT20-30C6 (SEQ ID NO:683) were individually mutated to alanine (Ala, A) and the mutated enzymes assayed for N-acetylglyphosate activity. Variants containing the substitutions D34A and H41A retained only about 2%-3% of the activity of the unmodified enzyme, while the variant containing the substitution H138A exhibited essentially no measurable GAT activity. On the other hand, variants containing the substitutions H138R and H138S retained low but measurable GAT activity (particularly at pHs greater than 6.8), suggesting that His (and nominally Arg and Ser) at position 138 may serve as an active-site base.

COR-AAD1-00001592

Table 12

Enzyme	0 KCl Bradford protein assay			
	k_{cat} (min^{-1})	K_m (mM)	k_{cat}/K_m (mM^{-1} min^{-1})	% k_{cat}/K_m of 20-30C6
20-30C6 (SEQ ID NO:683)	386	0.182	2122	100
20-30C6 H41 A	208	4.80	43	2.0
20-30C6 D34A	127	2.33	54	2.6
20-30C6 H138A	< 0.02	nd	nd	< 0.005
20-30C6 H138R	44.3	17.8	2.49	0.12
20-30C6 H138S	5.35	7.1	0.75	0.03

EXAMPLE 21: IMPROVING GAT EXPRESSION IN PLANTS

[0465] Plants, animals, and microbes are known to have specific codon preferences that affect the efficiency of amino acid incorporation during translation of gene transcripts. Rare codons could cause problems with tRNA recruitment during translation, which could then lead to lower accumulation of the encoded protein. The original parental *gat* genes were from bacteria such as *Bacillus licheniformis,* and, as such, may not have an optimal codon distribution for expression in plants. Evolved *gat* genes of the invention have successfully been expressed in plants (see, *e.g.,* Examples 9, 11, 13, and 17, above), yet an opportunity exists to improve protein production by increasing the translation efficiency in plants. One way to accomplish this is by substituting one or more codons in the *gat* coding sequence which are used infrequently in plants for codons for the same amino acid(s) which are more frequently used in plants, thereby generating silent mutations in the *gat* coding sequence with an unchanged sequence of the encoded protein.

[0466] Tables showing the frequency of codon usage in corn, cotton and soybeans (available, for example, from the website maintained by the Kazusa DNA Research Institute, Chiba, Japan) were compared to generate the following table (Table 13) showing codons which are, in general, more frequently or less frequently utilized in either monocot or dicot plants.

Table 13

Amino acid			Codons more frequently utilized in plants				Codons less frequently utilized in plants			
Alanine	Ala	A	GCA	GCC	GCT		GCG			
Cysteine	Cys	C	TGC	TGT						
Aspartic acid	Asp	D	GAC	GAT						
Glutamic acid	Glu	E	GAA	GAG						
Phenylalanine	Phe	F	TTC	TTT						
Glycine	Gly	G	GGA	GGT			GGC	GGG		
Histidine	His	H	CAC	CAT						
Isoleucine	Ile	I	ATC	ATT			ATA			
Lysine	Lys	K	AAA	AAG						
Leucine	Leu	L	TTG	CTC	CTG	CTT	CTA	TTA		
Methionine	Met	M	ATG							
Asparagine	Asn	N	AAC	AAT						
Asparagine	Asn	N	AAC	AAT						
Proline	Pro	P	CCA	CCT			CCC	CCG		
Glutamine	Gln	Q	CAA	CAG						

COR-AAD1-00001593

(continued)

Amino acid			Codons more frequently utilized in plants				Codons less frequently utilized in plants			
Arginine	Arg	R	AGA	AGG			CGA	CGC	CGG	CGT
Serine	Ser	S	AGT	TCA	TCC	TCT	AGC	TCG		
Threonine	Thr	T	ACA	ACC	ACT		ACG			
Valine	Val	V	GTG	GTT			GTA	GTC		
Tryptophan	Trp	W	TGG							
Tyrosine	Tyr	Y	TAC	TAT						

[0467] A second way to increase plant expression of microbial genes is to increase the G+C content near the initiating methionine residue. Naturally-occurring coding sequences in plants tend to contain two or three G and/or C residues immediately downstream of the ATG initiation codon (Joshi et al.(1997) Plant Mol. Biol. 35:993-1001). Introducing into the *gat* coding sequence one or two CG-rich codons immediately downstream of the ATG intitiation codon may create a more plant-like coding sequence and thus may enhance its expression in plants. Substitution of the second codon (isoleucine, ATA) for an alanine codon (GCG) resulted in a Ile2Ala variant with reduced k_{cat} compared to the unmodified enzyme. On the other hand, insertion of an alanine codon (either GCG or GCT) between the codons for Met at codon position 1 and Ile at codon position 2 resulted in a *gat* coding sequence encoding a GAT enzyme containing an Ala residue inserted between the Met at position 1 and the Ile at position 2. An exemplary GAT enzyme variant containing two alanines inserted between Met1 and Ile2 denoted, identified as 22-15B4 M1MAA (to signify the insertion of two Ala residues immediately following the Met at position 1) and having the protein sequence SEQ ID NO:948, exhibited a reduced k_{cat} compared to the unmodified enzyme 22-15B4 (SEQ ID NO:789). An exemplary GAT enzyme containing one alanine inserted between Met1 and Ile2, denoted 22-15B4 M1MA (to signify the insertion of an Ala residue immediately following the Met at position 1), having the protein sequence SEQ ID NO:946, exhibited essentially unaltered kinetics compared to the unmodified enzyme 22-15B4.

[0468] A general strategy for improving GAT expression in plants was developed. Evolved *gat* coding sequences may be altered by replacing codons less frequently utilized in plants for codons more frequently utilized in plants, for example according to the table above. Codons less frequently utilized in plants (*e.g.*, according to the above table) should generally be avoided. In this manner at least one codon (such as, at least three codons, at least five codons, or at least least ten codons), may be changed in the *gat* coding sequence from codon(s) less frequently utitlized in plants to codon(s) more frequently utitlized in plants. The codons which are replaced may be located at the 5' end of the coding sequence (*e.g.*, within the first 10 codons, within the first 20 codons, within the first 50 codons, or within the first 100 codons) of the *gat* coding sequence. Alternatively, the codons which are replaced may be located throughout the *gat* coding sequence. The more frequently utilized codons furthermore may be chosen to avoid more than about 5-10 (such as, *e.g.*, more than about 5, more than about 6, more than about 7, more than about 8, more than about 9 or more than about 10) consecutive occurrences of G+C or of A+T within the coding sequence. The coding sequence may also be altered to contain one or two CG-rich codons immediately downstream of the ATG intitiation codon, such as, for example, by inserting an Ala codon (*e.g.*, a frequently utilized Ala codon) immediately downstream of and adjacent to the initiating Met codon of the *gat* coding sequence.

[0469] Table 14 provides exemplary *gat* coding sequences altered as described above.

Table 14

original *gat* coding sequence	altered *gat* coding sequence	codon changes made	encoded protein
GAT20-8H12 /4604 (SEQ ID NO:738)	GAT4604SR (SEQ ID NO:932)	Ser 62 TCG -> TCT Arg 111 CGG -> AGG	20-8H12 (SEQ ID NO:739)
GAT22-18C5/4609 (SEQ ID NO:794)	GAT4609SR (SEQ ID NO:933)	Ser 62 TCG -> TCT Arg 111 CGG -> AGG	22-18C5 (SEQ ID NO:795)
GAT22-16D8/4610 (SEQ ID NO:792)	GAT4610R (SEQ ID NO:934)	Arg 111 CGG -> AGG	22-16D8 (SEQ ID NO:793)

COR-AAD1-00001594

(continued)

original *gat* coding sequence	altered *gat* coding sequence	codon changes made	encoded protein
GAT22-15B4/4611 (SEQ ID NO:788)	GAT461 1 R (SEQ ID NO:935)	Arg 111 CGG -> AGG	22-15B4 (SEQ ID NO:789)
GAT24-5H5/4614 (SEQ ID NO:848)	GAT4614SR (SEQ ID NO:936)	Ser 62 TCG -> TCT Arg 111 CGG -> AGG	24-5H5 (SEQ ID NO:849)
GAT24-5H5/4614 (SEQ ID NO:848)	GAT4614VSR (SEQ ID NO:937)	Val 4 GTG -> GTA Ser 62 TCG -> TCT Arg 111 CGG -> AGG	24-5H5 (SEQ ID NO:849)
GAT23-2H11/4615 (SEQ ID NO:836)	GAT4615R (SEQ ID NO:938)	Arg 111 CGG -> AGG	23-2H11 (SEQ ID NO:837)
GAT24-15C314616 (SEQ ID NO:862)	GAT4616R (SEQ ID NO:939)	Arg 111 CGG -> AGG	24-15C3 (SEQ ID NO:863)
GAT23-6H10/4617 (SEQ ID NO:844)	GAT4617R (SEQ ID NO:940)	Arg 111 CGG -> AGG	23-6H10 (SEQ ID NO:845)
GAT25-8H7/4618 (SEQ ID NO:906)	GAT4618SR (SEQ ID NO:941)	Ser 62 TCG -> TCT Arg 111 CGG -> AGG	25-8H7 (SEQ ID NO:907, 957)
GAT25-8H7/4618 (SEQ ID NO:906)	GAT4618VSR (SEQ ID NO:942)	Val 4 GTG -> GTA Ser 62 TCG -> TCT Arg 111 CGG -> AGG	25-8H7 (SEQ ID NO:907)
GAT25-19C8/4619 (SEQ ID NO:912)	GAT4619SR (SEQ ID NO:943)	Ser 62 TCG -> TCT Arg 111 CGG -> AGG	25-19C8 (SEQ ID NO:913)
GAT25-19C8/4619 (SEQ ID NO:912)	GAT4619VSR (SEQ ID NO:944)	Val 4 GTG -> GTA Ser 62 TCG -> TCT Arg 111 CGG -> AGG	25-19C8 (SEQ ID NO:913)
GAT22-15B4/4611 (SEQ ID NO:788)	GAT4611A (SEQ ID NO:945)	Ala codon inserted between Met1 and Ile2	22-15B4 M1MA (SEQ ID NO:946)
GAT22-15B4/4611 (SEQ ID NO:788)	GAT4611AA (SEQ ID NO:947)	2 Ala codons inserted between Met1 and Ile2	22-15B4 M1MAA (SEQ ID NO:948)
GAT25-8H7/4618 (SEQ ID NO:906)	GAT4618A (SEQ ID NO:949)	Ala codon inserted between Met1 and Ile2	25-8H7 M1 MA (SEQ ID NO:950)
GAT25-8H7/4618 (SEQ ID NO:906)	GAT4620 (SEQ ID NO:951)	Ala codon inserted between Met1 and Ile2 and first 7 codons = more frequently utilized codons; plus Ser 62 TCG -> TCT Arg 111 CGG -> AGG	25-8H7 M1MA (SEQ ID NO:950)
GAT25-8H7/4618 (SEQ ID NO:906)	GAT4621 (SEQ ID NO:952)	Ala codon inserted between Met1 and Ile2 and more frequently utilized codons throughout	25-8H7 M1 MA (SEQ ID NO:950)

[0470] A binary vector with a dMMV-*gat*-UBQ3 cassette in the T-DNA was transformed into competent *Agrobacterium tumefaciens* strain C58 cells by electroporation (McCormac *et al.*, Mol Biotechnol. 9:155-159, 1998). After growth on LB + 40ug/ml kanamycin plates for 2 days at 28°C, colonies were inoculated into LB + 40ug/ml kanamycin liquid medium and shaken overnight at 28°C. The *Agrobacterium* cells were collected by centrifugation at 4000 g for 10 minutes and then resuspended in a volume of 10mM MgSO4 equivalent to the initial culture volume. This bacterial suspension was forced or "infiltrated" into the intercellular spaces *of Nicotiana benthamiana* leaves using a 1 ml plastic syringe (with no

COR-AAD1-00001595

needle). By infiltrating 200-300 μl of bacterial suspension into each spot (typically 3-4cm^2 in infiltrated area), 4 or more spots could be arranged on a single leaf still attached to the plant. In some cases the gat-containing *Agrobacterium* strain was diluted 5:1 or 10:1 with a second *Agrobacterium* strain lacking *gat* prior to infiltration. This dilution step has the effect of reducing the overall expression of the *gat* gene in the plant cells, thereby preventing saturation and allowing easier visualization of expression differences between variants and constructs. After 3 days the leaf material was ground, extracted in aqueous buffer, and centrifuged. The supernatant, containing the soluble proteins, was subjected to SDS-PAGE, and the gel was blotted and probed with an antiGAT polyclonal antibody.

[0471] The level of GAT protein accumulated in tobacco leaves infiltrated with the GAT4620 gene was comparable to the level of protein accumulated in leaves transformed with the unmodified GAT25-8H7/4618 gene. Tobacco leaves harboring the GAT4621 gene, on the other hand, exhibited about two-fold greater GAT protein accumulation, as a percent of total protein, compared to leaves expressing the unmodified GAT25-8H7/4618 gene.

EXAMPLE 22: T1 STUDIES OF GLYPHOSATE-RESISTANT SOYBEAN EXPRESSING GAT TRANSGENES

[0472] Soybean plants expressing GAT transgene 18-28D9c (SEQ ID NO:824) were produced using the methods described in Example 17. T1 seed was collected from glyphosate sprayed T0 plants. T1 seed were germinated under greenhouse conditions in RediEarth^{R}360 medium, available from Scotts, Marysville, OH, and sprayed at V2-V3 stage with either 2X or 4X Glyphosate (RoundUp ULTRA MAX™, available from Monsanto, St. Louise, MO) as per methods described in Example 17. Plants were scored after 10 days and leaf discoloration scores taken as described in Example 17. The T1 greenhouse spray data correlated well with previous greenhouse results at the T0 plant stage. T2 seed was collected for further studies.

EXAMPLE 23: PRODUCTION OF GLYPHOSATE AND SULFONAMIDE RESISTANT SOYBEANS EXPRESSING GAT AND HRA TRANSGENES

[0473] Soybean plants expressing GAT & HRA, high resistance allele of acetolactate synthase (U.S. Patent Nos. 5,605,011, 5,378,824, 5,141,870, and 5013,659), genes were produced using the methods described in Example 17. The HRA gene was used as selectable marker gene for transformation. The selection agent was chlorsulfuron at a concentration of 100 ng/ml. The selectable marker gene was comprised of the S-adenosyl-L-methionine synthetase (SAMS) promoter from *Glycine max* (U.S. 2003/226166), HRA coding sequence from *Glycine max* and acetolactate synthase terminator from *Glycine max*. The selectable marker gene was either linked to or co-bombarded with a GAT construct consisting of a synthetic constitutive promoter (U.S. Patent Nos. 6,072,050 and 6,555,673) or the maize Histone 2B promoter (U.S. Patent No. 6,177,611), a GAT variant (18-28D9c (SEQ ID NO:824)) and the Pin II terminator (Gyheung an et al., *Plant Cell* 1:115:122 (1989)). Transgenic plants were generated as described in Example 17. Levels of glyphosate resistance were determined as described in Example 17 using plant discoloration scores after 2X or 4X glyphosate application rates. The results shown in Table 15 demonstrate that different constitutive promoters driving GAT variant (18-28D9c (SEQ ID NO:824)) confer glyphosate resistance in T0 plants.

Table 15. Resistance Scores at 10 days after treatment with 4X glyphosate.

EVENTS	# EVENTS TESTED WITH 4x	% EVENTS @ 7-8 SCORE	% @ 3-6 SCORE
PHP20163a (45) (SEQ ID NO:824) (SCP1 PROMOTER)	58	15.5% (9)	77.6%
PHP20558a (11) (SEQ ID NO:824) (H2B PROMOTER)	26	34.6% (9)	42.3%

EXAMPLE 24: T1 PRE-EMERGENCE STUDIES OF SOYBEANS EXPRESSING GAT AND HRA TRANSGENES

[0474] T1 seed generated from experiments as described in Example 17, were planted in pots of Tama Silt loam in the greenhouse. Pots were immediately sprayed with a pre-emergence application of chlorimuron, rimsulfuron or tribenuron at a rate of 70 gms a.i./hectare. Germinating plants were evaluated 10 days post spray application based on plant discoloration scores described in Example 17. All HRA and GAT events survived all pre-emergence spray applications with a rating of 9 (uninjured). These results demonstrate pre-emergence resistance to sulfonamide chemistry in soybeans.

COR-AAD1-00001596

EXAMPLE 25: T1 POST-EMERGENCE STUDIES OF SOYBEANS EXPRESSING GAT AND HERA TRANSGENES

[0475] T1 seed generated from experiments as described in Example 17 were germinated in RediEarth^{R} 360 medium in the greenhouse. Plants were sprayed at the V2-V3 stage (14 days after potting) with thifensulfuron, chlorimuron, rimsulfuron or tribenuron (70, 70, 35, 35 gm a.i./hectare, respectively). Plants were evaluated 10 days post application based on plant discoloration scores described in Example 17. Results are shown in Table 16.

Table 16. Resistance Scores at 10 days after Post-Emergence treatment with Sulfonamide Chemistry.

	Average Resistance scores from GAT (SEQ ID NO:824)/SAMS promoter-HRA events
Unsprayed control	9
Chlorimuron (70gm a.i./ha)	7.75
Rimsulfuron (35gm a.i./ha)	2.21
Tribenuron (35 gm a.i./ha)	3.83
Thifensulfuron (70 gm a.i./ha)	7.81

[0476] Events having a plant discoloration rating 7 or 8 after thifensulfuron spray were sprayed with either a 2X or 4X application of glyphosate after 10 days as per methods described in Example 17. Events were evaluated based on discoloration scores described in Example 17. All thifensulfuron tolerant events survived the glyphosate spray with score of 7 or 8 (results not shown). These results demonstrate 100% correlation of thifensulfuron tolerance with glyphosate tolerance under greenhouse conditions conferred by HRA and GAT genes, respectively, at 70 gm a.i./hectare thifensulfuron and 2X glyphosate, respectively.

EXAMPLE 26: T3 STUDIES OF GLYPHOSATE-RESISTANT MAIZE PLANTS EXPRESSING GAT TRANSGENES

[0477] Maize plants expressing GAT transgenes 20-H812 (SEQ ID NO:738) and 20-16A3 (SEQ ED NO:638) were produced using the methods described in Example 13. Plants were scored after 10 days and leaf discoloration scores taken as described in Example 13. Specifically, plants were sprayed at V4 leaf stage. The plants were thinned to equal spacing and stand counts after application of spray treatments. Commercially available NK603 (Monsanto, St. Louis, MO) was used as a control. Resistance scores are shown in Table 18. Plant height measurements were also taken 10 days after treatment and are shown in Table 18.

Table 18. Resistance Scores at 10 days after treatment with glyphosate

	Resistance Scores (1-9 scale)				
CONSTRUCTS	# events tested	No Glyphosate Treatment Control	1X (26 oz/A) UltraMax	4X (104 oz/A) UltraMax	8X (208 oz/A) UltraMax
19900 (SEQ ID NO:738)	2	9	8.8	7.9	5.4
19902 (SEQ ID NO:638)	4	9	8.4	8.0	5.9
NK603	1	9	8.4	7.9	5.5

Table 19. Plant Height (in inches) 10 days after treatment with glyphosate

CONSTRUCTS	# events tested	No Glyphosate Treatment Control	1X (26 oz/A) UltraMax	4X (104 oz/A) UltraMax	8X(208 oz/A) UltraMax
19900 (SEQ ID NO:738)	2	18.5	16.6	16.4	16.2
19902 (SEQ ID NO:638)	4	19.5	16.7	17.0	16.2
NK603	1	20.1	17.6	17.5	17.3

COR-AAD1-00001597

EXAMPLE 27: T3 YIELD STUDIES OF GLYPHOSATE-RESISTANT MAIZE EXPRESSING GAT TRANSGENES

[0478] T3 seed from Example 15 was used to generate T3 plants for the generation of glyphosate field tolerance data on hybrids. The experiment was conducted at Viluco, Chile with four (4) replications using a split-plot design. Specifically, 3 entries were included. Two of the entries comprised maize plants expressing GAT variant transgenes 17-15H3 (SEQ ID NO:549). A glyphosate-resistant control NK603, which is commercially available from Monsanto, was the third entry. All entries were treated in the field with four different glyphosate spray treatments (0X, 4X at V4, 8X at V4, and 4X at V4 and 4X at V8) for each event. Plants were scored 10 days after treatment for plant height comparisons as described in Example 13. The T3 field spray data correlated well with the results previously obtained in the field as reported in Example 15. Specifically, all entries sprayed with 1X and 4X glyphosate were similar in height to unsprayed controls. At the higher 4X at V4 and 4X at V8 rates, the GAT entries were temporarily set back between 12 and 17 % in height and the NK603 entry was set back 6%; however, later in the season (during reproductive maturity) the height of glyphosate-treated entries was the same as in the unsprayed entries. Moreover, yields among glyphosate-treated entries were neither numerically nor statistically reduced from unsprayed entries ($LSD_{0.05}$= 11.8 bu./acre, average yield per entry = 243 bu./acre). Similar results were observed in preliminary agronomic trials with T2 plants of the same events that were planted in Johnston, IA and York, NE (data not shown).

EXAMPLE 28: T2 STUDIES OF GLYPHOSATE-RESISTANT MAIZE EXPRESSING GAT TRANSGENES

[0479] Experiments were conducted on GAT positive and GAT negative iso-lines. Maize plants expressing GAT transgenes 18-28D9b (SEQ ID NO:814), 17-15H3 (SEQ ID NO:549), 20-8H12 (SEQ ID NO:738), 20-16A3 (SEQ ID NO:638), were produced using the methods described in Example 17. T2 plants were examined. GAT positive T2 plants were sprayed at V4 with 1X (26 oz/A ULTRA MAX™). GAT negative plants were PCR sampled at V4. GAT positive plants were removed from the row. No glyphosate was applied to the GAT negative plants. Plants were thinned to create equal spacing among plants within each row. Four (4) replications were performed. Grain from five (5) ears harvested from the middle of each row was dried and weighed. As shown in Table 20 no yield reduction was detected for any of the constructs.

Table 20. Yield data.

Construct	# of events	Yield GAT positive sprayed with 1X (26 oz/A) ULTRA MAX™ at V4	YIELD GAT negative no glyphosate applied
PHP19286 (SEQ ID NO:814)	40	1.65 lbs/5 ears	1.57 lbs/5 ears
PHP19288 (SEQ ID NO:549)	40	1.64 lbs/5 ears	1.60 lbs/5 ears
PHP19900 (SEQ ID NO:738)	6	1.20 lbs/5 ears	1.23 lbs/5 ears
PHP19902 (SEQ ID NO:638)	4	1.19 lbs/5 ears	1.21 lbs/5 ears

EXAMPLE 29: AMINO ACID SUBSTRATES OF GAT POLYPEPTIDES

[0480] GAT activity of several GAT polypeptides of the present invention was evaluated with respect to a number of amino acid substrates. The GAT polypeptide, AcCoA and amino substrate were incubated in 25 mM Hepes, pH 6.8, 10% ethylene glycol in the wells of a 96-well polystyrene plate. After 30 minutes, the reactions were stopped by the addition of 30 μl of 10 mM 5, 5'-dithiobis-2-nitrobenzoate (DTNB) in 500 mM Tris, pH 7.5. After 2 minutes, absorbance was read at 412 nm in a Spectramax Plus plate reader (Molecular Devices, Sunnyvale, CA).

[0481] In addition to glyphosate, native GAT polypeptide 401 (SEQ ID NO: 6) (or B6 (SEQ ID NO: 7), in the case of phosphoserine) exhibited detectable activity with 12 amino acids. The native GAT polypeptide was about as active with L-aspartate, about 4.7 times more active with L-serine, and about 2 times more active with phospho-L-serine than with glyphosate. When compared to native GAT polypeptide, non-native GAT polypeptides 17-15H3 (SEQ ID NO: 601) and 25-8H7 (SEQ ID NO: 907) exhibited a 40-fold increase in activity with aspartate, but loss of activity with respect to serine and phosphoserine.

[0482] In addition to aspartate and serine, activity with native GAT polypeptide at 3% or more of that toward glyphosate when present at 1 mM was observed with the following L-amino acids: histidine (10%), tyrosine (18%), threonine (250%),

COR-AAD1-00001598

valine (12%), glutamate (51%), asparagine (27%), glutamine (32%), alanine (33%), glycine (21 %) and cysteine (50%). Activity with the other protein amino acids was either undetected or less than 3% that of GAT activity towards glyphosate as the substrate. No detectable activity was observed with respect to the native GAT polypeptide on the N-methyl derivatives of L-aspartate (2 mM), L-alanine (10 mM) and glycine (i.e., sarcosine, 10 mM). The percentages refer to percent activity relative to activity of the GAT polypeptide towards the substrate, glyphosate. Some of the data is shown below in Table 21.

Table 21: GAT activity with respect to Glyphosate, Aspartate and Serine

GAT variant	Substrate	$k_{cat} \pm$ SE, min^{-1}	$K_M \pm$ SE MM	k_{cat}/K_M $min^{-1}mM^{-1}$	k_{cat}/K_M
401	Glyph	5.35 ± 0.043	1.27 ± 0.0144	4.21	Fold imp.
17-15H3		1150 ± 27.6	0.251 ± 0.0041	4573	1086
25-8H7		1480 ± 65.4	0.05	29600	7031
					% of glyph
401	Aspartate	24.1	6.7	3.6	85.5
17-15H3		435 ± 11.3	2.95 ± 0.162	148	3.24
25-8H7		702 ± 12.4	4.56 ± 0.112	154	0.520
401	Serine	854	43	19.8	471
17-15H3		242 ±15.8	60.1 ± 1.68	4.04	0.0882
25-8H7		388 ± 18.5	154 ± 10.7	2.53	0.00855

EXAMPLE 30: EFFECT OF PH ON GAT ACTIVITY

[0483] The pH optima of k_{cat} and K_M for wild-type enzyme B6 (SEQ ID NO: 7) and GAT polypeptide 17-15H3 (SEQ ID NO: 601) were determined using the spectrophotometric assay described in Example 7 except that assay buffer was 50 mM Hepes and 10% ethylene glycol, titrated to a range of pH values. Protein concentrations were determined by the UV absorbance assay described in Example 19. The effect of pH on K_M and Kcat is shown in Figure 18 for clones B6 (SEQ ID NO: 7) and 17-15H3 (SEQ ID NO: 601).

[0484] While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be clear to one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention. For example, all the techniques, methods, compositions, apparatus and systems described above may be used in various combinations. The invention is intended to include all methods and reagents described herein, as well as all polynucleotides, polypeptides, cells, organisms, plants, crops, etc., that are the products of these novel methods and reagents.

[0485] All publications, patents, patent applications, or other documents cited in this application are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication, patent, patent application, or other document were individually indicated to be incorporated by reference for all purposes.

FURTHER EMBODIMENTS OF THE INVENTION:

[0486]

1. An isolated or recombinant polynucleotide comprising:

(a) a nucleotide sequence encoding an amino acid sequence that can be optimally aligned with a sequence selected from the group consisting of SEQ ID N0:300, SEQ ID N0:445 and SEQ ID NO:457 to generate a similarity score of at least 460, using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1,
wherein the amino acid sequence comprises one or more amino acid residues which meet at least one restriction selected from the group consisting of:

(i) at positions 18 and 38, a Z5 amino acid residue;

COR-AAD1-00001599

(ii) at position 62, a Z1 amino acid residue;
(iii) at position 124, a Z6 amino acid residue; and
(iv) at position 144, a Z2 amino acid residue,
wherein:

Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V;
Z2 is an amino acid residue selected from the group consisting of F, W, and Y;
Z5 is an amino acid residue selected from the group consisting of D and E;
Z6 is an amino acid residue selected from the group consisting of C, G, and P; or

(b) a complementary nucleotide sequence thereof.

2. The isolated or recombinant polynucleotide of item 1, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 2, 4, 15, 19, 26, 28, 31, 45, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 123, 129, 139, and/or 145 the amino acid residue is B1; and
(b) at positions 3, 5, 8, 10, 11, 14, 17, 24, 27, 32, 37, 47, 48, 49, 52, 57, 58, 61, 63, 68, 69, 79, 80, 82, 83, 89, 92, 100, 101, 104, 119, 120, 125, 126, 128, 131, and/or 143 the amino acid residue is B2;
wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

3. The isolated or recombinant polynucleotide of item 1, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at positions 2, 4, 15, 19, 26, 28, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 129, 139, and/or 145 the amino acid residue is Z1;
(b) at positions 31 and/or 45 the amino acid residue is Z2;
(c) at position 8 the amino acid residue is Z3;
(d) at position 89 the amino acid residue is Z3 or Z6;
(e) at positions 82, 92, 101 and/or 120 the amino acid residue is Z4;
(f) at positions 3, 11, 27 and/or 79 the amino acid residue is Z5;
(g) at position 18 the amino acid residue is Z4 or Z5;
(h) at position 123 the amino acid residue is Z1 or Z2;
(i) at positions 12, 33, 35, 39, 53, 59, 112, 132, 135, 140, and/or 146 the amino acid residue is Z1 or Z3;
(j) at position 30 the amino acid residue is Z1;
(k) at position 6 the amino acid residue is Z6;
(l) at position 81 the amino acid residue is Z2 or Z4;
(m) at position 113 the amino acid residue is Z3;
(n) at position 138 the amino acid residue is Z4;
(o) at position 142 the amino acid residue is Z2;
(p) at positions 57 and/or 126 the amino acid residue is Z3 or Z4;
(q) at position 5, 17, and 61 the amino acid residue is Z4;
(r) at position 24 the amino acid residue is Z3;
(s) at position 104 the amino acid residue is Z5;
(t) at positions 52, and/or 69 the amino acid residue is Z3;
(u) at positions 14 and/or 119 the amino acid residue is Z5;
(v) at positions 10, 32, 63, and/or 83 the amino acid residue is Z5 ;
(w) at positions 48 and/or 80 the amino acid residue is Z6;
(x) at position 40 the amino acid residue is Z1 or Z2 ;
(y) at position 96 the amino acid residue is Z3 or Z5;
(z) at position 65 the amino acid residue is Z3, Z4, or Z6;
(aa) at positions 84 and/or 115 the amino acid residue is Z3;
(ab) at position 93 the amino acid residue is Z4;
(ac) at position 130 the amino acid residue is Z2;
(ad) at position 58 the amino acid residue is Z3, Z4 or Z6;

COR-AAD1-00001600

(ae) at position 47 the amino acid residue is Z4 or Z6;
(af) at positions 49 and/or 100 the amino acid residue is Z3 or Z4;
(ag) at position 68 the amino acid residue is Z4 or Z5;
(ah) at position 143 the amino acid residue is Z4;
(ai) at position 131 the amino acid residue is Z5;
(aj) at positions 125 and/or 128 the amino acid residue is Z5;
(ak) at position 67 the amino acid residue is Z3 or Z4;
(al) at position 60 the amino acid residue is Z5; and
(am) at position 37 the amino acid residue is Z4 or Z6;
wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting of N, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

4. The isolated or recombinant polynucleotide of item 3, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) - (am), at least 90% conform to the amino acid residue restrictions specified in (a) - (am).

5. The isolated or recombinant polynucleotide of any of items 1-4, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 13, 20, 36, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and
(b) at positions 16, 21, 22, 23, 25, 29, 34, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99, 102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2;
wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

6. The isolated or recombinant polynucleotide of any of items 1-4 wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 13, 20, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and
(b) at positions 16, 21, 22, 23, 25, 29, 34, 36, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99,102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2;
wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

7. The isolated or recombinant polynucleotide of any of items 1-4 wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 20, 42, 50, 72, 75, 76, 78, 94, 98, 110, 121, and/or 141 the amino acid residue is Z1;
(b) at positions 13, 46, 56, 70, 107, 117, and/or 118 the amino acid residue is Z2;
(c) at positions 23, 55, 71, 77, 88, and/or 109 the amino acid residue is Z3;
(d) at positions 16, 21, 41, 73, 85, 99, and/or 111 the amino acid residue is Z4;
(e) at positions 34 and/or 95 the amino acid residue is Z5;
(f) at position 22, 25, 29, 43, 44, 66, 74, 87, 102, 108, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is Z6;

wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting of N, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

8. The isolated or recombinant polynucleotide of item 7, wherein the amino acid sequence comprises at position 36, an amino acid residue selected from the group consisting of Z1 and Z3.

9. The isolated or recombinant polynucleotide of item 7 or item 8, wherein the amino acid sequence comprises at

COR-AAD1-00001601

position 64, an amino acid residue selected from the group consisting of Z1 and Z2.

