Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Published online 6 June 2008 Nucleic Acids Research, 2008, Vol. 36, No. 12 e76
doi:10.1093/nar/gkn369

Promoter library designed for fine-tuned gene expression in *Pichia pastoris*

Franz S. Hartner[1,2], Claudia Ruth[1], David Langenegger[3], Sabrina N. Johnson[4], Petr Hyka[3], Geoffrey P. Lin-Cereghino[4], Joan Lin-Cereghino[4], Karin Kovar[3], James M. Cregg[5] and Anton Glieder[1,*]

[1]Institute of Molecular Biotechnology, Graz University of Technology, [2]Research Centre Applied Biocatalysis, Petersgasse 14/2, 8010 Graz, Austria, [3]Institute of Biotechnology, Zurich University of Applied Sciences (ZHAW), Grüental, 8820 Wädenswil, Switzerland, [4]Department of Biological Sciences, University of the Pacific, Stockton, CA 95211 and [5]Keck Graduate Institute of Applied Life Sciences, 535 Watson Drive, Claremont, CA 91711, USA

Received March 19, 2008; Revised May 6, 2008; Accepted May 23, 2008

ABSTRACT

Although frequently used as protein production host, there is only a limited set of promoters available to drive the expression of recombinant proteins in *Pichia pastoris*. Fine-tuning of gene expression is often needed to maximize product yield and quality. However, for efficient knowledge-based engineering, a better understanding of promoter function is indispensable. Consequently, we created a promoter library by deletion and duplication of putative transcription factor-binding sites within the *AOX1* promoter (P_{AOX1}) sequence. This first library initially spanned an activity range between ~6% and >160% of the wild-type promoter activity. After characterization of the promoter library employing a green fluorescent protein (GFP) variant, the new regulatory toolbox was successfully utilized in a 'real case', i.e. the expression of industrial enzymes. Characterization of the library under repressing, derepressing and inducing conditions displayed at least 12 *cis*-acting elements involved in P_{AOX1}-driven high-level expression. Based on this deletion analysis, novel short artificial promoter variants were constructed by combining *cis*-acting elements with basal promoter. In addition to improving yields and quality of heterologous protein production, the new P_{AOX1} synthetic promoter library constitutes a basic toolbox to fine-tune gene expression in metabolic engineering and sequential induction of protein expression in synthetic biology.

INTRODUCTION

The methylotrophic yeast *Pichia pastoris* is a well-established yeast expression host with rising popularity. Several hundred proteins from almost all kingdoms of life have been expressed in this yeast, either intracellularly or secreted into the medium (1). As a yeast, *P. pastoris* combines the advantages of eukaryotes, like efficient secretion and eukaryotic post-translational modifications (e.g. proteolytic processing, glycosylation and disulfide bond formation), with fast growth on economic, salt-based media, with little risk of contamination with endotoxins or oncogenic or viral DNAs (2). In addition, *P. pastoris* belongs to the group of Crabtree negative yeasts and has a very low maintenance energy demand, which makes it well suited for high cell density (HCD) fermentation (3–5). Along with these favourable physiologic characteristics, the availability of strong and tightly regulated promoters makes this yeast a very attractive host for recombinant protein production. In fact, in HCD processes, up to 13 g/l of secreted protein (6) or 22 g/l intracellular protein have been obtained (7).

Although several promoters, either constitutive or inducible, are available for this yeast [P_{AOX1} (8), P_{AOX2} (9), P_{FLD1} (10), P_{DAS} (8), P_{PEX8} (P_{PER3}) (11), P_{YPT1} (12), P_{GAP} (13), P_{TEF1} (14), P_{PGK1} (15), P_{ICL1} (16)], the promoter of the alcohol oxidase I gene (*AOX1*) has been employed in most studies and applications (1). *AOX1* is the predominantly expressed form of two alcohol oxidases (*AOX1* and *AOX2*) in the *P. pastoris* cell. Besides being the first well-characterized promoter of this yeast and its availability in a commercial kit (Invitrogen, Carlsbad, CA, USA), P_{AOX1}'s extraordinary strength and tight regulation by carbon sources are the main reasons for its predominant use. Being tightly repressed by glucose, its transcription is highly induced upon shift to methanol as sole carbon source (8,17). Other carbon sources which repress transcription from the *AOX1* gene are glycerol and ethanol, while very low derepression of transcription occurs upon carbon starvation (8,17). In contrast to the alcohol oxidase genes from the other prominent methylotrophic yeasts (*MOX* from *Hansenula polymorpha*, *AOD1* from

*To whom correspondence should be addressed. Tel: +43 316 873 4074; Fax: +43 316 873 4071; Email: a.glieder@tugraz.at

© 2008 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Candida boidinii and *MOD1* and *MOD2* from *Pichia methanolica*), regulation of both alcohol oxidase genes in *P. pastoris* is strictly regulated. Although glucose repression of methanol utilization pathway genes takes place in all these yeasts, the presence of methanol is not absolutely necessary for high-level induction of alcohol oxidases in the other methylotrophic yeasts studied so far (17).

Despite its industrial importance and multiple manuscripts analyzing its expression under different fermentation strategies (18,19), little is known about the regulation of the *AOX1* promoter at the molecular level, either in terms of the *cis*- or *trans*-acting elements. Deletion analysis of P_{AOX2}, the promoter of the minor alcohol oxidase gene *AOX2*, resulted in the identification of three *cis*-acting elements in that promoter, two upstream repression sites (URS1 and URS2) and one upstream activation sequence (UAS) (9). Although some sequence similarity between the *AOX2* UAS to regions within the *AOX1* promoter was detected, no conserved function has been demonstrated so far. Lin-Cereghino *et al.* (20) identified a positive acting transcription factor (TF) (*MXR1*), which bound a P_{AOX1} region of approximately 250 base pairs. Mxr1p displayed sequence similarity to the *Saccharomyces cerevisiae* zinc-finger TF Adr1p, which is involved in regulation of the glucose-repressed *ADH2* gene and some peroxisomal proteins in baker's yeast (21,22). *Cis*-acting elements sensitive to methanol induction were also identified in the promoter sequences of the *H. polymorpha MOX* (23) and the *C. boidinii FDH* genes (24). Similar sequences exist in the *P. pastoris AOX1* promoter sequence.

Furthermore, it has been demonstrated that the endogenous Adr1p is involved in the induction of the *MOX* promoter upon shift from glucose to glycerol in *S. cerevisiae* by directly interacting with the promoter nucleotides between −245 and −112 (25). Taken together, focusing on putative transcription factor binding sites (TFBSs) within *P. pastoris* P_{AOX1} should give some insight into how this promoter is regulated and provide elements for a modular synthetic promoter design, even if the function of the respective TFs and/or their binding sites have been changed in *P. pastoris* during evolution.

In this study, we followed three major aims: (i) for the first time, we employed a comprehensive deletion analysis to identify the individual effects of small sequence elements and to surmise, which TFs might be involved in P_{AOX1} regulation; (ii) we created a simple set of promoters with regulatory properties different from P_{AOX1}, which are more desirable for heterologous protein expression and cell engineering and (iii) we demonstrated that even the extraordinarily strong *AOX1* promoter still has considerable potential for improvement by engineering.

MATERIALS AND METHODS

Sequence analysis of the *AOX1* promoter

The P_{AOX1} sequence from pPICZ B (downloaded from www.invitrogen.com) was used to search for putative TFBSs. TF analysis was done with MatInspector (26,27) release professional 6.1 January 2003 within the GenomatixSuite 1.6.1 at Genomatix Software GmbH Servers using the Matrix Family Library Version 3.1.1 April 2003 group ALL fungi.lib (www.genomatix.de). Beside this analysis, sequence modifications which were already described in literature (9,23,28) and their effects upon promoter derepression and induction were compared.

Chemicals and media

Oligonucleotides were purchased from either MWG Biotech AG (Ebersberg, Germany) or Invitrogen Corp. (Carlsbad, CA, USA). Sterile water was purchased from Fresenius Kabi Austria (Graz, Austria). Unless otherwise stated, all chemicals were purchased from Carl Roth GmbH (Karlsruhe, Germany), Becton, Dickinson and Company (Franklin Lakes, NJ, USA) and Sigma-Aldrich (St Louis, MO, USA). Media for small-scale experiments in deep-well plates were prepared essentially as described earlier (29). Complex media contained 10 g/l yeast extract, 20 g/l peptone and either 20 g/l glucose (YPD) or 10 g/l glycerol (YPG). Media for plates were solidified by addition of agar to 1.5% w/v. All five minimal media used contained per litre 200 mM potassium phosphate buffer at pH 6, yeast nitrogen base 13.4 g and D-biotin 0.0004 g, and differed with respect to the substrate concentration of 10 or 0.2 g/l of glucose, 0.5 g/l of glycerol or 0.5, 1 or 5% (v/v) of methanol for BMD (1%), BMD (0.2%), BMG (0.5%), BMM, BMM2 or BMM10, respectively.

All seed cultures (and inoculum cultures) were prepared using the buffered glycerol-complex medium (BMGY, Invitrogen), which contained per litre glycerol 10 g, yeast extract 10 g, peptone 20 g, 100 mM potassium phosphate buffer (pH 6.0), yeast nitrogen base without amino acids 13.4 g and biotin 0.0004 g.

The defined mineral medium used for the main fedbatch cultures contained per litre $CaSO_4{\cdot}2H_2O$ 0.17 g, K_2SO_4 2.86 g, KOH 0.64 g, $MgSO_4{\cdot}7H_2O$ 2.3 g, EDTA 0.2 g, H_3PO_4 7.225 g, $FeSO_4{\cdot}7H_2O$ 0.2826 g, H_3BO_3 0.00087 g, $CuSO_4$ 0.0167 g, $MnSO_4{\cdot}H_2O$ 0.013 g, $Na_2MoO_4{\cdot}2H_2O$ 0.00087 g, $ZnCl_2$ 0.087 g, NaI 0.00035 g, $CoCl_2{\cdot}6H_2O$ 0.004 g, H_2SO_4 (Synopharm, Basel, Switzerland) 0.04 g and biotin 0.00087 g. In total, per litre either 784.2 g glucose monohydrate or 773.24 g methanol was completed with 0.0024 g biotin and 12 ml PTM1 mineral salts solution [i.e. containing per litre 5.0 ml of 69% H_2SO_4 (Synopharm), $CuSO_4$ 3.84 g, NaI 0.08 g, $MnSO_4{\cdot}H_2O$ 3.0 g, $Na_2MoO_4{\cdot}2H_2O$ 0.2 g, H_3BO_3 0.02 g, $CoCl_2{\cdot}6H_2O$ 0.92 g, $ZnCl_2$ 20.0 g and $FeSO_4{\cdot}7H_2O$ 65.0 g].

Strains and plasmids

Standard molecular biology procedures were performed according to ref. (30). *Escherichia coli* DH5α-T1^R or TOP10F' (Invitrogen Corp.) were used for all *E. coli* cloning experiments.

