Structures and mechanisms of glycosyl hydrolases

The wealth of information provided by the recent structure determinations of many different glycosyl hydrolases shows that the substrate specificity and the mode of action of these enzymes are governed by exquisite details of their three-dimensional structures rather than by their global fold.

Structure 15 September 1995, **3**:853–859

Carbohydrates show wide stereochemical variation and can be assembled in so many different fashions that there are over 10^{12} possible isomers for a reducing hexasaccharide [1]. Living organisms take advantage of this diversity by using oligosaccharides and polysaccharides for a multitude of biological functions, from storage and structure to highly specific signalling roles. Selective hydrolysis of glycosidic bonds is therefore crucial for energy uptake, cell wall expansion and degradation, and turnover of signalling molecules. As a consequence of saccharide diversity, there is great variety amongst the enzymes that hydrolyze glycosidic bonds, the *O*-glycosyl hydrolases (EC 3.2.1.x). Heritable deficiencies in glycosyl hydrolases, for example lactose intolerance [2] and a number of lysosomal storage diseases [3], are among the most frequent genetically based syndromes in man.

Mechanisms

Enzymatic hydrolysis of the glycosidic bond takes place via general acid catalysis that requires two critical residues: a proton donor and a nucleophile/base [4,5] (Fig. 1). This hydrolysis occurs via two major mechanisms giving rise to either an overall retention, or an inversion, of anomeric configuration [4]. In both the retaining and the inverting mechanisms, the position of the proton donor is identical, in other words it is within hydrogen-bonding distance of the glycosidic oxygen. In retaining enzymes, the nucleophilic catalytic base is in close vicinity of the sugar anomeric carbon. This base, however, is more distant in inverting enzymes which must accommodate a water molecule between the base and the sugar. This difference results in an average distance between the two catalytic residues of ~5.5 Å in retaining enzymes as opposed to ~10 Å in inverting enzymes [6].

Lysozymes were the first glycosyl hydrolases to have their three-dimensional (3D) structures solved [7,8]. The two catalytic amino acids were identified as aspartate and glutamate residues. In most glycosyl hydrolases studied since, only aspartate and/or glutamate residues have been found to perform catalysis. Recent data, however, suggest that other residues may sometimes be involved in glycosidic bond cleavage. Typical examples are viral neuraminidase and bacterial sialidase, where the transition state is thought to be stabilized with the help of a tyrosine [9,10].

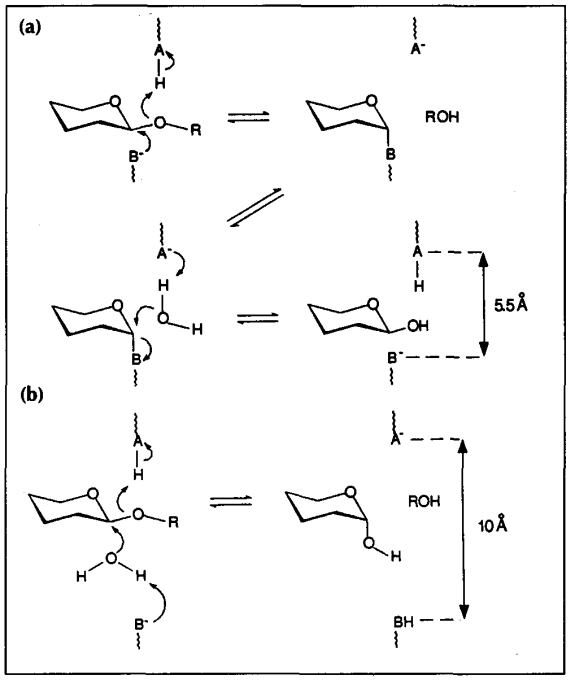

Fig. 1. The two major mechanisms of enzymatic glycosidic bond hydrolysis as first proposed by Koshland [4]. **(a)** The retaining mechanism, in which the glycosidic oxygen is protonated by the acid catalyst (AH) and nucleophilic assistance to aglycon departure is provided by the base B^-. The resulting glycosyl enzyme is hydrolyzed by a water molecule and this second nucleophilic substitution at the anomeric carbon generates a product with the same stereochemistry as the substrate. **(b)** The inverting mechanism, in which protonation of the glycosidic oxygen and aglycon departure are accompanied by a concomitant attack of a water molecule that is activated by the base residue (B^-). This single nucleophilic substitution yields a product with opposite stereochemistry to the substrate.