10. The isolated or recombinant polynucleotide of any items 1-9 wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at position 2 the amino acid residue is I or L;
(b) at position 3 the amino acid residue is E;
(c) at position 4 the amino acid residue is V or I;
(d) at position 5 the amino acid residue is K;
(e) at position 6 the amino acid residue is P;
(f) at position 8 the amino acid residue is N;
(g) at position 10 the amino acid residue is E;
(h) at position 11 the amino acid residue is D or E;
(i) at position 12 the amino acid residue is T;
(j) at position 14 the amino acid residue is E or D;
(k) at position 15 the amino acid residue is L;
(l) at position 17 the amino acid residue is H;
(m) at position 18 the amino acid residue is R, E or K;
(n) at position 19 the amino acid residue is I or V;
(o) at position 24 the amino acid residue is Q;
(p) at position 26 the amino acid residue is M, L, V or I;
(q) at position 27 the amino acid residue is E;
(r) at position 28 the amino acid residue is A or V;
(s) at position 30 the amino acid residue is M;
(t) at position 31 the amino acid residue is Y or F;
(u) at position 32 the amino acid residue is E or D;
(v) at position 33 the amino acid residue is T or S;
(w) at position 35 the amino acid residue is L;
(x) at position 37 the amino acid residue is R, G, E or Q;
(y) at position 39 the amino acid residue is A or S;
(z) at position 40 the amino acid residue is F or L ;
(aa) at position 45 the amino acid residue is Y or F;
(ab) at position 47 the amino acid residue is R or G;
(ac) at position 48 the amino acid residue is G;
(ad) at position 49 the amino acid residue is K, R, or Q;
(ae) at position 51 the amino acid residue is I or V;
(af) at position 52 the amino acid residue is S;
(ag) at position 53 the amino acid residue is I or V;
(ah) at position 54 the amino acid residue is A;
(ai) at position 57 the amino acid residue is H or N;
(aj) at position 58 the amino acid residue is Q, K, R or P;
(ak) at position 59 the amino acid residue is A;
(al) at position 60 the amino acid residue is E;
(am) at position 61 the amino acid residue is H or R;
(an) at position 63 the amino acid residue is E or D;
(ao) at position 65 the amino acid residue is E, P or Q;
(ap) at position 67 the amino acid residue is Q or R;
(aq) at position 68 the amino acid residue is K or E;
(ar) at position 69 the amino acid residue is Q;
(as) at position 79 the amino acid residue is E;
(at) at position 80 the amino acid residue is G;
(au) at position 81 the amino acid residue is Y, H or F;
(av) at position 82 the amino acid residue is R;
(aw) at position 83 the amino acid residue is E or D;
(ax) at position 84 the amino acid residue is Q;
(ay) at position 86 the amino acid residue is A;
(az) at position 89 the amino acid residue is G, T or S;
(ba) at position 90 the amino acid residue is L;

COR-AAD1-00001602

(bb) at position 91 the amino acid residue is L, I or V;
(bc) at position 92 the amino acid residue is R or K;
(bd) at position 93 the amino acid residue is H;
(be) at position 96 the amino acid residue is E or Q;
(bf) at position 97 the amino acid residue is I;
(bg) at position 100 the amino acid residue is K or N ;
(bh) at position 101 the amino acid residue is K or R;
(bi) at position 103 the amino acid residue is A or V;
(bj) at position 104 the amino acid residue is D;
(bk) at position 105 the amino acid residue is M, L or I;
(bl) at position 106 the amino acid residue is L;
(bm) at position 112 the amino acid residue is T or A;
(bn) at position 113 the amino acid residue is S or T;
(bo) at position 114 the amino acid residue is A;
(bp) at position 115 the amino acid residue is S;
(bq) at position 119 the amino acid residue is K or R;
(br) at position 120 the amino acid residue is K or R;
(bs) at position 123 the amino acid residue is F or L;
(bt) at position 125 the amino acid residue is E;
(bu) at position 126 the amino acid residue is Q or H;
(bv) at position 128 the amino acid residue is E or D;
(bw) at position 129 the amino acid residue is V or I;
(bx) at position 130 the amino acid residue is F;
(by) at position 131 the amino acid residue is D or E;
(bz) at position 132 the amino acid residue is T;
(ca) at position 135 the amino acid residue is V;
(cb) at position 138 the amino acid residue is H;
(cc) at position 139 the amino acid residue is I;
(cd) at position 140 the amino acid residue is L or M;
(ce) at position 142 the amino acid residue is Y;
(cf) at position 143 the amino acid residue is K or R;
(cg) at position 145 the amino acid residue is L or I; and
(ch) at position 146 the amino acid residue is T.

11. The isolated or recombinant polynucleotide of item 10, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) - (ch), at least 90% conform to the amino acid residue restrictions specified in (a) - (ch).

12. The isolated or recombinant polynucleotide of item 1, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at positions 9, 76, 94 and 110 the amino acid residue is A;
(b) at positions 29 and 108 the amino acid residue is C;
(c) at position 34 the amino acid residue is D;
(d) at position 95 the amino acid residue is E;
(e) at position 56 the amino acid residue is F;
(f) at positions 43, 44, 66, 74, 87, 102, 116, 122, 127 and 136 the amino acid residue is G;
(g) at position 41 the amino acid residue is H;
(h) at position 7 the amino acid residue is I;
(i) at position 85 the amino acid residue is K;
(j) at positions 20, 42, 50, 78 and 121 the amino acid residue is L;
(k) at positions 1 and 141 the amino acid residue is M;
(l) at positions 23 and 109 the amino acid residue is N;
(m) at positions 22, 25, 133, 134 and 137 the amino acid residue is P;
(n) at position 71 the amino acid residue is Q;
(o) at positions 16, 21, 73, 99 and 111 the amino acid residue is R;
(p) at position 55 the amino acid residue is S;
(q) at position 77 the amino acid residue is T;

COR-AAD1-00001603

(r) at position 107 the amino acid residue is W; and
(s) at position 13, 46, 70 and 118 the amino acid residue is Y.

13. The isolated or recombinant polynucleotide of any of items 1-12, wherein the amino acid sequence comprises an amino acid residue selected from the group consisting of:

(a) at position 36 the amino acid residue is M, L, or T;
(b) at position 72 the amino acid residue is L or I;
(c) at position 75 the amino acid residue is M or V;
(d) at position 64 the amino acid residue is L, I, or F;
(e) at position 88 the amino acid residue is T or S;
(f) at position 117 the amino acid residue is Y or F.

14. The isolated or recombinant polynucleotide of any of items 1-13, wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

(a) at position 14 the amino acid residue is D;
(b) at position 18 the amino acid residue is E;
(c) at position 26 the amino acid residue is M or V;
(e) at position 30 the amino acid residue is I;
(f) at position 32 the amino acid residue is D;
(g) at position 36 the amino acid residue is M or T;
(i) at position 37 the amino acid residue is C;
(j) at position 38 the amino acid residue is D;
(j) at position 53 the amino acid residue is V;
(k) at position 58 the amino acid residue is R;
(1) at position 61 the amino acid residue is R;
(m) at position 62 the amino acid residue is L;
(n) at position 64 the amino acid residue is I or F;
(o) at position 65 the amino acid residue is P;
(p) at position 72 the amino acid residue is I;
(q) at position 75 the amino acid residue is V;
(r) at position 88 the amino acid residue is T;
(s) at position 89 the amino acid residue is G;
(t) at position 91 the amino acid residue is L;
(u) at position 98 the amino acid residue is I;
(v) at position 105 the amino acid residue I;
(w) at position 112 the amino acid residue is A;
(x) at position 124 the amino acid residue is G or C;
(y) at position 128 the amino acid residue is D;
(z) at position 140 the amino acid residue is M;
(aa) at position 143 the amino acid residue is R; and
(ab) at position 144 the amino acid residue is W.

15. The isolated or recombinant polynucleotide of item 14, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) through (ab), at least 80% conform to the amino acid residue restrictions specified in (a) through (ab).

16. The isolated or recombinant polynucleotide of any of items 1-15, wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

(a) at position 41 the amino acid residue is H;
(b) at position 138 the amino acid residue is H;
(c) at position 34 the amino acid residue is N; and
(d) at position 55 the amino acid residue is S.

17. An isolated or recombinant polynucleotide comprising a nucleotide sequence selected from the group consisting of:

COR-AAD1-00001604

(a) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:577;
(b) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:578;
(c) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:621;
(d) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:579;
(e) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:602;
(f) a nucleotide sequence encoding an amino acid sequence that is at least 95% identical to SEQ ID NO:697;
(g) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:721;
(h) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:613;
(i) a nucleotide sequence encoding an amino acid sequence that is at least 89% identical to SEQ ID NO:677;
(j) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:584;
(k) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:707;
(l) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:616;
(m) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:612;
(n) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:590;
(o) a nucleotide sequence that is complementary to any one of (a) through (n).

18. The isolated or recombinant polynucleotide of item 17, wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

(i) at positions 18 and 38, a Z5 amino acid residue;
(ii) at position 62, a Z1 amino acid residue;
(iii) at position 124, a Z6 amino acid residue; and
(iv) at position 144, a Z2 amino acid residue,
wherein:

Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V;
Z2 is an amino acid residue selected from the group consisting of F, W, and Y;
Z5 is an amino acid residue selected from the group consisting of D and E;
Z6 is an amino acid residue selected from the group consisting of C, G, and P.

19. The isolated or recombinant polynucleotide of item 17, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 2, 4, 15, 19, 26, 28, 31, 45, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 123, 129, 139, and/or 145 the amino acid residue is B1; and
(b) at positions 3, 5, 8, 10, 11, 14, 17, 24, 27, 32, 37, 47, 48, 49, 52, 57, 58, 61, 63, 68, 69, 79, 80, 82, 83, 89, 92, 100, 101, 104, 119, 120, 125, 126, 128, 131, and/or 143 the amino acid residue is B2;

wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

20. The isolated or recombinant polynucleotide of item 17, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at positions 2, 4, 15, 19, 26, 28, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 129, 139, and/or 145 the amino acid residue is Z1;
(b) at positions 31 and/or 45 the amino acid residue is Z2;
(c) at position 8 the amino acid residue is Z3;
(d) at position 89 the amino acid residue is Z3 or Z6;
(e) at positions 82, 92, 101 and/or 120 the amino acid residue is Z4;
(f) at positions 3, 11, 27 and/or 79 the amino acid residue is Z5;
(g) at position 18 the amino acid residue is Z4 or Z5;
(h) at position 123 the amino acid residue is Z1 or Z2;
(i) at positions 12, 33, 35, 39, 53, 59, 112, 132, 135, 140, and/or 146 the amino acid residue is Z1 or Z3;
(j) at position 30 the amino acid residue is Z1;
(k) at position 6 the amino acid residue is Z6;
(l) at position 81 the amino acid residue is Z2 or Z4;

COR-AAD1-00001605

(m) at position 113 the amino acid residue is Z3;
(n) at position 138 the amino acid residue is Z4;
(o) at position 142 the amino acid residue is Z2;
(p) at positions 57 and/or 126 the amino acid residue is Z3 or Z4;
(q) at position 5, 17, and 61 the amino acid residue is Z4;
(r) at position 24 the amino acid residue is Z3;
(s) at position 104 the amino acid residue is Z5;
(t) at positions 52, and/or 69 the amino acid residue is Z3;
(u) at positions 14 and/or 119 the amino acid residue is Z5;
(v) at positions 10, 32, 63, and/or 83 the amino acid residue is Z5 ;
(w) at positions 48 and/or 80 the amino acid residue is Z6;
(x) at position 40 the amino acid residue is Z1 or Z2 ;
(y) at position 96 the amino acid residue is Z3 or Z5;
(z) at position 65 the amino acid residue is Z3, Z4, or Z6;
(aa) at positions 84 and/or 115 the amino acid residue is Z3;
(ab) at position 93 the amino acid residue is Z4;
(ac) at position 130 the amino acid residue is Z2;
(ad) at position 58 the amino acid residue is Z3, Z4 or Z6;
(ae) at position 47 the amino acid residue is Z4 or Z6;
(af) at positions 49 and/or 100 the amino acid residue is Z3 or Z4;
(ag) at position 68 the amino acid residue is Z4 or Z5;
(ah) at position 143 the amino acid residue is Z4;
(ai) at position 131 the amino acid residue is Z5;
(aj) at positions 125 and/or 128 the amino acid residue is Z5;
(ak) at position 67 the amino acid residue is Z3 or Z4;
(al) at position 60 the amino acid residue is Z5; and
(am) at position 37 the amino acid residue is Z4 or Z6;

wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting of N, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

21. The isolated or recombinant polynucleotide of item 20, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) - (am), at least 90% conform to the amino acid residue restrictions specified in (a) - (am).

22. The isolated or recombinant polynucleotide of any of items 17-21, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 13, 20, 36, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and
(b) at positions 16, 21, 22, 23, 25, 29, 34, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99, 102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2;

wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

23. The isolated or recombinant polynucleotide of any of items 17-22 wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 13, 20, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and
(b) at positions 16, 21, 22, 23, 25, 29, 34, 36, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99,102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2;

wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

COR-AAD1-00001606

24. The isolated or recombinant polynucleotide of any of items 17-22 wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 20, 42, 50, 72, 75, 76, 78, 94, 98, 110, 121, and/or 141 the amino acid residue is Z1;
(b) at positions 13, 46, 56, 70, 107, 117, and/or 118 the amino acid residue is Z2;
(c) at positions 23, 55, 71, 77, 88, and/or 109 the amino acid residue is Z3;
(d) at positions 16, 21, 41, 73, 85, 99, and/or 111 the amino acid residue is Z4;
(e) at positions 34 and/or 95 the amino acid residue is Z5;
(f) at position 22, 25, 29, 43, 44, 66, 74, 87, 102, 108, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is Z6;

wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting ofN, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

25. The isolated or recombinant polynucleotide of item 24, wherein the amino acid sequence comprises at position 36, an amino acid residue selected from the group consisting of Z1 and Z3.

26. The isolated or recombinant polynucleotide of item 24 or item 25, wherein the amino acid sequence comprises at position 64, an amino acid residue selected from the group consisting of Z1 and Z2.

27. The isolated or recombinant polynucleotide of any items 17-26 wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at position 2 the amino acid residue is I or L;
(b) at position 3 the amino acid residue is E;
(c) at position 4 the amino acid residue is V or I;
(d) at position 5 the amino acid residue is K;
(e) at position 6 the amino acid residue is P;
(f) at position 8 the amino acid residue is N;
(g) at position 10 the amino acid residue is E;
(h) at position 11 the amino acid residue is D or E;
(i) at position 12 the amino acid residue is T;
(j) at position 14 the amino acid residue is E or D;
(k) at position 15 the amino acid residue is L;
(1) at position 17 the amino acid residue is H;
(m) at position 18 the amino acid residue is R, E or K;
(n) at position 19 the amino acid residue is I or V;
(o) at position 24 the amino acid residue is Q;
(p) at position 26 the amino acid residue is M, L, V or I;
(q) at position 27 the amino acid residue is E;
(r) at position 28 the amino acid residue is A or V;
(s) at position 30 the amino acid residue is M;
(t) at position 31 the amino acid residue is Y or F;
(u) at position 32 the amino acid residue is E or D;
(v) at position 33 the amino acid residue is T or S;
(w) at position 35 the amino acid residue is L;
(x) at position 37 the amino acid residue is R, G, E or Q;
(y) at position 39 the amino acid residue is A or S;
(z) at position 40 the amino acid residue is F or L ;
(aa) at position 45 the amino acid residue is Y or F;
(ab) at position 47 the amino acid residue is R or G;
(ac) at position 48 the amino acid residue is G;
(ad) at position 49 the amino acid residue is K, R, or Q;
(ae) at position 51 the amino acid residue is I or V;
(af) at position 52 the amino acid residue is S;
(ag) at position 53 the amino acid residue is I or V;

COR-AAD1-00001607

(ah) at position 54 the amino acid residue is A;
(ai) at position 57 the amino acid residue is H or N;
(aj) at position 58 the amino acid residue is Q, K, R or P;
(ak) at position 59 the amino acid residue is A;
(al) at position 60 the amino acid residue is E;
(am) at position 61 the amino acid residue is H or R;
(an) at position 63 the amino acid residue is E or D;
(ao) at position 65 the amino acid residue is E, P or Q;
(ap) at position 67 the amino acid residue is Q or R;
(aq) at position 68 the amino acid residue is K or E;
(ar) at position 69 the amino acid residue is Q;
(as) at position 79 the amino acid residue is E;
(at) at position 80 the amino acid residue is G;
(au) at position 81 the amino acid residue is Y, H or F;
(av) at position 82 the amino acid residue is R;
(aw) at position 83 the amino acid residue is E or D;
(ax) at position 84 the amino acid residue is Q;
(ay) at position 86 the amino acid residue is A;
(az) at position 89 the amino acid residue is G, T or S;
(ba) at position 90 the amino acid residue is L;
(bb) at position 91 the amino acid residue is L, I or V;
(bc) at position 92 the amino acid residue is R or K;
(bd) at position 93 the amino acid residue is H;
(be) at position 96 the amino acid residue is E or Q;
(bf) at position 97 the amino acid residue is I;
(bg) at position 100 the amino acid residue is K or N ;
(bh) at position 101 the amino acid residue is K or R;
(bi) at position 103 the amino acid residue is A or V;
(bj) at position 104 the amino acid residue is D;
(bk) at position 105 the amino acid residue is M, L or I;
(bl) at position 106 the amino acid residue is L;
(bm) at position 112 the amino acid residue is T or A;
(bn) at position 113 the amino acid residue is S or T;
(bo) at position 114 the amino acid residue is A;
(bp) at position 115 the amino acid residue is S;
(bq) at position 119 the amino acid residue is K or R;
(br) at position 120 the amino acid residue is K or R;
(bs) at position 123 the amino acid residue is F or L;
(bt) at position 125 the amino acid residue is E;
(bu) at position 126 the amino acid residue is Q or H;
(bv) at position 128 the amino acid residue is E or D;
(bw) at position 129 the amino acid residue is V or I;
(bx) at position 130 the amino acid residue is F;
(by) at position 131 the amino acid residue is D or E;
(bz) at position 132 the amino acid residue is T;
(ca) at position 135 the amino acid residue is V;
(cb) at position 138 the amino acid residue is H;
(cc) at position 139 the amino acid residue is I;
(cd) at position 140 the amino acid residue is L or M;
(ce) at position 142 the amino acid residue is Y;
(cf) at position 143 the amino acid residue is K or R;
(cg) at position 145 the amino acid residue is L or I; and
(ch) at position 146 the amino acid residue is T.

28. The isolated or recombinant polynucleotide of item 27, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) - (ch), at least 90% conform to the amino acid residue restrictions specified in (a) - (ch).

COR-AAD1-00001608

29. The isolated or recombinant polynucleotide of item 17, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at positions 9, 76, 94 and 110 the amino acid residue is A;
(b) at positions 29 and 108 the amino acid residue is C;
(c) at position 34 the amino acid residue is D;
(d) at position 95 the amino acid residue is E;
(e) at position 56 the amino acid residue is F;
(f) at positions 43, 44, 66, 74, 87, 102, 116, 122, 127 and 136 the amino acid residue is G;
(g) at position 41 the amino acid residue is H;
(h) at position 7 the amino acid residue is I;
(i) at position 85 the amino acid residue is K;
(j) at positions 20, 42, 50, 78 and 121 the amino acid residue is L;
(k) at positions 1 and 141 the amino acid residue is M;
(l) at positions 23 and 109 the amino acid residue is N;
(m) at positions 22, 25, 133, 134 and 137 the amino acid residue is P;
(n) at position 71 the amino acid residue is Q;
(o) at positions 16, 21, 73, 99 and 111 the amino acid residue is R;
(p) at position 55 the amino acid residue is S;
(q) at position 77 the amino acid residue is T;
(r) at position 107 the amino acid residue is W; and
(s) at position 13, 46, 70 and 118 the amino acid residue is Y.

30. The isolated or recombinant polynucleotide of any of items 1-29, wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

(a) at position 14 the amino acid residue is D;
(b) at position 18 the amino acid residue is E;
(c) at position 26 the amino acid residue is M or V;
(e) at position 30 the amino acid residue is I;
(f) at position 32 the amino acid residue is D;
(g) at position 36 the amino acid residue is M or T;
(i) at position 37 the amino acid residue is C;
(j) at position 38 the amino acid residue is D;
(j) at position 53 the amino acid residue is V;
(k) at position 58 the amino acid residue is R;
(l) at position 61 the amino acid residue is R;
(m) at position 62 the amino acid residue is L;
(n) at position 64 the amino acid residue is I or F;
(o) at position 65 the amino acid residue is P;
(p) at position 72 the amino acid residue is I;
(q) at position 75 the amino acid residue is V;
(r) at position 88 the amino acid residue is T;
(s) at position 89 the amino acid residue is G;
(t) at position 91 the amino acid residue is L;
(u) at position 98 the amino acid residue is I;
(v) at position 105 the amino acid residue I;
(w) at position 112 the amino acid residue is A;
(x) at position 124 the amino acid residue is G or C;
(y) at position 128 the amino acid residue is D;
(z) at position 140 the amino acid residue is M;
(aa) at position 143 the amino acid residue is R; and
(ab) at position 144 the amino acid residue is W.

31. The isolated or recombinant polynucleotide of item 30, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) through (ab), at least 80% conform to the amino acid residue restrictions specified in (a) through (ab).

COR-AAD1-00001609

32. The isolated or recombinant polynucleotide of any of items 17-31, wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

(a) at position 41 the amino acid residue is H;
(b) at position 138 the amino acid residue is H;
(c) at position 34 the amino acid residue is N; and
(e) at position 55 the amino acid residue is S.

33. An isolated or recombinant polynucleotide encoding a protein having glyphosate-N-acetyltransferase activity, said polynucleotide selected from the group consisting of: SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952.

34. The isolated or recombinant polynucleotide of item 1, wherein a polypeptide encoded by said polynucleotide has glyphosate-N-acetyltransferase activity.

35. The isolated or recombinant polynucleotide of item 1, wherein a polypeptide encoded by said polynucleotide has glyphosate-N-acyltransferase activity.

36. The isolated or recombinant polynucleotide of item 34, wherein the polypeptide catalyzes the acetylation of glyphosate with a k_{cat}/K_m of at least 10 mM^{-1} min^{-1} for glyphosate.

37. The isolated or recombinant polynucleotide of item 34, wherein the polypeptide catalyzes the acetylation of aminomethylphosphonic acid.

38. A nucleic acid construct comprising the isolated or recombinant polynucleotide of item 1.

39. The nucleic acid construct of item 38, comprising a promoter operably linked to the isolated or recombinant polynucleotide of item 1, where the promoter is heterologous with respect to said polynucleotide and effective to cause sufficient expression of the encoded polypeptide to enhance the glyphosate tolerance of a plant cell transformed with said nucleic acid construct.

40. The nucleic acid construct of item 38, wherein the isolated or recombinant polynucleotide sequence of item 1 functions as a selectable marker.

41. The nucleic acid construct of item 38, wherein the construct is a vector.

42. The vector of item 41 further comprising a second polynucleotide sequence encoding a second polypeptide that confers a detectable phenotypic trait upon a cell or organism expressing the second polypeptide at an effective level.

43. The vector of item 42, wherein the detectable phenotypic trait functions as selectable marker.

44. The vector of item 42, wherein the detectable phenotypic trait is selected from the group consisting of herbicide resistance and a visible marker.

45. The vector of item 41, wherein the vector further comprises a T-DNA sequence.

46. The vector of item 41, wherein the isolated or recombinant polynucleotide is operably linked to a regulatory sequence.

COR-AAD1-00001610

47. The vector of item 41, wherein the vector is a plant transformation vector.

48. A cell comprising at least one isolated or recombinant polynucleotide of item 1, wherein the polynucleotide is heterologous to the cell.

49. The cell of item 48, wherein the isolated or recombinant polynucleotide is operably linked to a regulatory sequence.

50. A cell transduced by the vector of item 41.

51. The cell of item 49 or 50, wherein the cell is a transgenic plant cell.

52. The transgenic plant cell of item 51, wherein the transgenic plant cell expresses an exogenous polypeptide with glyphosate-N-acetyltransferase activity.

53. The transgenic plant cell of item 51, wherein the transgenic plant cell expresses an exogenous polypeptide with glyphosate-N-acyltransferase activity.

54. A transgenic plant or transgenic plant explant comprising the cell of item 52.

55. A transgenic plant or transgenic plant explant comprising the cell of item 53.

56. The transgenic plant or transgenic plant explant of item 54, wherein said plant or plant explant expresses a polypeptide with glyphosate-N-acetyltransferase activity.

57. The transgenic plant or transgenic plant explant of item 55, wherein said plant or plant explant expresses a polypeptide with glyphosate-N-acyltransferase activity.

58. The transgenic plant or transgenic plant explant of item 56, wherein said plant or plant explant is a crop plant selected from the group of genera consisting of: *Eleusine, Lollium, Bambusa, Brassica, Dactylis, Sorghum, Pennisetum, Zea, Oryza, Triticum, Secale, Avena, Hordeum, Saccharum, Coix, Glycine* and *Gossypium.*

59. The transgenic plant or transgenic plant explant of item 57, wherein said plant or plant explant is a crop plant selected from the group of genera consisting of: *Eleusine, Lollium, Bambusa, Brassica, Dactylis, Sorghum, Pennisetum, Zea, Oryza, Triticum, Secale, Avena, Hordeum, Saccharum, Coix, Glycine* and *Gossypium.*

60. The transgenic plant or transgenic plant explant of item 56, wherein said plant or plant explant is *Arabidopsis.*

61. The transgenic plant or transgenic plant explant of item 57, wherein said plant or plant explant is *Arabidopsis.*

62. The transgenic plant or transgenic plant explant of item 56, wherein said plant or plant explant is *Gossypium.*

63. The transgenic plant or transgenic plant explant of item 57, wherein said plant or plant explant is *Gossypium.*

64. The transgenic plant or transgenic plant explant of item 56, wherein said plant or plant explant exhibits enhanced resistance to glyphosate as compared to a wild type plant of the same species, strain or cultivar.

65. The transgenic plant or transgenic plant explant of item 57, wherein said plant or plant explant exhibits enhanced resistance to glyphosate as compared to a wild type plant of the same species, strain or cultivar.

66. A seed produced by the plant of item 56.

67. A seed produced by the plant of item 57.

68. A transgenic plant which contains a heterologous gene, said gene selected from the group consisting of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694,

COR-AAD1-00001611

696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952 which encodes a glyphosate-N-acetyltransferase having a k_{cat}/K_m of at least 10 mM^{-1} min^{-1} for glyphosate, wherein the plant exhibits tolerance to glyphosate applied at a level effective to inhibit the growth of the same plant lacking the heterologous gene, without significant yield reduction due to herbicide application.

69. The transgenic plant of item 68, wherein the glyphosate-N-acetyltransferase catalyzes the acetylation of aminomethylphosphonic acid.

70. An isolated or recombinant polypeptide comprising:

(a) an amino acid sequence that can be optimally aligned with a sequence selected from the group consisting of SEQ ID NO:300, SEQ ID NO:445 and SEQ ID NO:457 to generate a similarity score of at least 460, using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1,
wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

(i) at positions 18 and 38, a Z5 amino acid residue;
(ii) at position 62, a Z1 amino acid residue;
(iii) at position 124, a Z6 amino acid residue; and
(iv) at position 144, a Z2 amino acid residue,
wherein:

Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V;
Z2 is an amino acid residue selected from the group consisting of F, W, and Y;
Z5 is an amino acid residue selected from the group consisting of D and E;
Z6 is an amino acid residue selected from the group consisting of C, G, and P; or

(b) a complementary nucleotide sequence thereof.

71. The isolated or recombinant polypeptide of item 70, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 2, 4, 15, 19, 26, 28, 31, 45, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 123, 129, 139, and/or 145 the amino acid residue is B1; and
(b) at positions 3, 5, 8, 10, 11, 14, 17, 24, 27, 32, 37, 47, 48, 49, 52, 57, 58, 61, 63, 68, 69, 79, 80, 82, 83, 89, 92, 100, 101, 104, 119, 120, 125, 126, 128, 131, and/or 143 the amino acid residue is B2;
wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

72. The isolated or recombinant polypeptide of item 70, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at positions 2, 4, 15, 19, 26, 28, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 129, 139, and/or 145 the amino acid residue is Z1;
(b) at positions 31 and/or 45 the amino acid residue is Z2;
(c) at position 8 the amino acid residue is Z3;
(d) at position 89 the amino acid residue is Z3 or Z6;
(e) at positions 82, 92, 101 and/or 120 the amino acid residue is Z4;
(f) at positions 3, 11, 27 and/or 79 the amino acid residue is Z5;
(g) at position 18 the amino acid residue is Z4 or Z5;
(h) at position 123 the amino acid residue is Z1 or Z2;
(i) at positions 12, 33, 35, 39, 53, 59, 112, 132, 135, 140, and/or 146 the amino acid residue is Z1 or Z3;

COR-AAD1-00001612

(j) at position 30 the amino acid residue is Z1;
(k) at position 6 the amino acid residue is Z6;
(l) at position 81 the amino acid residue is Z2 or Z4;
(m) at position 113 the amino acid residue is Z3;
(n) at position 138 the amino acid residue is Z4;
(o) at position 142 the amino acid residue is Z2;
(p) at positions 57 and/or 126 the amino acid residue is Z3 or Z4;
(q) at position 5, 17, and 61 the amino acid residue is Z4;
(r) at position 24 the amino acid residue is Z3;
(s) at position 104 the amino acid residue is Z5;
(t) at positions 52, and/or 69 the amino acid residue is Z3;
(u) at positions 14 and/or 119 the amino acid residue is Z5;
(v) at positions 10, 32, 63, and/or 83 the amino acid residue is Z5 ;
(w) at positions 48 and/or 80 the amino acid residue is Z6;
(x) at position 40 the amino acid residue is Z1 or Z2 ;
(y) at position 96 the amino acid residue is Z3 or Z5;
(z) at position 65 the amino acid residue is Z3, Z4, or Z6;
(aa) at positions 84 and/or 115 the amino acid residue is Z3;
(ab) at position 93 the amino acid residue is Z4;
(ac) at position 130 the amino acid residue is Z2;
(ad) at position 58 the amino acid residue is Z3, Z4 or Z6;
(ae) at position 47 the amino acid residue is Z4 or Z6;
(af) at positions 49 and/or 100 the amino acid residue is Z3 or Z4;
(ag) at position 68 the amino acid residue is Z4 or Z5;
(ah) at position 143 the amino acid residue is Z4;
(ai) at position 131 the amino acid residue is Z5;
(aj) at positions 125 and/or 128 the amino acid residue is Z5;
(ak) at position 67 the amino acid residue is Z3 or Z4;
(al) at position 60 the amino acid residue is Z5; and
(am) at position 37 the amino acid residue is Z4 or Z6;
wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting ofN, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

73. The isolated or recombinant polypeptide of item 72, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) - (am), at least 90% conform to the amino acid residue restrictions specified in (a) - (am).

74. The isolated or recombinant polypeptide of any of items 70-73, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 13, 20, 36, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and
(b) at positions 16, 21, 22, 23, 25, 29, 34, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99, 102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2;
wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

75. The isolated or recombinant polypeptide of any of items 70-73, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 13, 20, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1, and
(b) at positions 16, 21, 22, 23, 25, 29, 34, 36, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99,102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2;
wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

COR-AAD1-00001613

76. The isolated or recombinant polypeptide of any of items 70-73, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 20, 42, 50, 72, 75, 76, 78, 94, 98, 110, 121, and/or 141 the amino acid residue is Z1;
(b) at positions 13, 46, 56, 70, 107, 117, and/or 118 the amino acid residue is Z2;
(c) at positions 23, 55, 71, 77, 88, and/or 109 the amino acid residue is Z3;
(d) at positions 16, 21, 41, 73, 85, 99, and/or 111 the amino acid residue is Z4;
(e) at positions 34 and/or 95 the amino acid residue is Z5;
(f) at position 22, 25, 29, 43, 44, 66, 74, 87, 102, 108, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is Z6;
wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting of N, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

77. The isolated or recombinant polypeptide of item 76, wherein the amino acid sequence comprises at position 36, an amino acid residue selected from the group consisting of Z1 and Z3.

78. The isolated or recombinant polypeptide of item 76 or item 77, wherein the amino acid sequence comprises at position 64, an amino acid residue selected from the group consisting of Z1 and Z2.

79. The isolated or recombinant polypeptide of any items 70-78 wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at position 2 the amino acid residue is I or L;
(b) at position 3 the amino acid residue is E;
(c) at position 4 the amino acid residue is V or I;
(d) at position 5 the amino acid residue is K;
(e) at position 6 the amino acid residue is P;
(f) at position 8 the amino acid residue is N;
(g) at position 10 the amino acid residue is E;
(h) at position 11 the amino acid residue is D or E;
(i) at position 12 the amino acid residue is T;
(j) at position 14 the amino acid residue is E or D;
(k) at position 15 the amino acid residue is L;
(1) at position 17 the amino acid residue is H;
(m) at position 18 the amino acid residue is R, E or K;
(n) at position 19 the amino acid residue is I or V;
(o) at position 24 the amino acid residue is Q;
(p) at position 26 the amino acid residue is M, L, V or I;
(q) at position 27 the amino acid residue is E;
(r) at position 28 the amino acid residue is A or V;
(s) at position 30 the amino acid residue is M;
(t) at position 31 the amino acid residue is Y or F;
(u) at position 32 the amino acid residue is E or D;
(v) at position 33 the amino acid residue is T or S;
(w) at position 35 the amino acid residue is L;
(x) at position 37 the amino acid residue is R, G, E or Q;
(y) at position 39 the amino acid residue is A or S;
(z) at position 40 the amino acid residue is F or L ;
(aa) at position 45 the amino acid residue is Y or F;
(ab) at position 47 the amino acid residue is R or G; (ac) at position 48 the amino acid residue is G;
(ad) at position 49 the amino acid residue is K, R, or Q; (ae) at position 51 the amino acid residue is I or V; (af) at position 52 the amino acid residue is S;
(ag) at position 53 the amino acid residue is I or V; (ah) at position 54 the amino acid residue is A;
(ai) at position 57 the amino acid residue is H or N; (aj) at position 58 the amino acid residue is Q, K, R or P;
(ak) at position 59 the amino acid residue is A;
(al) at position 60 the amino acid residue is E;

COR-AAD1-00001614

(am) at position 61 the amino acid residue is H or R; (an) at position 63 the amino acid residue is E or D; (ao) at position 65 the amino acid residue is E, P or Q; (ap) at position 67 the amino acid residue is Q or R; (aq) at position 68 the amino acid residue is K or E; (ar) at position 69 the amino acid residue is Q;
(as) at position 79 the amino acid residue is E;
(at) at position 80 the amino acid residue is G;
(au) at position 81 the amino acid residue is Y, H or F; (av) at position 82 the amino acid residue is R;
(aw) at position 83 the amino acid residue is E or D; (ax) at position 84 the amino acid residue is Q;
(ay) at position 86 the amino acid residue is A;
(az) at position 89 the amino acid residue is G, T or S; (ba) at position 90 the amino acid residue is L;
(bb) at position 91 the amino acid residue is L, I or V; (bc) at position 92 the amino acid residue is R or K; (bd) at position 93 the amino acid residue is H;
(be) at position 96 the amino acid residue is E or Q;
(bf) at position 97 the amino acid residue is I;
(bg) at position 100 the amino acid residue is K or N ;
(bh) at position 101 the amino acid residue is K or R;
(bi) at position 103 the amino acid residue is A or V;
(bj) at position 104 the amino acid residue is D;
(bk) at position 105 the amino acid residue is M, L or I;
(bl) at position 106 the amino acid residue is L;
(bm) at position 112 the amino acid residue is T or A;
(bn) at position 113 the amino acid residue is S or T;
(bo) at position 114 the amino acid residue is A;
(bp) at position 115 the amino acid residue is S;
(bq) at position 119 the amino acid residue is K or R;
(br) at position 120 the amino acid residue is K or R;
(bs) at position 123 the amino acid residue is F or L;
(bt) at position 125 the amino acid residue is E;
(bu) at position 126 the amino acid residue is Q or H;
(bv) at position 128 the amino acid residue is E or D;
(bw) at position 129 the amino acid residue is V or I;
(bx) at position 130 the amino acid residue is F;
(by) at position 131 the amino acid residue is D or E;
(bz) at position 132 the amino acid residue is T;
(ca) at position 135 the amino acid residue is V;
(cb) at position 138 the amino acid residue is H;
(cc) at position 139 the amino acid residue is I;
(cd) at position 140 the amino acid residue is L or M;
(ce) at position 142 the amino acid residue is Y;
(cf) at position 143 the amino acid residue is K or R;
(cg) at position 145 the amino acid residue is L or I; and
(ch) at position 146 the amino acid residue is T.