The cycle-3-GFP (31) coding sequence was amplified from pTRACER-CMV2 (Invitrogen Corp.) using Phusion™ polymerase (Finnzymes Oy, Espoo, Finland) with primers GFPZeo forw and wtGFP-XhoI-r according to the supplier's manual. The PCR product was TOPO

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

cloned into vector pCR4blunt-TOPO (Invitrogen Corp.) according to the supplier's manual. Unfavourable BamHI and XhoI restriction sites within the GFP coding sequence were removed using primers BamHI-del-f and XhoI-del-f using the QuickChange® Multi Site-Directed Mutagenesis Kit (Stratagene Corp.) according to the supplied protocol. The GFP gene was subsequently cloned into vector pGAPZ A (Invitrogen Corp.) employing EcoRI and XhoI as restriction enzymes. The *AOX1* promoter was synthesized according to the sequence of the plasmid pPICZ B and cloned into vector pUC18. The plasmid was cut with BglII and EcoRI, the promoter fragment was purified via agarose gel electrophoresis and the Wizard® SV Gel and PCR Cleanup System (Promega Corp., Madison, WI, USA) and ligated into the BglII- and EcoRI-cut vector pGAPZA-GFP. The resulting vector was named pPICZ-GFP.

Generation of promoter mutants

Short deletions or insertions within the *AOX1* promoter of pPICZ-GFP were generated by site-directed mutagenesis (SDM) using PfuUltra™ polymerase (Stratagene Inc.) according to the two-step SDM protocol (32). All other modifications which were not introduced by SDM were generated by overlap extension PCR (oePCR). 5′ arms of P_{AOX1} were synthesized by PCR using P(AOX1)forw as forward primer and appropriate reverse primers, 3′ arms using the appropriate forward primers and P(AOX1)rev as reverse primer. Truncated *AOX1* promoter variants were generated using an appropriate forward primer and P(AOX1)rev as reverse primer.

All arms were synthesized using 12 ng of pPICZ B as template and 10 pmol of each primer. The reaction also contained 200 μM of each dNTP and 0.6 U *Pwo* DNA polymerase (MWG Biotech GmbH) in appropriate buffer conditions in a final volume of 50 μl.

PCR was performed in an GeneAmp® PCR System 2700 (Applied Biosystems, Foster City, CA, USA) for 30 cycles (95°C, 1 min; 55°C, 1 min; 68°C, 1 min 30 s) with an initial denaturation step of 5 min at 95°C and a final extension step of 10 min at 68°C. All arms were purified by agarose gel electrophoresis and gel extraction prior use for oePCR. The reaction contained 10 ng of each purified arm, 200 μM of each dNTP and 0.6 U *Pwo* DNA polymerase in appropriate buffer conditions in a final volume of 50 μl. PCR was performed according to the manufacturer's suggestion. After 10 cycles, 20 μl of a mixture containing 10 pmol of the outer primers P(AOX1)forw and P(AOX1)rev, again 200 μM of each dNTP and 0.6 U *Pwo* DNA polymerase in appropriate buffer conditions were added. The PCR was continued as programmed after addition of this amplification mixture.

The obtained PCR products representing P_{AOX1} variants with the desired size of about 898–947 bp were purified on an agarose gel and cloned into the pGAPZ-GFP backbone replacing P_{GAP}.

A basal Leu2 promoter was generated by mixing 25 pmol of oligos Leu2basal1fw, Leu2basal 1rv, Leu2basal 2fw and Leu2basal 2rv in a total volume of 20 μl. The mixture was incubated at 95°C for 2 min followed by a slow cooling to room temperature. Three microlitres of the mixture was ligated into BglII/EcoRI cut pGAPZ-GFP backbone vector thereby replacing P_{GAP}. The resulting vector was named pLeu2basal-GFP. Basal promoter variants AOX194 and AOX176 were generated by amplifying the basal fragment from vector pPICZ-GFP using reverse primer AOX1basalrv and forward primers AOX1basalfwn and AOX176fw, respectively. After purification, the PCR products were cut with BglII/EcoRI and ligated into vector pGAPZ-GFP, thereby replacing P_{GAP}. The resulting vectors were named pAOX194-GFP and pAOX176-GFP. Truncated promoter fragments added to basal promoter variants were amplified from pPICZ-GFP using reverse primer AOX1oe176rv and forward primers P(AOX1)forw, AOX1fw-SacBgl and 415BglIIn, thereby generating fragments with a 5′ end at −940, −731 and −411, respectively. After purification, the PCR fragments were cut with BglII/BspTI and ligated into vectors containing basal promoters Leu2, AOX176 and AOX194. The resulting vectors were named pLeu2-176-GFP-full, pLeu2-176-GFP-SacI, pLeu2-176-GFP-d1.0, pAOX176-GFP-full, pAOX176-GFP-SacI, pAOX176-GFP-d1.0, pAOX194-GFP-full, pAOX176-GFP-SacI and pAOX176-GFP-d1.0 (Supplementary Figure 1A). The promoter variants truncated at −203 were amplified from plasmids pLeu2-176-GFP-full and pAOX176-GFP-full using primer pairs 214BglII/LEU2basalrv and 214BglII/AOX1basalrv, respectively. After purification, the PCR products were cut with BglII/EcoRI and ligated into vector pGAPZ-GFP, thereby replacing P_{GAP}. The resulting vectors were named pLeu2-176-GFP-214 and pAOX176-GFP-214.

Synthetic promoter variants were generated by amplifying the basal AOX176 promoter with forward primers 201-214AOX176, 737-38AOX176, ADR-AOX176, Stre-AOX176, Mat-AOX176 and Gcr-AOX176 and reverse primer AOX1basalrv. The PCR fragments were purified, cut with BglII/EcoRI and ligated into vector pGAPZ-GFP, thereby replacing P_{GAP}. The resulting plasmids were named pAOX176-GFP-201-214, pAOX176-GFP-737, pAOX176-GFP-Adr1, pAOX176-GFP-Stre, pAOX176-GFP-Mat and pAOX176-GFP-Gcr1. All primers are listed in Supplementary Table 1.

Generation of plasmids for the expression of alternative reporter genes

An NdeI restriction recognition site was introduced by SDM 5′ of the BglII site at the 5′ end of the *AOX1* promoter in the plasmid pPICZαB-HRP (a kind gift of Frances Arnold, Caltech) containing the gene coding for the mature protein of horseradish peroxidase, isoenzyme C1A (HRP) fused to the *S. cerevisiae* α-factor signal sequence and the wild-type P_{AOX1} (33). This modification enabled the exchange of promoter variants in this plasmid employing NdeI and HindIII. A SbfI and an AscI restriction site were introduced employing the QuickChange® Multi Site-Directed Mutagenesis Kit (Stratagene Corp.) according to the supplied protocol. The final plasmid was called pHRP-AOX1. The newly introduced NdeI

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

restriction site can also be used for plasmid linearization for *Pichia* transformation.

The *Pa*HNL5α gene (34), where the native hydroxynitrile lyase, isoenzyme 5 (HNL5) leader signal was replaced by the *S. cerevisiae* mating factor α leader sequence was cloned cloned into pGAPZ-A (Invitrogen Corp.) employing the EcoRI restriction sites flanking this gene construct on both ends. The resulting plasmid with the open reading frame in correct orientation was named pGAPPaHNL5α. All promoter variants were introduced into pGAPPaHNL5α employing the BglII and the EcoRI restriction sites. The resulting vectors were named pHNL5-AOX1 for the wild-type promoter and for promoter mutants the respective mutant name was added at the end.

The gene for the intracellular *Hevea brasiliensis* hydroxynitrile lyase gene (*Hb*HNL) was amplified from plasmid pHNL-400 (7) by PCR using primers HNL-Ecofw and HNL-Ecorv and ligated into pPICZ B (Invitrogen Corp.) employing EcoRI sites. The resulting plasmid was named pSHNL-AOX1. The *AOX1* promoter was replaced with promoter mutants via BglII and HindIII.

The series using the β-galactosidase reporter was constructed as follows. The 3.3-kb fragment of pLG178F [(35), kindly provided by Leonard Guarente, MIT, Cambridge, MA, USA], containing the *S. cerevisiae CYC1* TATA box fused to a partial *lacZ* coding sequence, was inserted into the same restriction sites in pBLHIS (36) to create pGC140. The 350 bp EcoRI–HindIII fragment, containing the *AOX1* transcription termination sequence of pHILD2 (Invitrogen), was then inserted into the same site of pGC140 to produce pGC141. A 1.3-kb PCR fragment, containing the *S. cerevisiae LEU2* TATA box fused to the 5′ end of the *lacZ* coding sequence, was amplified from pMC2019 [(37), kindly provided by Malcolm Casadaban, University of Chicago, Chicago, IL, USA] using the primers LEU1X and LEUR5. The PCR product was digested with XhoI and EcoRV and used to replace the XhoI–EcoRV fragment, containing the *CYC1* TATA element, of pGC141 to construct pGC146. The full-length *lacZ* coding sequence was cloned into pGC146 by amplifying a 2.1-kb fragment from pMC2019 with primers LACZX5 and LACZX3 and inserting it into the EcoRV and EcoRI sites of pGC146 to produce pGC187. The pGC187 contains the basal promoter of *S. cerevisiae LEU2*, with no upstream promoter region, fused to the complete coding sequence of *lacZ* followed by the *AOX1* transcription termination region. For amplification of *AOX1* promoter regions and cloning of sufficiency constructs, pHILD-2 was used as the template in PCR. The pGC179 was produced by inserting the 245-bp *AOX1* promoter region created by PCR with the primers SUFFS1 and SUFFX into the SpeI and XhoI sites of pGC187. The pGC180 was created by digesting pGC187 with SpeI and XhoI, blunt ending both ends with Klenow fragment and ligating the ends together. The pGC181 was produced by inserting the 742-bp *AOX1* promoter region created by PCR with the primers SUFFS0 and SUFFX into the SpeI and XhoI sites of pGC187. The pGC183 was constructed by inserting the 70-bp *AOX1* promoter region created by PCR with the primers SUFFS3 and SUFFX into the SpeI and XhoI sites of pGC187. pGC191 was produced by inserting the 45-bp *AOX1* promoter region created by PCR with the primers SUFF170S and SUFFAUS214X into the SpeI and XhoI sites of pGC179, replacing the 245-bp *AOX1* fragment.

Pichia transformation

Pichia pastoris was transformed according to a condensed protocol (38) after linearization of the plasmid with BglII and subsequent purification and concentration using the Wizard® SV Gel and PCR Cleanup System (Promega Corp.). The mixture was transferred to ice-cold electrotransformation cuvettes (0.2 cm, Bio-Rad, Munich, Germany) and pulsed at 200 Ω, 25 μF and 1.5 kV. 0.5 ml of ice-cold 1 M sorbitol was added immediately. The suspension was transferred to a sterile 12 ml PP-tube (Greiner, Frickenhausen, Germany). After regeneration for 1 h at 30°C without shaking, 0.5 ml of YPD was added followed by incubation for an additional hour. After regeneration, aliquots were plated on selection plates.

Promoter library characterization

A master plate containing representative strains bearing the promoter variant-GFP expression cassette was generated by inoculating 300 μl YPG with single colonies of the strains. After a 24 h incubation period at 28°C, 320 r.p.m., 80% humidity, 150 μl 60% glycerol was added and further incubated for 30 min. Then the library was transferred in 100 μl aliquots to 96-well microtitre plates and stored at −80°C. These glycerol stock master plates were used for all subsequent experiments.

A pre-culture in BMD (0.2%) was grown for 48 h to obtain uniform cell densities in all individual wells. The glycerol stock master plates were replica stamped to a preculture plate with 300 μl BMD (0.2%). The plate was incubated after inoculation at 28°C, 80% humidity and 320 r.p.m. for 48 h.

For determination of the promoter activity under repressing and derepressing conditions, three parallel deep-well plates containing 500 μl of appropriate growth medium [BMD (1%), BMG (0.5%)] was inoculated with 10 μl of the pre-cultures. Growth and GFP expression was followed by measuring optical density at 595 nm and fluorescence intensity for up to 70 h.