Enzymatic hydrolysis of certain oligosaccharides may also take advantage of the natural chemistry of the substrate. Participation of the substrate C2 acetamido group in catalysis by lysozymes and some retaining chitinases has been suggested from several underrated sources. Hydrolysis by hen egg white lysozyme (HEWL) of *N*-acetylchitobioside substrates with either a C2 hydroxyl or C2 acetamido group has been studied in detail [11].

© Current Biology Ltd ISSN 0969-2126

Halozyme EX2091
Merck v. Halozyme
PGR2025-00053

Removal of the C2 acetamido group leads to a reduction in k_{cat} of at least 100-fold, a factor which is similar, if not greater, than the effect of mutation of the proposed catalytic base Asp52. Perhaps the most serious evidence for the role of the acetamido substituent at C2, both in the enhancement of the catalytic rate and the retention of the anomeric configuration during hydrolysis, comes from solution studies on *N*-acetylglucosamine-containing saccharides [12,13]. Hydrolysis of these sugars, in solution, not only goes faster than that of C2 hydroxyl-substituted sugars, but with overall retention of anomeric configuration, suggesting that, 'on enzyme', a catalytic base, such as Asp52 in HEWL, is not necessarily required. These observations may also help to explain the apparent anomaly that some enzymes known to catalyze the cleavage of β1,4 bonds adjacent to C2 acetamido-substituted saccharides, whose structures have recently been determined, appear to lack a suitable catalytic base located in an appropriate position. Soluble lytic transglycosylase [14], goose lysozyme [15], hevamine [16], and others fall into this category.

Glycosyl hydrolases have developed various ways to lower the energy barrier of the hydrolysis reaction, such as substrate distortion into a sofa or half-chair conformation [17,18]. It is believed that protonation of the glycosidic bond is accompanied by a substantial lengthening of this bond [19]. Oligosaccharide complexes of the endoglucanase V from *Humicola insolens* suggest that this enzyme has evolved a 'stretched' subsite at the catalysis point in order to favour binding of this elongated transition state over that of the ground state substrate molecule (G Davies, S Tolley, B Henrissat, C Hjort and M Schülein, unpublished data).

Families and folds

A classification of glycosyl hydrolases into more than 45 families, on the basis of similarities in amino acid sequence, has recently been proposed [20,21], with the prospect that this may facilitate the derivation of useful information on the structure and function of these enzymes. Underlying this classification was the idea that proteins in a given family would have a fold sufficiently similar to allow homology modelling. In this classification, enzymes with different substrate specificities are sometimes found in the same family, indicating an evolutionary divergence to acquire new specificities, as is found in, for example, families 1, 13 and 16. On the other hand, enzymes that hydrolyze the same substrate are sometimes found in different families. For example, cellulases are found in 11 families. In other words, the sequence, and hence structural, classification differs significantly from that of the International Union of Biochemistry (IUB) nomenclature of enzymes, which is based mostly on substrate specificity.

Because the 3D structures of proteins are more highly conserved than their sequences, several sequence-based families may have related folds. For instance, a structural similarity was suggested for family 11 xylanases and

Table 1. Structures and mechanisms in various families of glycosyl hydrolases.