80. The isolated or recombinant polypeptide of item 79, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) - (ch), at least 90% conform to the amino acid residue restrictions specified in (a) - (ch).

81. The isolated or recombinant polypeptide of item 70, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at positions 9, 76, 94 and 110 the amino acid residue is A;
(b) at positions 29 and 108 the amino acid residue is C;
(c) at position 34 the amino acid residue is D;
(d) at position 95 the amino acid residue is E;
(e) at position 56 the amino acid residue is F;
(f) at positions 43, 44, 66, 74, 87, 102, 116, 122, 127 and 136 the amino acid residue is G;
(g) at position 41 the amino acid residue is H;
(h) at position 7 the amino acid residue is I;
(i) at position 85 the amino acid residue is K;

COR-AAD1-00001615

(j) at positions 20, 42, 50, 78 and 121 the amino acid residue is L;
(k) at positions 1 and 141 the amino acid residue is M;
(l) at positions 23 and 109 the amino acid residue is N;
(m) at positions 22, 25, 133, 134 and 137 the amino acid residue is P;
(n) at position 71 the amino acid residue is Q;
(o) at positions 16, 21, 73, 99 and 111 the amino acid residue is R;
(p) at position 55 the amino acid residue is S;
(q) at position 77 the amino acid residue is T;
(r) at position 107 the amino acid residue is W; and
(s) at position 13, 46, 70 and 118 the amino acid residue is Y.

82. The isolated or recombinant polypeptide of any of items 70-81, wherein the amino acid sequence comprises an amino acid residue selected from the group consisting of:

(a) at position 36 the amino acid residue is M, L, or T;
(b) at position 72 the amino acid residue is L or I;
(c) at position 75 the amino acid residue is M or V;
(d) at position 64 the amino acid residue is L, I, or F;
(e) at position 88 the amino acid residue is T or S;
(f) at position 117 the amino acid residue is Y or F.

83. The isolated or recombinant polypeptide of any of items 70-83, wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

(a) at position 14 the amino acid residue is D;
(b) at position 18 the amino acid residue is E;
(c) at position 26 the amino acid residue is M or V;
(e) at position 30 the amino acid residue is I;
(f) at position 32 the amino acid residue is D;
(g) at position 36 the amino acid residue is M or T;
(i) at position 37 the amino acid residue is C;
(j) at position 38 the amino acid residue is D;
(j) at position 53 the amino acid residue is V;
(k) at position 58 the amino acid residue is R;
(1) at position 61 the amino acid residue is R;
(m) at position 62 the amino acid residue is L;
(n) at position 64 the amino acid residue is I or F;
(o) at position 65 the amino acid residue is P;
(p) at position 72 the amino acid residue is I;
(q) at position 75 the amino acid residue is V;
(r) at position 88 the amino acid residue is T;
(s) at position 89 the amino acid residue is G;
(t) at position 91 the amino acid residue is L;
(u) at position 98 the amino acid residue is I;
(v) at position 105 the amino acid residue I;
(w) at position 112 the amino acid residue is A;
(x) at position 124 the amino acid residue is G or C;
(y) at position 128 the amino acid residue is D;
(z) at position 140 the amino acid residue is M;
(aa) at position 143 the amino acid residue is R; and
(ab) at position 144 the amino acid residue is W.

84. The isolated or recombinant polypeptide of item 84, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) through (ab), at least 80% conform to the amino acid residue restrictions specified in (a) through (ab).

85. The isolated or recombinant polypeptide of any of items 70-84, wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

COR-AAD1-00001616

(a) at position 41 the amino acid residue is H;
(b) at position 138 the amino acid residue is H;
(c) at position 34 the amino acid residue is N; and
(f) at position 55 the amino acid residue is S.

86. An isolated or recombinant polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence that is at least 98% identical to SEQ ID NO:577;
(b) an amino acid sequence that is at least 97% identical to SEQ ID NO:578;
(c) an amino acid sequence that is at least 97% identical to SEQ ID NO:621;
(d) an amino acid sequence that is at least 98% identical to SEQ ID NO:579;
(e) an amino acid sequence that is at least 98% identical to SEQ ID NO:602;
(f) an amino acid sequence that is at least 95% identical to SEQ ID NO:697;
(g) an amino acid sequence that is at least 96% identical to SEQ ID NO:721;
(h) an amino acid sequence that is at least 97% identical to SEQ ID NO:613;
(i) an amino acid sequence that is at least 89% identical to SEQ ID NO:677;
(j) an amino acid sequence that is at least 96% identical to SEQ ID NO:584;
(k) an amino acid sequence that is at least 98% identical to SEQ ID NO:707;
(l) an amino acid sequence that is at least 98% identical to SEQ ID NO:616;
(m) an amino acid sequence that is at least 96% identical to SEQ ID NO:612;
(n) an amino acid sequence that is at least 98% identical to SEQ ID NO:590;
(o) an amino acid sequence that is complementary to any one of (a) through (n).

87. The isolated or recombinant polypeptide of item 86, wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

(i) at positions 18 and 38, a Z5 amino acid residue;
(ii) at position 62, a Z1 amino acid residue;
(iii) at position 124, a Z6 amino acid residue; and
(iv) at position 144, a Z2 amino acid residue,
wherein:

Z1 is an amino acid residue selected from the group consisting of A, I, L, M, and V;
Z2 is an amino acid residue selected from the group consisting of F, W, and Y;
Z5 is an amino acid residue selected from the group consisting of D and E;
Z6 is an amino acid residue selected from the group consisting of C, G, and P.

88. The isolated or recombinant polypeptide of item 86, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 2, 4, 15, 19, 26, 28, 31, 45, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 123, 129, 139, and/or 145 the amino acid residue is B1; and
(b) at positions 3, 5, 8, 10, 11, 14, 17, 24, 27, 32, 37, 47, 48, 49, 52, 57, 58, 61, 63, 68, 69, 79, 80, 82, 83, 89, 92, 100, 101, 104, 119, 120, 125, 126, 128, 131, and/or 143 the amino acid residue is B2;
wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

89. The isolated or recombinant polypeptide of item 86, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at positions 2, 4, 15, 19, 26, 28, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 129, 139, and/or 145 the amino acid residue is Z1;
(b) at positions 31 and/or 45 the amino acid residue is Z2;
(c) at position 8 the amino acid residue is Z3;
(d) at position 89 the amino acid residue is Z3 or Z6;
(e) at positions 82, 92, 101 and/or 120 the amino acid residue is Z4;
(f) at positions 3, 11, 27 and/or 79 the amino acid residue is Z5;

COR-AAD1-00001617

(g) at position 18 the amino acid residue is Z4 or Z5;
(h) at position 123 the amino acid residue is Z1 or Z2;
(i) at positions 12, 33, 35, 39, 53, 59, 112, 132, 135, 140, and/or 146 the amino acid residue is Z1 or Z3;
(j) at position 30 the amino acid residue is Z1;
(k) at position 6 the amino acid residue is Z6;
(l) at position 81 the amino acid residue is Z2 or Z4;
(m) at position 113 the amino acid residue is Z3;
(n) at position 138 the amino acid residue is Z4;
(o) at position 142 the amino acid residue is Z2;
(p) at positions 57 and/or 126 the amino acid residue is Z3 or Z4;
(q) at position 5, 17, and 61 the amino acid residue is Z4;
(r) at position 24 the amino acid residue is Z3;
(s) at position 104 the amino acid residue is Z5;
(t) at positions 52, and/or 69 the amino acid residue is Z3;
(u) at positions 14 and/or 119 the amino acid residue is Z5;
(v) at positions 10, 32, 63, and/or 83 the amino acid residue is Z5 ;
(w) at positions 48 and/or 80 the amino acid residue is Z6;
(x) at position 40 the amino acid residue is Z1 or Z2 ;
(y) at position 96 the amino acid residue is Z3 or Z5;
(z) at position 65 the amino acid residue is Z3, Z4, or Z6;
(aa) at positions 84 and/or 115 the amino acid residue is Z3;
(ab) at position 93 the amino acid residue is Z4;
(ac) at position 130 the amino acid residue is Z2;
(ad) at position 58 the amino acid residue is Z3, Z4 or Z6;
(ae) at position 47 the amino acid residue is Z4 or Z6;
(af) at positions 49 and/or 100 the amino acid residue is Z3 or Z4;
(ag) at position 68 the amino acid residue is Z4 or Z5;
(ah) at position 143 the amino acid residue is Z4;
(ai) at position 131 the amino acid residue is Z5;
(aj) at positions 125 and/or 128 the amino acid residue is Z5;
(ak) at position 67 the amino acid residue is Z3 or Z4;
(al) at position 60 the amino acid residue is Z5; and
(am) at position 37 the amino acid residue is Z4 or Z6;

wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting of N, Q, S, and T; Z4 is an amino acid selected from the group consisting of R, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

90. The isolated or recombinant polypeptide of item 89, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) - (am), at least 90% conform to the amino acid residue restrictions specified in (a) - (am).

91. The isolated or recombinant polypeptide of any of items 86-90, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 13, 20, 36, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or 141 the amino acid residue is B1; and
(b) at positions 16, 21, 22, 23, 25, 29, 34, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99, 102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2;

wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

92. The isolated or recombinant polypeptide of any of items 86-91 wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 13, 20, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121, and/or

COR-AAD1-00001618

141 the amino acid residue is B1; and
(b) at positions 16, 21, 22, 23, 25, 29, 34, 36, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99,102, 108, 109, 111, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is B2;
wherein B1 is an amino acid selected from the group consisting of A, I, L, M, F, W, Y, and V; and B2 is an amino acid selected from the group consisting of R, N, D, C, Q, E, G, H, K, P, S, and T.

93. The isolated or recombinant polypeptide of any of items 86-91 wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 90% conform to the following restrictions:

(a) at positions 1, 7, 9, 20, 42, 50, 72, 75, 76, 78, 94, 98, 110, 121, and/or 141 the amino acid residue is Z1;
(b) at positions 13, 46, 56, 70, 107, 117, and/or 118 the amino acid residue is Z2;
(c) at positions 23, 55, 71, 77, 88, and/or 109 the amino acid residue is Z3;
(d) at positions 16, 21, 41, 73, 85, 99, and/or 111 the amino acid residue is Z4;
(e) at positions 34 and/or 95 the amino acid residue is Z5;
(f) at position 22, 25, 29, 43, 44, 66, 74, 87, 102, 108, 116, 122, 127, 133, 134, 136, and/or 137 the amino acid residue is Z6;
wherein Z1 is an amino acid selected from the group consisting of A, I, L, M, and V; Z2 is an amino acid selected from the group consisting of F, W, and Y; Z3 is an amino acid selected from the group consisting of N, Q, S, and T; Z4 is an amino acid selected from the group consisting ofR, H, and K; Z5 is an amino acid selected from the group consisting of D and E; and Z6 is an amino acid selected from the group consisting of C, G, and P.

94. The isolated or recombinant polypeptide of item 93, wherein the amino acid sequence comprises at position 36, an amino acid residue selected from the group consisting of Z1 and Z3.

95. The isolated or recombinant polypeptide of item 93 or item 94, wherein the amino acid sequence comprises at position 64, an amino acid residue selected from the group consisting of Z1 and Z2.

96. The isolated or recombinant polypeptide of any items 86-95 wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at position 2 the amino acid residue is I or L;
(b) at position 3 the amino acid residue is E;
(c) at position 4 the amino acid residue is V or I;
(d) at position 5 the amino acid residue is K;
(e) at position 6 the amino acid residue is P;
(f) at position 8 the amino acid residue is N;
(g) at position 10 the amino acid residue is E;
(h) at position 11 the amino acid residue is D or E;
(i) at position 12 the amino acid residue is T;
(j) at position 14 the amino acid residue is E or D;
(k) at position 15 the amino acid residue is L;
(1) at position 17 the amino acid residue is H;
(m) at position 18 the amino acid residue is R, E or K;
(n) at position 19 the amino acid residue is I or V;
(o) at position 24 the amino acid residue is Q;
(p) at position 26 the amino acid residue is M, L, V or I;
(q) at position 27 the amino acid residue is E;
(r) at position 28 the amino acid residue is A or V;
(s) at position 30 the amino acid residue is M;
(t) at position 31 the amino acid residue is Y or F;
(u) at position 32 the amino acid residue is E or D;
(v) at position 33 the amino acid residue is T or S;
(w) at position 35 the amino acid residue is L;
(x) at position 37 the amino acid residue is R, G, E or Q;
(y) at position 39 the amino acid residue is A or S;
(z) at position 40 the amino acid residue is F or L ;
(aa) at position 45 the amino acid residue is Y or F;
(ab) at position 47 the amino acid residue is R or G;

COR-AAD1-00001619

(ac) at position 48 the amino acid residue is G;
(ad) at position 49 the amino acid residue is K, R, or Q;
(ae) at position 51 the amino acid residue is I or V;
(af) at position 52 the amino acid residue is S;
(ag) at position 53 the amino acid residue is I or V;
(ah) at position 54 the amino acid residue is A;
(ai) at position 57 the amino acid residue is H or N;
(aj) at position 58 the amino acid residue is Q, K, R or P;
(ak) at position 59 the amino acid residue is A;
(al) at position 60 the amino acid residue is E;
(am) at position 61 the amino acid residue is H or R;
(an) at position 63 the amino acid residue is E or D;
(ao) at position 65 the amino acid residue is E, P or Q;
(ap) at position 67 the amino acid residue is Q or R;
(aq) at position 68 the amino acid residue is K or E;
(ar) at position 69 the amino acid residue is Q;
(as) at position 79 the amino acid residue is E;
(at) at position 80 the amino acid residue is G;
(au) at position 81 the amino acid residue is Y, H or F;
(av) at position 82 the amino acid residue is R;
(aw) at position 83 the amino acid residue is E or D;
(ax) at position 84 the amino acid residue is Q;
(ay) at position 86 the amino acid residue is A;
(az) at position 89 the amino acid residue is G, T or S;
(ba) at position 90 the amino acid residue is L;
(bb) at position 91 the amino acid residue is L, I or V;
(bc) at position 92 the amino acid residue is R or K;
(bd) at position 93 the amino acid residue is H;
(be) at position 96 the amino acid residue is E or Q;
(bf) at position 97 the amino acid residue is I;
(bg) at position 100 the amino acid residue is K or N ;
(bh) at position 101 the amino acid residue is K or R;
(bi) at position 103 the amino acid residue is A or V;
(bj) at position 104 the amino acid residue is D;
(bk) at position 105 the amino acid residue is M, L or I;
(bl) at position 106 the amino acid residue is L;
(bm) at position 112 the amino acid residue is T or A;
(bn) at position 113 the amino acid residue is S or T;
(bo) at position 114 the amino acid residue is A;
(bp) at position 115 the amino acid residue is S;
(bq) at position 119 the amino acid residue is K or R;
(br) at position 120 the amino acid residue is K or R;
(bs) at position 123 the amino acid residue is F or L;
(bt) at position 125 the amino acid residue is E;
(bu) at position 126 the amino acid residue is Q or H;
(bv) at position 128 the amino acid residue is E or D;
(bw) at position 129 the amino acid residue is V or I;
(bx) at position 130 the amino acid residue is F;
(by) at position 131 the amino acid residue is D or E;
(bz) at position 132 the amino acid residue is T;
(ca) at position 135 the amino acid residue is V;
(cb) at position 138 the amino acid residue is H;
(cc) at position 139 the amino acid residue is I;
(cd) at position 140 the amino acid residue is L or M;
(ce) at position 142 the amino acid residue is Y;
(cf) at position 143 the amino acid residue is K or R;
(cg) at position 145 the amino acid residue is L or I; and
(ch) at position 146 the amino acid residue is T.

COR-AAD1-00001620

97. The isolated or recombinant polypeptide of item 96, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) - (ch), at least 90% conform to the amino acid residue restrictions specified in (a) - (ch).

98. The isolated or recombinant polypeptide of item 86, wherein of the amino acid residues in the amino acid sequence that correspond to the following positions, at least 80% conform to the following restrictions:

(a) at positions 9, 76, 94 and 110 the amino acid residue is A;
(b) at positions 29 and 108 the amino acid residue is C;
(c) at position 34 the amino acid residue is D;
(d) at position 95 the amino acid residue is E;
(e) at position 56 the amino acid residue is F;
(f) at positions 43, 44, 66, 74, 87, 102, 116, 122, 127 and 136 the amino acid residue is G;
(g) at position 41 the amino acid residue is H;
(h) at position 7 the amino acid residue is I;
(i) at position 85 the amino acid residue is K;
(j) at positions 20, 42, 50, 78 and 121 the amino acid residue is L;
(k) at positions 1 and 141 the amino acid residue is M;
(l) at positions 23 and 109 the amino acid residue is N;
(m) at positions 22, 25, 133, 134 and 137 the amino acid residue is P;
(n) at position 71 the amino acid residue is Q;
(o) at positions 16, 21, 73, 99 and 111 the amino acid residue is R;
(p) at position 55 the amino acid residue is S;
(q) at position 77 the amino acid residue is T;
(r) at position 107 the amino acid residue is W; and
(s) at position 13, 46, 70 and 118 the amino acid residue is Y.

99. The isolated or recombinant polypeptide of any of items 70-98, wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

(a) at position 14 the amino acid residue is D;
(b) at position 18 the amino acid residue is E;
(c) at position 26 the amino acid residue is M or V;
(e) at position 30 the amino acid residue is I;
(f) at position 32 the amino acid residue is D;
(g) at position 36 the amino acid residue is M or T;
(i) at position 37 the amino acid residue is C;
(j) at position 38 the amino acid residue is D;
(j) at position 53 the amino acid residue is V;
(k) at position 58 the amino acid residue is R;
(1) at position 61 the amino acid residue is R;
(m) at position 62 the amino acid residue is L;
(n) at position 64 the amino acid residue is I or F;
(o) at position 65 the amino acid residue is P;
(p) at position 72 the amino acid residue is I;
(q) at position 75 the amino acid residue is V;
(r) at position 88 the amino acid residue is T;
(s) at position 89 the amino acid residue is G;
(t) at position 91 the amino acid residue is L;
(u) at position 98 the amino acid residue is I;
(v) at position 105 the amino acid residue I;
(w) at position 112 the amino acid residue is A;
(x) at position 124 the amino acid residue is G or C;
(y) at position 128 the amino acid residue is D;
(z) at position 140 the amino acid residue is M;
(aa) at position 143 the amino acid residue is R; and
(ab) at position 144 the amino acid residue is W.

COR-AAD1-00001621

100. The isolated or recombinant polypeptide of item 99, wherein of the amino acid residues in the amino acid sequence that correspond to the positions specified in (a) through (ab), at least 80% conform to the amino acid residue restrictions specified in (a) through (ab).

101. The isolated or recombinant polypeptide of any of items 86-100, wherein the amino acid sequence comprises one or more amino acid residues selected from the group consisting of:

(a) at position 41 the amino acid residue is H;
(b) at position 138 the amino acid residue is H;
(c) at position 34 the amino acid residue is N; and
(e) at position 55 the amino acid residue is S.

102. The isolated or recombinant polypeptide of item 69, wherein the polypeptide catalyzes the acetylation of glyphosate with a k_{cat}/K_m of at least 10 mM^{-1} min^{-1} for glyphosate.

103. The isolated or recombinant polypeptide of item 102, wherein the polypeptide catalyzes the acetylation of glyphosate with a k_{cat}/K_m of at least 100 mM^{-1} min^{-1} for glyphosate.

104. The isolated or recombinant polypeptide of item 103, wherein the polypeptide catalyzes the acetylation of aminomethylphosphonic acid.

105. A polypeptide having GAT activity said polypeptide selected from the group consisting of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972 characterized by:

(a) a K_m for glyphosate of at least about 2 mM or less;
(b) a K_m for acetyl CoA of at least about 200 μM or less; and
(c) a k_{cat} equal to at least about 6/minute.

106. An isolated or recombinant polypeptide having glyphosate-N-acyltransferase activity, said polypeptide selected from the group consisting of: SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972.

107. A method of producing a glyphosate-resistant transgenic plant or plant cell comprising:

(a) transforming a plant or plant cell with a polynucleotide encoding a glyphosate-N-acetyltransferase said polynucleotide selected from the group consisting of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708,

COR-AAD1-00001622

710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952; and

(b) optionally regenerating a transgenic plant from the transformed plant cell.

108. A method for selecting a plant or plant cell containing a nucleic acid construct, comprising:

(a) providing a transgenic plant or plant cell containing a nucleic acid construct, wherein the nucleic acid construct comprises a nucleotide sequence that encodes a glyphosate-N-acetyltransferase, said nucleotide sequence selected from the group consisting of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952; and

(b) growing the transgenic plant or plant cell in the presence of glyphosate under conditions where the glyphosate-N-acetyltransferase is expressed at an effective level, whereby the transgenic plant or plant cell grows at a rate that is discernibly greater than the plant or plant cell would grow if it did not contain the nucleic acid construct.

109. A method for selectively controlling weeds in a field containing a crop comprising:

(a) planting a field with crop seeds or plants which are glyphosate-tolerant as a result of being transformed with a gene encoding a glyphosate N-acetyltransferase, said gene selected from the groups consisting of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952; and

(b) applying to the crop and weeds in the field a sufficient amount of glyphosate to control the weeds without significantly affecting the crop.

110. The method of item 109, where the crop is cotton, corn, or soybean.

111. A transgenic plant or transgenic plant explant having an enhanced tolerance to glyphosate, wherein the plant or plant explant expresses a polypeptide with glyphosate-N-acetyltransferase activity, said polypeptide selected from the group consisting of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917,

COR-AAD1-00001623

919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972; and at least one polypeptide imparting glyphosate tolerance by an additional mechanism.

112. A transgenic plant or transgenic plant explant, wherein the plant or plant explant expresses a polypeptide with glyphosate-N-acetyltransferase activity, said polypeptide selected from the group consisting of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972; and at least one polypeptide imparting tolerance to an additional herbicide.

113. A transgenic plant or transgenic plant explant having an enhanced tolerance to glyphosate, wherein the plant or plant explant expresses a polypeptide with glyphosate-N-acetyltransferase activity, said polypeptide selected from the group consisting of SEQ ID NO: 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 621, 623, 625, 627, 629, 631, 633, 635, 637, 639, 641, 643, 645, 647, 649, 651, 653, 655, 657, 659, 661, 663, 665, 667, 669, 671, 673, 675, 677, 679, 681, 683, 685, 687, 689, 691, 693, 695, 697, 699,701, 703, 705, 707, 709, 711, 713, 715, 717, 719,721, 723, 725, 727, 729, 731, 733, 735, 737, 739, 741, 743, 745, 747, 749, 751, 753, 755, 757, 759, 761, 763, 765, 767, 769, 771, 773, 775, 777, 779, 781, 783, 785, 787, 789, 791, 793, 795, 797, 799, 801, 803, 805, 807, 809, 811, 813, 815, 817, 819, 821, 823, 825, 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972; at least one polypeptide imparting glyphosate tolerance by an additional mechanism, and at least one polypeptide imparting tolerance to an additional herbicide.

114. A method for controlling weeds in a field containing a crop comprising:

(a) planting the field with crop seeds or plants which are transformed with a gene encoding a glyphosate-N-acetyltransferase, said gene being selected from the group consisting of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952; and at least one gene encoding a polypeptide imparting glyphosate tolerance by an additional mechanism; and
(b) applying to the crop and weeds in the field an effective application of glyphosate sufficient to inhibit growth of the weeds in the field without significantly affecting the crop.

115. A method for selectively controlling weeds in a field containing a crop comprising:

(a) planting the field with crop seeds or plants which are transformed with a gene encoding a glyphosate-N-acetyltransferase, said gene selected from the group consisting of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702,

COR-AAD1-00001624

704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952; and at least one gene encoding a polypeptide imparting tolerance to an additional herbicide, and;

(b) applying to the crop and weeds in the field a simultaneous or chronologically staggered application of glyphosate and the additional herbicide which is sufficient to inhibit growth of the weeds in the field without significantly affecting the crop.

116. A method for selectively controlling weeds in a field containing a crop comprising:

(a) planting the field with crop seeds or plants which are transformed with a gene encoding a glyphosate-N-acetyltransferase, said gene selected from the group consisting of SEQ ID NO: 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 788, 790, 792, 794, 796, 798, 800, 802, 804, 806, 808, 810, 812, 814, 816, 818, 820, 822, 824, 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952; at least one gene encoding a polypeptide imparting glyphosate tolerance by an additional mechanism and at least one gene encoding a polypeptide imparting tolerance to an additional herbicide, and;

(b) applying to the crop and weeds in the field a simultaneous or chronologically staggered application of glyphosate and the additional herbicide which is sufficient to inhibit growth of the weeds in the field without significantly affecting the crop.

117. An isolated or recombinant polynucleotide comprising a nucleotide sequence selected from the groups consisting of:

(a) a nucleotide sequence encoding an amino acid sequence that is at least 96% identical to SEQ ID NO:919;
(b) a nucleotide sequence encoding an amino acid sequence that is at least 97% identical to SEQ ID NO:929;
(c) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:847;
(d) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:851;
(e) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:853;
(f) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:855;
(g) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:857;
(h) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:861;
(i) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:871;
(j) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:875;
(k) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:881;
(l) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:885;
(m) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:887;
(n) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:889;
(o) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:893;
(p) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:897;
(q) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:899;
(r) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:909;
(s) a nucleotide sequence encoding an amino acid sequence that is at least 98% identical to SEQ ID NO:911;
(t) a nucleotide sequence encoding an amino acid sequence that is at least 99% identical to SEQ ID NO:837;
(u) a nucleotide sequence encoding an amino acid sequence that is at least 99% identical to SEQ ID NO:841;
(v) a nucleotide sequence encoding an amino acid sequence that is at least 99% identical to SEQ ID NO:865;
(w) a nucleotide sequence encoding an amino acid sequence that is at least 99% identical to SEQ ID NO:869; and
(x) a nucleotide sequence encoding an amino acid sequence that is at least 99% identical to SEQ ID NO:879.

COR-AAD1-00001625

118. An isolated or recombinant polynucleotide comprising a nucleotide sequence encoding an amino acid sequence that is at least 95% identical to SEQ ID NO:929 and which comprises a Gly or an Asn residue at the amino acid position corresponding to position 33 of SEQ ID NO:929.

119. An isolated or recombinant polynucleotide encoding a polypeptide having glyphosate-N-acetyltransferase activity, said polynucleotide being selected from the group consisting of: SEQ ID NO: 832, 834, 836, 838, 840, 842, 844, 846, 848, 850, 852, 854, 856, 858, 860, 862, 864, 866, 868, 870, 872, 874, 876, 878, 880, 882, 884, 886, 888, 890, 892, 894, 896, 898, 900, 902, 904, 906, 908, 910, 912, 914, 916, 918, 920, 922, 924, 926, 928, 930,932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 947, 949, 951, and 952.

120. An isolated or recombinant polypeptide comprising an amino acid sequence selected from the groups consisting of:

(a) an amino acid sequence that is at least 96% identical to SEQ ID NO:919;
(b) an amino acid sequence that is at least 97% identical to SEQ ID NO:929;
(c) an amino acid sequence that is at least 98% identical to SEQ ID NO:847;
(d) an amino acid sequence that is at least 98% identical to SEQ ID NO:851;
(e) an amino acid sequence that is at least 98% identical to SEQ ID NO:853;
(f) an amino acid sequence that is at least 98% identical to SEQ ID NO:855;
(g) an amino acid sequence that is at least 98% identical to SEQ ID NO:857;
(h) an amino acid sequence that is at least 98% identical to SEQ ID NO:861;
(i) an amino acid sequence that is at least 98% identical to SEQ ID NO:871;
(j) an amino acid sequence that is at least 98% identical to SEQ ID NO:875;
(k) an amino acid sequence that is at least 98% identical to SEQ ID NO:881;
(l) an amino acid sequence that is at least 98% identical to SEQ ID NO:885;
(m) an amino acid sequence that is at least 98% identical to SEQ ID NO:887;
(n) an amino acid sequence that is at least 98% identical to SEQ ID NO:889;
(o) an amino acid sequence that is at least 98% identical to SEQ ID NO:893;
(p) an amino acid sequence that is at least 98% identical to SEQ ID NO:897;
(q) an amino acid sequence that is at least 98% identical to SEQ ID NO:899;
(r) an amino acid sequence that is at least 98% identical to SEQ ID NO:909;
(s) an amino acid sequence that is at least 98% identical to SEQ ID NO:911;
(t) an amino acid sequence that is at least 99% identical to SEQ ID NO:837;
(u) an amino acid sequence that is at least 99% identical to SEQ ID NO:841;
(v) an amino acid sequence that is at least 99% identical to SEQ ID NO:865;
(w) an amino acid sequence that is at least 99% identical to SEQ ID NO:869; and
(x) an amino acid sequence that is at least 99% identical to SEQ ID NO:879.

121. An isolated or recombinant polypeptide comprising an amino acid sequence that is at least 95% identical to SEQ ID NO:929 and which comprises a Gly or an Asn residue at the amino acid position corresponding to position 33 of SEQ ID NO:929.

122. An isolated or recombinant polypeptide comprising residues 2-146 of an amino acid sequence selected from the group consisting of: SEQ ID NO: 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972.

123. The isolated or recombinant polypeptide of item 122, comprising an amino acid sequence selected from the group consisting of: SEQ ID NO: 833, 835, 837, 839, 841, 843, 845, 847, 849, 851, 853, 855, 857, 859, 861, 863, 865, 867, 869, 871, 873, 875, 877, 879, 881, 883, 885, 887, 889, 891, 893, 895, 897, 899, 901, 903, 905, 907, 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 948, 950, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, and 972.

124. A method of increasing the expression of a polynucleotide in a plant or plant cell, the method comprising:

(a) obtaining a first polynucleotide comprising a nucleic acid sequence encoding a polypeptide comprising an amino acid sequence that can be optimally aligned with a sequence selected from the group consisting of SEQ

COR-AAD1-00001626

ID NO:300, SEQ ID NO:445 and SEQ ID NO:457 to generate a similarity score of at least 460 using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1, the nucleic acid sequence comprising an initiating methionine ATG codon;
(b) inserting one or two GCG or GCT codons into the nucleic acid sequence immediately downstream and adjacent to the initiating methionine ATG codon, resulting in a modified polynucleotide; and
(c) transforming the plant or plant cell with the modified polynucleotide,
wherein the plant or plant cell transformed with the modified polynucleotide accumulates a larger amount of encoded polypeptide compared to a plant or plant cell transformed with the first polynucleotide.

125. The method of item 124, further comprising identifying one or more codons in the nucleic acid sequence which are less frequently utilized in plants, and substituting the codons which are less frequently utilized in plants for codons encoding the same amino acids which are more frequently utilized in plants.

Claims

1. An isolated or recombinant polynucleotide comprising a nucleotide sequence encoding a polypeptide having glyphosate-N acetyltransferase activity and comprises an amino acid sequence comprising at least 90% sequence identity to SEQ ID NO: 950.

2. The isolated or recombinant polynucleotide of claim 1, wherein said polypeptide comprises an amino acid sequence that shares sequence identity with SEQ ID NO: 950 of at least 95%, 96%, 97%, 98%, or 99%.

3. The isolated or recombinant polynucleotide of any of the preceding claims, wherein said polypeptide catalyzes the acetylation of aminomethylphosphonic acid, or wherein said polypeptide also has glyphosate-N-acyltransferase activity.

4. A nucleic acid construct comprising the isolated or recombinant polynucleotide of any of the preceding claims, optionally further comprising a promoter operably linked to said polynucleotide, wherein said promoter is heterologous with respect to said polynucleotide and wherein the glyphosate tolerance of a plant cell transformed with said nucleic acid construct is increased in comparison to a plant cell that has not been transformed with said nucleic acid construct.

5. A cell comprising at least one polynucleotide of any of claims 1-4, wherein said polynucleotide is heterologous to the cell; wherein optionally said cell is a plant cell.

6. A transgenic plant or plant explant comprising at least one polynucleotide of any of claims 1-4, wherein said polynucleotide is heterologous to the explant or the plant.

7. The transgenic plant of claim 6, wherein:

(a) said transgenic plant is a crop plant in a genus selected from the following group: *Brassica, Linum, Zea, Glycine,* and *Gossypium;* and/or
(b) said plant exhibits enhanced resistance to glyphosate as compared to a wild type plant of the same species, strain or cultivar that does not comprise at least one polynucleotide of any of claims 1-4.

8. The transgenic plant of any of claims 6 or 7, wherein said plant further comprises at least one polypeptide imparting tolerance to an additional herbicide, which polypeptide optionally comprises:

(a) a sulfonamide-tolerant acetolactate synthase;
(b) a sulfonamide-tolerant acetohydroxy acid synthase;
(c) an imidazolinone-tolerant acetolactate synthase; or,
(d) an imidazolinone-tolerant acetohydroxy acid synthase;
and wherein preferably said acetolactate synthase contains an HRA mutation.

9. A seed produced by the plant of any one of claims 6 to 8, wherein said seed comprises at least one polynucleotide of any of claims 1-4.

10. A polypeptide encoded by a polynucleotide of any of claims 1-4.

COR-AAD1-00001627

11. A method for controlling weeds in a field containing a crop comprising:

(a) planting the field with herbicide-tolerant crop seeds or plants of any one of claims 7-9; and
(b) applying to any crop and weeds in the field a sufficient amount of one or more herbicides to control weeds without significantly affecting the crop.

12. The method of claim 11, wherein glyphosate is applied, wherein a sulfonylurea herbicide is applied, or wherein glyphosate and a sulfonylurea herbicide are applied.

13. The method of claim 12, wherein said sulfonylurea herbicide comprises at least one of chlorimuron, chlorsulfuron, rimsulfuron, thifensulfuron, or tribenuron; preferably wherein said sulfonylurea herbicide is tribenuron and wherein step (b) further comprises applying thifensulfuron to the field.