For determination of the promoter activity under methanol inducing conditions, three parallel deep-well plates containing 300 μl of appropriate growth medium, [BMD (1%) or BMG (0.5%)] was inoculated with 10 μl of the pre-cultures and incubated for 60 h. After this time, 20 μl were taken for determination of the GFP fluorescence and OD measurements. Induction was performed by adding 250 μl BMM2/well followed by a further incubation for 72 h. Additional methanol was fed by adding 50 μl BMM10 after 10, 24, 36, 48 and 60 h. Twenty microlitre samples were taken every 12 h before methanol addition for determination of fluorescence intensity and optical density.

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Quantification of GFP fluorescence and cell density measurements

A total of 20 or 50 µl of culture were added to 180 or 150 µl of water in fluorescence microtitre plates with clear bottom (Greiner Bio-One GmbH, Kremsmünster, Austria). Expression of GFP in *P. pastoris* was analysed by fluorescence measurements in a Spectramax Gemini XS plate reader (Molecular Devices Corp., Sunnyvale, CA, USA) with excitation at 395 nm and emission at 507 nm. Optical density was measured at 600 nm using a Spectramax Plus 384 plate reader (Molecular Devices Corp.). The GFP fluorescence was calculated by subtracting the blank value (*P. pastoris* X-33 grown and measured under the same conditions) from the sample value. The measured values were corrected for the dilution during measurement and for the dilution due to addition of the methanol induction medium.

Copy number determination by quantitative PCR

Copy number of the expression cassette was determined by quantitative real-time PCR using the *P. pastoris ARG4* gene as reference gene. Quantitative real-time PCR was performed using Power SYBR® Green PCR Master Mix (Applied Biosystems) in an ABI PRISM 7300 Real Time PCR System (Applied Biosystems). GFP-RTfw/GFP-RTrv and ARG4-RTfw/ARG4-RTrv were used as primers at concentrations of 200 nM with 1 ng of genomic DNA as template. Temperature conditions were 10 min at 95°C, 40 cycles at 15 s at 95°C and 60 s at 60°C followed by a dissociation step (15 s 95°C, 30 s 60°C, 15 s 95°C) at the end of the last cycle.

Activity assays for additional reporter systems

The HNL activity and HRP activity was determined essentially as described earlier (29), with just a few modifications: instead of an endpoint assay for the HRP activity, the activity was determined with a kinetic assay following the absorption at 405 nm for 5 min. The 15 µl of supernatant was transferred to a microtitre plate and the reaction was started with 140 µl of the substrate solution as described earlier (33). For the *Hb*HNL activity 100 µl culture were centrifuged and the pellet was lysed with 50 µl "Y-PER reagent" (Pierce, Rockford, IL, USA) for 30 min at 25°C. After centrifugation, 20 µl of supernatant was diluted 1:5 with 5 mM potassium phosphate buffer, pH 6.5. A 50 µl of the diluted sample was taken for the activity assay as described previously (29) for the secreted *Pa*HNL5 (100 µl 0.1 M phosphate buffer was used instead of 130 µl).

Fedbatch cultures in bioreactor and process monitoring

All cultivation processes were performed in a 14 l stirred tank bioreactor (MAVAG, Neunkirch, Switzerland). The initial batch working volume was 6 l and increased to total 12 l at the time of harvesting. The set-points of all control variables were maintained the same during the entire process duration and, thus, the cultivations were accomplished under the conditions of constant temperature 30 C, 18 l/min airflow (i.e. without any oxygen enrichment), 1400 r.p.m. agitation, 0.5 bar overpressure and pH 5. Ammonia solution 25% and phosphoric acid 8.5% were used to automatically control the pH.

The process consisted of a phase of biomass growth (i.e. in both batch and exponentially fed fedbatch cultures) and production phase with constant feed addition. The cultivation began at time 0 h with a batch (containing 33 g/l glucose monohydrate), during which the pO_2 value continuously decreased (i.e. no control of the pO_2 set-point applied). After some 12–14 h, the pO_2 rose again rapidly, and at this point the exponential addition of glucose solution was started and continued over 13.5 h in accordance with the function $f(t) = 4.655\ e^{0.2 \cdot t}$ (in gram of carbon per hour). During the subsequent production phase, either glucose or methanol as sole substrate was added at a constant rate of 29.54 g of carbon h^{-1} for at least 32.5 h.

Biomass concentration was determined gravimetrically as cell dry weight (CDW). Samples were centrifuged for 5 min at 14 000 r.p.m. (5417C, Nethler, Hamburg, Germany) in pre-weighted 2 ml Eppendorf tubes, washed and re-centrifuged, and subsequently dried for 24 h at 105°C (Heraeus Instruments, Zurich, Switzerland) to constant weight. Concentrations of glucose, methanol and other metabolites were analysed by HPLC using an LC-20AB device equipped with autosampler SIL-20A, thermostated column cell CTO-20A and RID-10A refractometer detector (all Shimadzu). Aminex HPX-87H column of $300 \times 7.8\ mm^2$ I.D. (Biorad) was operated at 40°C, flow rate of 0.6 ml/min under isocratic conditions with 2.5 mmol/l H_2SO_4. The injection volume was 25 µl.

RESULTS

Sequence analysis of the *AOX1* promoter

To carry out a comprehensive analysis of the *AOX1* promoter, we first summarized all sequences previously demonstrated to be involved in promoter regulation by methanol and for which we could find a homologous region in the P_{AOX1} sequence. In addition, we performed an *in silico* analysis of the P_{AOX1} sequence for putative TFBSs for eukaryotic TFs (Figure 1A and Supplementary Table 2).

The putative TFBSs found within the *AOX1* promoter sequence can be divided into three general groups by the effects of the TFs:

- glucose repression or regulation of glucose-repressed genes, especially those related to oxygen consumption (e.g. Adr1p and the Hap2/3/4/5p complex);
- stress response (e.g. Msn2p/Msn4p and HSF);
- developmental programs and general TFs (e.g. abaA, Rap1p, QA-1F, Mat1-Mc).

Based on the sequences identified, we first generated a deletion library where small internal sequence stretches, including putative TFBSs, were deleted. The library contained both small deletions of 4–11 nt and larger stretches of up to 66 bp, which were surmised to encompass several putative TFBSs.

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

A

```
                    -960                                                    -881
                    AGATCTAACA TCCAAAGACG AAAGGTTGAA TGAAACCTTT TTGCCATCCG ACATCCACAG
                    TCTAGATTGT AGGTTTCTGC TTTCCAACTT ACTTTGGAAA AACGGTAGGC TGTAGGTGTC
                                                                           Hap1
-880                                                                        -801
GTCCATTCTC ACACATAAGT GCCAAACGCA ACAGGAGGGG ATACACTAGC AGCAGACCGT TGCAAACGCA GGACCTCCAC
CAGGTAAGAG TGTGTATTCA CGGTTTGCGT TGTCCTCCCC TATGTGATCG TCGTCTGGCA ACGTTTGCGT CCTGGAGGTG
                                                                              HSF
-800                                                                        -721
TCCTCTTCTC CTCAACACCC ACTTTTGCCA TCGAAAAACC AGCCCAGTTA TTGGGCTTGA TTGGAGCTCG CTCATTCCAA
AGGAGAAGAG GAGTTGTGGG TGAAAACGGT AGCTTTTTGG TCGGGTCAAT AACCCGAACT AACCTCGAGC GAGTAAGGTT
   HSF                                                Hap2/3/4/5  (2x)         abaA
-720                                                                        -641
TTCCTTCTAT TAGGCTACTA ACACCATGAC TTTATTAGCC TGTCTATCCT GGCCCCCCTG GCGAGGTTCA TGTTTGTTTA
AAGGAAGATA ATCCGATGAT TGTGGTACTG AAATAATCGG ACAGATAGGA CCGGGGGGAC CGCTCCAAGT ACAAACAAAT
                                                                 STRE
-640                                                                        -561
TTTCCGAATG CAACAAGCTC CGCATTACAC CCGAACATCA CTCCAGATGA GGGCTTTCTG AGTGTGGGGT CAAATAGTTT
AAAGGCTTAC GTTGTTCGAG GCGTAATGTG GGCTTGTAGT GAGGTCTACT CCCGAAAGAC TCACACCCCA GTTTATCAAA
                                Rap1                                     Adr1
-560                                                                        -481
CATGTTCCCC AAATGGCCCA AAACTGACAG TTTAAACGCT GTCTTGGAAC CTAATATGAC AAAAGCGTGA TCTCATCCAA
GTACAAGGGG TTTACCGGGT TTTGACTGTC AAATTTGCGA CAGAACCTTG GATTATACTG TTTTCGCACT AGAGTAGGTT
-480                                                                        -401
GATGAACTAA GTTTGGTTCG TTGAAATGCT AACGGCCAGT TGGTCAAAAA GAAACTTCCA AAAGTCGGCA TACCGTTTGT
CTACTTGATT CAAACCAAGC AACTTTACGA TTGCCGGTCA ACCAGTTTTT CTTTGAAGGT TTTCAGCCGT ATGGCAAACA
                                                                 HSF
-400                                                                        -321
CTTGTTTGGT ATTGATTGAC GAATGCTCAA AAATAATCTC ATTAATGCTT AGCGCAGTCT CTCTATCGCT TCTGAACCCC
GAACAAACCA TAACTAACTG CTTACGAGTT TTTATTAGAG TAATTACGAA TCGCGTCAGA GAGATAGCGA AGACTTGGGG
-320                                                                        -241
GGTGCACCTG TGCCGAAACG CAAATGGGGA AACACCCGCT TTTTGGATGA TTATGCATTG TCTCCACATT GTATGCTTCC
CCACGTGGAC ACGGCTTTGC GTTTACCCCT TTGTGGGCGA AAAACCTACT AATACGTAAC AGAGGTGTAA CATACGAAGG
                                            Hap2/3/4/5       Mat1MC                Gcr1
-240                                                                        -161
AAGATTCTGG TGGGAATACT GCTGATAGCC TAACGTTCAT GATCAAAATT TAACTGTTCT AACCCCTACT TGACAGCAAT
TTCTAAGACC ACCCTTATGA CGACTATCGG ATTGCAAGTA CTAGTTTTAA ATTGACAAGA TTGGGGATGA ACTGTCGTTA
 Gcr1                                          QA-1F
-160                                                                         -81
ATATAAACAG AAGGAAGCTG CCCTGTCTTA AACCTTTTTT TTTATCATCA TTATTAGCTT ACTTTCATAA TTGCGACTGG
TATATTTGTC TTCCTTCGAC GGGACAGAAT TTGGAAAAAA AAATAGTAGT AATAATCGAA TGAAAGTATT AACGCTGACC
-80                                                                           -1
TTCCAATTGA CAAGCTTTTG ATTTTAACGA CTTTTAACGA CAACTTGAGA AGATCAAAAA ACAACTAATT ATTCGAAACG
AAGGTTAACT GTTCGAAAAC TAAAATTGCT GAAAATTGCT GTTGAACTCT TCTAGTTTTT TGTTGATTAA TAAGCTTTGC
```