Family	Enzyme	Organism	EC number	PDB* code	Mechanism	Reference
1	β-glucosidase	*Trifolium repens*	3.2.1.21	–	retaining	[31]
2	β-galactosidase	*Escherichia coli*	3.2.1.23	1BGL	retaining	[32]
5	endoglucanase A	*Clostridium cellulolyticum*	3.2.1.4	–	retaining	[33]
6	cellobiohydrolase II	*Trichoderma reesei*	3.2.1.91	3CBH	inverting	[28]
	endoglucanase	*Thermononospora fusca*	3.2.1.4	1TML	inverting	[34]
7	cellobiohydrolase I	*Trichoderma reesei*	3.2.1.91	1CEL	retaining	[35]
	endoglucanase I	*Humicola insolens*	3.2.1.4	–	retaining	[Davies, unpublished]
9	endoglucanase D	*Clostridium thermocellum*	3.2.1.4	–	inverting	[36]
10	xylanase A	*Streptomyces lividans*	3.2.1.8	1XAS	retaining	[37]
11	xylanase	*Bacillus circulans*	3.2.1.8	1BCX	retaining	[38]
13	α-amylase	*Aspergillus oryzae*	3.2.1.1	6TAA	retaining	[39]
14	β-amylase	*Glycine max*	3.2.1.2	1BTC	inverting	[40]
15	glucoamylase	*Aspergillus awamori*	3.2.1.3	3GLY	inverting	[41]
16	β-1,3-1,4-glucanase	*Bacillus sp.*	3.2.1.73	1BYH	retaining	[42]
17	β-1,3-1,4-glucanase	*Hordeum vulgare*	3.2.1.73	1GHR	unknown†	[43]
18	chitinase	*Serratia marcescens*	3.2.1.14	1CTN	retaining	[44]
19	chitinase	*Hordeum vulgare*	3.2.1.14	1BAA	inverting	[45]
20	chitobiase	*Serratia marcescens*	3.2.1.52	–	retaining	‡
22	lysozyme	Hen egg white	3.2.1.17	1HEL	retaining	[7]
23	lysozyme	Goose	3.2.1.17	153L	unknown	[15]
24	lysozyme	Bacteriophage T4	3.2.1.17	1LYD	unknown	[46]
33	sialidase	*Salmonella typhimurium*	3.2.1.18	2SIL	retaining	[10]
34	neuraminidase	Influenza virus B	3.2.1.18	1NSB	retaining	[47]
45	endoglucanase V	*Humicola insolens*	3.2.1.4	1ENG	inverting	[48]

*Protein Data Bank. †This family is predicted to have a retaining mechanism [24,25]. ‡I Tews, Z Dauter, KS Wilson & CE Vorgias, [abstract 038], 4th European Workshop on Crystallography of Biological Macromolecules, Como, Italy, May 1995.

family 12 cellulases [22], whereas family 7 cellulases have been found to have an arrangement of catalytic residues and a fold similar to those of the β-1,3-glucanases and β-1,3-1,4-glucanases of family 16 [23]. More recently, families 1, 2, 5, 10, 17, 30, 35, 39 and 42 were proposed to have evolved from a common ancestor [24,25]. All of these family groupings substantiate the strict conservation of the catalytic machinery and mechanism during evolution. A grouping indicating structural similarity for various lysozymes (families 22, 23 and 24) and family 19 plant chitinases has also been proposed [26]. This grouping, however, includes enzymes known to operate with retention of configuration, such as HEWL, and some with a substantially different arrangement of catalytic amino acids, such as family 19 plant chitinases, which are inverting enzymes [27].

Table 1 reports the various glycosyl hydrolases families for which at least one 3D structure has been determined, together with the mechanism of glycosidic bond hydrolysis, in cases where it is known. The substrates for these enzymes are shown in Figure 2. Many glycosyl hydrolases have a modular structure consisting of a catalytic domain and one or more non-catalytic domains, some of which are involved in substrate binding, but most of which have unknown functions. Figure 3 shows the main folds found in the catalytic domains of selected glycosyl hydrolases and, for families 6 and 7, the comparison of the structures of cellobiohydrolases with those of the corresponding endoglucanases (see below).