14. The method of any one of claims 11 to 13, wherein step (b) is performed before step (a).

15. A method of producing a glyphosate-resistant transgenic plant comprising:

(a) transforming a plant cell with the polynucleotide of any of claims 1-4;
(b) selecting a plant cell which is resistant to glyphosate by growing plant cells in the presence of a concentration of glyphosate that inhibits the growth of a plant cell which does not comprise a polynucleotide of any of claims 1-4; and
(c) regenerating a transgenic plant from said plant cell.

COR-AAD1-00001628

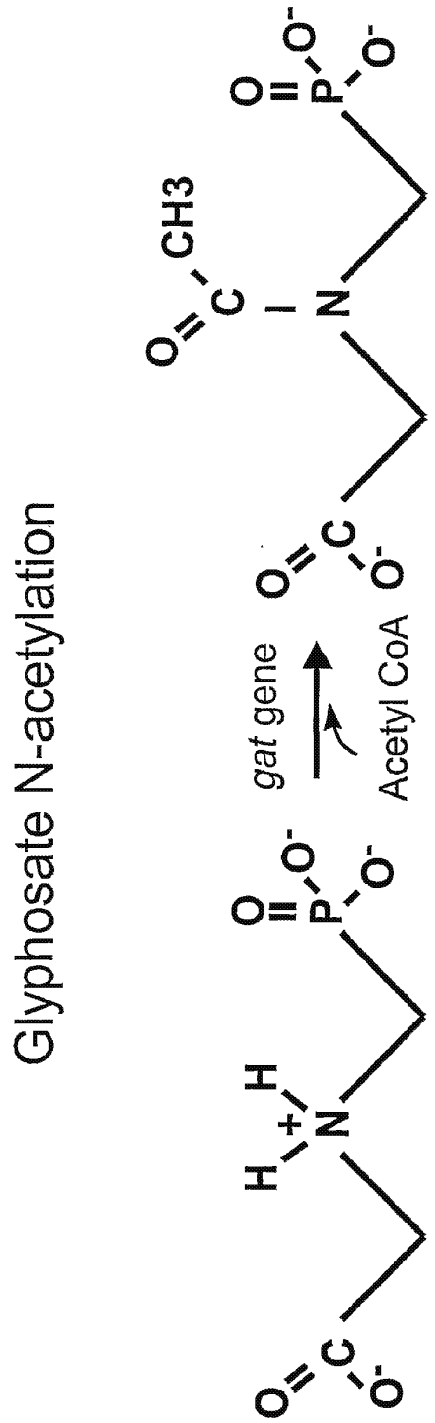


Fig. 1

COR-AAD1-00001629

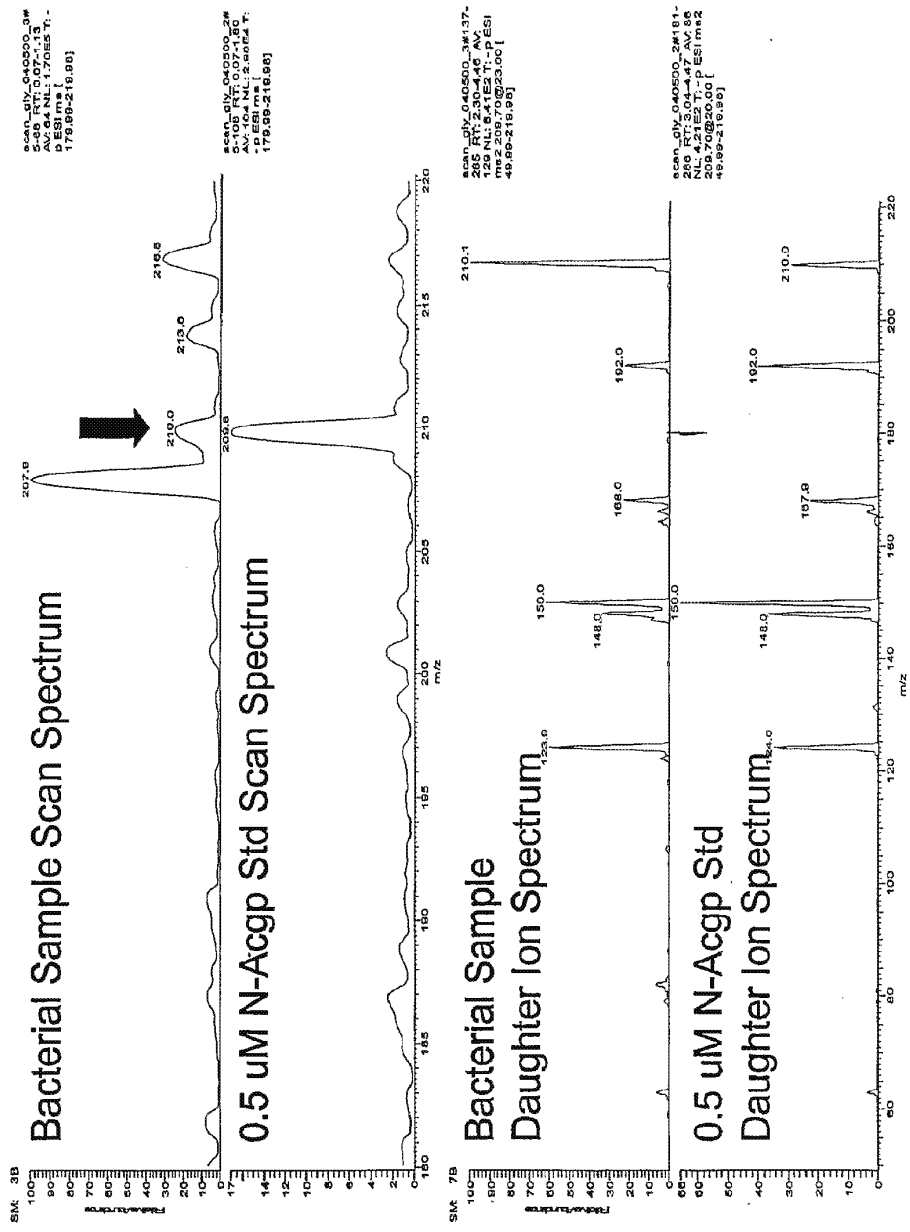

Bacterial Sample Scan Spectrum
0.5 uM N-Acgp Std Scan Spectrum
Bacterial Sample
Daughter Ion Spectrum
0.5 uM N-Acgp Std
Daughter Ion Spectrum
m/z

Fig. 2

Nucleotide identity between sequence pairs

	B6	DS3	ST401	NHA-2	NH5-2	yitI
B6	***	95.0	93.2	94.8	92.8	62.0
DS3		***	95.5	99.8	95.0	61.8
ST401			***	95.2	99.5	62.4
NHA-2				***	94.8	61.5
NH5-2					***	62.9
yitI						***

Fig. 3

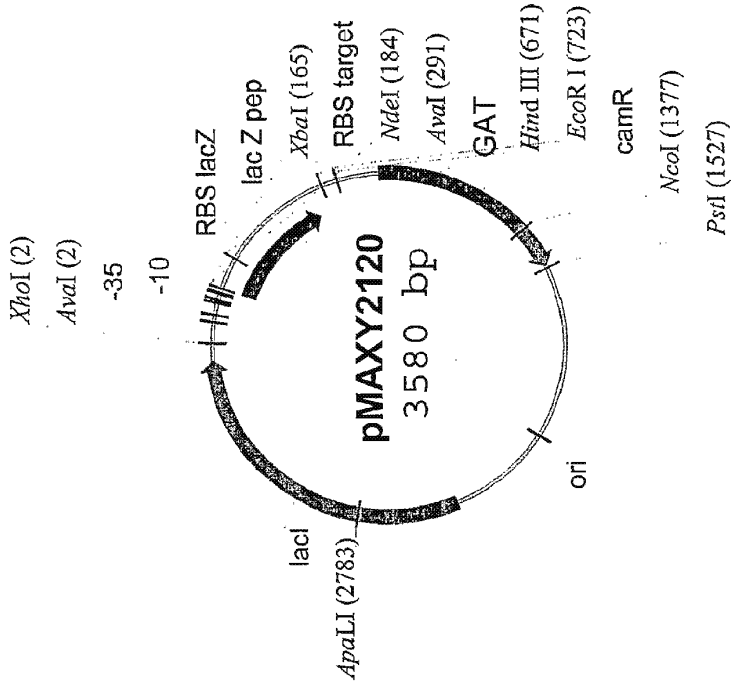
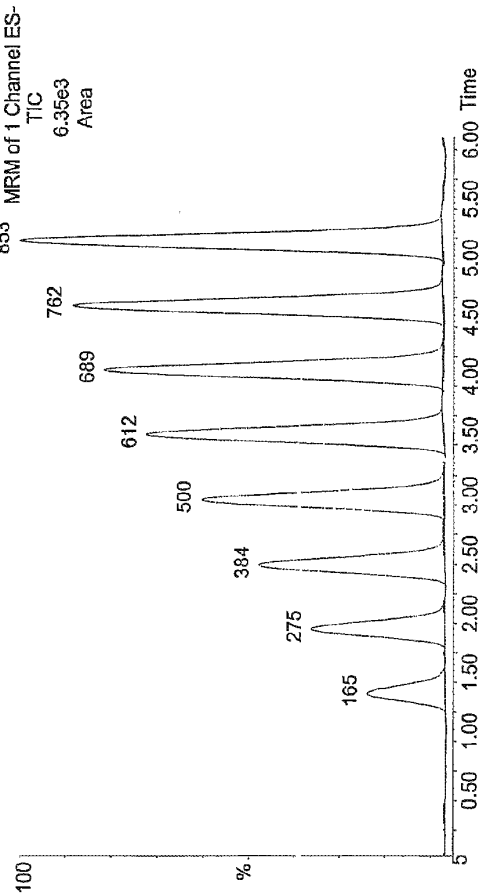


Fig. 4

COR-AAD1-00001632

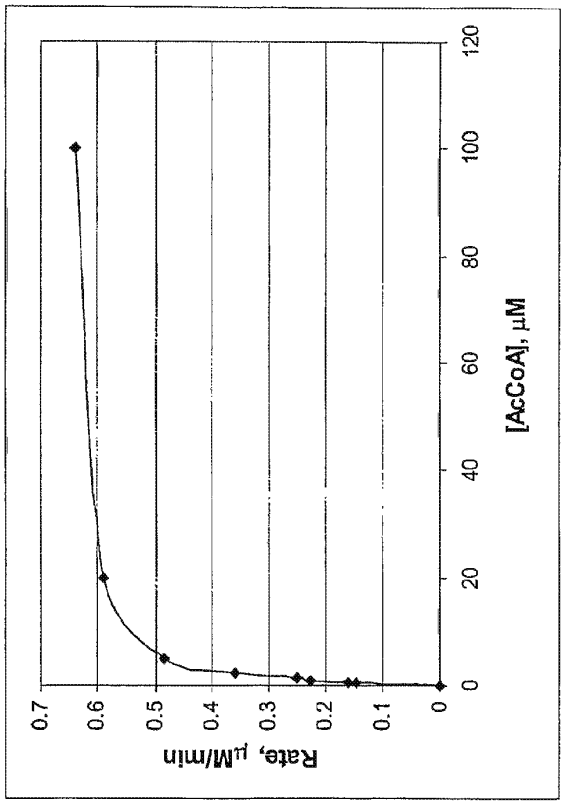


Purified B6 GAT catalyzes N-acetylglyphosate production

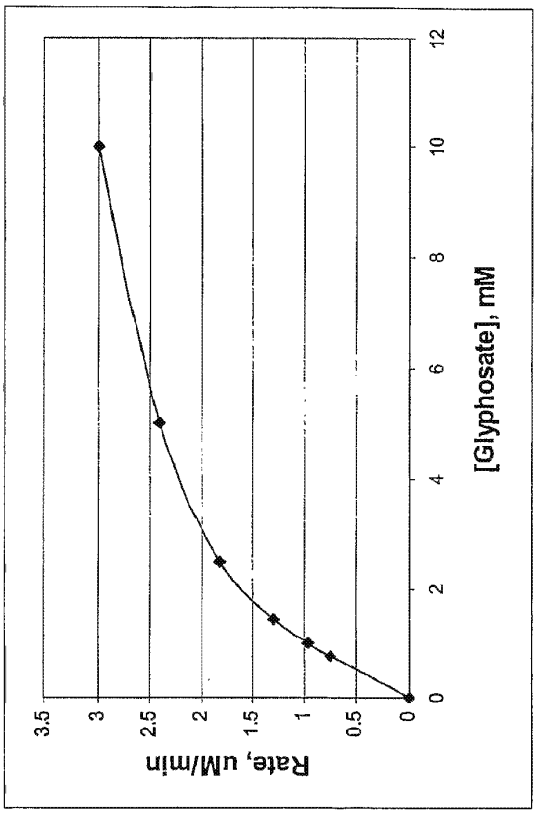
Fig. 5

COR-AAD1-00001633

Fig. 6



B6 GAT K_M for Acetyl CoA = 2 μM



B6 GAT K_M for glyphosate = 2.9 mM

Fig. 7

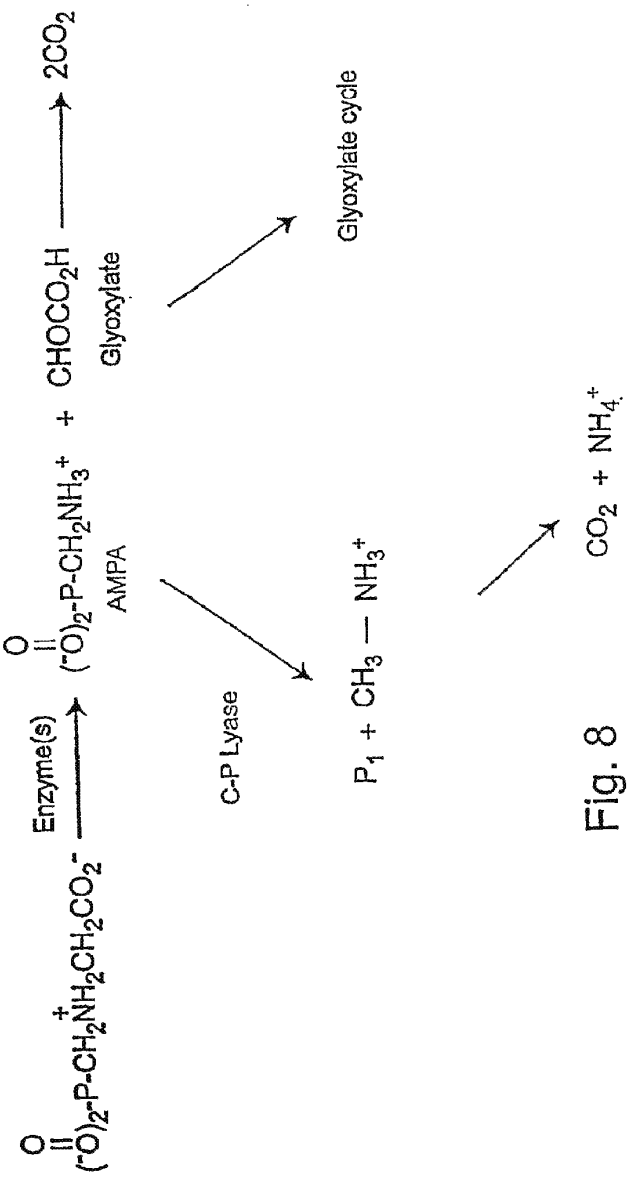

Enzyme(s)
C-P Lyase
AMPA
Glyoxylate
Glyoxylate cycle
$2CO_2$
$P_1 + CH_3 — NH_3^+$
$CO_2 + NH_4^+$

Fig. 8

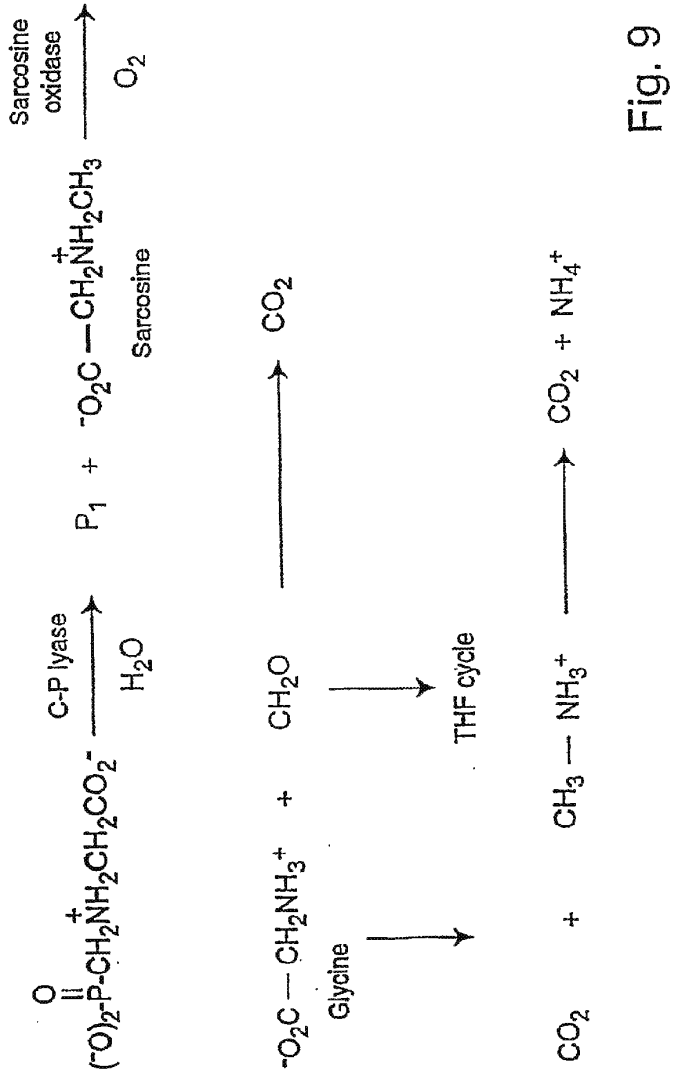


Fig. 9

COR-AAD1-00001637

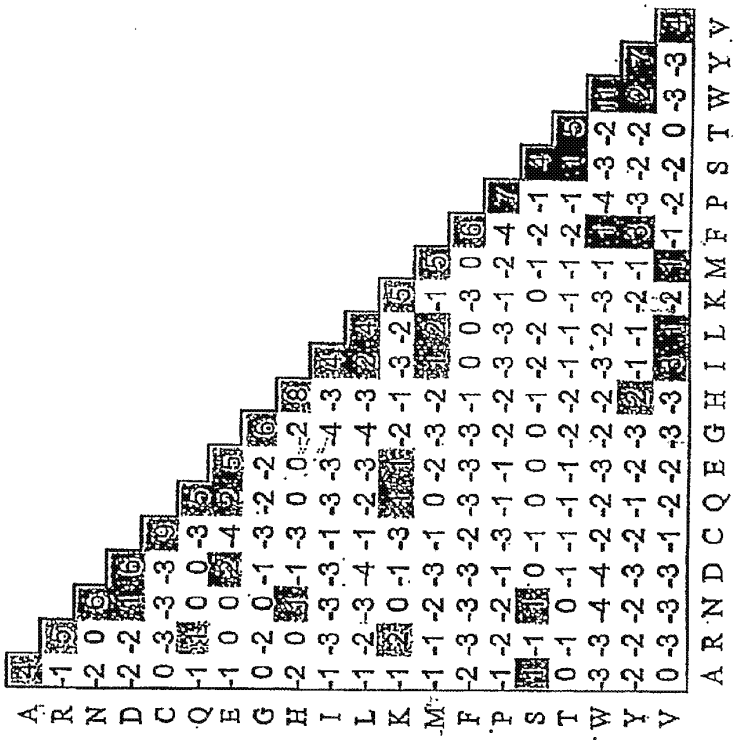


Fig. 10

COR-AAD1-00001638

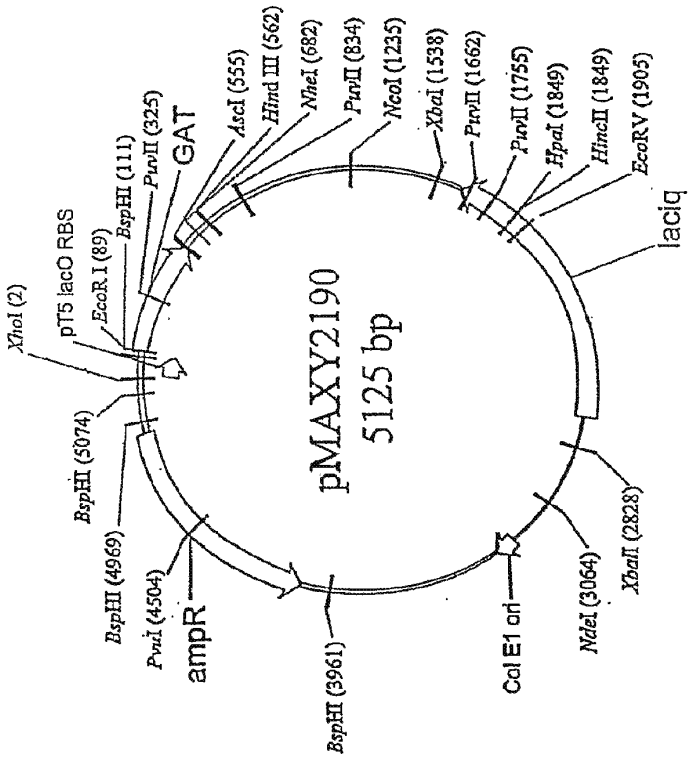
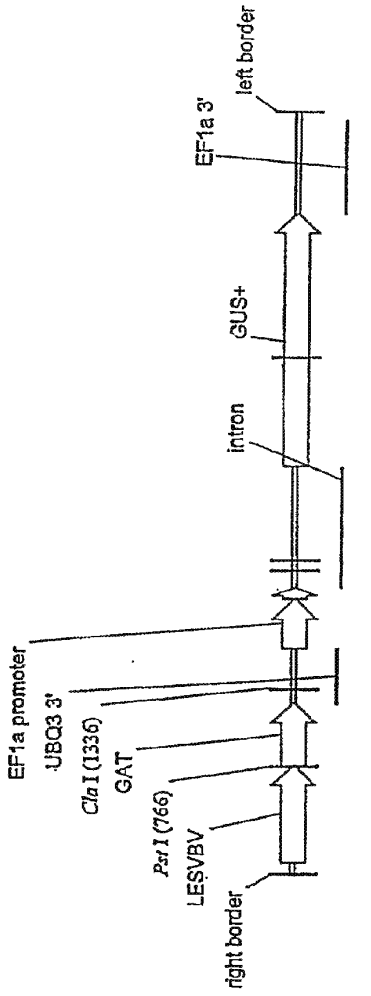


Fig. 11

COR-AAD1-00001639



GAT T-DNA
5546 bp

Fig. 12

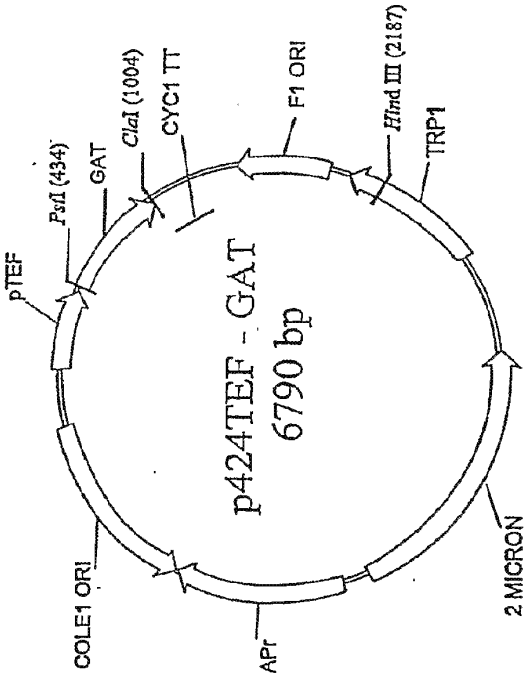


Fig. 13

COR-AAD1-00001641

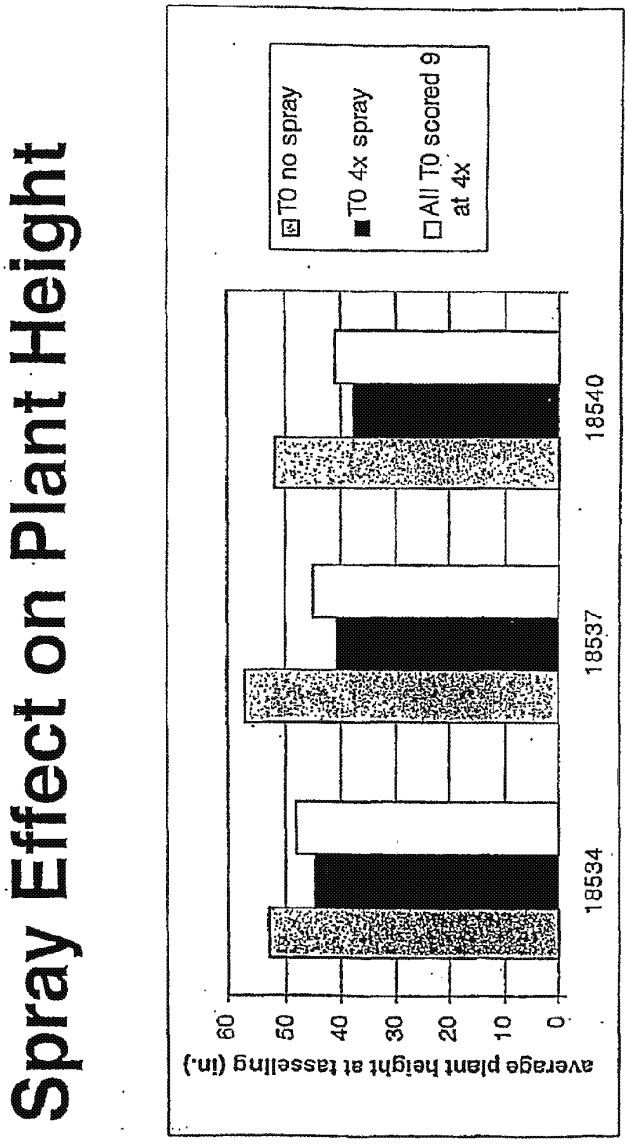


Fig. 14

COR-AAD1-00001642

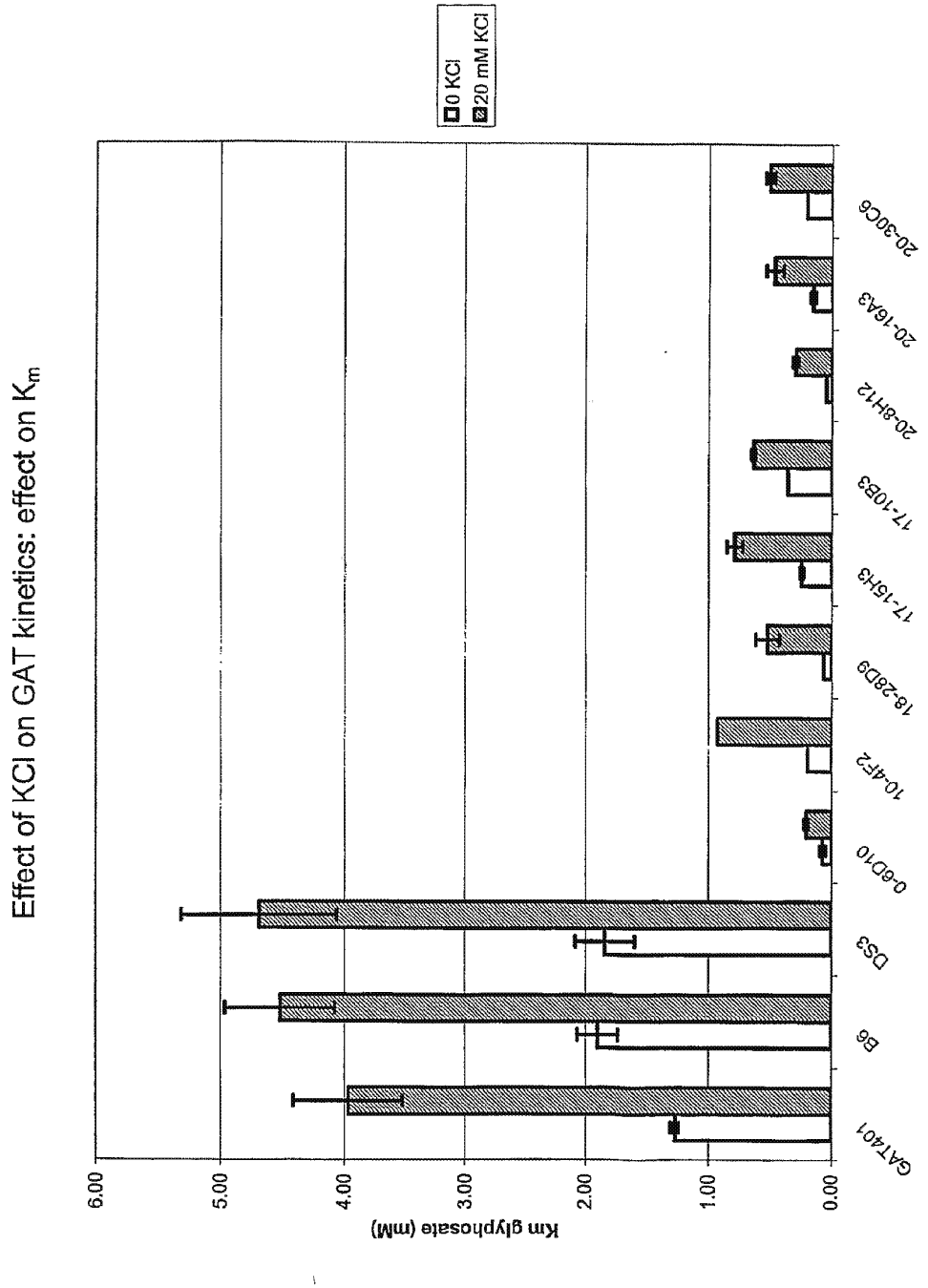


Fig. 15A

COR-AAD1-00001643

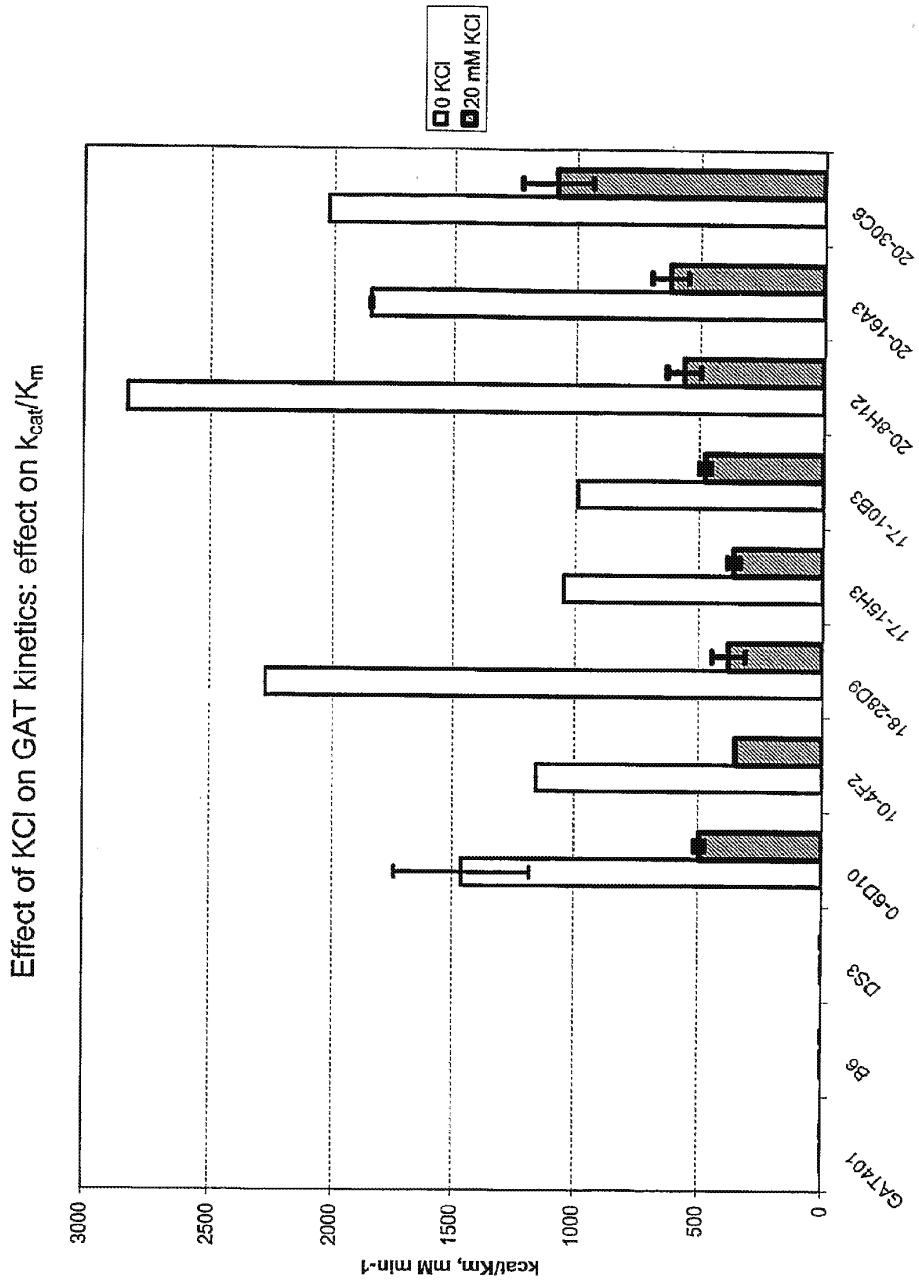

Effect of KCl on GAT kinetics: effect on K_{cat}/K_m
0 KCl
20 mM KCl
kcat/Km, mM min-1
3000
2500
2000
1500
1000
500
0
GAT401
B6
DS3
0-6D10
10-4F2
18-28D9
17-15H3
17-10B3
20-8H12
20-16A3
20-30C6
GAT enzyme