B

```
                                                                                          relative
                                                                                          promoter
                                                                                          activity
-----------------------------------CCAAGATAGGCTATTTTTGTCGCA-TAAATTTTT--G-TCTCGGA      AOX2 (UAS region)
-------ATCACCTTCCAAGAT---------------------AGGCTATTTTTGTCGCA-TAAATTTTT--G-TCTCGGA      AOX2
-257   -250      -240      -230      -220      -210      -200      -190      -177
CCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCTAACGTTCATGATCAAAATTTAACTGTTCTAACC      WT          100 ± 6%
            Gcr1                                                  QA-1F
CCACATTGTATG.....AAGATTCTGGTGGGAATACTGCTGATAGCCTAACGTTCATGATCAAAATTTAACTGTTCTAACC      dGcr1       86 ± 9%
CCAC.....................ACTGCTGATAGCCTAACGTTCATGATCAAAATTTAACTGTTCTAACC      d6          44 ± 10%
CCAC.....................ACTGCTGATAGCC..ACGTTCATGATCAAAATTTAACTGTTCTAACC      d6*         55 ± 12%
CCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCC..ACGTTCATGATCAAAATTTAACTGTTCTAACC      d737-38     6.0 ± 0.3%
CCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCATACGTTCATGATCAAAATTTAACTGTTCTAACC      737AT       5.6 ± 0.5%
CCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCGCACGTTCATGATCAAAATTTAACTGTTCTAACC      737GC       29 ± 1%
CCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGC......TTCATGATCAAAATTTAACTGTTCTAACC      d736-41     26 ± 12%
CCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCTAACGTTCATG...............TTCTAACC      dQA-1F      102 ± 21%
CCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCTAACGTT..............AACTGTTCTAACC      db          48 ± 7%
CCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGC......TTCATG...............TTCTAACC      dQA-1Fzus   14 ± 3%
CCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGC.............................TCTAACC      d7          36 ± 6%
```

C

```
                                                                                          relative
                                                                                          promoter
                                                                                          activity
-784-780      -770      -760      -750      -740      -730      -720      -705
ACCCACTTTGCCATCGAAAAACCAGCCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTCCAATTCCTTCTATTAGGCT     WT          100 ± 6%
ACCCAC.....................TTAGGCT     d1          133 ± 23%
ACCCAC.....................AGTTATTGGGCTTGATTGGAGCTCGCTCATTCCAATTCCTTCTATTAGGCT     d1-1        111 ± 24%
ACCCACTTTGCCATCGAAAAACCAGCCC.............TCGCTCATTCCAATTCCTTCTATTAGGCT     d1-2        112 ± 22%
ACCCACTTTGCCATCGAAAAACCAGCCCAGTTATTGGGCTTGATTGGAGC.....................TTAGGCT     d1-3        92 ± 2%
ACCCAC.....................AGCTCGCTCATTCCAATTCCTTCTATTAGGCT     d1-SacI     88 ± 8%
ACCCACTTTGCCATCGAAAAACCAGCCCAGTT.....GCTTGATTGGAGCTCGCTCATTCCAATTCCTTCTATTAGGCT     dHap2345_1  106 ± 14%
ACCCACTTTGCCATCGAAAAACCAGCCCAGTTATTGGGCTTG.....AGCTCGCTCATTCCAATTCCTTCTATTAGGCT     dHap2345_2  75 ± 13%
          Hap2/3/4/5  Hap2/3/4/5     abaA
```

Figure 1. (**A**) Sequence of the *P. pastoris AOX1* promoter sequence (www.invitrogen.com). Putative fungal TFBSs as predicted by the MatInspector software are underlined. The frequently used BglII restriction site and the SacI restriction site are written italic. Putative regulatory sequences as predicted by Ohi *et al.* (9) based on P_{AOX2} sequence analysis are double underlined. Putative regulatory sequences as predicted by Kumagai *et al.* (28) based on *P. pastoris ZZA1* sequence analysis are written bold. Putative regulatory sequences as predicted by Gödecke *et al.* (23) based on *H. polymorpha MOX* sequence analysis are marked in grey. The TATA box is marked in dark grey. (**B**) P_{AOX1} deletions and their effects on *AOX1*-driven expression using methanol as sole carbon source during the induction phase following growth on glucose and glucose depletion. The *AOX2* UAS as proposed by Ohi *et al.* (9) is also aligned. Identical nucleotides between the *AOX2* and the *AOX1* sequence are written bold. Deletions performed are shown as dots instead of the nucleotides. Putative TFBSs predicted using MatInspector are underlined. The TTCCAA sequence proposed by Kumagai *et al.* (28) is italicized. Mutations introduced into the sequence are written bold and underlined. Relative promoter activity is given as mean % of the wild-type ± SD of three independent measurements. (**C**) P_{AOX1} deletions in the d1 region and their effects on *AOX1*-driven expression using methanol as sole carbon source during the induction phase following growth on glucose and glucose depletion. Deletions performed are shown as dots instead of the nucleotides. The first putative Hap2/3/4/5p-binding site and the putative abaA-binding site are underlined. The second putative Hap2/3/4/5p-binding site is written bold. Relative promoter activity is given as mean percentage of the wild-type ± SD of three independent measurements.

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Internal deletions within the *AOX1* promoter

In *P. pastoris*, *AOX1* expression is tightly repressed by glucose and glycerol, while methanol is absolutely necessary for high-level induction of this gene (39). In the absence of any carbon source, the promoter is derepressed, showing a small increase in transcriptional activity which is still much lower than the activity found in the presence of methanol. To characterize the *cis*-acting elements in P_{AOX1} that are necessary for this behaviour, a promoter library, fused to cycle-3-GFP as a reporter, was constructed. The library, consisting of *AOX1* promoter variants with internal deletions (as indicated in Supplementary Figure 2 and Supplementary Table 3), was used to inoculate 96-well deep-well plates (29). The library was cultivated in dextrose or glycerol containing media, i.e. BMD (1%) and BMG (0.5%), respectively. A 60 h growth phase was followed by a 72 h methanol induction phase as described in Materials and methods section. In neither the presence of glucose nor glycerol was any significant GFP expression detected for any strain with *AOX1* promoter deletion variants (as exemplified for some variants shown in Figure 2). Once the glucose or glycerol was used up, the wild-type *AOX1* promoter became slightly derepressed, which resulted in GFP expression and an increase in fluorescence. In some variants, a significant upregulation of GFP expression compared to the wild-type was detected during this time of derepression. After 60 h of growth and glucose depletion, the strongest derepression variants (d6 and d6*) showed more than a 4-fold higher GFP fluorescence intensity than the clone with P_{AOX1} wild-type driven expression. A similar but even more enhanced effect was observed with multicopy transformants of these constructs, where high expression of GFP could be easily monitored (data not shown). This increase was observed only under derepressing conditions and not in the presence of glucose or glycerol. The same mutations (d6 and d6*) resulted in a decreased fluorescence intensity (compared to the wild-type) after 70 h of methanol induction in deep-well plates.

After methanol addition, most promoter variants were induced to some degree, causing an increase in GFP fluorescence. A range of promoter activity from ~6 to 166% compared to the wild-type activity was observed with this library consisting of 46 promoter variants (including wild-type) having either one deletion or insertion or a combination of 2–3 such mutations (Figure 3 and Supplementary Table 3).

Most remarkably, a deletion of nucleotides −209 and −210 (TA), referred to as d737-38 in Figure 1B, within the P_{AOX1} sequence resulted in a drop of GFP fluorescence to ~6% of wild-type. The same drop in activity occurred when the TA base pair was interchanged to AT (737AT). This decrease was even more dramatic than the change caused by the deletion of the putative TATA box (−159 to −154) within the d8 region (18% residual activity). When the TA base pair was changed to GC (i.e. 737GC), the fluorescence intensity was five times higher, resulting in ~30% of the wild-type promoter driven expression activity. In addition, a deletion of nucleotides −184 to −218, called d7, which spans this region, resulted in a promoter with 36% of the wild-type activity. As a fourth observation, when the d737-38 deletion was introduced in addition to the d6 mutation, thereby generating mutant d6*, no change in activity under derepressing conditions and a comparatively less dramatic change in activity under induction conditions occurred (55% activity) (Figure 1B).

Within these two regions, sequences similar to the putative binding sites for *S. cerevisiae* TF Gcr1p and the *Neurospora crassa* QA-1F were found. Gcr1p is a major player in the regulation and coordination of glycolytic genes in baker's yeast and absolutely necessary for their high-level expression (40,41). In addition, the sequence to which this factor binds contains the central motif CTTCC, proposed by Kumagai *et al.* (28) to be involved in the

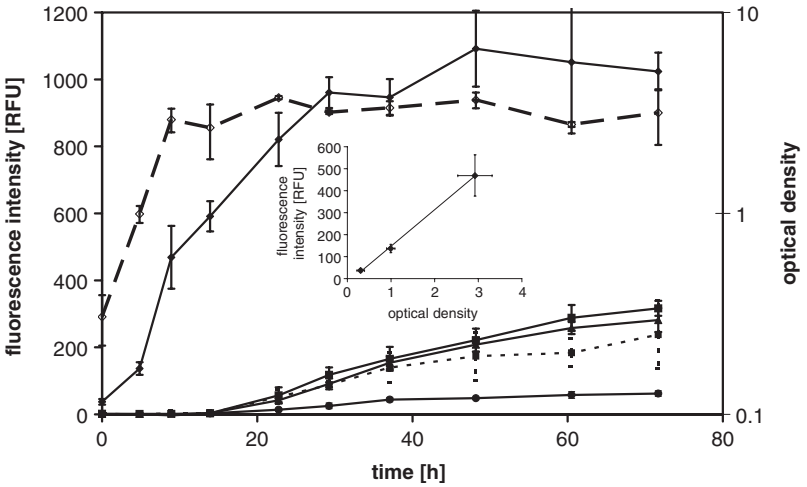

Figure 2. Time course of GFP expression driven by WT, d6 and d6* *AOX1* promoter variants and the *GAP* promoter in glucose containing media and during the glucose depletion phase in deep-well plates. The insert in the figure shows the correlation of the P_{GAP} driven GFP expression and biomass as represented by the optical density of the culture. Closed rhomboids, P_{GAP}; closed circles, P_{AOX1}; closed triangles, P_{AOX1}d6*; closed squares, P_{AOX1}d6; closed squares and dashed line, P_{AOX1}2xMat; open rhomboids, optical density of P_{GAP} strain. Fluorescence data and optical density represents mean ± SD of single measurements of four independent cultivations. Optical density curve of the P_{GAP} clone is shown as a representative.

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

regulation of the *AOX1* promoter. However, deletion of this core CTTCC motif (dGcr1) resulted in only minor decrease in promoter activity upon methanol induction, while deletion of a larger region (d6) resulted in a 55% decrease in activity of the *AOX1* promoter. However, deletion of the CTTCC motif resulted in an increase in promoter activity comparable to the d6 deletion under derepression conditions. In addition, the QA-1F protein is the activator protein of the glucose-repressed quinic acid cluster in *N. crassa* (42). Deletion of the core sequence of this putative regulatory region (dQA-1F) had no effect, while a deletion positioned 4 nt to the 5′ of this sequence (db) reduced fluorescence intensity ~50%. Other deletions 5′ to the QA-1F binding site (dQA-1Fzus, d7) resulted in further decreases in promoter activity as shown in Figure 1B. This region, in which no putative TFBS were predicted by the MatInspector analysis showed a high similarity to the UAS region proposed by Ohi *et al.* (9).

Taken together, the above results demonstrate that we closely located a major *cis*-acting element within the region between −253 and −184 of the *AOX1* promoter sequence positioned around −210 and also minor *cis*-acting elements surrounding the major −210 region. Remarkably, the expression effects of the individual deletions are largely dependent on their extension and exact location (Figure 1B).