Active-site topologies

Although many protein folds are represented in the 22 families for which a 3D structure is known, the overall topologies of the active sites fall into only three general classes, regardless of whether the enzyme is inverting or retaining. These three topologies (Fig. 4) can, in principle, be built on the same fold, with the same catalytic residues.

Pocket or crater

This topology (Fig. 4a) is optimal for the recognition of a saccharide non-reducing extremity and is encountered in monosaccharidases such as β-galactosidase, β-glucosidase, sialidase and neuraminidase, and in exopolysaccharidases such as glucoamylase and β-amylase. Such exopolysaccharidases are adapted to substrates having a large number of available chain ends, such as native starch granules, whose radial structure exposes all the non-reducing chain ends at the surface. On the other hand, these enzymes are not very efficient for fibrous substrates such as native cellulose, which has almost no free chain ends.

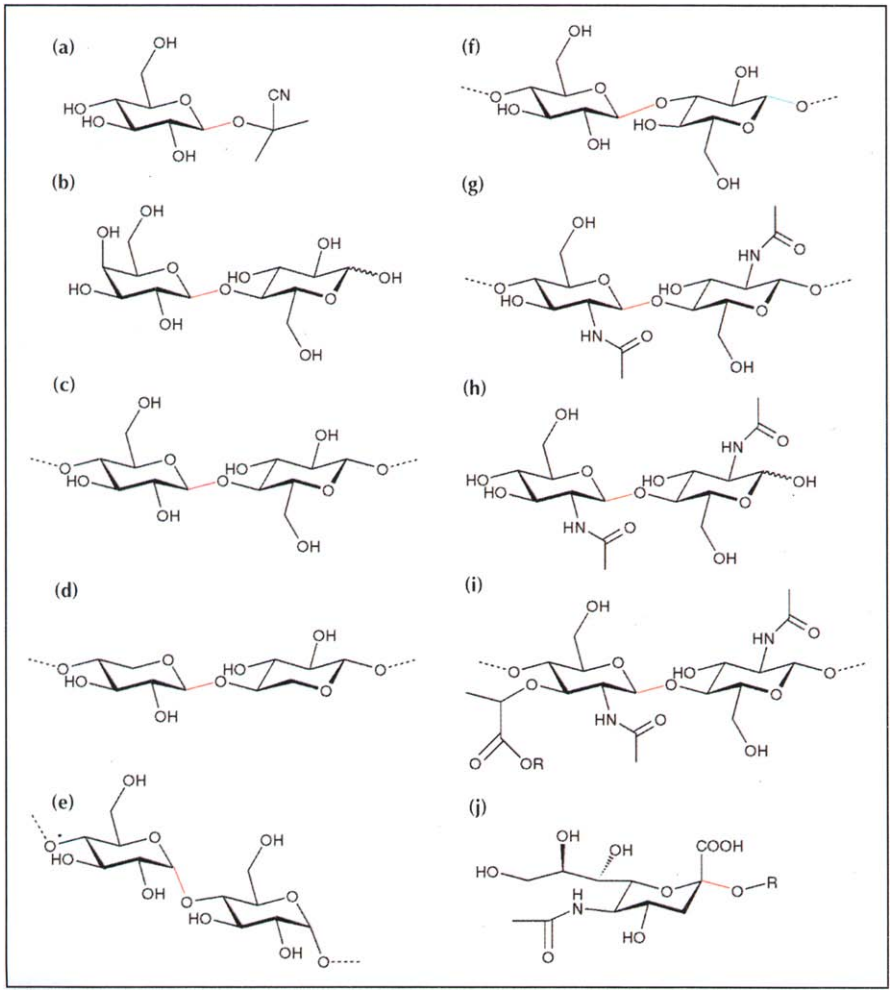

Fig. 2. Structures of the substrates for the enzymes given in Table 1. The hydrolyzed bond is shown in red. **(a)** Linamarin (substrate for cyanogenic β-glucosidase, family 1). **(b)** Lactose (β-galactosidase, family 2). **(c)** Cellulose (endoglucanases and cellobiohydrolases, families 5, 6, 7, 9 and 45). **(d)** Xylan (xylanases, families 10 and 11). **(e)** Amylose (α-amylase, family 13; β-amylase, family 14; and glucoamylase, family 15). **(f)** Mixed β-1,3-1,4-glucan with the bond hydrolyzed by β-1,3-glucanases (family 16) in red and that cleaved by β-1,3-1,4-glucanases (family 17) in cyan. **(g)** Chitin (chitinases in families 18 and 19, family 22 lysozymes). **(h)** Chitobiose (chitobiase, family 20). **(i)** Bacterial cell wall polymer consisting of alternating *N*-acetylmuramic acid and *N*-acetylglucosamine units (lysozymes in families 22, 23 and 24). **(j)** Glycoconjugate-linked sialic acid (sialidase, family 33 and neuraminidase, family 34).