Fig. 15B

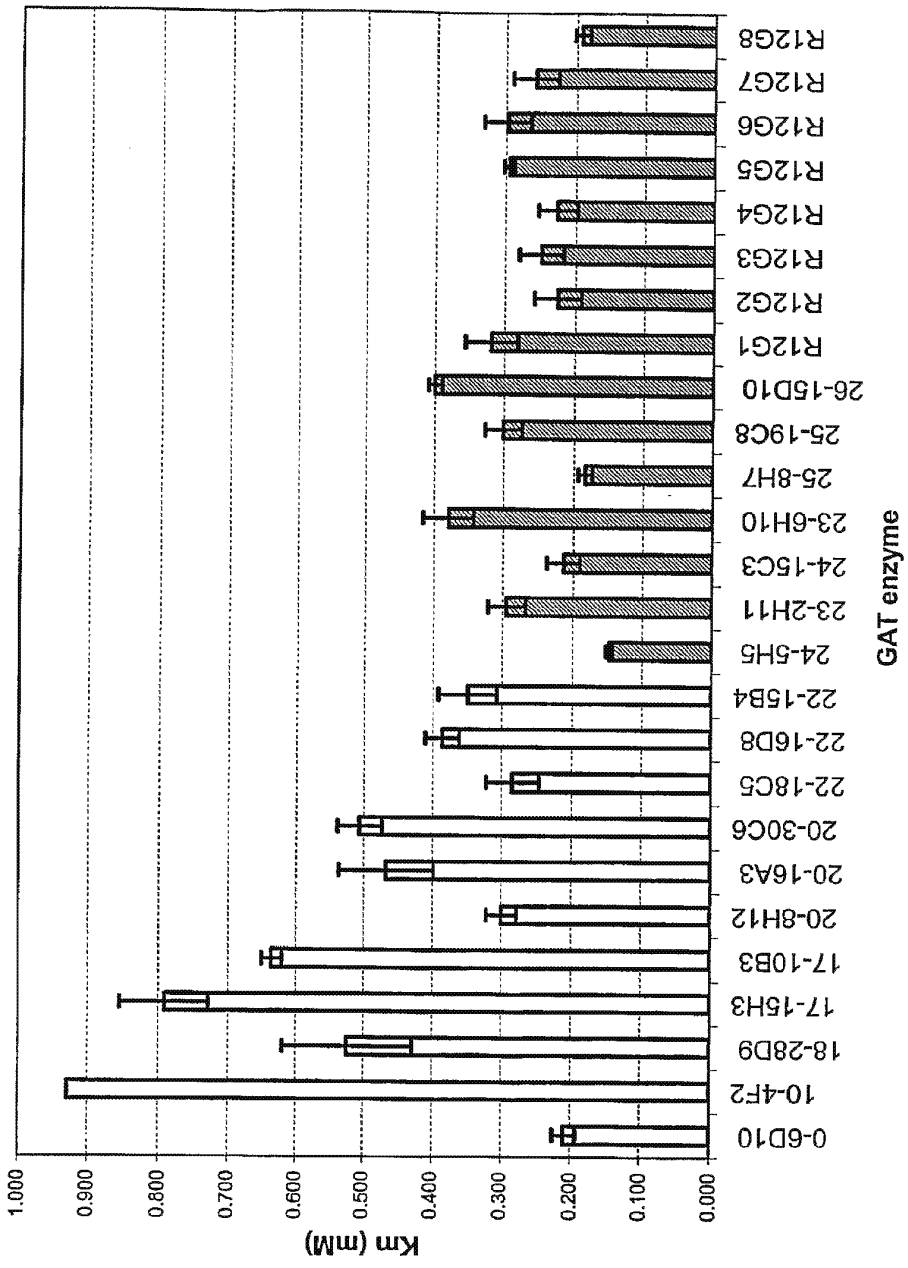


Fig. 16A

COR-AAD1-00001645

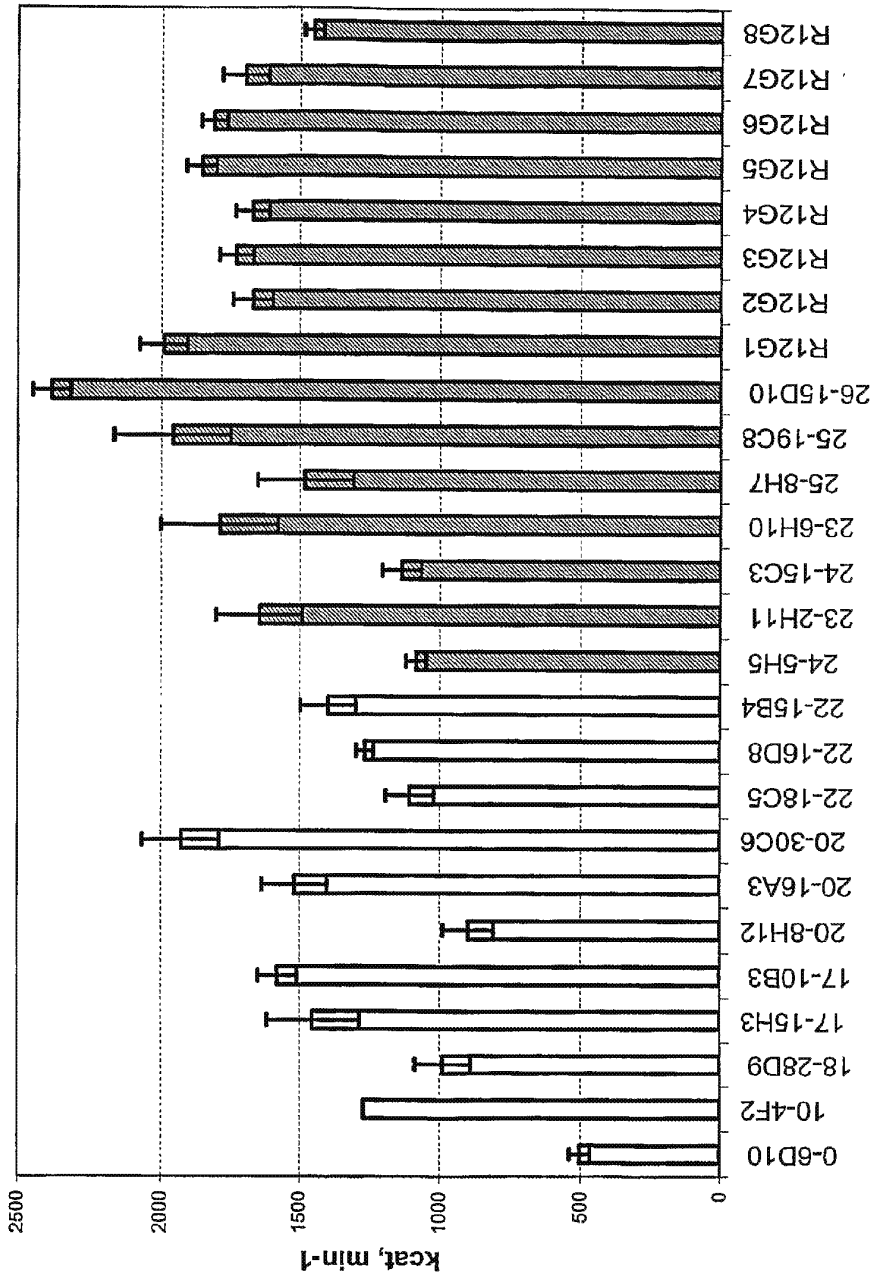


Fig. 16B

COR-AAD1-00001646

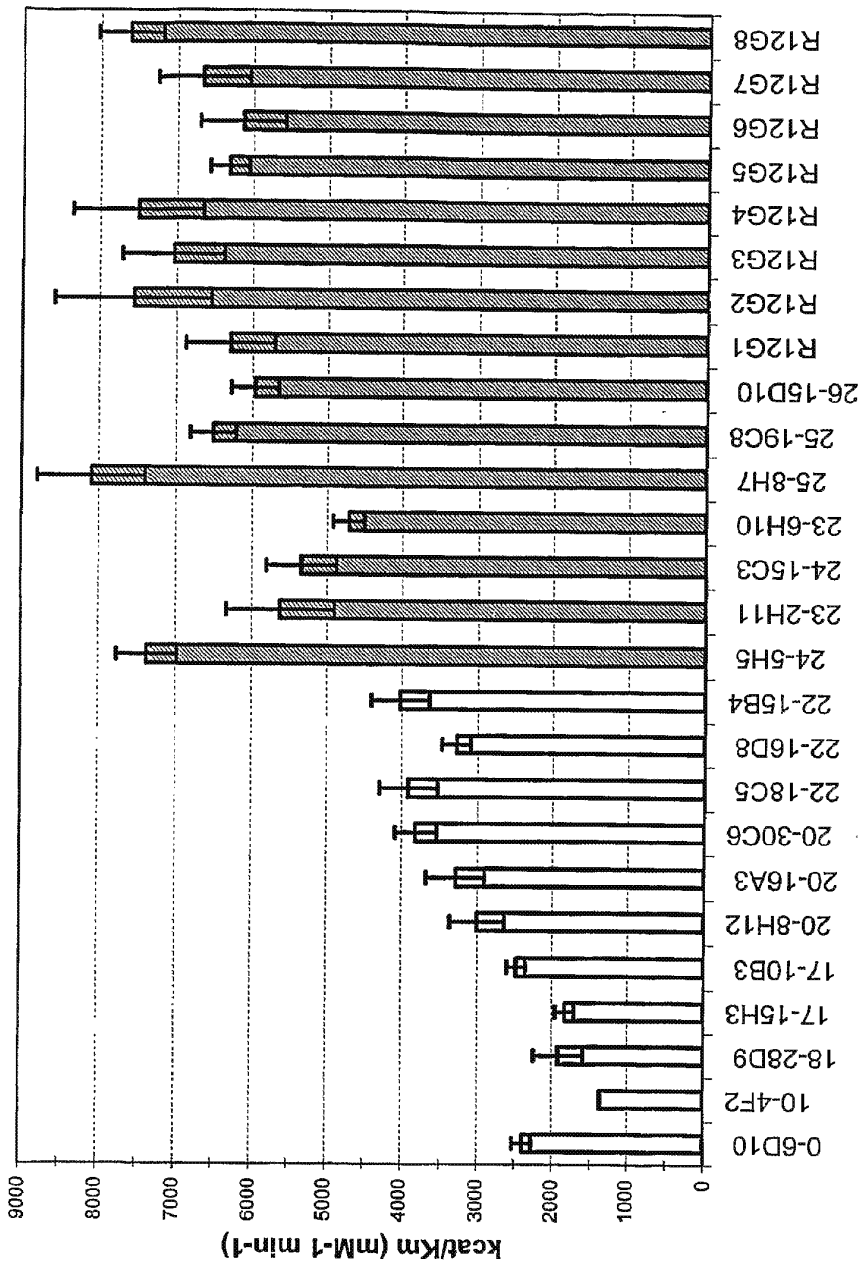


Fig. 16C

COR-AAD1-00001647

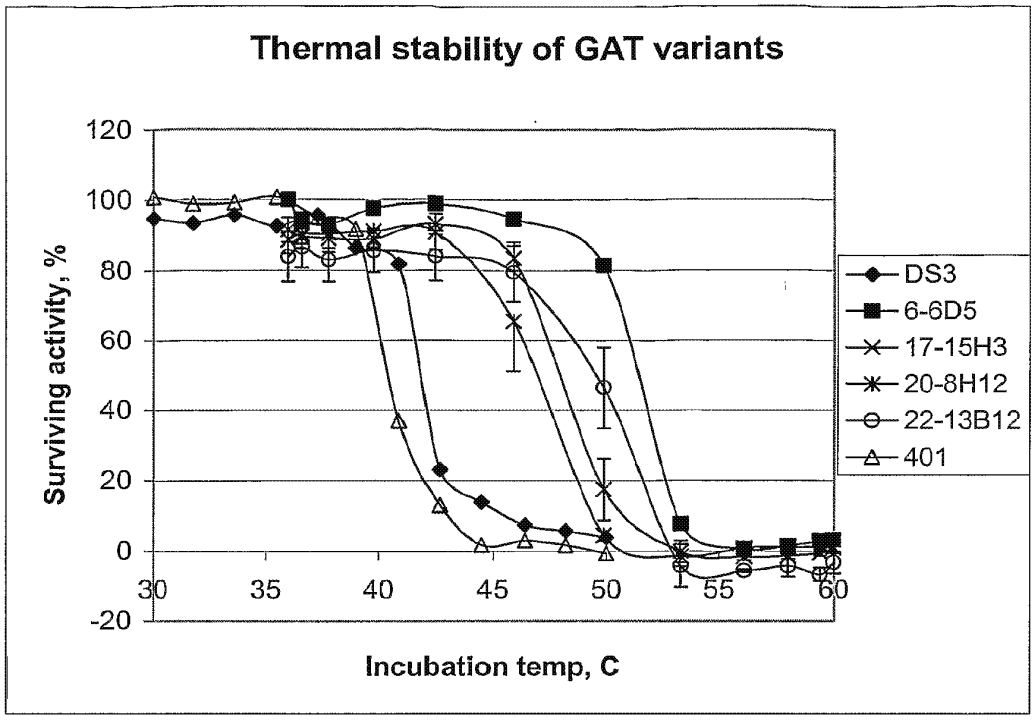


FIGURE 17

COR-AAD1-00001648

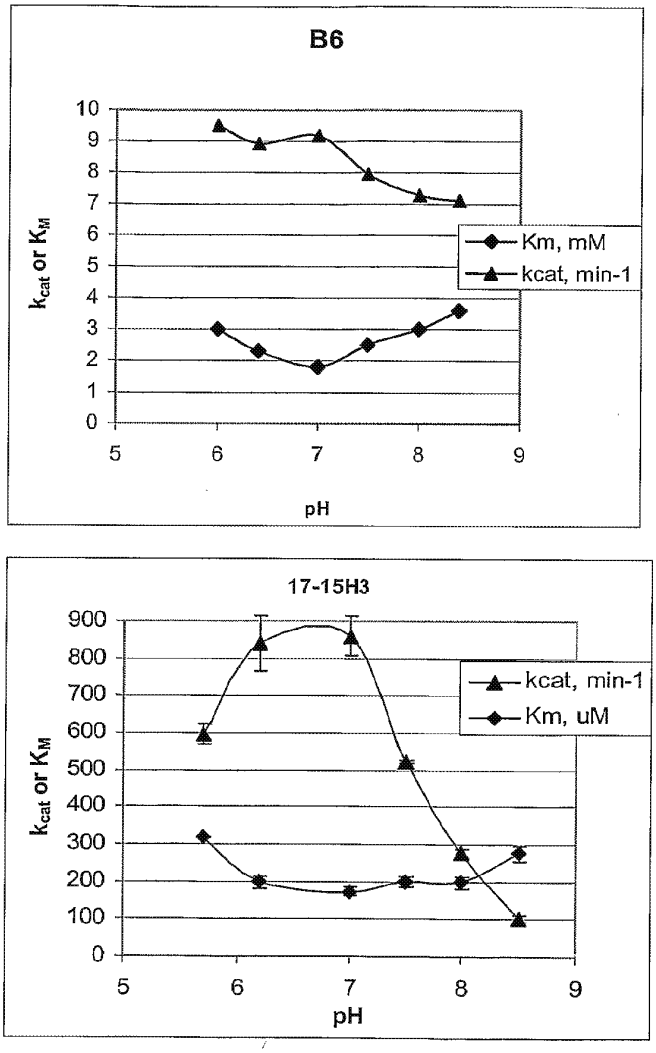


FIGURE 18

COR-AAD1-00001649

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 12 17 9273

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	WO 02/36782 A2 (MAXYGEN INC [US]; PIONEER HI BRED INT [US]; CASTLE LINDA A [US]; SIEHL) 10 May 2002 (2002-05-10) * the whole document * -----	1-15	INV. C12N15/54 C12N15/82 C12N9/10 A01H5/00
X,P	WO 03/092360 A2 (VERDIA INC [US]; PIONEER HI BRED INT [US]) 13 November 2003 (2003-11-13) * sequence 282 * -----	1-15	

TECHNICAL FIELDS SEARCHED (IPC)

C12N
A01H

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	7 November 2012	Maddox, Andrew

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

1

COR-AAD1-00001650

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 12 17 9273

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

07-11-2012

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0236782 A2	10-05-2002	AR 035595 A1	16-06-2004
		AR 064756 A2	22-04-2009
		AR 070289 A2	25-03-2010
		AU 2018102 A	15-05-2002
		BG 107758 A	30-07-2004
		BR 0115046 A	12-04-2005
		CA 2425956 A1	10-05-2002
		CN 1531594 A	22-09-2004
		CN 101684458 A	31-03-2010
		CN 102212534 A	12-10-2011
		CZ 20031120 A3	12-11-2003
		EP 1399566 A2	24-03-2004
		HR P20030439 A2	31-12-2008
		IL 155599 A	27-09-2011
		JP 2004534505 A	18-11-2004
		JP 2008206519 A	11-09-2008
		JP 2010142234 A	01-07-2010
		MX PA03003810 A	15-10-2004
		NZ 526148 A	30-09-2005
		PL 366144 A1	24-01-2005
		RS 32703 A	15-12-2006
		SK 5222003 A3	01-12-2004
		WO 0236782 A2	10-05-2002
		ZA 200303138 A	25-04-2005
WO 03092360 A2	13-11-2003	AR 039501 A1	23-02-2005
		AR 069485 A2	27-01-2010
		AU 2003234328 A1	17-11-2003
		WO 03092360 A2	13-11-2003

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

COR-AAD1-00001651

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 0029596 A **[0006]**
- US 517933 A **[0133] [0166]**
- US 09710686 A **[0133]**
- US 09710686 B **[0166]**
- US 6015891 A **[0169]**
- US 4683202 A, Mullis **[0176] [0248]**
- US 5426039 A, Wallace **[0176] [0248]**
- US 245354 P **[0252]**
- US 5683439 A **[0262]**
- US 5689049 A **[0264]**
- US 5689051 A **[0264]**
- US 09723398 P **[0264]**
- US 60098230 P **[0264]**
- US 5955310 A, Widner **[0281]**
- US 20020058249 A **[0287]**
- US 5605793 A, Stemmer **[0291]**
- US 5811238 A, Stemmer **[0291]**
- US 5830721 A, Stemmer **[0291]**
- US 5834252 A, Stemmer **[0291]**
- US 5837458 A, Minshull **[0291]**
- WO 9522625 A, Stemmer and Crameri **[0291]**
- WO 9633207 A, Stemmer and Lipschutz **[0291]**
- WO 9720078 A, Stemmer and Crameri **[0291]**
- WO 9735966 A, Minshull and Stemmer **[0291]**
- WO 9941402 A, Punnonen **[0291]**
- WO 9941383 A, Punnonen **[0291]**
- WO 9941369 A, Punnonen **[0291]**
- WO 9941368 A, Punnonen **[0291]**
- EP 752008 A, Stemmer and Crameri **[0291]**
- EP 0932670 A, Stemmer **[0291]**
- WO 9923107 A, Stemmer **[0291]**
- WO 9921979 A, Apt **[0291]**
- WO 9831837 A, del Cardayre **[0291] [0296]**
- WO 9827230 A, Patten and Stemmer **[0291]**
- WO 9813487 A, Stemmer **[0291]**
- WO 0000632 A **[0291]**
- WO 0009679 A **[0291]**
- WO 9842832 A, Arnold **[0291]**
- WO 9929902 A, Arnold **[0291]**
- WO 9841653 A, Vind **[0291]**
- WO 9841622 A, Borchert **[0291]**
- WO 9842727 A, Pati and Zarling **[0291]**
- WO 0018906 A, Patten **[0291]**
- WO 0004190 A, del Cardayre **[0291] [0296]**
- WO 0042561 A, Crameri **[0291] [0292] [0297]**
- WO 0042559 A, Selifonov and Stemmer **[0291] [0292] [0297] [0298]**
- WO 0042560 A, Selifonov **[0291] [0292] [0297] [0298]**
- WO 0123401 A, Welch **[0291] [0297]**
- WO 0164864 A **[0291] [0299]**
- US SN09407800 A **[0292]**
- US SN09166188 A **[0292]**
- US 6379964 B **[0292]**
- US 6376246 B **[0292]**
- US 6436675 B **[0292]**
- US SN09618579 A **[0292]**
- US SN60186482 P **[0292]**
- US 5965408 A **[0302]**
- US 5756316 A **[0310]**
- US 5783431 A, Peterson **[0312]**
- US 5824485 A, Thompson **[0312]**
- US 5958672 A, Short **[0312]**
- US 5939250 A, Short **[0314]**
- WO 9910539 A **[0314] [0315]**
- WO 9858085 A **[0314]**
- WO 0046344 A **[0316]**
- US 5830650 A, Crea, R **[0316]**
- US 5798208 A **[0316]**
- EP 0527809 B1 **[0316]**
- US 5591616 A **[0349] [0375]**
- US 5550318 A **[0350] [0407]**
- US 5981840 A **[0350] [0372]**
- WO 8802405 A **[0351]**
- WO 9943838 A **[0357]**
- US 6072050 A **[0357]**
- US 5659026 A **[0357]**
- US 5608149 A **[0357]**
- US 5608144 A **[0357]**
- US 5604121 A **[0357]**
- US 5569597 A **[0357]**
- US 5466785 A **[0357]**
- US 5399680 A **[0357]**
- US 5268463 A **[0357]**
- US 5608142 A **[0357]**
- US 6177611 B **[0357]**
- US 5814618 A **[0358]**
- US 5789156 A **[0358]**
- US 5633636 A **[0361]**
- US 104706 A **[0361]**
- US 5837876 A **[0361]**
- US 5750386 A **[0361]**
- US 5459252 A **[0361]**
- US 5401836 A **[0361]**
- US 5110732 A **[0361]**
- US 5023179 A **[0361]**
- US 6225529 B **[0362]**
- WO 0012733 A **[0362]**

COR-AAD1-00001652

- US 5889191 A **[0371]**
- US 5889190 A **[0371]**
- US 5866785 A **[0371]**
- US 5589367 A **[0371]**
- US 5316931 A **[0371]**
- US 5563055 A **[0372]**
- US 4945050 A **[0372]**
- US 5879918 A **[0372]**
- US 5886244 A **[0372]**
- US 5932782 A **[0372]**
- WO 0028058 A **[0372]**
- US 5240855 A **[0372]**
- US 5322783 A **[0372]**
- US 5324646 A **[0372]**
- US 5736369 A **[0372]**
- US 5187073 A **[0375]**
- US 5366892 A **[0377]**
- US 5747450 A **[0377]**
- US 5737514 A **[0377]**
- US 5723756 A **[0377]**
- US 5593881 A **[0377]**
- US 5981722 A **[0377]**
- US 6232529 B **[0377] [0378]**
- US 5990389 A **[0377]**
- US 5885801 A **[0377]**
- US 5885802 A **[0377]**
- US 5703409 A **[0377]**
- WO 9820122 A **[0377]**
- US 05341001 A **[0377]**
- US 00542901 A **[0377]**
- US 5792931 A **[0378]**
- US 5952544 A **[0378]**
- WO 9411516 A **[0378]**
- US 5602321 A **[0378]**
- US 5583210 A **[0378]**
- WO 9961619 A **[0378]**
- WO 0017364 A **[0378]**
- WO 9925821 A **[0378] [0379]**
- WO 9925854 A **[0379]**
- WO 9925840 A **[0379]**
- WO 9925855 A **[0379]**
- WO 9925853 A **[0379]**
- US 6248876 B1 **[0404]**
- US 5627061 B1 **[0404]**
- US 5804425 B1 **[0404]**
- US 5633435 B1 **[0404]**
- US 5145783 B1 **[0404]**
- US 4971908 B1 **[0404]**
- US 5312910 B1 **[0404]**
- US 5188642 B1 **[0404]**
- US 4940835 B1 **[0404]**
- US 5866775 B1 **[0404]**
- US 6225114 B1 **[0404]**
- US 6130366 B1 **[0404]**
- US 5310667 B1 **[0404]**
- US 4535060 B1 **[0404]**
- US 4769061 B1 **[0404]**
- US 5633448 B1 **[0404]**
- US 5510471 B1 **[0404]**
- US RE36449 B1 **[0404]**
- US RE37287 E **[0404]**
- US 5491288 E **[0404]**
- WO 9704103 A **[0404]**
- WO 0066746 A **[0404]**
- WO 0166704 A **[0404]**
- WO 0066747 A **[0404]**
- US 5776760 A **[0404]**
- US 5463175 A **[0404]**
- US 6245968 B1 **[0405]**
- US 6268549 B1 **[0405]**
- US 6069115 B1 **[0405]**
- WO 9923886 A **[0405]**
- US 5605011 A **[0406]**
- US 5013659 A **[0406]**
- US 5141870 A **[0406]**
- US 5767361 A **[0406]**
- US 5731180 A **[0406]**
- US 5304732 A **[0406]**
- US 4761373 A **[0406]**
- US 5331107 A **[0406]**
- US 5928937 A **[0406]**
- US 5378824 A **[0406]**
- WO 9633270 A **[0406]**
- US 5969213 A **[0407]**
- US 5489520 A **[0407]**
- US 5874265 A **[0407]**
- US 5919675 A **[0407]**
- US 5561236 A **[0407]**
- US 5648477 A **[0407]**
- US 5646024 A **[0407]**
- US 6177616 B1 **[0407]**
- US 5879903 B1 **[0407]**
- US 6288306 B1 **[0408]**
- US 6282837 B1 **[0408]**
- US 5767373 B1 **[0408]**
- WO 0112825 A **[0408]**
- US 5981849 A **[0442]**

Non-patent literature cited in the description

- New weed control opportunities: Development of soybeans with a Round UP Ready™ gene" and Vasil (1996) "Phosphinothricin-resistant crops. **PADGETTE et al.** Herbicide-Resistant Crops. CRC Press, 1996, 54-8485-91 **[0002]**
- **HATTORI et al.** *Mol. Gen. Genet.,* 1995, vol. 246, 419 **[0002]**
- **SHIOTA et al.** *Plant Physiol.,* 1994, vol. 106, 17 **[0002]**

COR-AAD1-00001653

- **AONO et al.** *Plant Cell Physiol.,* 1995, vol. 36, 1687 **[0002]**
- **DATTA et al.** *Plant Mol. Biol.,* 1992, vol. 20, 619 **[0002]**
- Fundamental Immunology. Raven Press, 1998 **[0058]**
- **CHRISPEELS.** *Ann. Rev. Plant Phys. Plant Mol. Biol.,* 1991, vol. 42, 21-53 **[0060]**
- **RAIKHEL.** *Plant Phys.,* 1992, vol. 100, 1627-1632 **[0060]**
- A model of evolutionary change in proteins. **DAYHOFF et al.** Atlas of Protein Sequence and Structure. Natl. Biomed. Res. Found, 1978, vol. 5, 345-352 **[0080]**
- **HENIKOFF et al.** *Proc. Nat'l. Acad. Sci. USA,* 1992, vol. 89, 10915-10919 **[0080]**
- **ALTSCHUL et al.** *Nucl. Acids Res.,* 1997, vol. 25, 3389-3402 **[0080]**
- **WALKER.** Protein Protocols on CD-ROM. Humana Press, 1998 **[0136] [0280]**
- **STEWART et al.** Solid-Phase Peptide Synthesis. WH Freeman Co, 1969 **[0137]**
- **MERRIFIELD.** *J. Am. Chem. Soc.,* 1963, vol. 85, 2149-2154 **[0137]**
- **COLIGAN.** Current Protocols in Immunology. Wiley/Greene, 1991 **[0140]**
- **HARLOW ; LANE.** Antibodies: A Laboratory Manual. Cold Spring Harbor Press, 1989 **[0140]**
- Basic and Clinical Immunology. Lange Medical Publications **[0140]**
- **GODING.** Monoclonal Antibodies: Principles and Practice. Academic Press, 1986 **[0140]**
- **KOHLER ; MILSTEIN.** *Nature,* 1975, vol. 256, 495-497 **[0140]**
- **HUSE et al.** *Science,* 1989, vol. 246, 1275-1281 **[0140]**
- **WARD et al.** *Nature,* 1989, vol. 341, 544-546 **[0140]**
- Antibody Engineering. Freeman and Company, 1995 **[0141]**
- **MCCAFFERTY et al.** Antibody Engineering, A Practical Approach. IRL at Oxford Pres, 1996 **[0141]**
- **PAUL.** Antibody Engineering Protocols. Humana Press, 1995 **[0141]**
- **HARLOW ; LANE.** Antibodies, A Laboratory Manual. Cold Spring Harbor Publications, 1988 **[0149]**
- **NESS et al.** *Nature Biotech.,* 1999, vol. 17, 893-96 **[0162]**
- **ZHANG et al.** *Gene,* 1991, vol. 105, 61-72 **[0168]**
- **MURRAY et al.** *Nucl. Acids Res.,* 1989, vol. 17, 477-508 **[0169]**
- **DALPHIN et al.** *Nucl. Acids Res.,* 1996, vol. 24, 216-218 **[0169]**
- **BEAUCAGE et al.** *Tetrahedron Letters,* 1981, vol. 22, 1859-69 **[0173]**
- **MATTHES et al.** *EMBO J.,* 1984, vol. 3, 801-05 **[0173]**
- **CARRUTHERS et al.** *Cold Spring Harbor Symp. Quant. Biol.,* 1982, vol. 47, 411-418 **[0175]**
- **ADAMS et al.** *J. Am. Chem. Soc.,* 1983, vol. 105, 661 **[0175]**
- Guide to Molecular Cloning Techniques. **BERGER ; KIMMEL.** Methods in Enzymology. Academic Press, Inc, vol. 152 **[0176] [0248]**
- **SAMBROOK et al.** Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Laboratory, 1989, vol. 1-3 **[0176] [0248]**
- Current Protocols in Molecular Biology. Greene Publishing Associates, Inc. and John Wiley & Sons, Inc, 2000 **[0176]**
- PCR Protocols: A Guide to Methods and Applications. Academic Press Inc, 1990 **[0176] [0248]**
- **ARNHEIM ; LEVINSON.** *Chemical and Engineering News,* 01 October 1990, 36-47 **[0176]**
- *The Journal Of NIH Research,* 1991, vol. 3, 81-94 **[0176] [0248]**
- **KWOH et al.** *Proc. Nat'l. Acad. Sci. USA,* 1989, vol. 86, 1173 **[0176]**
- **GUATELLI et al.** *Proc. Nat'l. Acad. Sci. USA,* 1990, vol. 87, 1874 **[0176] [0248]**
- **LOMELL et al.** *J. Clin. Chem.,* 1989, vol. 35, 1826 **[0176]**
- **LANDEGREN et al.** *Science,* 1988, vol. 241, 1077-1080 **[0176] [0248]**
- **VAN BRUNT.** *Biotechnology,* 1990, vol. 8, 291-294 **[0176] [0248]**
- **WU ; WALLACE.** *Gene,* 1989, vol. 4, 560 **[0176] [0248]**
- **BARRINGER et al.** *Gene,* 1990, vol. 89, 117 **[0176]**
- **SOOKNANAN ; MALEK.** *Biotechnology,* 1995, vol. 13, 563-564 **[0176] [0248]**
- **CHENG et al.** *Nature,* 1994, vol. 369, 684-685 **[0176] [0248]**
- Overview of principles of hybridization and the strategy of nucleic acid probe assays. **TIJSSEN.** Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes. Elsevier, 1993 **[0224]**
- **HAMES ; HIGGINS.** Gene Probes 1. IRL Press at Oxford University Press, 1995 **[0224]**
- **HAMES ; HIGGINS.** Gene Probes 2. IRL Press at Oxford University Press, 1995 **[0224]**
- Molecular Biomethods Handbook. Humana Press, Inc, 1998 **[0228]**
- **SORG, U. et al.** *Nucleic Acids Res.,* 11 September 1991, vol. 19 (17 **[0242]**
- Current Protocols in Molecular Biology. Current Protocols, 1999 **[0248]**
- **ARNHEIM ; LEVINSON.** *C&EN,* 01 October 1990, 36-47 **[0248]**
- **KWOH et al.** *Proc. Natl. Acad. Sci. USA,* 1989, vol. 86, 1173 **[0248]**
- **LOMELL et al.** *J. Clin. Chem,* 1989, vol. 35, 1826 **[0248]**
- **BARRINGER.** *Gene,* 1990, vol. 89, 117 **[0248]**
- **FRESHNEY.** Culture of Animal Cells: A Manual of Basic Technique. Wiley- Liss, 1994 **[0249]**