Furthermore, we identified several other regions where either short deletions of the core region of putative TFBSs or larger deletions spanning this region resulted in a strong decrease in promoter activity upon methanol addition. The responsible short putative binding sites and the corresponding larger deletion (d)-regions include Rap1p like sequence (34% residual activity) with the corresponding d2 region (33%), the Adr1p like sequence (30%) and d3 (70%), Hsf like sequences (60%) and d4 (46%) and Hap2/3/4/5p (67%), Mat1-Mc like sequences (42%) and d5 (40%). While the first three combinations are directly overlapping, i.e. the core sequence deletion of the putative TFBS is integrated into the larger deletion, this is not the case for the combination d5 and Mat1-Mc, which are only partly overlapping (Supplementary Figure 2). Taken together, we identified five regions based on the prediction of putative TFBSs where a single deletion of such a region of 2–59 bp resulted in a significant reduction in promoter activity, in most cases >50%.

While most variants of the *AOX1* promoter resulted in decreased expression, a few mutants actually enhanced its already strong expression in methanol containing medium. Notably, deletions in the region called d1, where three putative TFBSs were found, caused higher expression (Figure 1C). Two of these sequence motifs are overlapping and show sequence similarity to the *S. cerevisiae* Hap2/3/4/5p complex binding site (22). The second one of these binding sites is overlapping with a putative *Aspergillus nidulans abaA* binding site (43). The core consensus sequence of the Hap2/3/4/5p complex, CCAAT, is also present twice within this *abaA* motif. Deletion of the whole d1 region resulted in an increase of promoter activity of about 30%. To further delineate this *cis*-acting region, we narrowed the d1 deletion by dividing it into three parts, called d1-1 to d1-3 (Figure 1C). A positive effect of the deletions d1-1 and d1-2 of ~110% could be observed, while deletion of region d1-3 resulted in a reduction of promoter activity of ~10%. Deletion of the core sequence of the first putative Hap2/3/4/5p-binding site also resulted in a small increase, while deletion of the second resulted in a decrease in activity of ~75%. The effect of the d1 deletion seems not to be the sum of the effects of deletions of subfragments of d1. In case of the d1-3 deletion a new putative Hap2/3/4/5 site was generated due to the deletion (including the abaA site), which might explain that we did not see the expected increase of the promoter activity in this case.

Addition of putative *cis*-acting elements

After we made deletions of putative TFBSs, we next examined if we can utilize the results of our deletion analysis to engineer new promoter variants. Therefore, we duplicated putative positive *cis*-acting elements at their natural location. We chose three such elements: nucleotides 201–214 (−203 to −190), Mat1-Mc (−271 to −254) and Rap1p (−615 to −601). While duplicating the putative Rap1p-binding site resulted in no significant change in promoter activity under the conditions tested, doubling the putative Mat1-Mc-binding site resulted in a ~3-fold improvement of promoter activity under derepressing conditions (data not shown). Duplicating the 201–214 regions resulted in an increase in promoter activity of 139% of the wild-type in presence of methanol (Supplementary Table 3). This rise in transcriptional activity, together with the increase produced by the deletion of region d1, is the best promoter activity observed by single modifications within the *AOX1* promoter sequence. Both mutations, deleting d1 and doubling 201–214, resulted also in an increase in activity under derepression conditions, but less significant compared to the d6 and d6* modifications.

Next, we asked if we could further improve the *AOX1* promoter by combining both mutations, deletion of d1 and doubling 201–214. In doing so, we achieved an increase in promoter activity of >60% under methanol induction conditions (Figure 3 and Supplementary Table 3), which is remarkable for two reasons: (i) we started with an already very strong promoter and (ii) we did not find additivity in other cases as e.g. in the case of d6 and d6* variants or combinations of d1, d2, d6, d6*, dHap2/3/4/5p and the second dHsf deletion (Supplementary Table 3) by simple combination of the mutations.

Alternative reporters

To test the transferability of the results obtained with the laboratory reporter GFP to a real case scenario, we selected five promoter variants (wt, d1, d2, d6* and d2d6) with different activities to express three additional reporters: the first is the hydroxynitrile lyase isoenzyme 5 of *Prunus amygdalus* (almond) (*Pa*HNL5α) (34). This FAD-containing enzyme can be produced and secreted by *P. pastoris* in the range of ~1 g/l in bioreactor cultures. The second enzyme, HRP (33) has so far been produced by *P. pastoris* only in the lower microgram per litre range. For both of these secreted enzymes, the native signal sequence was replaced by the *S. cerevisiae* mating factor α signal sequence. The third enzyme, the hydroxynitrile

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

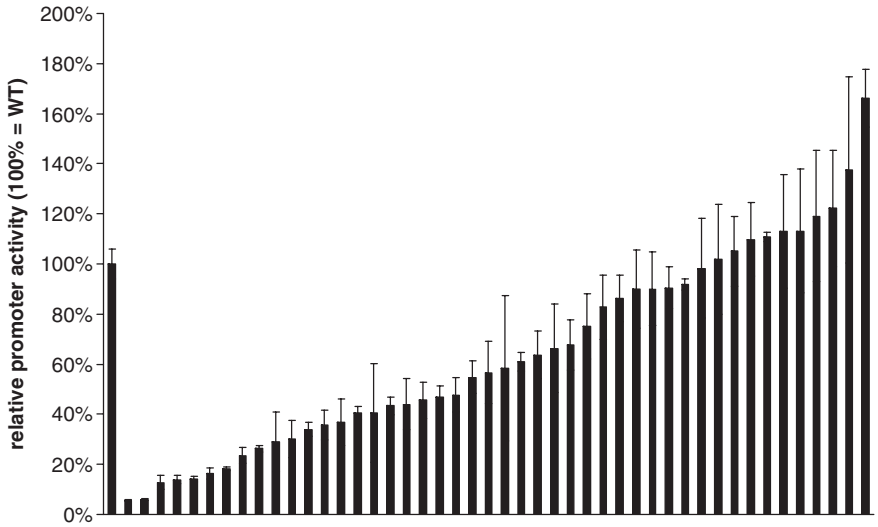

Figure 3. Promoter activity as represented by the fluorescence intensity of reporter clones after cultivation in 96-well deep-well plates and methanol induction for 72 h. The data represent mean ± SD of single measurements of three independent cultivations in separate deep-well plates.

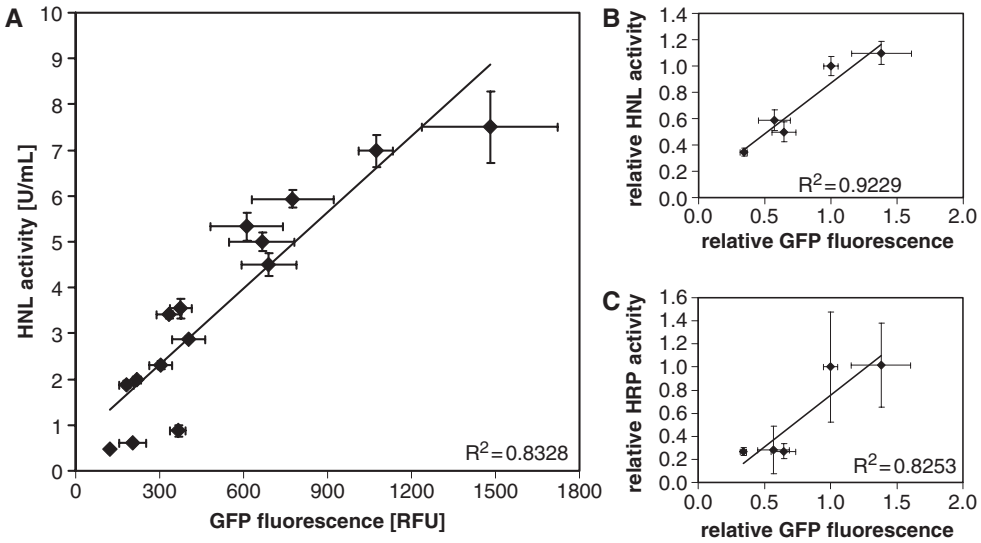

Figure 4. Correlation of P_{AOX1}-driven reporter enzyme expression and GFP expression driven by the same promoter variants. (**A**) *Pa*HNL5α activity of five different promoter variants at three different time points during methanol induction (after 24, 28 and 72 h of methanol induction) compared as a function of the GFP expression using the same promoter variants at the same time points. HNL activity is given in U/ml in the culture supernatants. Data represent mean ± SD of four independent cultivations. (**B**) Relative *Hb*HNL activity of cell lysates compared to the wild-type promoter transformant after 72 h of methanol induction. The relative *Hb*HNL activities were normalized to the activity obtained with the wild-type P_{AOX1} and are represented by the mean ± SD of four independent cultivations. (**C**) Relative HRP activity of culture supernatants compared to the wild-type promoter transformant after 72 h of methanol induction. The relative HRP activities were normalized to the activity obtained with the wild-type P_{AOX1} and are represented by the mean ± SD of four independent cultivations. The GFP fluorescence intensities or relative GFP fluorescence are represented by the mean ± SD of three independent cultivations.

lyase from *Hevea brasiliensis* (tropical rubber tree) (*Hb*HNL) had been produced by *P. pastoris* to a level of more than 20 g/l previously (7). All three are industrially important enzymes. For these three enzymes and the five promoter variants the same trend was observed as for the GFP expression strains (Figure 4).

For HCD cultivations, Mut^S strains, where the *AOX1* gene is disrupted, are often preferred since they are easier to handle in terms of oxygen consumption, heat production and control of methanol concentration in the bioreactor. Thus, we tested the three promoter variants (wt, d1 and d6*), including the derepressed d6* variant, for the GFP and the *Pa*HNL5α expression in a Mut^S phenotype background (KM71H). The promoter activities found in the mutant library were fairly independent of the reporter used and the methanol utilization phenotype of the strain that harboured them (Supplementary Figure 3).

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Basal promoter fragments

As a first step towards the creation of useful fully artificial *P. pastoris* promoters by using modular building blocks, several *cis*-acting elements described above were added to different basal promoters. Since the core promoter containing the transcription initiation site is also a key *cis*-acting element crucial for efficient transcription (44), we first analysed three different core promoter fragments. The first one was the basal *S. cerevisiae LEU2* promoter fragment (20), and the other two were *AOX1* promoter fragments which where truncated at positions −10 from the 5′ end of the TATA box, called AOX176, and at −25, called AOX194. First we tested these three core promoters by fusing them to full-length *AOX1* promoter fragments (−170 to −940) and truncated *AOX1* promoter variants with a 5′ end at position −731 (truncSacI), −411(trunc411) and −203 (trunc203). Therefore, P_{AOX1} was cut at position −10 from the TATA box for the *LEU2* and the AOX176 fragment, and cut at −25 from the TATA box for the AOX194 variant. In doing so, we introduced three additional nucleotides at −10 of the TATA position to generate a BspTI restriction site (for AOX176) and we changed two outer nucleotides in the TTTAAC hexanucleotide to CTTAAG, again to generate the BspTI site for AOX194 (Supplementary Figure 1B). In addition, we constructed reporter genes with the *LEU2*-variants and the *E. coli lacZ* as a reporter gene. Table 1 clearly demonstrates the superior performance of the basal AOX fragments in comparison to the *S. cerevisiae LEU2* fragment. AOX176 and AOX194, when fused to the full *AOX1* promoter regions, showed a greater induction than the *LEU2* in methanol medium. While no promoter activity of any mutated P_{AOX1} variant could be found as long as glucose (or glycerol) was present in the medium, all three basal promoter variants showed some basal activity in presence of glucose (Figure 5). Since there was no difference between the AOX194 and the AOX176 fragment in terms of methanol-sensitive induction, we chose the AOX176 as a core fragment for further studies. In addition, the AOX176 core promoter showed the lowest basal GFP expression on methanol of the three variants tested (data not shown), making it the best suited to study induction effects by glucose depletion or addition of methanol.