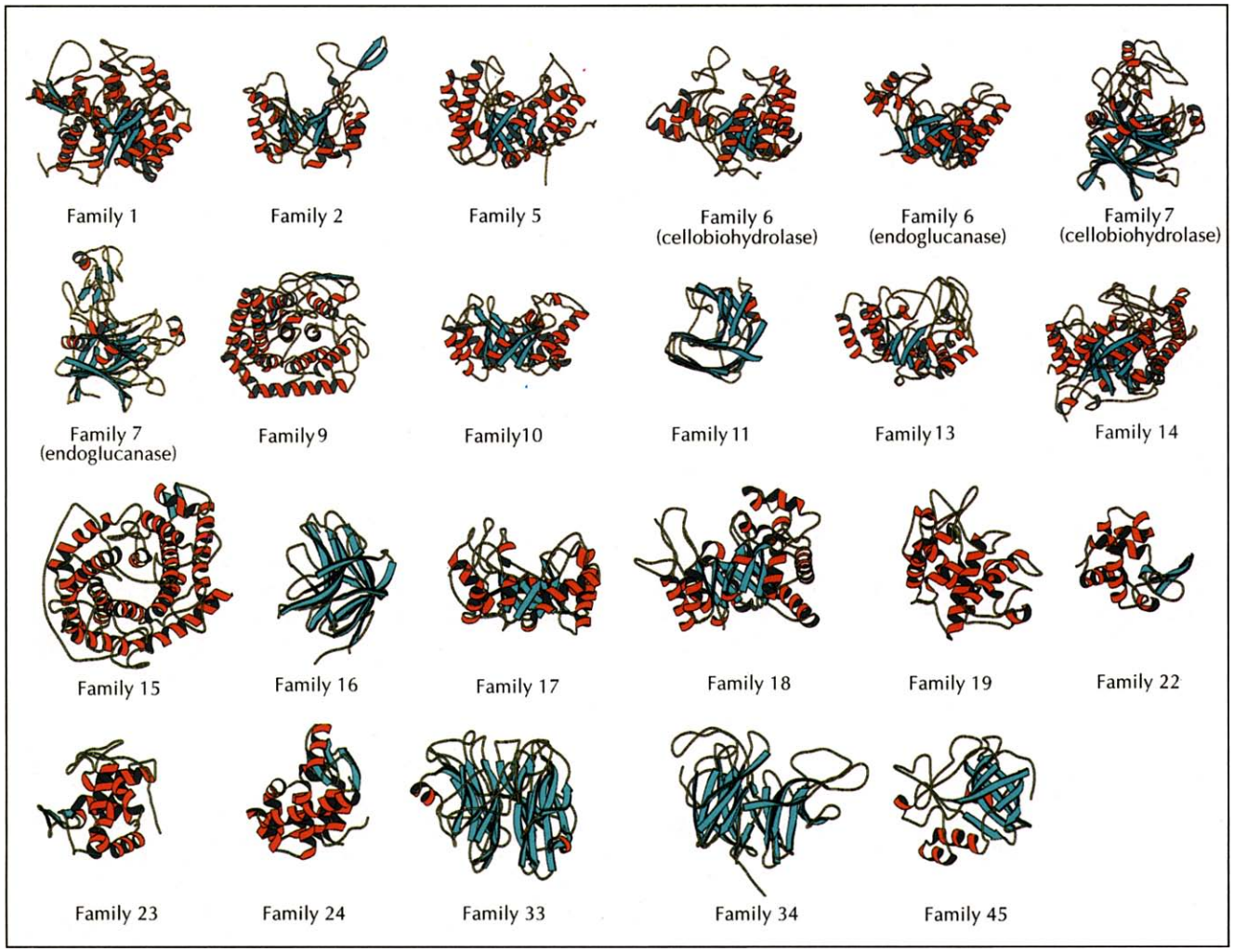

Fig. 3. Ribbon representation of the main fold of the catalytic domain in various glycosyl hydrolase families (see Table 1 and text). β strands are shown in cyan and α helices in red. (Figure produced using the program MOLSCRIPT [49].)

Cleft or groove

This 'open' structure (Fig. 4b) allows a random binding of several sugar units in polymeric substrates and is commonly found in endo-acting polysaccharidases such as lysozymes, endocellulases, chitinases, α-amylases, xylanases, β-1,3-1,4-glucanases and β-1,3-glucanases.

Tunnel

This topology (Fig. 4c) arises from the previous one when the protein evolves long loops that cover part of the cleft. Found so far only in cellobiohydrolases, the resulting tunnel enables a polysaccharide chain to be threaded through it [28]. A comparison of the cellobiohydrolases in families 6 and 7 with the corresponding endoglucanases is shown in Figure 3. The loops that cause the catalytic centres of cellobiohydrolases to lie within enclosed tunnels can be seen clearly and, for the family 6 enzymes, are also illustrated by the surface representations in Figures 4b and 4c. This topology allows these enzymes to release the product while remaining firmly bound to the polysaccharide chain, thereby creating the conditions for processivity (Fig. 5). It remains unclear at present whether the substrate initially penetrates the active site in an 'exo' fashion by one of the two entrances of the tunnel or whether the loops that close the active site can 'open' occasionally to allow a random binding followed by the processive action. In either case it should be noted that, depending on the mechanism (inverting or retaining) and the exact position of the cleavage point with respect to the several subsites, the directionality of the enzyme motion along the chain may change. For instance, cellobiohydrolase II of *Trichoderma reesei* proceeds towards the reducing end of cellulose, whereas the reverse was suggested for cellobiohydrolase I from the same organism [23]. Processivity is probably a key factor for the efficient enzymatic degradation of insoluble microcrystalline cellulose.

Concluding remarks

Orengo *et al.* [29] have shown that certain protein folds (superfolds) occur more often than others. More precisely, only nine superfolds are sufficient to describe the folding in ~30% of all proteins and it is thought that the total number of protein folds is not more than a few thousand [29,30]. So far, from the 22 families of glycosyl hydrolases for which a 3D structure has been