COR-AAD1-00001654

- **PAYN et al.** Plant Cell and Tissue Culture in Liquid Systems. John Wiley & Sons, Inc, 1992 **[0250]**
- Plant Cell, Tissue and Organ Culture: Fundamental Methods. Springer Lab Manual. Springer-Verlag, 1995 **[0250] [0279]**
- The Handbook of Microbiological Media. CRC Press, 1993 **[0250] [0279]**
- **VAN HEEKE ; SCHUSTER.** *J. Biol. Chem.,* 1989, vol. 264, 5503-5509 **[0256]**
- **GRANT et al.** *Methods in Enzymology,* 1987, vol. 153, 516-544 **[0257]**
- **LOGAN ; SHENK.** *Proc. Nat'l. Acad. Sci. USA,* 1984, vol. 81, 3655-3659 **[0258]**
- **ROGERS et al.** *Meth. Enzymol.,* 1987, vol. 153, 253-277 **[0260]**
- **SCHARDL et al.** *Gene,* 1987, vol. 61, 1-11 **[0260]**
- **BERGER et al.** *Proc. Nat'l. Acad. Sci. U.S.A.,* 1989, vol. 86, 8402-8406 **[0260]**
- **BORONAT et al.** *Plant Sci.,* 1986, vol. 47, 95-102 **[0264]**
- **REINA et al.** *Nucleic Acids Res.,* 1990, vol. 18 (21), 6426 **[0264]**
- **KLOESGEN et al.** *Mol. Gen. Genet.,* 1986, vol. 203, 237-244 **[0264]**
- **HERRARA-ESTRELLA et al.** *Nature,* 1983, vol. 303, 209-213 **[0265]**
- **ODELL et al.** *Nature,* 1985, vol. 313, 810-812 **[0265] [0357]**
- **DEIKMAN ; FISCHER.** *EMBO J.,* 1988, vol. 7, 3315-3327 **[0265]**
- **MESSING et al.** Genetic Engineering in Plants. 1983, 221-227 **[0266]**
- **BUCHMAN ; BERG.** *Mol. Cell Biol.,* 1988, vol. 8, 4395-4405 **[0268]**
- **CALLIS et al.** *Genes Dev.,* 1987, vol. 1, 1183-1200 **[0268]**
- The Maize Handbook. Springer, 1994 **[0268] [0384]**
- **SCHARF et al.** *Results Probl. Cell Differ.,* 1994, vol. 20, 125-62 **[0270]**
- **BITTNER et al.** *Methods in Enzymol,* 1987, vol. 153, 516-544 **[0270]**
- **DAVIS et al.** *Basic Methods in Molecular Biology* **[0273]**
- **WILSON et al.** *Cell,* 1984, vol. 37, 767 **[0277]**
- **PORATH et al.** *Protein Expression and Purification,* 1992, vol. 3, 263-281 **[0277]**
- **FRESHNEY.** Culture of Animal Cells: A Manual of Basic Technique. Wiley-Liss, 1994 **[0279]**
- **DOYLE ; GRIFFITHS.** Mammalian Cell Culture: Essential Techniques. John Wiley and Sons, 1997 **[0279]**
- **HUMASON.** Animal Tissue Techniques. W.H. Freeman and Company, 1979 **[0279]**
- **RICCIARDELLI et al.** *In vitro Cell Dev. Biol.,* 1989, vol. 25, 1016-1024 **[0279]**
- **PAYNE et al.** *Plant Cell and Tissue Culture in Liquid Systems,* 1992 **[0279]**
- Plant Gene Transfer and Expression Protocols. Humana Press, 1984 **[0279]**
- Plant Molecular Biology. Bios Scientific Publishers, 1993 **[0279] [0354]**
- Life Science Research Cell Culture Catalogue. Sigma-Aldrich, Inc, 1998 **[0279]**
- The Plant Culture Catalogue. Sigma-Aldrich, Inc, 1997 **[0279]**
- **SANDANA.** Bioseparation of Proteins. Academic Press, Inc, 1997 **[0280]**
- **BOLLAG et al.** Protein Methods. Wiley-Liss, 1996 **[0280] [0283]**
- **WALKER.** The Protein Protocols Handbook. Humana Press, 1996 **[0280] [0283]**
- **HARRIS ; ANGAL.** Protein Purification Applications: A Practical Approach. IRL Press at Oxford, 1990 **[0280]**
- **HARRIS ; ANGAL.** Protein Purification Methods: A Practical Approach. IRL Press at Oxford **[0280]**
- **SCOPES.** Protein Purification: Principles and Practice. Springer Verlag, 1993 **[0280]**
- **JANSON ; RYDEN.** Protein Purification: Principles, High Resolution Methods and Applications. Wiley-VCH, 1998 **[0280]**
- **CHANG ; COHEN.** *Mol. Gen. Genet.,* 1979, vol. 168, 111 **[0281]**
- **YOUNG ; SPIZIZIN.** *J. Bacteriol.,* 1961, vol. 81, 823 **[0281]**
- **DAVIDOFF-ABELSON.** *J. Mol. Biol.,* 1971, vol. 56, 209 **[0281]**
- **SHIGEKAWA ; DOWER.** *Biotechniques,* 1988, vol. 6, 742 **[0281]**
- **KOEHLER ; THORNE.** *J. Bacteriol,* 1987, vol. 169, 5271 **[0281]**
- **TYMMS.** In vitro Transcription and Translation Protocols: Methods in Molecular Biology. Garland Publishing, 1995, vol. 37 **[0284]**
- **SOONG, N. et al.** *Nat. Genet.,* 2000, vol. 25 (4), 436-39 **[0288]**
- **STEMMER et al.** *Tumor Targeting,* 1999, vol. 4, 1-4 **[0288]**
- **NESS et al.** *Nature Biotech.,* 1999, vol. 17, 893-896 **[0288]**
- **CHANG et al.** *Nature Biotech.,* 1999, vol. 17, 793-797 **[0288]**
- **MINSHULL ; STEMMER.** *Current Opinion in Chemical Biology,* 1999, vol. 3, 284-290 **[0288]**
- **CHRISTIANS et al.** *Nature Biotech.,* 1999, vol. 17, 259-264 **[0288]**
- **CRAMERI et al.** *Nature,* 1998, vol. 391, 288-291 **[0288]**
- **CRAMERI et al.** *Nature Biotech.,* 1997, vol. 15, 436-438 **[0288]**
- **ZHANG et al.** *Proc. Nat'l. Acad. Sci. USA,* 1997, vol. 94, 4504-4509 **[0288]**
- **PATTEN et al.** *Current Opinion in Biotech.,* 1997, vol. 8, 724-733 **[0288]**

COR-AAD1-00001655

- **CRAMERI et al.** *Nature Med,* 1996, vol. 2, 100-103 **[0288]**
- **CRAMERI et al.** *Nature Biotech.,* 1996, vol. 14, 315-319 **[0288]**
- **GATES et al.** *J. Mol. Biol.,* 1996, vol. 255, 373-386 **[0288]**
- Sexual PCR and Assembly PCR. **STEMMER.** The Encyclopedia of Molecular Biology. VCH Publishers, 1996, 447-457 **[0288]**
- **CRAMERI ; STEMMER.** *BioTechniques,* 1995, vol. 18, 194-195 **[0288]**
- **STEMMER et al.** *Gene,* 1995, vol. 164, 49-53 **[0288]**
- **STEMMER.** *Science,* 1995, vol. 270, 1510 **[0288]**
- **STEMMER.** *Biol Technology,* 1995, vol. 13, 549-553 **[0288]**
- **STEMMER.** *Nature,* 1994, vol. 370, 389-391 **[0288]**
- **STEMMER.** *Proc. Nat'l. Acad. Sci. USA,* 1994, vol. 91, 10747-10751 **[0288] [0294]**
- **LING et al.** Approaches to DNA mutagenesis: an overview. *Anal Biochem.,* 1997, vol. 254 (2), 157-178 **[0289]**
- **DALE et al.** Oligonucleotide-directed random mutagenesis using the phosphorothioate method. *Methods Mol. Biol.,* 1996, vol. 57, 369-374 **[0289]**
- **SMITH.** In vitro mutagenesis. *Ann. Rev. Genet.,* 1985, vol. 19, 423-462 **[0289]**
- **BOTSTEIN ; SHORTLE.** Strategies and applications of in vitro mutagenesis. *Science,* 1985, vol. 229, 1193-1201 **[0289]**
- **CARTER.** Site-directed mutagenesis. *Biochem. J.,* 1986, vol. 237, 1-7 **[0289]**
- The efficiency of oligonucleotide directed mutagenesis. **KUNKEL.** Nucleic Acids & Molecular Biology. Springer Verlag, 1987 **[0289]**
- **KUNKEL.** Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA,* 1985, vol. 82, 488-492 **[0289]**
- **KUNKEL et al.** Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods in Enzymol.,* 1987, vol. 154, 367-382 **[0289]**
- **BASS et al.** Mutant Trp repressors with new DNA-binding specificities. *Science,* 1988, vol. 242, 240-245 **[0289]**
- *Methods in Enzymol.,* 1983, vol. 100, 468-500 **[0289]**
- *Methods in Enzymol.,* 1987, vol. 154, 329-350 **[0289]**
- **ZOLLER ; SMITH.** Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment. *Nucleic Acids Res.,* 1982, vol. 10, 6487-6500 **[0289]**
- **ZOLLER ; SMITH.** Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. *Methods in Enzymol.,* 1983, vol. 100, 468-500 **[0289]**
- **ZOLLER ; SMITH.** Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. *Methods in Enzymol.,* 1987, vol. 154, 329-350 **[0289]**
- **TAYLOR et al.** The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA. *Nucl. Acids Res.,* 1985, vol. 13, 8749-8764 **[0289]**
- **TAYLOR et al.** The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA. *Nucl. Acids Res.,* 1985, vol. 13, 8765-8787 **[0289]**
- **NAKAMAYE ; ECKSTEIN.** Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis. *Nucl. Acids Res.,* 1986, vol. 14, 9679-9698 **[0289]**
- **SAYERS et al.** Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis. *Nucl. Acids Res.,* 1988, vol. 16, 791-802 **[0289]**
- **SAYERS et al.** Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide. *Nucl. Acids Res.,* 1988, vol. 16, 803-814 **[0289]**
- **KRAMER et al.** The gapped duplex DNA approach to oligonucleotide-directed mutation construction. *Nucl. Acids Res.,* 1984, vol. 12, 9441-9456 **[0289]**
- **KRAMER ; FRITZ.** Oligonucleotide-directed construction of mutations via gapped duplex DNA. *Methods in Enzymol.,* 1987, vol. 154, 350-367 **[0289]**
- **KRAMER et al.** Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations. *Nucl. Acids Res.,* 1988, vol. 16, 7207 **[0289]**
- **FRITZ et al.** Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro. *Nucl. Acids Res.,* 1988, vol. 16, 6987-6999 **[0289]**
- **KRAMER et al.** Point Mismatch Repair. *Cell,* 1984, vol. 38, 879-887 **[0290]**
- **CARTER et al.** Improved oligonucleotide site-directed mutagenesis using M13 vectors. *Nucl. Acids Res.,* 1985, vol. 13, 4431-4443 **[0290]**
- **CARTER.** Improved oligonucleotide-directed mutagenesis using M13 vectors. *Methods in Enzymol.,* 1987, vol. 154, 382-403 **[0290]**
- **EGHTEDARZADEH ; HENIKOFF.** Use of oligonucleotides to generate large deletions. *Nucl. Acids Res.,* 1986, vol. 14, 5115 **[0290]**
- **WELLS et al.** Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin. *Phil. Trans. R. Soc. Lond. A,* 1986, vol. 317, 415-423 **[0290]**
- **NAMBIAR et al.** Total synthesis and cloning of a gene coding for the ribonuclease S protein. *Science,* 1984, vol. 223, 1299-1301 **[0290]**
- **SAKAMAR ; KHORANA.** Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin. *Nucl. Acids Res.,* 1988, vol. 14, 6361-6372 **[0290]**

COR-AAD1-00001656

- **WELLS et al.** Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites. *Gene,* 1985, vol. 34, 315-323 **[0290]**
- **GRUNDSTRÖM et al.** Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis. *Nucl. Acids Res.,* 1985, vol. 13, 3305-3316 **[0290]**
- **ARNOLD.** Protein engineering for unusual environments. *Current Opinion in Biotechnology,* 1993, vol. 4, 450-455 **[0290]**
- Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis. *Proc. Natl. Acad. Sci. USA,* vol. 83, 7177-7181 **[0290]**
- *Methods in Enzymology,* vol. 154 **[0290]**
- **OSTERMEIER et al.** A combinatorial approach to hybrid enzymes independent of DNA homology. *Nature Biotech,* 1999, vol. 17, 1205 **[0303]**
- **OSTERMEIER et al.** Combinatorial Protein Engineering by Incremental Truncation. *Proc. Natl. Acad. Sci. USA,* 1999, vol. 96, 3562-67 **[0303]**
- **OSTERMEIER et al.** Incremental Truncation as a Strategy in the Engineering of Novel Biocatalysts. *Biological and Medicinal Chemistry,* 1999, vol. 7, 2139-44 **[0303]**
- **LEUNG et al.** *Technique,* 1989, vol. 1, 11-15 **[0305]**
- **CALDWELL et al.** *PCR Methods Applic,* 1992, vol. 2, 28-33 **[0305]**
- **REIDHAAR-OLSON et al.** *Science,* 1988, vol. 241, 53-57 **[0306]**
- **ARKIN ; YOUVAN.** *Proc. Nat'l. Acad. Sci. USA,* 1992, vol. 89, 7811-7815 **[0307]**
- **DELEGRAVE ; YOUVAN.** *Biotech. Res.,* 1993, vol. 11, 1548-1552 **[0308]**
- **JIRHOLT et al.** Exploiting sequence space: shuffling in vivo formed complementarity determining regions into a master framework. *Gene,* 1998, vol. 215, 471 **[0313]**
- **ARKIN ; YOUVAN.** Optimizing nucleotide mixtures to encode specific subsets of amino acids for semi-random mutagenesis. *Biotechnology,* 1992, vol. 10, 297-300 **[0316]**
- **REIDHAAR-OLSON et al.** Random mutagenesis of protein sequences using oligonucleotide cassettes. *Methods Enzymol.,* 1991, vol. 208, 564-86 **[0316]**
- **LIM ; SAUER.** The role of internal packing interactions in determining the structure and stability of a protein. *J. Mol. Biol.,* 1991, vol. 219, 359-76 **[0316]**
- **BREYER ; SAUER.** Mutational analysis of the fine specificity of binding of monoclonal antibody 51F to lambda repressor. *J. Biol. Chem.,* 1989, vol. 264, 13355-60 **[0316]**
- *Methods in Enzymology,* 1987, vol. 153 **[0347]**
- **WEISING et al.** *Ann. Rev. Genet,* 1988, vol. 22, 421-477 **[0347]**
- Direct DNA Transfer into Intact Plant Cells Via Microprojectile Bombardment. **TOMES et al.** Plant Cell, Tissue and Organ Culture, Fundamental Methods. Springer-Verlag, 1995, 197-213 **[0347]**
- **KLEIN et al.** Transformation of microbes, plants and animals by particle bombardment. *BiolTechnol.,* 1992, vol. 10 (3), 286-291 **[0347]**
- **PASZKOWSKI et al.** *EMBO J.,* 1984, vol. 3, 2717-2722 **[0348] [0372]**
- **FROMM et al.** *Proc. Natl. Acad. Sci.,* 1985, vol. 82, 5824 **[0348]**
- **KLEIN et al.** *Nature,* 1987, vol. 327, 70-73 **[0348] [0451]**
- **HORSCH et al.** *Science,* 1984, vol. 233, 496-498 **[0350]**
- **FRALEY et al.** *Proc. Natl. Acad. Sci.,* 1983, vol. 80, 4803 **[0350]**
- **LICHTENSTEIN ; FULLER.** Genetic Engineering. Academic Press, 1987, vol. 6 **[0351]**
- **LICHTENSTEIN, C. P. ; DRAPER, J.** DNA Cloning. IRI Press, 1985, vol. II **[0351]**
- **FREEMAN et al.** *Plant Cell Physiol.,* 1984, vol. 25, 1353 **[0351]**
- **KINDLE.** *Proc. Natl. Acad. Sci. USA,* 1990, vol. 87, 1228 **[0351]**
- **ZHOU et al.** *Methods in Enzymology,* 1983, vol. 101, 433 **[0352]**
- **D. HESS.** *Intern Rev. Cytol.,* 1987, vol. 107, 367 **[0352]**
- **LUO et al.** *Plant Mol. Biol. Reporter,* 1988, vol. 6, 165 **[0352]**
- **PENA et al.** *Nature,* 1987, vol. 325, 274 **[0352]**
- **NEUHAUS et al.** *Theor. Appl. Genet.,* 1987, vol. 75, 30 **[0352]**
- **BENBROOK et al.** Proceedings Bio Expo. Butterworth, 1986, 27-54 **[0352]**
- **KUCHLER, R.J.** Biochemical Methods in Cell Culture and Virology. Dowden, Hutchinson and Ross, Inc, 1977 **[0353]**
- Plant Gene Transfer and Expression Protocols. Methods in Molecular Biology. Humana Press, 1995, vol. 49 **[0354]**
- **PAYNE et al.** Plant Cell and Tissue Culture in Liquid Systems. John Wiley & Sons, Inc, 1992 **[0354]**
- Plant Cell, Tissue and Ozgan Culture. Fundamental Methods/Springer Lab Manual. Springer-Verlag, 1995 **[0354]**
- The Handbook of Microbiological Media. CRC Press **[0354]**
- Science Research Cell Culture Catalogue. Sigma-Aldrich, Inc, 1998 **[0354]**
- Plant Culture Catalogue and supplement. Sigma-Aldrich, Inc, 1997 **[0354]**
- **MCELROY et al.** *Plant Cell,* 1990, vol. 2, 163-171 **[0357]**
- **CHRISTENSEN et al.** *Plant Mol. Biol.,* 1989, vol. 12, 619-632 **[0357]**
- **CHRISTENSEN et al.** *Plant Mol. Biol.,* 1992, vol. 18, 675-689 **[0357]**
- **LAST et al.** *Theor. Appl. Genet.,* 1991, vol. 81, 581-588 **[0357]**

COR-AAD1-00001657

- **VELTEN et al.** *EMBO J.,* 1984, vol. 3, 2723-2730 **[0357]**
- **SCHENA et al.** *Proc. Natl. Acad. Sci. USA,* 1991, vol. 88, 10421-10425 **[0358]**
- **MCNELLIS et al.** *Plant J.*, 1998, vol. 14 (2), 247-257 **[0358]**
- **GATZ et al.** *Mol. Gen. Genet,* 1991, vol. 227, 229-237 **[0358]**
- **YAMAMOTO et al.** *Plant J.,* 1997, vol. 12 (2), 255-265 **[0359] [0360]**
- **KAWAMATA et al.** *Plant Cell Physiol.,* 1997, vol. 38 (7), 792-803 **[0359]**
- **HANSEN et al.** *Mol. Gen Genet,* 1997, vol. 254 (3), 337-343 **[0359]**
- **RUSSELL et al.** *Transgenic Res.,* vol. 6 (2), 157-168 **[0359]**
- **RINEHART et al.** *Plant Physiol.,* 1996, vol. 112 (3), 1331-1341 **[0359]**
- **VAN CAMP et al.** *Plant Physiol.,* 1996, vol. 112 (2), 525-535 **[0359]**
- **CANEVASCINI et al.** *Plant Physiol.,* 1996, vol. 112 (2), 513-524 **[0359]**
- **YAMAMOTO et al.** *Plant Cell Physiol.*, 1994, vol. 35 (5), 773-778 **[0359] [0360]**
- **LAM.** *Results Probl. Cell Differ.,* 1994, vol. 20, 181-196 **[0359]**
- **OROZCO et al.** *Plant Mol Biol.,* 1993, vol. 23 (6), 1129-1138 **[0359]**
- **MATSUOKA et al.** *Proc Natl. Acad. Sci. USA*, 1993, vol. 90 (20), 9586-9590 **[0359]**
- **GUEVARA-GARCIA et al.** *Plant J.,* 1993, vol. 4 (3), 495-505 **[0359]**
- **KWON et al.** *Plant Physiol.,* 1994, vol. 105, 357-67 **[0360]**
- **GOTOR et al.** *Plant J.,* 1993, vol. 3, 509-18 **[0360]**
- **OROZCO et al.** *Plant Mol. Biol.,* 1993, vol. 23 (6), 1129-1138 **[0360]**
- **MATSUOKA et al.** *Proc. Natl. Acad. Sci. USA,* 1993, vol. 90 (20), 9586-9590 **[0360]**
- **HIRE et al.** *Plant Mol. Biol.,* 1992, vol. 20 (2), 207-218 **[0361]**
- **KELLER et al.** *Cell,* 1991, vol. 3 (10), 1051-1061 **[0361]**
- **SANGER et al.** *Plant Mol. Biol.,* 1990, vol. 14 (3), 433-443 **[0361]**
- **MIAO et al.** *Plant Cell,* 1991, vol. 3 (1), 11-22 **[0361]**
- **BOGUSZ et al.** *Plant Cell,* 1990, vol. 2 (7), 633-641 **[0361]**
- *Plant Science (Limerick,* vol. 79 (1), 69-76 **[0361]**
- **TEERI et al.** *EMBO J.,* 1989, vol. 8 (2), 343-350 **[0361]**
- **KUSTER et al.** *Plant Mol. Biol.*, 1995, vol. 29 (4), 759-772 **[0361]**
- **CAPANA et al.** *Plant Mol. Biol.,* 1994, vol. 25 (4), 681-691 **[0361]**
- **THOMPSON et al.** *BioEssays,* 1989, vol. 10, 108 **[0362]**
- **HERRERA-ESTRELLA et al.** *Nature,* 1983, vol. 303, 209 **[0366]**
- **ODELL et al.** *Nature,* 1985, vol. 313, 810 **[0366]**
- **DEIKMAN ; FISCHER.** *EMBO J,* 1988, vol. 7, 3315 **[0366]**
- **MCELROY ; BRETTELL.** Foreign gene expression in transgenic cereals. *Trends Biotech.,* 1994, vol. 12, 62-68 **[0366]**
- **SHEEN et al.** *The Plant Journal,* 1995, vol. 8, 777 **[0368]**
- **HASHIMOTO et al.** *Plant Physiol.,* 1990, vol. 93, 857 **[0369]**
- Plant Protoplasts. 1994 **[0369]**
- **SAUNDERS et al.** *Applications of Plant In vitro Technology Symposium,* 1993, 16-18 **[0369]**
- **LYZNIK et al.** *BioTechniques,* 1991, vol. 10, 295 **[0369]**
- **DANIELL et al.** *Nature Biotech.*, 1998, vol. 16, 346 **[0370]**
- **O'NEILL et al.** *The Plant Journal,* 1993, vol. 3, 729 **[0370]**
- **MALIGA.** *TIBTECH,* 1993, vol. 11, 1 **[0370]**
- **WEISING et al.** *Ann. Rev. Genet.,* 1988, vol. 22, 421 **[0371]**
- **CROSSWAY et al.** *Biotechniques,* 1986, vol. 4, 320-334 **[0372]**
- **RIGGS et al.** *Proc. Natl. Acad. Sci. USA,* 1986, vol. 83, 5602-5606 **[0372]**
- Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment. **TOMES et al.** Plant Cell, Tissue, and Organ Culture: Fundamental Methods. Springer-Verlag, 1995 **[0372]**
- **MCCABE et al.** *Biotechnology,* 1988, vol. 6, 923-926 **[0372]**
- **WEISSINGER et al.** *Ann. Rev. Genet.,* 1988, vol. 22, 421-477 **[0372]**
- **SANFORD et al.** *Particulate Science and Technology,* 1987, vol. 5, 27-37 **[0372]**
- **CHRISTOU et al.** *Plant Physiol.,* 1988, vol. 87, 671-674 **[0372]**
- **MCCABE et al.** *BioTechnology,* 1988, vol. 6, 923-926 **[0372]**
- **FINER ; MCMULLEN.** *In vitro Cell Dev. Biol.,* 1991, vol. 27P, 175-182 **[0372]**
- **SINGH et al.** *Theor. Appl. Genet.,* 1998, vol. 96, 319-324 **[0372]**
- **DATTA et al.** *Biotechnology,* 1990, vol. 8, 736-740 **[0372]**
- **KLEIN et al.** *Proc. Natl. Acad. Sci. USA,* 1988, vol. 85, 4305-4309 **[0372]**
- **KLEIN et al.** *Biotechnology,* 1988, vol. 6, 559-563 **[0372]**
- **KLEIN et al.** *Plant Physiol.,* 1988, vol. 91, 440-444 **[0372]**
- **FROMM et al.** *Biotechnology,* 1990, vol. 8, 833-839 **[0372]**
- **HOOYKAAS-VAN SLOGTEREN et al.** *Nature (London,* 1984, vol. 311, 763-764 **[0372]**

COR-AAD1-00001658

- **BYTEBIER et al.** *Proc. Natl. Acad. Sci. USA,* 1987, vol. 84, 5345-5349 **[0372]**
- **DE WET et al.** The Experimental Manipulation of Ovule Tissues. 1985, 197-209 **[0372]**
- **KAEPPLER et al.** *Plant Cell Reports,* 1990, vol. 9, 415-418 **[0372]**
- **KAEPPLER et al.** *Theor. Appl. Genet.,* 1992, vol. 84, 560-566 **[0372]**
- **D'HALLUIN et al.** *Plant Cell,* 1992, vol. 4, 1495-1505 **[0372]**
- **LI et al.** *Plant Cell Reports,* 1993, vol. 12, 250-255 **[0372]**
- **CHRISTOU ; FORD.** Annals of Botany. 1995, vol. 75, 407-413 **[0372]**
- **OSJODA et al.** *Nature Biotechnology,* 1996, vol. 14, 745-750 **[0372]**
- **PASZKOWSKI et al.** *EMBO J,* 1984, vol. 3, 2717 **[0374]**
- **FROMM et al.** *Proc Nat'l Acad Sci USA,* 1985, vol. 82, 5824 **[0374]**
- **KLEIN et al.** *Nature,* 1987, vol. 327, 70 **[0374]**
- **WEEKS et al.** *Plant Physiol,* vol. 102, 1077 **[0374]**
- **HIEI et al.** *Plant J.,* 1994, vol. 6, 271 **[0375]**
- **LI et al.** *Science in China,* 1991, vol. 34, 54 **[0375]**
- **RAINERI et al.** *BioITechnology,* 1990, vol. 8, 33 **[0375]**
- **XU et al.** *Chinese J Bot,* 1990, vol. 2, 81 **[0375]**
- **FIROOZABADY ; KUEHNLE.** Plant Cell Tissue and Organ Culture Fundamental Methods. 1995 **[0376]**
- **GEISER et al.** *Gene,* 1986, vol. 48, 109 **[0377]**
- **VAN DAMME et al.** *Plant Mol. Biol.,* 1994, vol. 24, 825 **[0377]**
- **WILLIAMSON et al.** *Eur. J. Biochem.,* 1987, vol. 165, 99-106 **[0377]**
- **PEDERSEN et al.** *J. Biol. Chem.,* 1986, vol. 261, 6279 **[0377]**
- **KIRIHARA et al.** *Gene,* 1988, vol. 71, 359 **[0377]**
- **MUSUMURA et al.** *Plant Mol. Biol.,* 1989, vol. 12, 123 **[0377]**
- **JONES et al.** *Science,* 1994, vol. 266, 789 **[0378]**
- **MARTIN et al.** *Science,* 1993, vol. 262, 1432 **[0378]**
- **MINDRINOS et al.** *Cell,* 1994, vol. 78, 1089 **[0378]**
- **SCHUBERT et al.** *J. Bacteriol.,* 1988, vol. 170, 5837-5847 **[0378]**
- **GORDON-KAMM et al.** *The Plant Cell,* 1990, vol. 2, 603-618 **[0380]**
- Protoplasts Isolation and Culture. **EVANS et al.** Handbook of Plant Cell Culture. Macmillan Publishing Company, 1983, 124-176 **[0380]**
- Regeneration of Plants. **BINDING.** Plant Protoplasts. CRC Press, 1985, 21-73 **[0380]**
- **KLEE et al.** *Ann Rev of Plant Phys,* 1987, vol. 38, 467 **[0380]**
- **KAMO et al.** *Bot. Gaz,* 1985, vol. 146 (3), 324-334 **[0381]**
- **WEST et al.** *The Plant Cell,* 1993, vol. 5, 1361-1369 **[0381]**
- **DUNCAN et al.** *Planta,* 1985, vol. 165, 322-332 **[0381]**
- **HORSCH et al.** *Science,* 1985, vol. 227, 1229-1231 **[0383]**
- **FRALEY et al.** *Proc. Natl. Acad. Sci. U.S.A.,* 1983, vol. 80, 4803 **[0383]**
- **KLEE et al.** *Ann. Rev. of Plant Phys.,* 1987, vol. 38, 467-486 **[0384]**
- Methods for Plant Molecular Biology. Academic Press, Inc, 1988 **[0384]**
- Corn and Corn Improvement. American Society of Agronomy, 1988 **[0384]**
- **DODDS ; ROBERTS.** Experiments in Plant Tissue Culture. Cambridge University Press, 1995 **[0386]**
- **BECHTOLD N. ; ELLIS J. ; PELLETIER G.** planta Agrobacterium mediated gene transfer by infiltration of adult Arabidopsis thaliana plants. *CR Acad Sci Paris Life Sci,* 1993, vol. 316, 1194-1199 **[0387]**
- **DESFEUX, C. ; CLOUGH S.J. ; BENT A.F.** Female reproductive tissues are the primary target of Agrobacterium-mediated transformation by the Arabidopsis floral-dip method. *Plant Physiol.,* 2000, vol. 123, 895-904 **[0387]**
- *Methods in Molecular Biology,* 1995, vol. 49, 9-18 **[0431]**
- *Gene,* 1995, vol. 156, 119-122 **[0434]**
- **CLOUGH, SJ ; BENT, AF.** *Plant J.*, 1998, vol. 16 (6), 735-43 **[0440]**
- **SCOPES.** Protein Purification, Principles and Practice. Springer, 1994 **[0459]**
- **JOSHI et al.** *Plant Mol. Biol,* 1997, vol. 35, 993-1001 **[0467]**

COR-AAD1-00001659

(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2005-287415
(P2005-287415A)
(43)公開日　平成17年10月20日(2005.10.20)

(51)Int.Cl.[7]	F I		テーマコード（参考）
C12N 15/09	C12N 15/00	ZNAA	4B024
C12Q 1/68	C12Q 1/68	Z	4B063

審査請求 未請求　請求項の数 10　OL　（全 18 頁）

(21)出願番号　特願2004-108166 (P2004-108166)
(22)出願日　平成16年3月31日 (2004.3.31)

(71)出願人　301021533
独立行政法人産業技術総合研究所
東京都千代田区霞が関１－３－１
(72)発明者　中村　和憲
茨城県つくば市東１－１－１　独立行政法人産業技術総合研究所つくばセンター内
(72)発明者　鎌形　洋一
茨城県つくば市東１－１－１　独立行政法人産業技術総合研究所つくばセンター内
(72)発明者　木村　信忠
茨城県つくば市東１－１－１　独立行政法人産業技術総合研究所つくばセンター内
(72)発明者　北川　航
茨城県つくば市東１－１－１　独立行政法人産業技術総合研究所つくばセンター内
最終頁に続く

(54)【発明の名称】特定遺伝子の変異導入方法

(57)【要約】
【課題】特定の遺伝子にマーカー配列を導入したものを、この遺伝子に導入されたマーカー配列を含むプライマーとPCRで増幅することにより、この特定の遺伝子を特異的に検出する方法の提供。
【解決手段】目的遺伝子に変異を導入する方法であって、変異を導入しようとする位置よりも前方領域を増幅するためのプライマーセットのうち少なくとも一方のプライマーに変異の導入の目的配列が含まれるプライマーセット、及び後方領域を増幅するためのプライマーセットのうち少なくとも一方のプライマーに変異の導入の目的配列が含まれるプライマーセットをそれぞれ用いて、目的遺伝子を増幅し、得られる前方領域の増幅断片と後方領域の増幅断片とを連結させることを特徴とする、前記方法。
【選択図】なし

COR-AAD1-00001660

【特許請求の範囲】
【請求項1】
目的遺伝子に変異を導入する方法であって、変異を導入しようとする位置よりも前方領域を増幅するためのプライマーセットのうち少なくとも一方のプライマーに変異の導入の目的配列が含まれるプライマーセット、及び後方領域を増幅するためのプライマーセットのうち少なくとも一方のプライマーに変異の導入の目的配列が含まれるプライマーセットをそれぞれ用いて、目的遺伝子を増幅し、得られる前方領域の増幅断片と後方領域の増幅断片とを連結させることを特徴とする、前記方法。
【請求項2】
変異の導入箇所が少なくとも１箇所である請求項１記載の方法。
【請求項3】
変異の導入が、目的遺伝子への塩基配列の付加、又は目的遺伝子の塩基配列の置換である請求項１又は２記載の方法。
【請求項4】
塩基配列の置換が、アミノ酸をコードするコドンの同義コドンへの置換である請求項３記載の方法。
【請求項5】
アミノ酸が、セリン、ロイシン及び/又はアルギニンを含むものである請求項４記載の方法。
【請求項6】
請求項１～３のいずれか１項に記載の方法を用いて2,4-ジクロロフェノキシ酢酸ジオキシゲナーゼをコードする遺伝子の塩基配列の３’末端側に塩基配列GTCATCGGAが付加された、変異型tfdAm1遺伝子。
【請求項7】
2,4-ジクロロフェノキシ酢酸ジオキシゲナーゼをコードする遺伝子が、ラルストニア・ユートロファJMP134株由来のものである請求項６記載の変異型tfdAm1遺伝子。
【請求項8】
2,4-ジクロロフェノキシ酢酸ジオキシゲナーゼが、2,4-ジクロロフェノキシ酢酸を2,4-ジニトロフェノールへ変換する機能を有するものである請求項７記載の変異型tfdAm1遺伝子。
【請求項9】
請求項１～５のいずれか１項に記載の方法を用いて2,4-ジクロロフェノキシ酢酸ジオキシゲナーゼをコードする遺伝子の塩基配列の第274-288番目の塩基配列AACGTCAGTCTCGACがAATGTGTCGTTGGATに置換された、変異型tfdAm2遺伝子。
【請求項10】
請求項１～５のいずれか１項に記載の方法により変異が導入された変異型遺伝子、又は請求項６～９のいずれか１項に記載の変異型遺伝子を、変異が導入された塩基配列を含むプライマーをフォワード又はリバースプライマーとして用いて増幅し、得られる増幅産物から変異型遺伝子を検出する方法。

【発明の詳細な説明】
【技術分野】
【0001】
本発明は、特定遺伝子をPCRにより特異的に検出する方法として、その遺伝子に変異を導入する方法に関する。すなわち、本発明は「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」、また、「遺伝子配列を特異的検出の標的となる塩基配列(マーカー配列)に置換する方法」に関するものである。本発明の方法によって構築された遺伝子は、その付加または置換した配列を標的にしたPCRにより特異的検出が可能になる。また本法は簡便かつ迅速な部位特異的変異導入を行うことが出来る。
【背景技術】
【0002】

環境浄化を目的として、特定の微生物を環境中に導入する場合、微生物の有効性や安全性を判断するための指標として、環境中に多様多種な微生物が存在する中から、導入した微生物を検出・追跡する必要がある。そのために、微生物に固有に存在する遺伝子の塩基配列を標的にして、遺伝子の塩基配列に相同なオリゴヌクレオチドを利用したPCRやハイブリダイゼーションを行う方法がある。

【0003】

しかしながら、従来から知られている検出方法には次のような問題点がある。すなわち、微生物の遺伝子配列から設計した検出プライマーを用いたときは、同じ塩基配列を有する微生物種を検出することが出来るが、微生物種の中の特定の個体を識別することは出来ない。また環境中に類似の遺伝子配列が存在する可能性があり、特に環境中においては、導入した微生物が移動するために、環境中の全ての場所における類似の遺伝子配列が存在する可能性を予測することは困難である。

【0004】

さらに、環境に導入する特定遺伝子を含む微生物をDNAレベルで追跡する場合、以下の問題がある。すなわち、環境浄化を目的としてある微生物を環境中に導入しようとする場合において、汚染物質分解遺伝子を意図的にプラスミドなどの核外遺伝子上に導入するときは、プラスミドは接合伝達などによって他の微生物に水平伝搬する可能性がある。このため、宿主自身と共に組換えプラスミドも追跡しなければ、導入微生物の遺伝子を追跡することにはならない。また、分解遺伝子を意図的に染色体DNAに組み込んだ場合でも、やはりその遺伝子が他の微生物に伝播する可能性がある。