Synthetic promoters using small *cis*-acting elements

We were interested in utilizing short synthetic promoter variants with several *cis*-acting elements to design useful tailor-made promoter variants, which would optimize

Table 1. Fluorescence intensities of basal promoter fragments with added *AOX1* promoter parts

Construct	5′ promoter region added	Fluorescence intensities [RFU]			lacZ activity [10^3 U μg^{-1}]
		LEU2	AOX194	AOX176	LEU2
Full-length	−940 to −170[a]	310 ± 30	849 ± 120	845 ± 81	ND
truncSacI	−731 to −170[a]	345 ± 31	1095 ± 123	593 ± 158	5.82
trunc411	−411 to −170[a]	115 ± 18	175 ± 8	121 ± 13	5.11
trunc203	−203 to −170[a]	47 ± 3	ND	48 ± 4	0.54
Basal	No addition	17 ± 7	11 ± 2	7 ± 4	0.27

The fluorescence intensity of the wild-type AOX1 promoter is 889 ± 70.
[a]Constructs used for fusion with AOX194 were truncated at the 5′ end at –193.

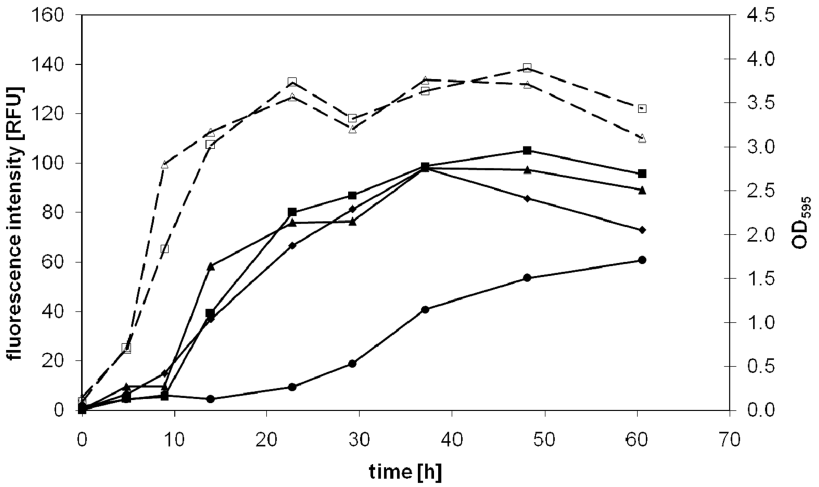

Figure 5. Time course of GFP expression driven by basal *S. cerevisiae LEU2*, two *P. pastoris AOX1* basal promoter variants (AOX176 and AOX194) and the wild-type *AOX1* promoter in glucose-containing media and during the glucose depletion phase in deep-well plates. Closed rhomboids, $P_{LEU2basal}$; closed circles, P_{AOX1}; closed triangles, P_{AOX194}; closed squares, P_{AOX176}; open squares, optical density of P_{AOX176}; open triangles, optical density of P_{AOX194}. Fluorescence data and optical density represents mean of single measurements of four independent cultivations.

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

gene expression for applications such as protein manufacturing (45,46), metabolic engineering (47–50) and synthetic biology (51–53). In addition, for some specific applications such as directed evolution of enzymes using *P. pastoris* (54), short promoters are preferable since they usually enable a more efficient PCR-based generation of expression cassettes or mutant libraries. Therefore, our next goal towards the design of fully synthetic promoter variants was to determine the effect of selected putative *cis*-acting sequences on the expression of the basal AOX176 region. Based on the results of the deletion libraries as well as the similarity of predicted sites to the matrix of the TFBS, we chose: (i) the 201–214 bp fragment, (ii) the 737 region, (iii) the STRE element, (iv) the putative binding site for Adr1p, (v) the putative binding site for Gcr1p and (vi) the putative binding site for Mat1-Mc to be inserted upstream of the basal promoters.

While the fragments 201–214, STRE and Mat1-Mc enhanced expression constitutively under both the conditions [the absence of glucose (derepression) and the presence of methanol (induction)], the other fragments displayed only an effect when methanol was used for induction (Figure 6). Adding the 737 element to a basal promoter resulted in a decrease of GFP expression when glucose is present; therefore, the induction ratio was further increased with this variant compared to the sole basal promoter element. The same effect was found when the 737 element was added to the basal *LEU2* promoter (data not shown). Therefore, the verified five *cis*-acting elements within the *AOX1* promoter cause different effects. While the elements 201–214, STRE and Mat1-Mc seemed to enhance transcription from the basal promoter equally in absence of glucose and presence of methanol, the 737 element had a repression effect in presence of glucose and an induction effect when methanol was present in the medium. Adr1p and Gcr1p, exerted effects only when methanol was present, but activity under derepressing conditions remained unchanged.

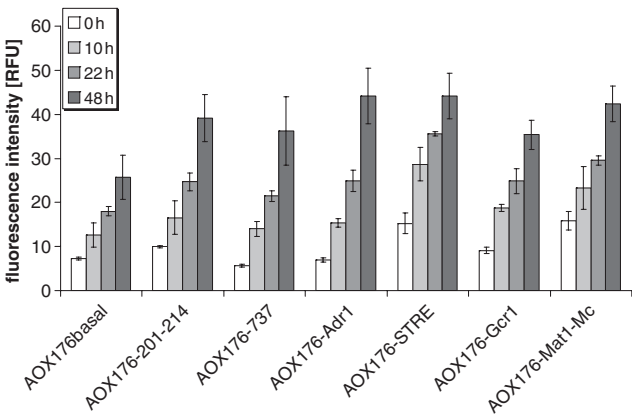

Figure 6. Promoter activity as represented by the fluorescence intensity of reporter clones after cultivation in 96-well deep-well plates and methanol induction for 72 h. The data represent mean ± SD of single measurements of three independent cultivations in separate deep-well plates. Samples were taken at 0, 10, 24 and 48 h after methanol addition. Values at 0 h represent promoter activities under derepression conditions. Fluorescence intensities of the wild-type *AOX1* promoter are: 0 h, 7 ± 2 RFU; 10 h, 110 ± 3 RFU; 24 h, 262 ± 32 RFU; 48 h, 457 ± 55 RFU.

Product formation in HCD fedbatch cultures

To evaluate the scalability and the potential for industrial application of our findings with new promoter variants showing interesting technological features, we analysed the d6* promoter in a laboratory-scale bioreactor experiment. In 96-well format, this promoter variant caused enhanced protein expression capability without the need of methanol for induction. In addition, we faced the challenge of keeping the derepression effect observed upon glucose depletion over a long lasting production phase in the bioreactor, where some cell growth and energy metabolism has to be maintained by feeding limiting amounts of glucose. In this part of our study, two multicopy strains in which HRP was expressed either under the control of the wild-type (wt) or the d6* *AOX1* promoter variant were cultured in a HCD fedbatch, and compared with respect to both biomass formation and extra-cellular secretion of the recombinant product. The following two fedbatch procedures were carried out on each strain: (i) glucose as the sole carbon and energy source was added during the entire process duration to determine product levels under derepressing conditions, where the carbon source was limiting, and (ii) after an initial biomass growth with glucose, the composition of the original feed solution was changed to pure methanol, which caused further induction of the heterologous protein expression under the control of the wild-type P_{AOX1} and most of its variants.

After 32.5 h, the fedbatch cultures grown with glucose addition during the entire process were stopped as the working volume limit was reached. As both processes (i.e. either with glucose or methanol) were designed in such a way that the same total amount of carbon was added to the bioreactor, only data obtained at this particular time (32.5 h) were appropriate for direct comparisons of the process performance. During both, the glucose and the methanol processes, product titres approx. twice as high were reached in cultures with the d6* *AOX1*-promoter variant (i.e. given as volumetric activity in U/l) than in those where wt was used (Figure 7). Furthermore, compared to glucose in methanol processes a titre enhancement was observed for each of the particular strains/promoters, i.e. at a factor of 1.7 and 2.1 for the d6* and wt strains respectively. Thus, the highest amount of product produced per biomass [i.e. ~107 U (g CDW)$^{-1}$ at the end of 51.5 h production period and 63.0 U (g CDW)$^{-1}$ during the first 32.5 h], which corresponds to an average specific product formation rate of 2.08 U (g/CDW)$^{-1}$ h^{-1} (51.5 h), 1.94 U (g CDW)$^{-1}$ h^{-1} (32.5 h) was demonstrated in the experiment with the d6* promoter operating under full induction by methanol at pH 5 and 30°C. However, even without induction by methanol 4133 U/l were obtained with the d6* promoter-based expression construct corresponding to 129% of the maximal titre of the wild-type P_{AOX1}-based strain in the presence of methanol and 2.5 times more than in case of derepression of the wild-type P_{AOX1} system. Due to

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

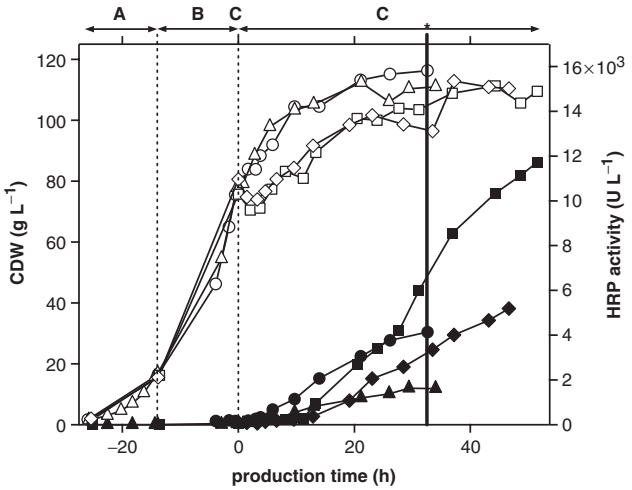

Figure 7. Time courses of biomass concentration and volumetric product (HRP) activity as determined in HCD-fedbatch cultures. Marked with an asterisk: vertical line at 32.5 h, the time when the final titres were compared. Open symbols, biomass concentrations determined as cell dry weight; closed symbols, product measured as volumetric activity in a photometric assay; squares and circles, for methanol and glucose processes with d6* AOX1-promoter variant, respectively; rhomboids and triangles, for methanol and glucose processes with wild-type promoter, respectively. The capital letters show the following process phases: A—batch growth with glucose; B—fedbatch growth over 13.5 h with exponential addition of glucose at a feed rate $F(t) = 4.655 \cdot e^{0.2 \cdot t}$ g C h^{-1}; C—production phase of about 44 h with a constant addition of substrate, i.e. at $F(t) = 29.54$ g C h^{-1}.

specific differences in the mode of operation between microscale and bioreactor experiments, batch-wise induction in microscale and continuous induction with limiting amounts in bioreactors, together with the fact that the d6* promoter variant offers faster induction kinetics at the beginning of the induction phase resulted in a higher overall productivity of the d6* promoter, even under methanol induction conditions (Figure 7).

DISCUSSION

A promoter library consisting of several *AOX1* promoter variants fused to GFP displayed a broad range of activities (from ~6% to >160%) with small increments between the individual promoter variants. Although we started with one of the most powerful promoters described to date, the P_{AOX1} (55), we obtained some variants which were even stronger. Several important *cis*-acting elements responsible for high P_{AOX1} activity were identified and characterized. Using our approach starting with computational TFBS prediction, we: (i) achieved a high success rate in changing the promoter strength by introducing mutations within the TFBSs and (ii) localized the *cis*-acting elements and their individual effects precisely by the first deletion series.