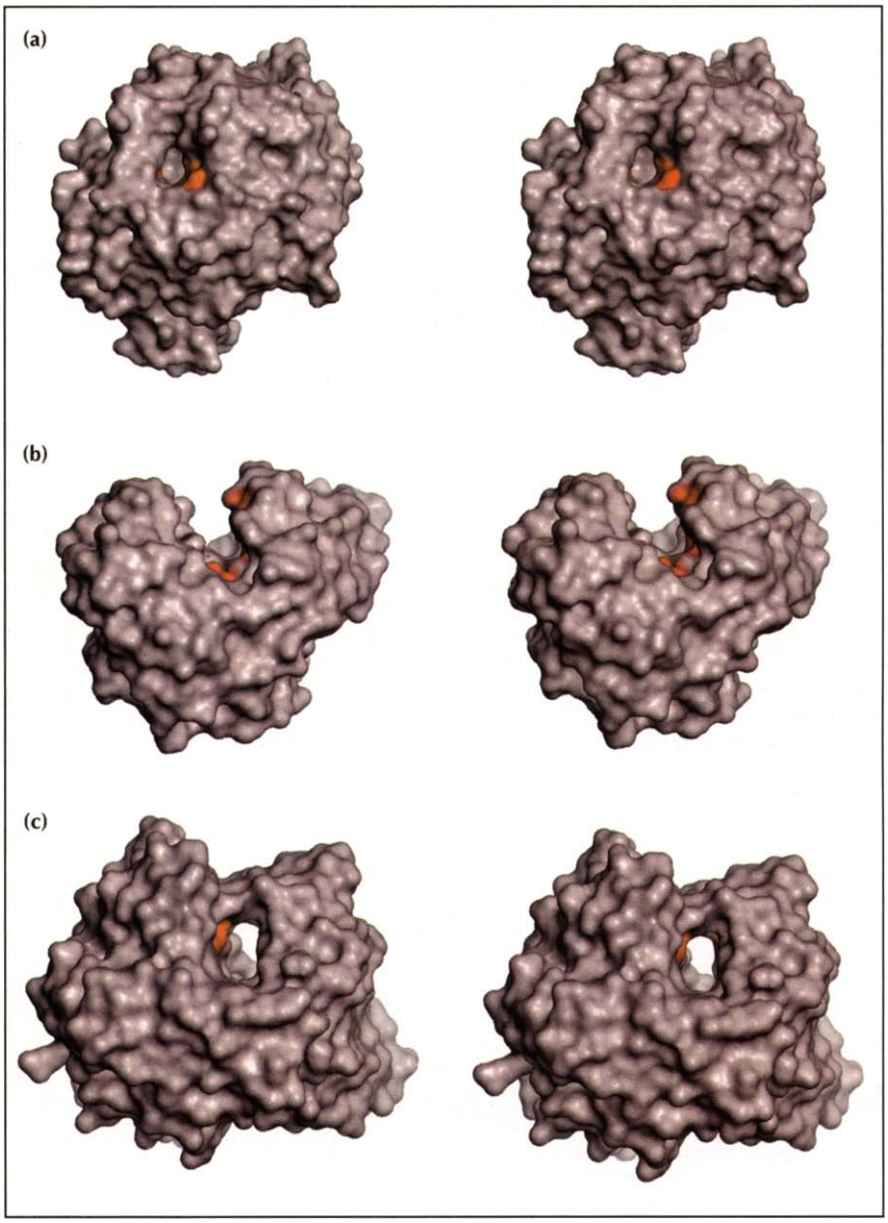

Fig. 4. The three types of active site found in glycosyl hydrolases. **(a)** The pocket (glucoamylase from *A. awamori*). **(b)** The cleft (endoglucanase E2 from *T. fusca*). **(c)** The tunnel (cellobiohydrolase II from *T. reesei*). The proposed catalytic residues are shaded in red. (Molecular surface diagrams were prepared using the MOLVIEWER program [M Hartshorn, unpublished program].)