【発明の開示】

【発明が解決しようとする課題】

【0005】

既知の遺伝子配列から設計した検出プライマーでは微生物種を検出することはできても、個体を識別することは出来ない。環境中には、互いに類似する遺伝子配列が存在する可能性があるため、導入した遺伝子のみを特異的に検出することが困難である。本発明は、上記のような問題点を鑑みてなされたものであり、その目的は、任意の遺伝子配列に任意の特異的検出の標的となる塩基配列を付加、または特異的検出の標的となる塩基配列に置換する方法を開発し、従来の種レベルの検出ではなく細胞、DNAレベルでの特異的検出を可能にすることである。すなわち、本発明は、特定の塩基配列を有する遺伝子を特異的に検出する方法を提供することを目的とする。

【課題を解決するための手段】

【0006】

本発明者は、上記課題を解決するために、鋭意研究を行った。そして、特定の微生物種の検出・追跡のために、その微生物の塩基配列にマーカーとなる配列を付加または置換した変異遺伝子を作製し、このマーカー遺伝子を含むプライマーとともにPCRを行うことにより簡便かつ迅速に、目的とする遺伝子の検出または追跡ができることを見出し、本発明を完成させた。

【0007】

すなわち、本発明は以下のとおりである。

(1)目的遺伝子に変異を導入する方法であって、変異を導入しようとする位置よりも前方領域を増幅するためのプライマーセットのうち少なくとも一方のプライマーに変異の導入の目的配列が含まれるプライマーセット、及び後方領域を増幅するためのプライマーセットのうち少なくとも一方のプライマーに変異の導入の目的配列が含まれるプライマーセットをそれぞれ用いて、目的遺伝子を増幅し、得られる前方領域の増幅断片と後方領域の増幅断片とを連結させることを特徴とする、前記方法。

【0008】

(2)変異の導入箇所が少なくとも１箇所である(1)記載の方法。

(3) 変異の導入が、目的遺伝子への塩基配列の付加、又は目的遺伝子の塩基配列の置換である(1)又は(2)記載の方法。

COR-AAD1-00001662

(4) 塩基配列の置換が、アミノ酸をコードするコドンの同義コドンへの置換である(2)記載の方法。

【0009】

(5) アミノ酸が、セリン、ロイシン及び/又はアルギニンを含むものである(4)記載の方法。

(6) 上記(1)～(3)のいずれか１項に記載の方法を用いて2,4-ジクロロフェノキシ酢酸ジオキシゲナーゼをコードする遺伝子の塩基配列の３’末端側に塩基配列GTCATCGGAが付加された、変異型tfdAm1遺伝子。

(7) 2,4-ジクロロフェノキシ酢酸ジオキシゲナーゼをコードする遺伝子が、ラルストニア・ユートロファJMP134株由来のものである(6)記載の変異型tfdAm1遺伝子。

【0010】

(8) 2,4-ジクロロフェノキシ酢酸ジオキシゲナーゼが、2,4-ジクロロフェノキシ酢酸を2,4-ジニトロフェノールへ変換する機能を有するものである(7)記載の変異型tfdAm1遺伝子。

(9) 上記(1)～(5)のいずれか１項に記載の方法を用いて2,4-ジクロロフェノキシ酢酸ジオキシゲナーゼをコードする遺伝子の塩基配列の第274-288番目の塩基配列AACGTCAGTCTCGACがAATGTGTCGTTGGATに置換された、変異型tfdAm2遺伝子。

【0011】

(10) 上記(1)～(5)のいずれか１項に記載の方法により変異が導入された変異型遺伝子、又は上記(6)～(9)のいずれか１項に記載の変異型遺伝子を、変異が導入された塩基配列を含むプライマーをフォワード又はリバースプライマーとして用いて増幅し、得られる増幅産物から変異型遺伝子を検出する方法。

【発明の効果】

【0012】

本発明により、特定の遺伝子を個体レベルで特異的に検出する方法が提供される。本発明の遺伝子の特異的検出方法は、環境浄化のために用いられる微生物を環境中で検出・追跡することができ、これにより、その微生物の効果・安全性を簡便かつ迅速に判断できる点で有用である。

【発明を実施するための最良の形態】

【0013】

以下、本発明の内容を詳細に説明する。

本発明は、特定の遺伝子を、マーカー配列を含むプライマーを用いて増幅することにより、上記特定の遺伝子にマーカー配列を導入し（変異を導入し）、この特定の遺伝子を特異的に検出する方法である。そして、本発明は汎用性の高い簡便な方法による遺伝子への特異的検出の標的となる塩基配列(マーカー配列)の導入方法を提供する。マーカー配列の導入方法として「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」と「遺伝子配列を特異的検出の標的となる塩基配列(マーカー配列)に置換する方法」という2種類の方法を提供するが、基本技術は同じであり、マーカー配列となる配列を含んだ複数セット（例えば2セット）のプライマーを用いてプライマーのセット数に対応した数のPCR（2セットのプライマーセットを用いた場合は2種類のPCR）を行い、それらのPCR産物を変性、会合、PCRによって結合することにより、特異的検出の標的となる塩基配列の導入が可能になる。

【0014】

例えば、2セットのプライマーセットを用いて2種類のPCRを行う場合は、マーカー配列を導入する目的遺伝子をその導入位置で前方（前半）領域、後方（後半）領域と二分し、個別にPCRによる増幅を行う。「前方領域」とは、目的遺伝子のセンス鎖を基準としたときに、変異導入位置からみてその上流側（5’側）の領域を意味し、「後方領域」とは、目的遺伝子のセンス鎖を基準としたときに、変異導入位置からみてその下流側（3’側）の領域を意味する。

【0015】

COR-AAD1-00001663

このときに前方領域を増幅するリバースプライマーと後方領域を増幅するフォワードプライマーのいずれか一方又は両方にそれぞれ相補的なマーカー配列を連結しておく。2種類の増幅産物、すなわち鋳型となる遺伝子の前半部の増幅断片と後半部の増幅断片を混合、変性、再会合し、DNA断片の前半領域と後半領域の相補的な付加又は置換する塩基配列で二つの断片を結合し、前半領域のフォワードプライマーと後半領域のリバースプライマーを用いて結合させたDNA断片の全長について増幅反応を行い、変異配列が導入された遺伝子の全長を増幅する。このようにして得られた変異型の遺伝子は、付加又は置換した塩基配列を標的としてPCRによる特異的検出が可能になる。

【0016】

本発明において、「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」は、任意の遺伝子を鋳型に任意の付加配列を連結したセットのプライマーを用いて例えば2種類のPCRを行い、さらに2種類のPCR産物を結合するPCRを行うことによって遺伝子に任意のマーカー配列を付加する方法である。

【0017】

特異的検出の際には、PCRに用いるフォワード・リバースプライマーのいずれか一方、または両方のプライマーにマーカー配列を含ませることで、上記遺伝子の特異的検出が可能になる。またプライマーセットを複数種用い、複数回のPCRを行うことにより、マーカー配列を持つDNAの検出感度をより高めることができる。

【0018】

(1) 本発明の「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」は、組み換え対象となる遺伝子がpUC19などのプラスミドにクローン化されており、その遺伝子とその周辺領域のDNA塩基配列が決定されていることが必要である。

本方法は既知の遺伝子に任意のDNA塩基配列をマーカー配列として付加し、その配列を標的として特異的検出を可能にするものである。

【0019】

特異的検出の標的として導入するDNA塩基配列は、後のPCRによる特異的検出を可能にするため、数塩基以上必要である。

特異的検出の標的とする塩基配列（マーカー配列）を遺伝子のコード領域内に付加する場合、マーカー配列は、アミノ酸に翻訳される際に読み枠がずれないように、3、6、9など3の倍数の塩基番号を有する塩基の後ろに付加するように設計し、また付加するマーカー配列もアミノ酸に翻訳される3塩基の塩基配列を単位として設計・調製する。さらに付加する塩基配列は、コードするアミノ酸自体が酵素の活性や立体構造に影響が少ないと考えられる配列を用いることが好ましい。

【0020】

マーカー配列は上記の条件をふまえ、6、9、12、15、18塩基程度が適当であるが、これに制限されるものではない。

塩基配列を付加する位置はその導入によって元の遺伝子が機能を失ったり、低下したりしないように、検討する必要がある。タンパク質のN末またはC末領域に導入することが最もそれらの危険性を回避できると言える。

【0021】

本発明において、2組のプライマーセットを用いた場合を説明する（図１）。

塩基配列を付加する位置を決定した後、その位置で遺伝子を前半領域、後半領域の2つの領域に分け、それぞれを増幅する2セットのPCR
プライマーを設計し、個々にPCRを行い、それぞれの増幅産物を得る。この時、前半領域を増幅するフォワードプライマー をA、リバースプライマー をB、また後半領域を増幅するフォワードプライマー
をC、リバースプライマー をDとする。プライマーBの5’側にはマーカー配列と相補的な配列を付加しておき、さらにその5’側には後半領域で増幅される部分のコード鎖の5’部を5～20塩基程度、好ましくは10塩基程度付加しておく。このように、マーカー配列の5’側に10塩基程度付加された配列を、本発明においては「クランプ配列」という。

【0022】
プライマー Cの5’側にはマーカーとなるDNA塩基配列を付加しておき、さらにその5’側には前半領域で増幅される部分のコード鎖の3’部を5～20塩基程度、好ましくは10塩基程度、クランプ配列を付加しておく。それぞれのプライマーの本体部位(鋳型DNAと相同な、マーカー配列の3’側)を20塩基、マーカー配列を9塩基で設計した場合、A、D プライマーは20 塩基、B、C プライマーは20+9+10=39
塩基となる(図1参照)。
【0023】
本発明の方法で行うPCRに用いるDNA
ポリメラーゼには、Taq ポリメラーゼなどのPCR増幅産物の5’側にアデニン塩基を付加するようなものを使用するのではなく、何ら修飾をせず、鋳型の塩基配列に対して正確性の高い活性を有するポリメラーゼを使用することが好ましい。そのような正確性の高いポリメラーゼとしては、例えばPfuポリメラーゼ、Pyrobest
DNAポリメラーゼなどが挙げられる。
【0024】
ここで、A、Dプライマーがアニーリングする位置は遺伝子をコードする領域である必要はなく、遺伝子を連結しているベクター部位でも可能である。
B、Cプライマーに付加した10塩基程度の配列(クランプ部位)は導入する塩基配列とともに2つのPCR断片を結合する相補部位となる塩基配列であり、マーカー配列とクランプ部位で計15
塩基程度の長さの場合に効果的である。但し、マーカー配列が長い場合、クランプ部位は必須ではない。Bプライマーのクランプ部位はCプライマーの本体部位の一部と、Cプライマーのクランプ部位は
Bプライマーの本体部位の一部とそれぞれ相補的になっている(図2参照)。
【0025】
A、Bプライマーで増幅した遺伝子の前半領域のPCR産物を「A-B断片」とし、C、Dプライマーで増幅した遺伝子の後半領域のPCR産物を「C-D断片」とすると、増幅されたA-B断片及びC-D断片をそれぞれ精製した後、混合、熱変性、再会合を行なう。このとき前半領域を増幅したPCR産物のAプライマーが伸長して3’側にマーカー配列と相補部位が連結した分子（A-B断片のBプライマーのマーカー領域に対応する部分）と、後半領域を増幅したPCR産物のDプライマーが伸長して3’側にマーカー配列と相補部位が連結した分子（C-D断片のCプライマーのマーカー領域に対応する部分）が、それぞれマーカー配列とクランプ部位からなる相補部位で会合した部分二本鎖DNA分子が出来る(図1参照)。
【0026】
この再会合したDNA分子を鋳型にDNA
polymeraseとdNTPを加え、 DNAの伸長反応を行う。このときプライマーAとDの伸長産物が会合した分子は部分的に二本鎖になっており、それぞれの3’側から伸長反応(A伸長産物（A-B断片の3’側）がプライマーとなりD伸長産物が鋳型となる反応と、D伸長産物（C-D断片の3’側）がプライマーとなりA伸長産物が鋳型となる反応)が起こり、マーカー配列を有する遺伝子の全体の二本鎖DNAが合成される(図3）。このときBプライマーの伸長産物とCプライマーの伸長産物が会合し、部分的に二本鎖になった分子は、それぞれが5’側で結合しているため、伸長反応は起こらない(図1、3参照)。
【0027】
伸長反応を行ったDNAサンプルを鋳型に、プライマーA、Dを加え、マーカー配列を有する遺伝子の全体を増幅するためにPCRを行う(図1参照)。
上記の2つのPCR産物の混合、変性、再会合、伸長反応のステップは必須ではなく、二つのPCR産物を直接鋳型にしてPCRを用い、プライマーA、Dを加えて、リンキングPCRを行うことも可能である。
【0028】
上記のPCR産物を精製し、適当なベクター（例えばpUC19、pHSG396ベクターなど）にク

ローニングし、DNA塩基配列解析装置を用いてマーカー配列が導入されたことを確認し、遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)が付加された遺伝子を得る。

【0029】

(2)本発明の「遺伝子配列を特異的検出の標的となる塩基配列(マーカー配列)に置換する方法」は、上記の「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」と同様、組換えの対象となる遺伝子がpUC19などのプラスミドにクローン化されており、その遺伝子とその周辺領域のDNA塩基配列が決定されていることが必要である。

【0030】

本発明の方法は、既知の遺伝子の塩基配列をそのコードするアミノ酸に変化が起こらないような塩基配列に置換することにより、その塩基配列を標的として、遺伝子の特異的検出を可能にするものである。

特異的検出の標的とするために、塩基配列の置換はその翻訳されるアミノ酸に変化が無い様に設計する。すなわち遺伝暗号表を用い、マーカーとする位置のDNA塩基配列を出来るだけ異なるDNA塩基配列に置換しながら、同じアミノ酸をコードするコドンを選択する(表1参照)。

【0031】

【表1】

	U		C		A		G	
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
	UUA	Leu	UCA	Ser	UAA	***	UGA	***
	UUG	Leu	UCG	Ser	UAG	***	UGG	Trp
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly

【0032】

アミノ酸を置換しないように、特異的検出の標的とするために塩基配列に置換を導入する場合は、セリン(Ser)、ロイシン(Leu)、アルギニン(Arg)をコードする領域(枠で囲った部分)を用いることで置換率を高くすることが出来る。

特異的検出の標的とする塩基配列の位置は、その塩基配列の置換によって翻訳されるアミノ酸が変化しないように設計するので（同義コドンへの置換）、遺伝子のどの部位を選択する事も可能である。

【0033】

塩基配列の置換を行う位置は、そのコードするアミノ酸を考慮に入れて選択することが望ましい。アミノ酸の置換を伴わない塩基配列の置換を行う場合、ほとんどのアミノ酸が

COR-AAD1-00001666

コドンの3番目の塩基の置換のみ行うことが出来る。すなわち5アミノ酸をコードする15塩基を選択した場合、3塩基おきに5塩基の置換が可能である。一方、ロイシン(Leu)、アルギニン(Arg)
のコドンは1番目と3番目の塩基配列、すなわち 2塩基を置換することが可能であり、またセリン(Ser)のコドンは3塩基配列の全ての塩基配列を置換する事が可能である(表1参照)。すなわち、マーカーとしてアミノ酸置換を伴わない同義コドンへの塩基置換を行う場合、セリン(Ser)、ロイシン(Leu)、アルギニン(Arg)が隣接してコードされている領域をターゲットにすれば、置換率の高い、特異的検出に適した配列とする事が出来る。セリンを例に取ると、TCT、TCA、TCGでコードされるセリンはAGCに、TCC、TCA、TCGでコードされるセリンはAGTに、またそれらの逆の変異のように3塩基とも置換することが出来る。5アミノ酸をコードする15塩基を選択し、その中にセリン(Ser)とロイシン(Leu)またはアルギニン(Arg)が一つずつ含まれていた場合、少なくとも15塩基中8塩基の置換を導入する事が出来る(表1参照)。

【0034】

置換する塩基配列を決定したのち、遺伝子をその置換する塩基配列の部位を含まない前半領域、その配列部位を含まない後半領域の2つの部位に分け、それぞれのDNA
断片を増幅する2セットのPCR プライマーを設計し、個々にPCRを行い、それぞれの増幅産物を得る。この時前半領域を増幅するフォワードプライマー をA、リバースプライマー をE、また後半領域を増幅するフォワードプライマー をF、リバースプライマー をDとする。プライマーEの5’側には特異的検出の標的となる置換した塩基配列と相補的な配列を付加しておき、さらにその5’側には後半領域で増幅される部分のコード鎖の5’側を10 塩基程度付加しておく。プライマー
Fの5’側には特異的検出の標的となる置換した塩基配列を付加しておき、さらにその5’側には前半領域で増幅される部分のコード鎖の3’部を10 塩基程度付加しておく。それぞれのプライマーの本体部位を20
塩基、特異的検出の標的となる置換した塩基配列を15塩基の長さになるように設計した場合、A、D プライマーは20 塩基、E、F プライマーは20+15+10=45
塩基となる(図4参照)。

【0035】

本発明の方法により行うPCRには、上記の「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」と同様に、Taq
polymeraseのように産物の5’側にアデニン塩基を付加するようなDNA polymeraseではなく、何ら修飾をせず、鋳型の塩基配列に対して正確性の高い活性を有するポリメラーゼ（例えばPfuポリメラーゼ、
Pyrobestポリメラーゼ等）を使用する。

また「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」と同様に、A、Dプライマーがアニーリングする位置は、遺伝子の内部である必要はなく、遺伝子を連結しているベクター部位でも可能である。

【0036】

E、Fプライマーに付加した10
塩基程度の配列(クランプ部位)は特異的検出の標的となる塩基配列とともに、リンキングPCRの際に互いに結合する相補的な塩基配列であり、特異的検出の標的となる塩基配列とクランプ部位で計15
塩基程度で効果的であるが、特異的検出の標的となる塩基配列が長い場合はクランプ部位は必須ではない。Eプライマーのクランプ部はFプライマーの本体部位(鋳型DNAと相同な、マーカー配列の3’側)の一部と、Fプライマーのクランプ部は Eプライマーの本体部位の一部とそれぞれ相補的になっているはずである(図4参照)。

「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」と同様の手順により、それぞれのプライマーから増幅したPCR産物から特異的検出の標的となる塩基配列が導入された遺伝子を得ることが可能である。

COR-AAD1-00001667

【0037】
(3) 本発明の「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」、「遺伝子配列を特異的検出の標的となる塩基配列(マーカー配列)に置換する方法」は適切なプライマーを設計することにより同時に複数の箇所にマーカー配列を導入することが出来る。例えばマーカー配列を二カ所に導入する場合、鋳型遺伝子をマーカー配列を導入する位置で三カ所に分け、前部域、中部域、後部域を個別に増幅するプライマーによってPCR増幅する。この場合、前部域と中部域との間の変異導入位置からみれば、前部域が「前方領域」となり、中部域が「後方領域」となる。また、中部域と後部域との間の変異導入位置からみれば、中部域が「前方領域」となり、後部域が「後方領域」となる。
【0038】
それぞれの フォワードプライマー、リバースプライマーを
GとH、IとJ、KとLとすると、HとIプライマーに相補的なマーカー配列とクランプ部位、またJとKプライマーに相補的なマーカー配列とクランプ部位を付加しておくことで二カ所へのマーカー配列の導入が可能になる。このとき二カ所のマーカー配列はそれぞれ異なる塩基配列を用いるべきである。塩基配列の付加、置換の導入を同時に行うことも可能である。
本発明の「遺伝子配列を特異的検出の標的となる塩基配列(マーカー配列)に置換する方法」は、置換した塩基によってアミノ酸置換も起こるようにプライマーを設計することにより簡便な部位特異的変異(Site
directed mutagenesis)を行うことが可能である。
【0039】
(4) tfdAm1遺伝子/
tfdAm2遺伝子
本発明は、本発明の変異導入方法により、2,4-ジクロロフェノキシ酢酸
ジオキシゲナーゼ遺伝子（tfdA）のDNA塩基配列の第6番の塩基の後に、マーカーとなる塩基配列GTCATCGGAが導入された変異型tfdAm1遺伝子（配列番号１）を提供する。tfdA遺伝子は、ラルストニア・ユートロファ（Ralstonia
eutropha）JMP134 株由来の遺伝子であり、2,4-ジクロロフェノキシ酢酸 を2,4-ジニトロフェノール へ変換する機能を有する酵素をコードするものである。
【0040】
本発明はまた、本発明の変異導入方法により上記tfdAm1のAsn-Val-Ser-Leu-Asp（配列番号３）をコードするDNA（塩基番号274-288）の塩基配列AACGTCAGTCTCGAC（配列番号４）を、マーカーとなるAATGTGTCGTTGGAT（配列番号５）へ置換した変異型tfdAm2遺伝子（配列番号２）を提供する。
本発明の方法により構築されたマーカー配列を含む遺伝子は、その塩基配列を含むプライマーを利用したPCRにより特異的検出が可能である。
本発明の遺伝子は後述するように、導入したマーカー配列を含むプライマーを利用したPCR
を行うことにより、本発明の遺伝子を特異的に高精度、しかも簡便に検出することができる。
【0041】
なお、本発明においては、上記変異型遺伝子を適当なベクターに連結して組換えベクターを作成することができる。本発明で使用するベクターとしては、プラスミド、コスミド、バクテリオファージなど宿主中で複製可能なものであれば特に限定されない。またそのベクターが機能するものであれば大腸菌、放線菌など宿主も限定されない。
【0042】
また、本発明においては、上記ベクターを含む形質転換体を提供することができる。ベクターを導入する宿主は、細菌（大腸菌等）、酵母、動物細胞、昆虫細胞など、当分野において周知のものを任意に使用することができる。上記形質転換体は、公知のエレクトロ

ポレーション法、プロトプラスト法、塩化カルシウム法(スフェロプラスト法)などの形質転換手法により作製することができる。またベクターや宿主、高発現を誘導する誘導基質やプロモーター、オペレーター、エンハンサーなどの組み合わせを考慮し用いることによって、2,4-ジクロロフェノキシ酢酸ジオキシゲナーゼ活性を持つ酵素を産生する微生物を作製することが可能である。

本発明のtfdAm1、tfdAm2遺伝子を発現した大腸菌の細胞抽出物は、野生型の遺伝子を発現したものと同程度の2,4-ジクロロフェノキシ酢酸分解活性を示す。

【0043】

(5) 特異的検出の方法

本発明は、上記のように作製した変異遺伝子を用いて、導入されたマーカー配列を標的として特異的に検出する方法である。特異的検出に使用するマーカー配列は、PCRによって野生型の遺伝子と判別できるように設計されている。従って、検出は、検出の目的遺伝子を鋳型とし、マーカー配列を含む塩基配列を有するオリゴヌクレオチドを標識し、これをプライマーとして用いてPCR反応を行い、反応産物を確認すればよい。プライマーへの標識は、蛍光標識、放射標識のいずれを採用してもよい。蛍光物質としては、例えばAlexa系、Cy2、Cy3、Cy5、FITC、ローダミン、テキサスレッドなどが挙げられ、放射標識物質としては、^{3}H、^{13}C、^{15}N、^{32}P等が挙げられる。

【0044】

マーカー配列を含むPCRプライマーを設計し、それぞれ野生型遺伝子と変異遺伝子を鋳型としてPCRを行うと、野生型では変異遺伝子をターゲットにしたプライマーを用いたPCRを行っても、DNA断片の増幅が確認することができない。これに対し変異遺伝子は、変異配列をターゲットとしたプライマーによりDNA断片の増幅が確認される。

増幅の条件は、94℃で30秒の変性工程、60℃で30秒のアニーリング工程、及び74℃で増幅長500-bpあたり30秒の伸長工程を１サイクルとしてこれを30サイクル行う。但し、鋳型DNA及びプライマーを十分変性させるために、上記増幅サイクルの前に94℃で５～10分の変性工程を加えてもよい。さらに、増幅産物の検出を直ちに行わない場合は、非特異的な増幅が起こらないようにするために、増幅産物を4℃で保存する工程を加えることが好ましい。

このようにして、本発明のプライマーセットを用いて、目的のDNAを増幅することができる。

【実施例】

【0045】

塩基配列の付加

(1)本発明の「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」を用いてRalstonia eutropha JMP134株由来の2,4-Dichlorophenoxyacetic acid(2,4-D)を2,4-Dichlorophenolに変換する酵素遺伝子tfdA（M16730【GenBank】）への特異的検出の標的となる塩基配列の付加を行った(図5参照)。

特異的検出の標的となる塩基配列(マーカー配列)を付加する位置は酵素の活性が失われる危険性の少ないアミノ酸配列のN末側残基の2番目と3番目の間、塩基配列では塩基番号6番目の後ろとした。

【0046】

マーカー配列は、Val-Ile-GlyをコードするGTCATCGGAの9塩基とした。このアミノ酸配列は構造的に小さいアミノ酸残基で酵素の立体構造や活性に影響が少ないこと、アミノ酸の反応性が低いことを考慮した。また付加するDNA配列は遺伝子をPCRにより特異的に検出する場合に、GC含量や遺伝子上の配置がPCRプライマーの設計に都合が良いことを考慮した。また元のtfdA遺伝子でも用いられているコドンを採用している。

tfdAの前半領域約450-bpを増幅するプライマーセット、フォワードプライマーA、リバースプライマー B、また後半領域約1100-bpを増幅するプライマーセット、フォワ

ードプライマーC、リバースプライマー Dを設計し、プライマーB、Cにはそれぞれマーカー配列(プライマーCにGTCATCGGA、プライマーBにその相補配列)とクランプ部位を付加した(図参照)。プライマーAはtfdAがクローン化された位置の約450-bp
5’側のベクター部位の配列であり、プライマーDはtfdA終止コドンより約180-bp　3’側と相補的な塩基配列を含むプライマーである(図5参照)。

【0047】

プライマーA：5’-GCCTTTGAGTGAGCTGATACC-3’（配列番号６）
プライマーB：5’-TTGCGACGACTCCGATGACGCTCACTTTTTTGCTCCTTTTC-3’（配列番号７）
プライマーC：5’-AAAAGTGAGCGTCATCGGAGTCGTCGCAAATCCCCTTCATC-3’（配列番号８）
プライマーD：5’-CCAGCGTCATTTCGAGCACC-3’（配列番号９）

【0048】

二つのプライマーセットにより二つのPCR増幅産物を得て、それぞれ精製した。PCRのプログラムは変性95度30秒、アニーリング60度30秒、伸長反応は74度で増幅長500-bpあたり30秒に設定し、30サイクル行った。DNA
ポリメラーゼはTAKARA社のPyrobest polymeraseを用いた。

精製した二つのPCR産物を混合、変性、再会合、伸長反応を行い、最終的にA、Dプライマーを用いマーカー配列が導入された遺伝子全体をPCR増幅した(図6参照)。

【0049】

得られたPCR増幅配列をプラスミドベクターに連結し大腸菌に形質転換してクローン化し、DNA塩基配列解析装置にて塩基配列を決定し、マーカー配列が付加されていることを確認した。塩基配列を決定した全てのクローンにおいてその配列が付加されている事が確認され、マーカー配列を持ったtfdAm1遺伝子の取得に成功し、本手法の有用性が示された。

【実施例】

【0050】

塩基配列の置換

(2)本発明の「遺伝子配列を特異的検出の標的となる塩基配列(マーカー配列)に置換する方法」を用いて、上記で構築したtfdAm1遺伝子に塩基配列の置換による特異的検出の標的となる塩基配列の導入を行った。

マーカー配列とする塩基を置換する配列は、アミノ酸配列では92-96番目(野生型TfdAでは89-93)のアミノ酸配列Asn-Val-Ser-Leu-Asp（配列番号３）、DNAでは塩基番号274-288(野生型tfdAでは265-279)の塩基配列AACGTCAGTCTCGAC（配列番号４）であり、これをAATGTGTCGTTGGAT（配列番号５）に置換するプライマー設計を行った。この塩基置換により15塩基中8塩基を置換することが可能である。

【0051】

塩基置換を行う位置でtfdAm1遺伝子を前半領域、後半領域に分け、置換する塩基配列を含まないPCRプライマーセットを設計した。前半領域を増幅するフォワードプライマーには上記のAプライマー、後半領域のリバースプライマーには上記のDプライマーをそのまま用い、前半領域のリバースプライマーは塩基配列を置換する位置の5’側に新たにEプライマー、後半領域のフォワードプライマーは塩基配列を置換する位置の3’側に新たにFプライマーを設計し、その5’側にそれぞれ塩基配列の置換を行う相補的な塩基配列(マーカー配列)を付加し、さらにその5’側にクランプ部位となる10 塩基の塩基配列を付加した。

【0052】

以上のプライマーセットを用い、「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」で行った場合と同様に、前半領域、後半領域のPCR増幅、PCR産物の連結、マーカー配列を有する遺伝子のPCR増幅を行なった。得られたクローンをDNA塩基配列解析装置にて塩基配列を確認し、塩基置換の起こった変異体遺伝子tfdAm2を構築した。

上記のように得られたtfdAm1遺伝子は、野生型のtfdA遺伝子の塩基番号6の下流に9塩基長のマーカー配列を導入したものであり、野生型よりもDNAで9塩基、コードするアミノ酸

は3残基が付加されている(野生型tfdAでは861塩基287アミノ酸、tfdAm1では870塩基290アミノ酸)。tfdAm2遺伝子はtfdAm1にさらに塩基配列274-288の位置の塩基配列を置換してマーカー配列を導入したものであり、塩基長はtfdAm1と同一である。

【実施例】

【0053】

2,4-Dの分解試験

本実施例は、構築したtfdAm遺伝子の2,4-Dの分解活性を調査するために行ったものであり、次の手順による。

【0054】

(1) 粗酵素抽出液の調整

本発明に係る遺伝子を連結したベクターDNA

を導入した大腸菌(Escherichia coli DH5α)をシャーレから１白金耳採取し、Luria-Bertani(LB)培地3 ml を入れた試験管に添加する。これを30℃で6-12時間振盪培養(160

rpm)する。この培養液に終濃度5 mMのIsopropyl-β-D-thiogalactopyranoside (IPTG)を加え、30℃でさらに3-4時間培養し、酵素を誘導する。この培養液を6,000

rpm で5分間の遠心分離を行い、菌体を集めてTris-HClバッファー(pH8.0)で3回洗浄し、300μlのTris-HClバッファーに再懸濁し、超音波による細胞破砕を行った。これを遠心分離15,000

rpm で10分間行い、その上清を粗酵素抽出液とし、2,4-D ジオキシゲナーゼ活性(TfdA活性)の測定に用いた。

【0055】

(2) 2,4-D分解活性の測定

2,4-Dの分解活性の測定用反応は30℃で行い、終濃度1mM

α-ketoglutarate、 50 μM Ascorbate、 50μM $Fe(NH_4)_2(SO_4)_2$、

20 mM Tris-HCl、 1mM 2,4-Dの反応バッファーに体積比1/10量の上記調整の粗酵素抽出液を加えることにより開始し、等量のメタノールを加えることで終了した。その反応終了液を遠心分離15,000

rpm で10分間行い、上清中に残存する2,4-Dの濃度を高速液体クロマトグラフィー(HPLC)にて分析した。分析条件の固定相はOctadecyl

Silica (ODS)、移動相は水:アセトニトリル:リン酸=49.5:49.5:1で、流速1 ml/minで行なった。

【0056】

(3) 野生型TfdA、特異的検出の標的となる塩基配列を導入した変異型TfdAm1、TfdAm2の活性

それぞれの粗酵素抽出液によるSpecific

activity(μmol/min/mg of protein)測定を行い、野生型TfdAは1.20 (μmol/min/mg of protein) の活性を示し、TfdAm1、TfdAm2ではそれぞれ1.16、1.15(μmol/min/mg

of protein) の活性を示した。野生型の活性値を100とすると、TfdAm1、TfdAm2では相対的にそれぞれ97、96となり、酵素活性値に有意な差は無く、特異的検出の標的となる塩基配列導入による2,4-D代謝活性の変化はほとんど無いことが結論付けられた。

【実施例】

【0057】

DNA 塩基配列を導入した2,4-D遺伝子の検出試験

本発明で構築した特異的検出の標的となる塩基配列(マーカー配列)を導入した遺伝子tfdAm1、

tfdAm2はPCRによって野生型のtfdA遺伝子と判別できる様に設計されている。これを確認するためマーカー配列を含むPCRプライマーを設計し、それぞれ野生型tfdA遺伝子、tfdAm1

、

tfdAm2を鋳型にPCR検出を行った。その結果、野生型ではtfdAm1、 tfdAm2をターゲットにしたプライマー(それぞれプライマーIとJ、及びプライマーIとK)を用いたPCRにより、DNA

断片の増幅が確認されず、tfdAm1、
tfdAm2ではそれぞれをターゲットとしたプライマーでDNA断片の増幅が確認された。これらのことから構築したtfdAm1、 tfdAm2が野生型のtfdAとは区別してPCRにより増幅することが可能であり、特異的に検出できることを確認した(図7参照)。

【0058】

プライマーI：5'-AAAAAGTGAGCGTCATCGGAG-3'（配列番号10）
プライマーJ：5'-GCCGAGATGTCGTACCTGCG-3'（配列番号11）
プライマーK：5'-TTGCCATCCAACGACACATTC-3'（配列番号12）

【実施例】

【0059】

環境サンプルからのtfdAm遺伝子の検出試験

本発明によるマーカー配列を導入した遺伝子は、環境サンプル中においても特異的に検出可能であることを示す。例として、下水処理場より採取した活性汚泥に混合したtfdAm1遺伝子が特異的に検出できることを示す。茨城県土浦市の霞ヶ浦下水処理場3池より採取した活性汚泥10
ml (MLSS=4877mg/L)にtfdAm1をプラスミドベクター上に持つ大腸菌を約10^1-10^6細胞(10-1,000,000)加え、その混合液から全DNAを抽出し、それを鋳型としてPCRを行った。またコントロールとして野生型tfdAを有する大腸菌を加えた活性汚泥、さらに何も加えない活性汚泥より抽出した全DNAを用いて実験を行った。活性汚泥10ml(乾燥重量約0.05g)から抽出した全DNA量は約12.5μgで、その1/10,000にあたる1.25ngを鋳型に20μlの反応系でtfdAm1をターゲットにしたプライマーIとJを用いたPCRを行った。その結果、活性汚泥から抽出した全DNAを鋳型にした場合、野生型tfdAを有する大腸菌を加えた活性汚泥DNAを鋳型にした場合おいて、DNA断片の増幅は確認されなかった。一方、tfdAm1をプラスミドベクター上に持つ大腸菌を含む活性汚泥から抽出したDNA試料からは1,000細胞以上加えた全ての試料においてDNA断片の増幅が確認された(図8参照)。