We identified seven regions within the *AOX1* promoter where a deletion of either small core sequences of the putative TFBS or the deletion of a larger stretch covering the whole putative TFBS resulted in a decrease in promoter activity of >50%. Additionally, the prediction of putative TFBSs suggests possible target TFs, where a search for homologues in the *P. pastoris* genome might facilitate future analyses about P_{AOX1} regulation (and regulation of other genes of the methanol utilization pathway). While we expected to strongly downregulate the promoter activity upon deletion of the putative TATA box, we did not expect such a large number of regions where a mutation can decrease promoter activity by >50%, some of the effects being even more dramatic than in the case of the TATA box deletion. For instance, when we deleted (or mutated) nucleotides surrounding position −209 to −210, the promoter activity dropped to ~6–30% of the wild-type activity, depending on the mutation (Figure 1B). In addition, these results are consistent with previous studies where promoter activity dropped with stepwise truncations (20) and which effects observed in the case of larger deletions of the whole *AOX1* promoter sequence (56,57).

Our approach of predicting putative TFBSs followed by deletion or duplication of the respective region within the promoter sequence was highly successful in generating a library spanning a broad range of activities. This not only delivered mutants with strongly altered activity but also suggested which TFs might be involved in the regulation of the *AOX1* promoter. Our results also demonstrated the complex architecture of the *AOX1* promoter with several activator/repressor sites, which might explain its strong and tight regulation.

To identify TFs involved in methanol-sensitive gene regulation, past studies have focused on the behaviour of these types of promoters in heterologous hosts that are better characterized than most methylotrophic fungi. *Pichia pastoris AOX1* and *DAS1* and *H. polymorpha MOX* promoter regions were shown to promote expression of the reporter enzyme β-galactosidase (*lacZ* of *E. coli*) in *S. cerevisiae* (8,25). The regulation patterns of these genes in *S. cerevisiae* are similar, though not the same, as in their natural hosts: glucose represses gene expression, under carbon starvation conditions expression is slightly derepressed, and glycerol as carbon source induces expression [however, the latter is contradictory to what happens in *P. pastoris* (8)]. *MOX* promoter driven expression was induced by ethanol, methanol and oleic acid in *S. cerevisiae* (21). Furthermore, Adr1p, a positive effector of *ADH2* (alcohol dehydrogenase 2) and some peroxisomal proteins in *S. cerevisiae* (21), was shown to be an activator of the *MOX* promoter in *S. cerevisiae* when glucose is lacking in the medium (25). Interestingly, the only identified TF involved in *AOX1* regulation, Mxr1p, shows similarity at the structural level to the *S. cerevisiae* Adr1p (20). However, Mxr1p was identified to bind to a region between −170 and −411, but the putative Adr1p-binding site we found with MatInspector is located at −576 (5′ end of the binding site). Thus, the exact localization of the Mxr1p-binding site in the *AOX1* promoter remains to be elucidated.

We introduced deletions into the promoter sequence, which possibly modified positioning of *trans*-acting factors that work in a concerted fashion to achieve full induction (e.g. facilitating the binding of partner factors or bending of DNA). Thus, the observed effects are not necessarily

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

caused by the deletion of a *cis*-acting element, but may result simply from changes in spacing between authentic elements. Whether this characteristic P_{AOX1} regulation pattern (glucose repression, derepression in absence of carbon source and methanol induction) is caused by the presence of one or several repressor proteins, by inaccessibility of the promoter packed in nucleosomes or by the absence of activator proteins, or a combination of those possibilities remains to be elucidated in further studies.

However, a rough model can be presented based on our preliminary results in addition to other previous studies. Cells lacking *MXR1* function cannot induce the *AOX1* promoter at all. Mxr1p has many similarities to *S. cerevisiae* Adr1p and can be considered a homologue. Having several domains, Adr1p not only binds to and activates target promoters but also modifies the activity or expression level of other TFs which modulate the same target promoters. Adr1p can be considered both a global and local regulator of transcription. In parallel fashion, we hypothesize that Mxr1p not only binds the *AOX1* promoter but also can regulate at least one additional *P. pastoris* TF protein additionally involved in regulation of the *AOX1* promoter. Such a model helps to explain the presence of so many different binding sites in the same promoter.

Furthermore, we demonstrated the transferability of promoter strength of the new promoter variants to the expression of industrial enzymes that are produced by *P. pastoris* in a wide variety of concentrations either intracellularly (*Hb*HNL) or extracellularly (*Pa*HNL5α, HRP).

Since *P. pastoris* is one of the major protein production host, a detailed understanding of the regulation of its most commonly used promoter at a molecular level will have a dramatic impact on the optimization of protein production processes in the future. This should be supported further by the availability of the *P. pastoris* genome sequence, which could enable a straight-forward isolation and study of *trans*-acting factors involved in regulation of P_{AOX1}-driven gene expression. There is no general optimal promoter: the regulatory properties of the promoter variant, the target protein and the production process have to be seen as a whole. The new promoter variants are especially useful for tailoring a production strategy that is specific to the needs of the individual protein. For example, stronger variants would be useful for the expression of non-toxic proteins, as in the case of reporters where fewer copies of the expression cassette would be sufficient to obtain a maximum of expression. Weaker variants could be favourable for the expression of proteins, which cause deleterious effects to the host cell. Within this group, we include proteins which are toxic to the cell per se due to their biological activity (e.g. proteases, toxins) or which induce an unfolded protein response (UPR) or ER-associated degradation (ERAD) due to incomplete folding in the ER during secretion (58–61) or which face other bottlenecks where an excess of transcript would ultimately end up in less active protein. Another group of 'difficult' proteins that needs optimization of expression are integral membrane proteins, which are intrinsically hard to produce (62,63).

For the first time, we succeeded in the generation of promoter mutants which displayed increased promoter activity without the need for any methanol for strong expression while, in contrast to constitutive promoters like P_{GAP}, still are strongly repressed in presence of glucose. Such an expression profile was also found for short synthetic promoter variants, which were constructed using P_{AOX1}-derived *cis*-acting elements. Such promoters can be used to express proteins in carbon starvation phases like the stationary phase in batch cultures, under carbon-limited conditions like fedbatch or by continuous cultivation below the maximum specific growth rate, as demonstrated by HRP production employing a glucose-limited fedbatch process.

Briefly, we demonstrated that specific mutations introduced at putative TFBSs can be employed to generate a promoter library with a broad range of activities and thereby, identified regulatory regions that provide a basis for designing short, synthetic promoter variants for metabolic engineering and synthetic biology. Further work is needed to understand the biological role of the different *cis*-acting elements so that a complete characterization of the molecular basis of P_{AOX1} regulation can be achieved. Nevertheless, considering the results obtained during this study and the advantages of *P. pastoris* in recombinant protein production processes, especially proteins needed in the pharmaceutical, chemical and biofuel industry, we are convinced that our library will be broadly applicable to optimize production yields.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are indebted to Mrs Verena Looser for carefully reviewing the article and giving many useful suggestions on the cultivation and physiology aspects and to Mr Clemens Mayer for excellent technical support. Our thanks are also expressed to SATW (Swiss Academy of Engineering Sciences, Zurich, Switzerland), whose financial support under the scope of the Industrial Biotechnology initiative helped enable the bioreactor cultivations to be performed. Funding to pay the Open Access publication charges for this article was provided by SATW (Swiss Academy of Engineering Sciences, Zurich, Switzerland) Seidengasse 16, 8001 Zürich. Work in JL-C and GL-C's laboratory was supported by NIH AREA Grant No. GM65882-02.

Conflict of interest statement. A patent application covering the work described in this article has been filed by the Graz University of Technology. Graz University of Technology, who was the employer of the authors and inventors Franz Hartner and Anton Glieder could potentially gain financially from the publication of the studies described in this work. The authors do not see any other conflict of interest.