【0060】

これらの結果から、本発明で示した「遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法」また「遺伝子配列を特異的検出の標的となる塩基配列(マーカー配列)に置換する方法」によって構築された遺伝子が環境サンプル中からも特異的に検出が可能であることを示した。

【0061】

以上説明したように、本発明はマーカー配列を導入した遺伝子、具体的には、2,4-D
を分解し、2,4-ジクロロフェノールへ変換する酵素をコードする遺伝子、2,4-D ジオキシゲナーゼ遺伝子tfdA、および導入したDNA塩基配列とその導入方法を提供するものである。本発明に係る遺伝子は導入した特異的な塩基配列を含むプライマーを用いてPCR
を行うことにより、本発明の遺伝子を特異的かつ簡便に検出することが出来る。従って環境浄化を目的として、本発明の遺伝子を含む微生物を環境中に導入する場合、微生物の有効性や安全性を判断するための指標として、環境中に多様多種な微生物が存在する中から、導入した微生物や遺伝子を検出・追跡する事が出来る。

本発明において構築したtfdAm1遺伝子及びtfdAm2を以下に示す。

【0062】

tfdAm1塩基配列(下線部は付加配列を示す)（配列番号１）

GTGAGCGTCATCGGAGTCGTCGCAAATCCCCTTCATCCTCTTTTCGCCGCAGGGGTCGAAGACATCGACCTTCGAGAGGC
CTTGGGTTCGACCGAGGTCCGAGAGATCGAACGGCTAATGGACGAGAAGTCGGTGCTGGTGTTCCGGGGGCAGCCCCTGA
GTCAGGATCAGCAGATCGCCTTCGCGCGCAATTTCGGGCCACTCGAAGGCGGTTTCATCAAGGTCAATCAAAGACCTTCG
AGATTCAAGTACGCGGAGTTGGCGGACATCTCGAACGTCAGTCTCGACGGCAAGGTCGCGCAACGCGATGCGCGCGAGGT
GGTCGGGAACTTCGCGAACCAGCTCTGGCACAGCGACAGCTCCTTTCAGCAACCTGCTGCCCGCTACTCGATGCTCTCCG
CGGTGGTGGTTCCGCCGTCGGGCGGCGACACCGAGTTCTGCGACATGCGTGCGGCATACGACGCGCTGCCTCGGGACCTC
CAATCCGAGTTGGAAGGGCTGCGTGCCGAGCACTACGCACTGAACTCCCGCTTCCTGCTCGGCGACACCGACTATTCGGA
AGCGCAACGCAATGCCATGCCGCCGGTCAACTGGCCGCTGGTTCGAACCCACGCCGGCTCCGGGCGCAAGTTTCTCTTCA

TCGGCGCGCACGCGAGCCACGTCGAAGGCCTTCCGGTGGCCGAAGGCCGGATGCTGCTTGCGGAGCTTCTCGAGCACGCG
ACACAGCGGGAATTCGTGTACCGGCATCGCTGGAACGTGGGAGATCTGGTGATGTGGGACAACCGCTGCGTTCTTCACCG
CGGACGCAGGTACGACATCTCGGCCAGGCGTGAGCTGCGCCGGGCGACCACCCTGGACGATGCCGTCGTCTAG

【0063】

TfdAm2塩基配列(最初の下線部は付加配列、２番目の下線部は置換配列を示す)

（配列番号２）

GTGAGCGTCATCGGAGTCGTCGCAAATCCCCTTCATCCTCTTTTCGCCGCAGGGGTCGAAGACATCGACCTTCGAGAGGC
CTTGGGTTCGACCGAGGTCCGAGAGATCGAACGGCTAATGGACGAGAAGTCGGTGCTGGTGTTCCGGGGGCAGCCCCTGA
GTCAGGATCAGCAGATCGCCTTCGCGCGCAATTTCGGGCCACTCGAAGGCGGTTTCATCAAGGTCAATCAAAGACCTTCG
AGATTCAAGTACGCGGAGTTGGCGGACATCTCGAATGTGTCGTTGGATGGCAAGGTCGCGCAACGCGATGCGCGCGAGGT
GGTCGGGAACTTCGCGAACCAGCTCTGGCACAGCGACAGCTCCTTTCAGCAACCTGCTGCCCGCTACTCGATGCTCTCCG
CGGTGGTGGTTCCGCCGTCGGGCGGCGACACCGAGTTCTGCGACATGCGTGCGGCATACGACGCGCTGCCTCGGGACCTC
CAATCCGAGTTGGAAGGGCTGCGTGCCGAGCACTACGCACTGAACTCCCGCTTCCTGCTCGGCGACACCGACTATTCGGA
AGCGCAACGCAATGCCATGCCGCCGGTCAACTGGCCGCTGGTTCGAACCCACGCCGGCTCCGGGCGCAAGTTTCTCTTCA
TCGGCGCGCACGCGAGCCACGTCGAAGGCCTTCCGGTGGCCGAAGGCCGGATGCTGCTTGCGGAGCTTCTCGAGCACGCG
ACACAGCGGGAATTCGTGTACCGGCATCGCTGGAACGTGGGAGATCTGGTGATGTGGGACAACCGCTGCGTTCTTCACCG
CGGACGCAGGTACGACATCTCGGCCAGGCGTGAGCTGCGCCGGGCGACCACCCTGGACGATGCCGTCGTCTAG

【図面の簡単な説明】

【0064】

【図1】遺伝子配列に特異的検出の標的となる塩基配列(マーカー配列)を付加する方法の概要を示す図。

【図2】プライマーB、プライマーCの構造とクランプ部位及び相補領域を示す図。　プライマーB、プライマーCはそれぞれ別のPCR反応を行うが、そのお互いのPCR産物が相補領域を持つように設計する。

【図3】Aプライマー、Dプライマー伸長分子の結合を示す図。それぞれの分子が互いの鋳型、またプライマーとなって伸長反応が起こる。

【図4】遺伝子配列を特異的検出の標的となる塩基配列(マーカー配列)に置換する方法の概要を示す図。

【図5】tfdAへのマーカー配列を導入する位置及びプライマー設計、並びにtfdAへの特異的検出の標的となる塩基配列を導入するための設計を示す図。

【図6】tfdA遺伝子へのマーカー配列を導入する各過程のDNA断片の電気泳動写真。

【図7】導入したマーカー配列を利用したPCRによるtfdAm1、tfdAm2遺伝子の特異的検出結果を示す図。　特異的PCRにより増幅したDNA断片の電気泳動写真。野生型tfdA遺伝子の増幅はみられず、本発明の手法で構築したtfdAm1、tfdAm2遺伝子が特異的に検出可能であることを示した。

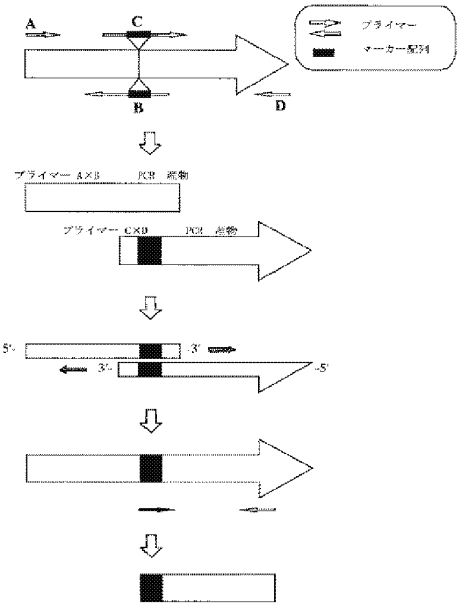
【図8】tfdAm1の活性汚泥中からの特異的検出結果を示す図。活性汚泥中のtfdAm1遺伝子の特異的PCR増幅断片の電気泳動写真。環境サンプル中に存在するマーカー配列を導入した遺伝子の特異的検出が可能であることを示した。

【配列表フリーテキスト】

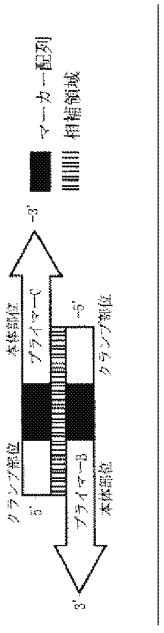
【0065】

配列番号３：合成ペプチド
配列番号４：合成DNA
配列番号５：合成DNA
配列番号６：合成DNA
配列番号７：合成DNA
配列番号８：合成DNA
配列番号９：合成DNA
配列番号10：合成DNA
配列番号11：合成DNA

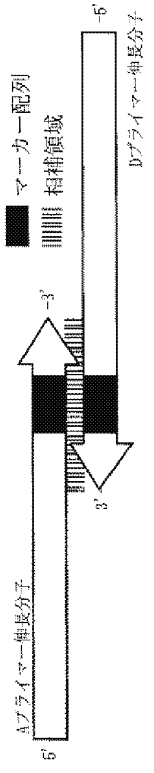
【図1】


A
C
B
D
プライマー
マーカー配列
プライマー AXB　PCR 産物
プライマー CXD　PCR 産物
5'
3'
3'
5'

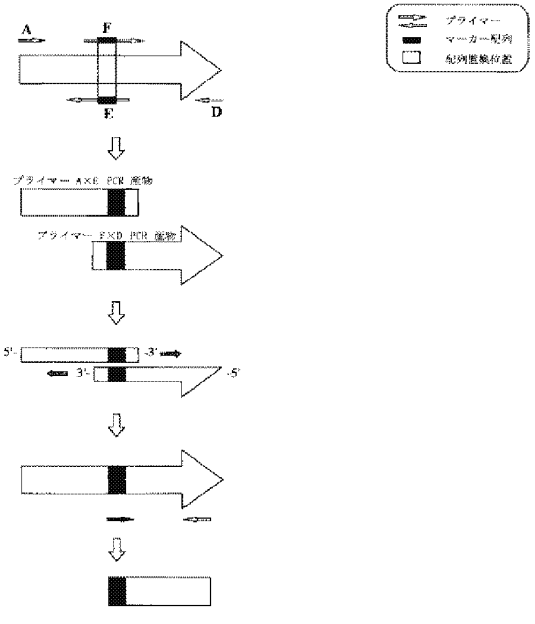
【図2】


マーカー配列
相補領域
3'
本体部位
プライマーC
5'
クランプ部位
クランプ部位
5'
プライマーB
本体部位
3'

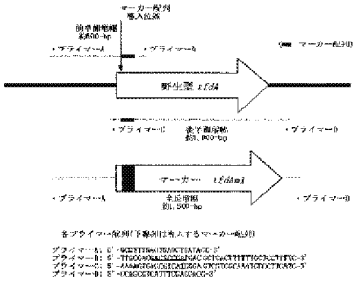
【図3】


マーカー配列
相補領域
Dプライマー伸長分子
5'
3'
3'
Aプライマー伸長分子
5'

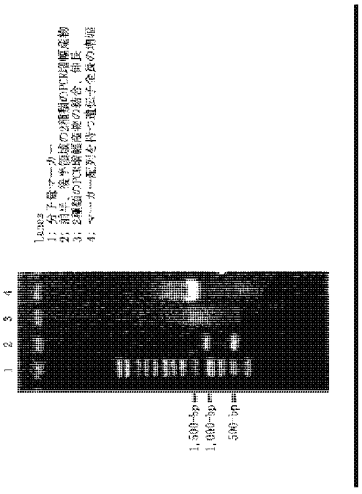
【図4】


プライマー
マーカー配列
配列置換位置
A
F
F
D
プライマー AXE PCR 産物
プライマー FXD PCR 産物
5'
3'
3'
5'

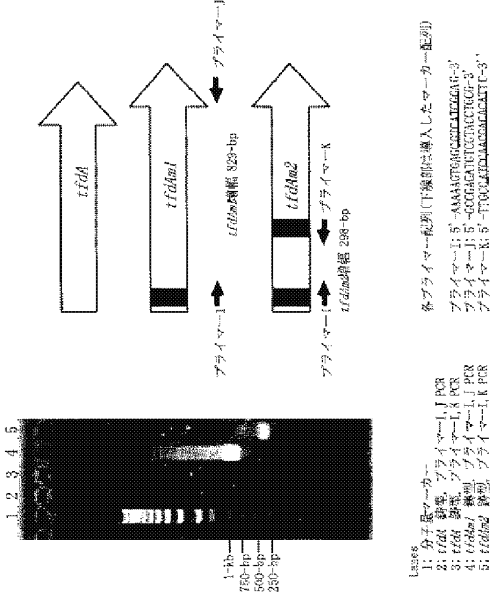
【図5】



【図6】



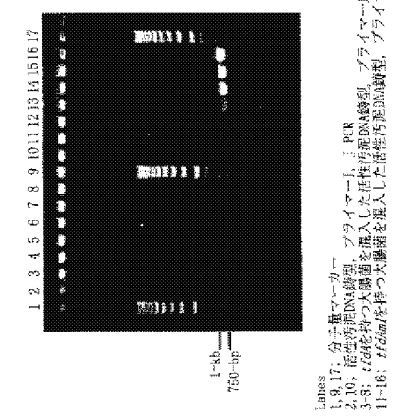
【図7】



【図8】

【配列表】

2005287415000001.app



Ｆターム(参考) 4B024 AA11 AA20 CA01 CA04 GA25 HA11
4B063 QA01 QA17 QQ42 QR08 QR42 QR62 QS25 QS36 QX02

Espacenet

Bibliographic data: JP2005287415 (A) — 2005-10-20

MUTATION-INTRODUCING METHOD OF SPECIFIC GENE

Inventor(s): NAKAMURA KAZUNORI; KAMAGATA YOICHI; KIMURA NOBUTADA; KITAGAWA KO ± (NAKAMURA KAZUNORI, ; KAMAGATA YOICHI, ; KIMURA NOBUTADA, ; KITAGAWA KO)

Applicant(s): NAT INST OF ADV IND & TECHNOL ± (NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL & TECHNOLOGY)

Classification: **- international:** ***C12N15/09; C12Q1/68;*** (IPC1-7): C12N15/09; C12Q1/68

- cooperative:

Application number: JP20040108166 20040331 Global Dossier

Priority number(s): JP20040108166 20040331

Abstract of JP2005287415 (A)

PROBLEM TO BE SOLVED: To provide a method for specifically detecting specific gene by introducing a marker sequence into a specific gene and amplifying, by PCR, the obtained gene with a primer containing the same marker sequence introduced into the gene. ;SOLUTION: The method for introducing mutation in the objective gene comprises amplifying the objective gene by using primer set containing, in at least one primer, the objective sequence of mutation introduction of the primer set for amplifying anterior region to the position where mutation to be introduced, and using primer set containing, in at least one primer, the objective sequence of mutation introduction of the primer set for amplifying posterior region, and connecting the obtained amplified fragments in the anterior region with the amplified fragments of the posterior region. ;COPYRIGHT: (C)2006,JPO&NCIPI

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/82 // A01H 1/00, 5/10

A1

(11) International Publication Number: **WO 98/08963**

(43) International Publication Date: 5 March 1998 (05.03.98)

(21) International Application Number: PCT/US97/15344

(22) International Filing Date: 29 August 1997 (29.08.97)

(30) Priority Data:
08/697,826 30 August 1996 (30.08.96) US

(60) Parent Application or Grant
(63) Related by Continuation
US Not furnished (CIP)
Filed on Not furnished

(71) Applicant *(for all designated States except US)*: REGENTS OF THE UNIVERSITY OF MINNESOTA [US/US]; Morrill Hall, 100 Church Street S.E., Minneapolis, MN 55455 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: GENGENBACH, Burle, G. [US/US]; 1994 Beacon Street, St. Paul, MN 55113 (US). SOMERS, David, A. [US/US]; 924 West County Road B2, Roseville, MN 55113 (US). EGLI, Margaret, A. [US/US]; 705 Heinel Circle, Roseville, MN 55113 (US). MARSHALL, Lorelei, C. [US/US]; 7 Glencrest Drive, Iowa City, IA 52246 (US). WYSE, Donald, L. [US/US]; 26370 Golden Court, Wyoming, MN 55092 (US). LUTZ, Sheila, M. [US/US]; Apartment 1, 271 Summit Avenue, St. Paul, MN 55102-2161 (US). VAN DEE, Kevin, L. [US/US]; 1202 Franklin Street, Iowa City, IA 52240 (US). PARKER, William, B. [US/US]; 19162 Highway 98, Boonville, MO 65233 (US).

(74) Agent: EMBRETSON, Janet, E.; Schwegman, Lundberg, Woessner & Kluth, P.O. Box 2938, Minneapolis, MN 55402 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: GENE COMBINATIONS FOR HERBICIDE TOLERANCE IN CORN

(57) Abstract

The present invention provides methods for preparing herbicide tolerant corn plants. Also provided are herbicide tolerant corn plants, as well as seeds and progeny derived from these plants.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

GENE COMBINATIONS FOR HERBICIDE TOLERANCE IN CORN

Background of the Invention

The use of selective herbicides for controlling specific weeds or plants in crops has become almost a universal practice. The market for these herbicides approaches a billion dollars annually. Even with this extensive use, weed control remains a significant and costly problem for the farmer.

Present day herbicides used singly or in so-called tank mixes require careful management to be effective. Time and method of application and stage of weed plant development are critical to getting good weed control with herbicides. Application of large amounts of preemergence herbicides can result in a commitment to grow the same crop in subsequent years because of chemical persistence in the soil which prevents rotation with a crop sensitive to that herbicide. Furthermore, some weed species are simply resistant to the available herbicides. Therefore, the development of effective herbicides increases in importance every year, especially as other weeds are controlled and thus reduce interplant competition.

Weed control in maize is currently accomplished by soil application of herbicides that are applied before the crop emerges and prior to the observation of a weed problem. The preemergence herbicides currently used adequately control most dicot and monocot (grass) weeds in maize. However, annual grass weeds such as wild proso millet and wooly cupgrass and perennial grass weeds commonly escape preemergence weed control. Preemergence herbicides require rainfall for activation, and under low rainfall conditions they fail to control grass weeds in maize. Furthermore, some preemergence herbicides persist in the soil and several have been detected as groundwater contaminants. The options for controlling these escape grass weeds are very limited.

A postemergence herbicide for grass weed control in maize would be very beneficial. An attractive alternative to developing new herbicides to combat this weed control problem in maize and/or to decrease the amount of herbicide carryover and groundwater contamination in maize fields from the

existing herbicides is to develop maize hybrids or varieties that are tolerant to other existing herbicides that normally kill all monocot (grass) species. The herbicide POAST™ (BASF Corp., Parsippany, New Jersey), the active ingredient of which is sethoxydim, kills most grasses, and is applied at lower rates than many preemergence herbicides. POAST™ is nonpersistent in the environment and therefore does not represent a groundwater contamination threat. POAST™ tolerant maize would provide the producer with increased weed management flexibility because POAST™ could be applied when a grass weed problem was detected without risk of damage to the crop and only to the areas with a weed problem. Therefore, postemergence control of local weed problems would further decrease the amount of herbicide applied compared to existing preemergence weed control strategies.

The sensitivity of maize to some herbicides is the result of the presence of herbicide sensitive forms of acetyl CoA carboxylase (ACCase) in those plants. ACCase is an enzyme involved in many important metabolic pathways in plant, animal and bacterial cells. Structurally, ACCases are biotinylated and are quite large enzymes consisting of one or more subunits. For example, most ACCases of animals, higher plants, and yeast are dimers of 420 to 700 kD native MW and contain subunits of 200 to 280 kD.

Two forms of ACCase, termed ACCase I and ACCase II, can be isolated from maize. These forms differ in their size, charge, cellular location, immunoreactivity with ACCase I antiserum, and sensitivity to herbicide inhibition. The predominate form, ACCase I, is plastid localized and is sensitive to herbicide inhibition.

Four ACCase genes have been identified in maize by Southern blot analyses (Lutz et al., Maize Genetics Conference (1995)). Restriction fragment length polymorphism (RFLP) analyses mapped one ACCase gene, termed *Acc1*, to chromosome 2 between umc131 and umc2b (Figure 1) in recombinant inbred lines from Tx303 x C059 (Egli et al., Maize Genetics Coop. Newsletter, 68, 92 (1994)). Mutations in the *Acc1* gene can result an ACCase that is resistant or tolerant to herbicide inactivation ((Marshall et al., Theor. Appl. Genet., 83,435

(1992); Egli et al., Plant Physiol., 101, 499 (1993); Egli et al., MNL, 66, 94 (1992)). However, the R0 plants which were regenerated from tissue cultures selected for resistance (or tolerance) to herbicides, by virtue of mutations in *Acc1*, exhibit only partial herbicide resistance, i.e., the symptoms induced by herbicide exposure are not prevented (see U.S. Patent No. 5,162,202).

Caffrey et al. (Maize Genetics Coop. Newsletter, 69, 3 (1995)) disclose that RFLP analyses of recombinant inbreds derived from the crosses Tx303 x CO159 and T232 x CM37 showed that one ACCase gene maps to chromosome 2 between umc131 and uox while another ACCase gene maps to chromosome 10 between ncsu2 and umc155 (Figure 2). Caffrey et al. further disclose that the ACCase gene on chromosome 10 appears to correspond to an herbicide resistance locus described by Van Dee et al. (Agro. Abs., page 198 (1992)). The authors propose that the ACCase gene on chromosome 10 encodes an herbicide sensitive ACCase that is localized to the plastid while the ACCase gene on chromosome 2 encodes an herbicide resistant ACCase that is non-plastid localized.

Thus, there is a need for a method to prepare a maize plant with resistance or tolerance to herbicides.

Summary of the Invention

The invention provides a method to prepare maize ("corn" or *Zea mays* L.) plants with resistance or tolerance to cyclohexanedione or aryloxyphenoxypropionate herbicides, i.e., maize plants with high tolerance to field application rates of herbicide. The herbicide resistance or tolerance is the result of the plants having at least one copy of each of two herbicide resistant or tolerant acetyl CoA carboxylase (ACCase) genes, one of which is encoded on chromosome 2 (*Acc1*) and the other of which is encoded on chromosome 10 (*Acc2*), i.e., the plant is a double heterozygous mutant (e.g., *Acc1-S2/+;+/Acc2-S5*). In contrast, when a plant has only one copy of either of the herbicide resistant ACCase genes, e.g., *Acc1-S2/+;+/+ or +/+;Acc2-S5/+*, the plant exhibits only partial tolerance to the herbicide, i.e., the expression of the

herbicide resistant ACCase gene does not prevent symptoms of herbicide damage under standard field-application rates of herbicides.

In commercial corn breeding practices, the double heterozygous genotype can be a hybrid variety planted by farmers. To prepare the double heterozygote, a homozygous double mutant parent line (e.g., *Acc1-S2/Acc1-S2;Acc2-S5/Acc2-S5*) is crossed to a normal (nonmutant) susceptible parent line. Thus, corn breeders can maintain fewer homozygous double mutant parent lines than would be needed with a system in which both parents must be homozygous for a single mutant gene to obtain a fully herbicide resistant maize plant.

Thus, the invention provides a method of imparting cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance to a corn plant. The method comprises crossing a first corn plant with a second corn plant so as to yield progeny plants. The first plant is homozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance. The second plant is homozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance. The progeny plant is heterozygous for the *Acc1* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and heterozygous for the *Acc2* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance. A preferred embodiment of the invention is a method of imparting herbicide tolerance to a hybrid plant.

The invention also provides a method of imparting tolerance to a corn plant to an agent which inhibits acetyl CoA carboxylase, wherein the agent is selected from the group consisting of 3-(2,4-dichlorophenyl)-perhydroindolizine-2,4-dione (Babczinski et al., Pesti. Sci., 33, 455 (1991)), 3-isopropyl-6-(N-[2,2-dimethylpropyl]-acetamido-1,3,5-triazine-2,4-(1H,3H)dione (Walker et al., Phytochem., 29, 3743 (1990)), soraphen A (Vahlenesiock et al., Curr. Genet., 25, 93 (1994)), and structural and/or functional analogs thereof. The method comprises crossing a first corn plant, which is homozygous for an allele of *Acc1* which imparts tolerance to the agent, with a second corn plant, which is homozygous for an allele of *Acc2* which imparts tolerance to the agent, so as to

yield progeny plants. The progeny plant is heterozygous for the *Acc1* allele which imparts agent tolerance and heterozygous for the *Acc2* allele which imparts agent tolerance.

Also provided is a method to prepare an herbicide resistant or tolerant corn plant. The method comprises crossing a first corn plant which comprises at least one herbicide resistant allele with a second corn plant which comprises at least one herbicide resistant allele which is not allelic to the herbicide resistant allele in the first plant, to yield a progeny plant which is a heterozygote for each allele.

The invention further provides a method of imparting cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance to a corn plant. The method comprises self pollinating a corn plant which comprises (i) an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and (ii) an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, so as to yield a progeny plant. Then a progeny plant is identified that is homozygous for the allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and is homozygous for the allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

Also provided is a method of preparing a plant which is a double heterozygote for alleles of *Acc1* and *Acc2* which impart cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance. The method comprises crossing a first corn plant with a second corn plant so as to yield progeny plants. The first plant is homozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and the second plant is homozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance. The progeny plant is heterozygous for the *Acc1* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and heterozygous for the *Acc2* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

The invention also provides a method of preparing a plant which is a double homozygote for alleles of *Acc1* and *Acc2* which impart cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance. The method comprises self pollinating a corn plant which comprises (i) an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and (ii) an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, so as to yield a progeny plant. A progeny plant is identified that is homozygous for the allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and is homozygous for the allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

Also provided are progeny and seed derived from the plants prepared by the methods described herein.

Yet another embodiment of the invention is an inbred or hybrid cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerant *Zea mays* plant. The genome of the inbred or hybrid plant of the invention is homozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, and is homozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

A further embodiment of the invention is an inbred or hybrid cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerant *Zea mays* plant, the genome of which is heterozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, and is heterozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

Progeny and seed derived from the inbred plants of the invention are also provided.

As used herein, the term "cyclohexanedione herbicide" includes, but is not limited to, 1,3-cyclohexanediones which exhibit general and selective herbicidal activity against plants. One such cyclohexanedione is sethoxydim {2-

[1-(ethoxyimino)-butyl]-5-[2-(ethylthio)propyl]-3-hydroxy-2-cyclohexen-1-one}. Sethoxydim is commercially available from BASF (Parsippany, New Jersey) under the designation POAST™.

Other herbicidal cyclohexanediones falling with in the scope of the invention include clethodim, (E,E)-(±)-2-[1-[[(3-chloro-2-propenyl)oxy] imino]propyl]-5-[2-(ethylthio)propyl]-3-hydroxy-2-cyclohexen-1-one; available as SELECT™ from Chevron Chemical (Valent) (Fresno, California); cloproxydim, (E,E)-2-[1-[[(3-chloro-2-propenyl)oxy]imino] butyl]-5-[2-(ethylthio)propyl]-3-hydroxy-2-cyclohexen-1-one; available as SELECTONE™ from Chevron Chemical (Valent) (Fresno, California); and tralkoxydim, 2-[1-(ethoxyimino)propyl]-3-hydroxy-5-mesitylcyclohex-2-enone, available as GRASP™ from Dow Chemical USA (Midland, Michigan), as well as other cyclohexanedione herbicides that are structurally related to the compounds described hereinabove.

As used herein, the term "aryloxyphenoxypropanoic acid herbicide" includes aryloxyphenoxypropanoic acids which exhibit general and selective herbicidal activity against plants. Such herbicides include, but are not limited to compounds wherein the aryloxy group may be phenoxy, pyridinyloxy or quinoxalinyl. One such herbicidal aryloxyphenoxypropanoic acid is haloxyfop, {2-[4-[[3-chloro-5-(trifluoromethyl)-2-pyridinyl]oxy]phenoxy]- propanoic acid}, which is available as VERDICT™ from Dow Chemical USA (Midland, Michigan). Another is diclofop, {(±)-2-[4-(2,4-dichlorophenoxy)-phenoxy]propanoic acid}, available as HOELON™ from Hoechst-Roussel Agri-Vet Company (Somerville, New Jersey). Other aryloxyphenoxypropanoic acid herbicides within the scope of the invention include fenoxyaprop, (±)-2-[4-[(6-chloro-2-benzoxazolyl)oxy] phenoxy]propanoic acid; available as WHIP™ from Hoechst-Roussel Agri-Vet Company (Somerville, New Jersey); fluazifop, (±)-2-[4-[[5-(trifluoromethyl)-2-pyridinyl]oxy]phenoxy]propanoic acid; available as FUSILADE™ from ICI Americas (Wilmington, Delaware); fluazifop-P, (R)-2-[4-[[5-(trifluoromethyl)-2-pyridinyl]oxy]phenoxy]propanoic acid; available as FUSILADE 2000™ from ICI Americas (Wilmington, Delaware); and

quizalofop, (±)-2-[4[(6-chloro-2-quinoxalinyl)- oxy]phenoxy]propanoic acid; available as ASSURE™ from E. I. DuPont de Nemours (Wilmington, Delaware), as well as other herbicidal compounds which are structurally related to the compounds described hereinabove.

As used herein, a plant that is "resistant or tolerant to inhibition by a herbicide or agent" is a plant that grows in an amount of herbicide or agent that normally inhibits growth of a corresponding susceptible plant, as determined by methodologies known to the art. For example, a maize plant of the invention (e.g., *Acc1-S2/Acc1-S2;Acc2-S5/Acc2-S5* or *Acc1-S2/+;Acc2-S5/+*) grows in an amount of cyclohexanedione that inhibits the growth of a corresponding susceptible maize plant (e.g., +/+;+/+, *Acc1-S2*/+;+/+ or +/+;*Acc2-S5*/+). In the alternative, the herbicide resistance of a homozygous backcross converted inbred plant of the invention (Inbred A *(Acc1-S2/Acc1-S2;Acc2-S5/Acc2-S5)* is compared to the herbicide resistance of a recurrent inbred susceptible plant (Inbred A (+/+;+/+)). A homozygous backcross converted inbred plant of the invention is a plant which has been repeatedly crossed to the recurrent inbred parent until the backcross converted inbred plant is substantially isogenic with the recurrent inbred parent except at *Acc1* and *Acc2* loci, and is then is self-pollinated (selfed) at least once.

As used herein, "substantially isogenic" means that the genomic DNA content of a homozygous backcross converted inbred plant is at least about 92%, preferably at least about 98%, and most preferably at least about 99%, identical to the genomic DNA content of a recurrent inbred parent of the backcross converted inbred plant.

Exemplary susceptible maize lines, e.g., lines which are sensitive to growth inhibition by cyclohexanedione or aryloxyphenoxypropanoic acid herbicides include, but are not limited to, A188, A641, A619, B73 and *Zea mays*, var. PI 3140. Exemplary maize lines which are a source of either *Acc1* or *Acc2* herbicide resistance alleles include, but are not limited to, *Zea mays*, var. DK 592_{SR}, *Zea mays*, var. DK 404_{SR}, 4400_{SR} and 7800_{SR}.

Brief Description of the Figures

Figure 1 depicts the genetic map of maize chromosome 2.

Figure 2 depicts the genetic map of maize chromosome 10.

Figure 3 depicts the DNA sequence (SEQ ID NO:9) of a 7470 base pair cDNA of a maize ACCase gene.

Figure 4 depicts the predicted amino acid sequence of a complete ACCase gene of maize (SEQ ID NO:10).

Figure 5 depicts the partial nucleotide sequence of a Type A_1 ACCase genomic clone (SEQ ID NO:11).

Figure 6 depicts the partial nucleotide sequence of clone 5A, a Type A ACCase genomic clone (SEQ ID NO:12).

Figure 7 depicts the partial nucleotide sequence of five Type A_2 ACCase genomic clones (A-F) (SEQ ID NOs 13, 14, 15, 16 and 17).

Figure 8 depicts the partial nucleotide sequence of three Type B ACCase clones (SEQ ID NOs 18, 19 and 20).

Detailed Description of the Invention

It is envisioned that a variety of corn breeding programs may be employed to introduce an herbicide tolerant gene or allele into a particular genetic background. Field crops are bred through techniques that take advantage of the plant's method of pollination. A plant is self-pollinating if pollen from one flower is transferred to the same or another flower of the same plant. A plant is cross pollinated if the pollen comes from a flower on a different plant.

Plants that have been self pollinated and selected for type for many generations become homozygous at almost all gene loci and produce a uniform populations of true breeding progeny. A cross between two homozygous plants from differing backgrounds or two homozygous lines produce a uniform population of hybrid plants that may be heterozygous for many gene loci. A cross of two plants that are each heterozygous at a number of gene loci will produce a population of hybrid plants that differ genetically and will not be uniform.

Corn plants can be bred by both self-pollination and cross pollination techniques. Corn has male flowers, located on the tassel, and female flowers,

located on the ear, on the same plant. Natural pollination occurs in corn when wind blows pollen from the tassels to the silks that protrude from the tops of incipient ears.

The development of corn hybrids requires the development of homozygous inbred lines, the crossing of these lines, and the evaluation of the crosses. Pedigree breeding and recurrent selection breeding methods are used to develop inbred lines from breeding populations. Breeding programs combine desirable traits from two or more inbred lines or various broad-based sources into breeding pools from which new inbred lines are developed by selfing and selection of desired phenotypes. The new inbreds are crossed with other inbred lines and the hybrids from these crosses are evaluated to determine which have commercial potential.

Pedigree breeding starts with the crossing of two genotypes, each of which may have one or more desirable characteristics that is lacking in the other or which complement the other. If the two original parents do not provide all of the desired characteristics, other sources can be included in the breeding population. In the pedigree method, superior plants are selfed and selected in successive generations. In the succeeding generations, the heterozygous condition gives way to homogenous lines as a result of self-pollination and selection. Typically, in the pedigree method of breeding, five or more generations of selfing and selection is practiced.

Backcrossing can be used to improve an inbred line. Backcrossing transfers a specific desirable trait from one inbred or other source to an inbred that lacks that trait (see below).

A hybrid corn variety is the cross of two inbred lines, each of which may have one or more desirable characteristics lacked by the other or which complement the other. The hybrid progeny of the first generation is designated F_1. In the development of hybrids, only the F_1 hybrid plants are sought. The F_1 is more vigorous than its inbred parents.

The development of a hybrid corn variety involves three steps: (1) the selection of superior plants from various germplasm pools; (2) the selfing of the