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

REFERENCES

1. Lin-Cereghino,J. and Cregg,J.M. (2000) Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.*, **24**, 45–66.
2. Romanos,M.A., Scorer,C.A. and Clare,J.J. (1992) Foreign gene expression in yeast: a review. *Yeast*, **8**, 423–488.
3. Jahic,M., Rotticci-Mulder,J., Martinelle,M., Hult,K. and Enfors,S.-O. (2002) Modeling of growth and energy metabolism of *Pichia pastoris* producing a fusion protein. *Bioprocess. Biosyst. Eng.*, **24**, 385–393.
4. Jungo,C., Rerat,C., Marison,I.W. and von Stockar,U. (2006) Quantitative characterization of the regulation of the synthesis of alcohol oxidase and of the expression of recombinant avidin in a *Pichia pastoris* Mut^+ strain. *Enzyme Microb. Technol.*, **39**, 936–944.
5. Zhang,W., Bevins,M.A., Plantz,B.A., Smith,L.A. and Meagher,M.M. (2000) Modeling *Pichia pastoris* growth on methanol and optimizing the production of a recombinant protein, the heavy-chain fragment C of botulinum neurotoxin, serotype A. *Biotechnol. Bioeng.*, **70**, 1–8.
6. Xiong,A.S., Yao,Q.H., Peng,R.H., Zhang,Z., Xu,F., Liu,J.G., Han,P.L. and Chen,J.M. (2006) High level expression of a synthetic gene encoding *Peniophora lycii* phytase in methylotrophic yeast *Pichia pastoris*. *Appl. Microbiol. Biotechnol.*, **72**, 1039–1047.
7. Hasslacher,M., Schall,M., Hayn,M., Bona,R., Rumbold,K., Luckl,J., Griengl,H., Kohlwein,S.D. and Schwab,H. (1997) High-level intracellular expression of hydroxynitrile lyase from the tropical rubber tree *Hevea brasiliensis* in microbial hosts. *Protein Expr. Purif.*, **11**, 61–71.
8. Tschopp,J.F., Brust,P.F., Cregg,J.M., Stillman,C.A. and Gingeras,T.R. (1987) Expression of the *lacZ* gene from two methanol-regulated promoters in *Pichia pastoris*. *Nucleic Acids Res.*, **15**, 3859–3876.
9. Ohi,H., Miura,M., Hiramatsu,R. and Ohmura,T. (1994) The positive and negative *cis*-acting elements for methanol regulation in the *Pichia pastoris AOX2* gene. *Mol. Gen. Genet.*, **243**, 489–499.
10. Shen,S., Sulter,G., Jeffries,T.W. and Cregg,J.M. (1998) A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene*, **216**, 93–102.
11. Liu,H., Tan,X., Russell,K.A., Veenhuis,M. and Cregg,J.M. (1995) *PER3*, a gene required for peroxisome biogenesis in *Pichia pastoris*, encodes a peroxisomal membrane protein involved in protein import. *J. Biol. Chem.*, **270**, 10940–10951.
12. Sears,I.B., O'Connor,J., Rossanese,O.W. and Glick,B.S. (1998) A versatile set of vectors for constitutive and regulated gene expression in *Pichia pastoris*. *Yeast*, **14**, 783–790.
13. Waterham,H.R., Digan,M.E., Koutz,P.J., Lair,S.V. and Cregg,J.M. (1997) Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene*, **186**, 37–44.
14. Ahn,J., Hong,J., Lee,H., Park,M., Lee,E., Kim,C., Choi,E. and Jung,J. (2007) Translation elongation factor 1-α gene from *Pichia pastoris*: molecular cloning, sequence, and use of its promoter. *Appl. Microbiol. Biotechnol.*, **74**, 601–608.
15. de Almeida,J.R., de Moraes,L.M. and Torres,F.A. (2005) Molecular characterization of the 3-phosphoglycerate kinase gene (*PGK1*) from the methylotrophic yeast *Pichia pastoris*. *Yeast*, **22**, 725–737.
16. Menendez,J., Valdes,I. and Cabrera,N. (2003) The *ICL1* gene of *Pichia pastoris*, transcriptional regulation and use of its promoter. *Yeast*, **20**, 1097–1108.
17. Hartner,F.S. and Glieder,A. (2006) Regulation of methanol utilisation pathway genes in yeasts. *Microb. Cell. Fact.*, **5**, 39.
18. Cos,O., Ramon,R., Montesinos,J.L. and Valero,F. (2006) Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris*: a review. *Microb. Cell. Fact.*, **5**, 17.
19. Zhang,W., Hywood Potter,K.J., Plantz,B.A., Schlegel,V.L., Smith,L.A. and Meagher,M.M. (2003) *Pichia pastoris* fermentation with mixed-feeds of glycerol and methanol: growth kinetics and production improvement. *J. Ind. Microbiol. Biotechnol.*, **30**, 210–215.
20. Lin-Cereghino,G.P., Godfrey,L., de la Cruz,B.J., Johnson,S., Khuongsathiene,S., Tolstorukov,I., Yan,M., Lin-Cereghino,J., Veenhuis,M., Subramani,S. *et al.* (2006) Mxr1p, a key regulator of the methanol utilization pathway and peroxisomal genes in *Pichia pastoris*. *Mol. Cell. Biol.*, **26**, 883–897.
21. Simon,M., Binder,M., Adam,G., Hartig,A. and Ruis,H. (1992) Control of peroxisome proliferation in *Saccharomyces cerevisiae* by *ADR1*, *SNF1* (*CAT1*, *CCR1*) and *SNF4* (*CAT3*). *Yeast*, **8**, 303–309.
22. Gancedo,J.M. (1998) Yeast carbon catabolite repression. *Microbiol. Mol. Biol. Rev.*, **62**, 334–361.
23. Gödecke,S., Eckart,M., Janowicz,Z.A. and Hollenberg,C.P. (1994) Identification of sequences responsible for transcriptional regulation of the strongly expressed methanol oxidase-encoding gene in *Hansenula polymorpha*. *Gene*, **139**, 35–42.
24. Komeda,T., Yurimoto,H., Kato,N., Sakai,Y. and Kondo,K. (2003) *Cis*-acting elements sufficient for induction of *FDH1* expression by formate in the methylotrophic yeast *Candida boidinii*. *Mol. Genet. Genomics*, **270**, 273–280.
25. Pereira,G.G. and Hollenberg,C.P. (1996) Conserved regulation of the *Hansenula polymorpha MOX* promoter in *Saccharomyces cerevisiae* reveals insights in the transcriptional activation by Adr1p. *Eur. J. Biochem.*, **238**, 181–191.
26. Quandt,K., Frech,K., Karas,H., Wingender,E. and Werner,T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.*, **23**, 4878–4884.
27. Cartharius,K., Frech,K., Grote,K., Klocke,B., Haltmeier,M., Klingenhoff,A., Frisch,M., Bayerlein,M. and Werner,T. (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*, **21**, 2933–2942.
28. Kumagai,M.H., Sverlow,G.G., della-Cioppa,G. and Grill,L.K. (1993) Conversion of starch to ethanol in a recombinant *Saccharomyces cerevisiae* strain expressing rice α-amylase from a novel *Pichia pastoris* alcohol oxidase promoter. *Biotechnology*, **11**, 606–610.
29. Weis,R., Luiten,R., Skranc,W., Schwab,H., Wubbolts,M. and Glieder,A. (2004) Reliable high-throughput screening with *Pichia pastoris* by limiting yeast cell death phenomena. *FEMS Yeast Res.*, **5**, 179–189.
30. Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (2007) *Current Protocols in Molecular Biology*. John Wiley and Sons, New York.
31. Crameri,A., Whitehorn,E.A., Tate,E. and Stemmer,W.P. (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.*, **14**, 315–319.
32. Wang,W. and Malcolm,B.A. (1999) Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange site-directed mutagenesis. *Biotechniques*, **26**, 680–682.
33. Morawski,B., Lin,Z., Cirino,P., Joo,H., Bandara,G. and Arnold,F.H. (2000) Functional expression of horseradish peroxidase in *Saccharomyces cerevisiae* and *Pichia pastoris*. *Protein Eng.*, **13**, 377–384.
34. Glieder,A., Weis,R., Skranc,W., Poechlauer,P., Dreveny,I., Majer,S., Wubbolts,M., Schwab,H. and Gruber,K. (2003) Comprehensive step-by-step engineering of an (*R*)-hydroxynitrile lyase for large-scale asymmetric synthesis. *Angew. Chem. Int. Ed. Engl.*, **42**, 4815–4818.
35. Guarente,L. and Mason,T. (1983) Heme regulates transcription of the *CYC1* gene of *S. cerevisiae* via an upstream activation site. *Cell*, **32**, 1279–1286.
36. Lin-Cereghino,G.P., Lin-Cereghino,J., Sunga,A.J., Johnson,M.A., Lim,M., Gleeson,M.A. and Cregg,J.M. (2001) New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of *Pichia pastoris*. *Gene*, **263**, 159–169.
37. Martinez-Arias,A.E. and Casadaban,M.J. (1983) Fusion of the *Saccharomyces cerevisiae leu2* gene to an *Escherichia coli* β-galactosidase gene. *Mol. Cell. Biol.*, **3**, 580–586.
38. Lin-Cereghino,J., Wong,W.W., Xiong,S., Giang,W., Luong,L.T., Vu,J., Johnson,S.D. and Lin-Cereghino,G.P. (2005) Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast *Pichia pastoris*. *Biotechniques*, **38**, 44, 46, 48.
39. Cregg,J.M., Madden,K.R., Barringer,K.J., Thill,G.P. and Stillman,C.A. (1989) Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*. *Mol. Cell. Biol.*, **9**, 1316–1323.

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

40. Barbara,K.E., Haley,T.M., Willis,K.A. and Santangelo,G.M. (2007) The transcription factor Gcr1 stimulates cell growth by participating in nutrient-responsive gene expression on a global level. *Mol. Genet. Genomics*, **277**, 171–188.
41. Uemura,H., Koshio,M., Inoue,Y., Lopez,M.C. and Baker,H.V. (1997) The role of Gcr1p in the transcriptional activation of glycolytic genes in yeast *Saccharomyces cerevisiae*. *Genetics*, **147**, 521–532.
42. Baum,J.A., Geever,R. and Giles,N.H. (1987) Expression of qa-1F activator protein: identification of upstream binding sites in the *qa* gene cluster and localization of the DNA-binding domain. *Mol. Cell. Biol.*, **7**, 1256–1266.
43. Andrianopoulos,A. and Timberlake,W.E. (1994) The *Aspergillus nidulans* abaA gene encodes a transcriptional activator that acts as a genetic switch to control development. *Mol. Cell. Biol.*, **14**, 2503–2515.
44. Smale,S.T. and Kadonaga,J.T. (2003) The RNA polymerase II core promoter. *Annu. Rev. Biochem.*, **72**, 449–479.
45. Weber,W., Bacchus,W., Daoud-El Baba,M. and Fussenegger,M. (2007) Vitamin H-regulated transgene expression in mammalian cells. *Nucleic Acids Res.*, **35**, e116.
46. Hartenbach,S., Baba,M.D., Weber,W. and Fussenegger,M. (2007) An engineered L-arginine sensor of *Chlamydia pneumoniae* enables arginine-adjustable transcription control in mammalian cells and mice. *Nucleic Acids Res.*, **35**, e136.
47. Pfleger,B.F., Pitera,D.J., Smolke,C.D. and Keasling,J.D. (2006) Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat. Biotechnol.*, **24**, 1027–1032.
48. Lu,C. and Jeffries,T. (2007) Shuffling of promoters for multiple genes to optimize xylose fermentation in an engineered S*accharomyces cerevisiae* strain. *Appl. Environ. Microbiol.*, **73**, 6072–6077.
49. Jensen,P.R. and Hammer,K. (1998) Artificial promoters for metabolic optimization. *Biotechnol. Bioeng.*, **58**, 191–195.
50. Mijakovic,I., Petranovic,D. and Jensen,P.R. (2005) Tunable promoters in systems biology. *Curr. Opin. Biotechnol.*, **16**, 329–335.
51. Kramer,B.P., Viretta,A.U., Daoud-El-Baba,M., Aubel,D., Weber,W. and Fussenegger,M. (2004) An engineered epigenetic transgene switch in mammalian cells. *Nat. Biotechnol.*, **22**, 867–870.
52. Cox,R.S. III, Surette,M.G. and Elowitz,M.B. (2007) Programming gene expression with combinatorial promoters. *Mol. Syst. Biol.*, **3**, 145.
53. Guido,N.J., Wang,X., Adalsteinsson,D., McMillen,D., Hasty,J., Cantor,C.R., Elston,T.C. and Collins,J.J. (2006) A bottom-up approach to gene regulation. *Nature*, **439**, 856–860.
54. Liu,Z., Pscheidt,B., Avi,M., Gaisberger,R., Hartner,F.S., Schuster,C., Skranc,W., Gruber,K. and Glieder,A. (2008) Laboratory evolved biocatalysts for stereoselective syntheses of substituted benzaldehyde cyanohydrins. *Chembiochem*, **9**, 58–61.
55. van der Klei,I.J., Yurimoto,H., Sakai,Y. and Veenhuis,M. (2006) The significance of peroxisomes in methanol metabolism in methylotrophic yeast. *Biochim. Biophys. Acta*, **1763**, 1453–1462.
56. Inan,M. (2000) Studies on the alcohol oxidase (*AOX1*) promoter of *Pichia pastoris*. *Ph.D. thesis*. Department of Food Science and Technology. University of Nebraska, Lincoln, NE.
57. Inan,M., Meagher,M.M. and Benson,A.K. (2004) patent no. 6699691.
58. Ellgaard,L., Molinari,M. and Helenius,A. (1999) Setting the standards: quality control in the secretory pathway. *Science*, **286**, 1882–1888.
59. Griffith,D.A., Delipala,C., Leadsham,J., Jarvis,S.M. and Oesterhelt,D. (2003) A novel yeast expression system for the overproduction of quality-controlled membrane proteins. *FEBS Lett.*, **553**, 45–50.
60. Sauer,M., Branduardi,P., Gasser,B., Valli,M., Maurer,M., Porro,D. and Mattanovich,D. (2004) Differential gene expression in recombinant *Pichia pastoris* analysed by heterologous DNA microarray hybridisation. *Microb. Cell. Fact.*, **3**, 17.
61. Hohenblum,H., Gasser,B., Maurer,M., Borth,N. and Mattanovich,D. (2004) Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol. Bioeng.*, **85**, 367–375.
62. Parekh,R.N. and Wittrup,K.D. (1997) Expression level tuning for optimal heterologous protein secretion in *Saccharomyces cerevisiae*. *Biotechnol. Prog.*, **13**, 117–122.
63. Higgins,M.K., Demir,M. and Tate,C.G. (2003) Calnexin co-expression and the use of weaker promoters increase the expression of correctly assembled Shaker potassium channel in insect cells. *Biochim. Biophys. Acta*, **1610**, 124–132.

Downloaded from https://academic.oup.com/nar/article/36/12/e76/1142282 by guest on 18 October 2022