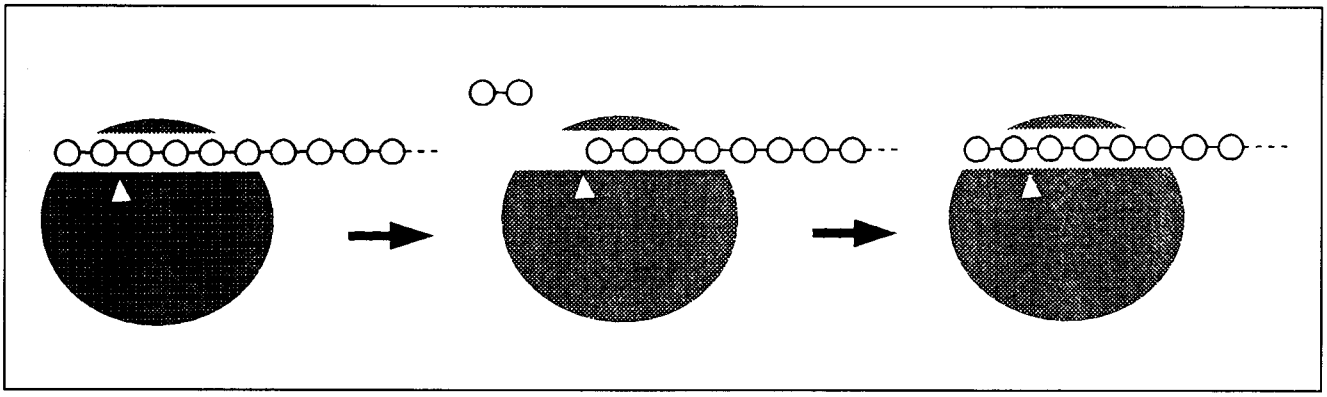

Fig. 5. The mechanism of processivity of cellobiohydrolases. Once a disaccharide product is liberated (shown as two linked circles), the enzyme remains bound to the polysaccharide chain by several subsites, the 'lid' closing the active site and, for retaining enzymes, the glycosyl enzyme. The two empty sites, and perhaps other factors such as loop movements, provide the driving force for enzyme motion along the chain (or chain threading along the enzyme's active site) by two sugar units. Hydrolysis can then proceed iteratively until enzyme movement is stopped by steric factors, or until the loops that close the active site move and release the polysaccharide chain.

determined, nine are of the TIM barrel type (Table 1). In addition, detailed sequence comparisons have suggested that several other families also share the same fold, implying that, for glycosyl hydrolases, the bias towards this fold could be even more important [24].

There are presently 52 families in the classification based on sequence similarities (B Henrissat and A Bairoch, unpublished data); 3D structures are known for about 40% of these families. The first enzyme structure to be solved, more than 30 years ago, was that of hen egg white lysozyme, a glycosyl hydrolase. At the present pace of structural investigations, the 3D fold for all of the remaining families could be determined within a few years.

Acknowledgements: We thank those who gave coordinates not yet available from the Protein Data Bank: Tracey Barrett (Mill Hill) for the cyanogenic β-glucosidase of *Trifolium repens*; Zygmunt and Ulla Derewenda (University of Alberta) for xylanase A from *Streptomyces lividans*; Valérie Ducros (CNRS Marseille) for endoglucanase A from *Clostridium cellulolyticum*; Michel Juy (CNRS Marseille) for endoglucanase D from *Clostridium thermocellum*; and Brian Matthews (University of Oregon) for β-galactosidase from *E. coli* and for goose lysozyme. We would like to thank Mike Hartshorn (University of York) for help with the MOLVIEWER program and the production of figures. This work has been funded by the BIO2 CT 94 3018 research grant from the European Community.

References

1. Laine, R.A. (1994). A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the *Isomer Barrier* to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* **4**, 759–767.
2. Semenza, G. & Auricchio, S. (1989). Small-intestinal disaccharidases. In *The Metabolic Basis of Inherited Diseases*. (Scrives, C.R., Beaudet, A.L. & Valle, V.S., eds), pp. 2975–2997, McGraw-Hill, NY.
3. Neufeld, E.F. (1991). Lysosomal storage diseases. *Annu. Rev. Biochem.* **60**, 257–280.
4. Koshland, D.E. (1953). Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev. Camb. Philos. Soc.* **28**, 416–436.
5. Sinnott, M.L. (1990). Catalytic mechanisms of enzymic glycosyl transfer. *Chem. Rev.* **90**, 1171–1202.
6. McCarter, J.D. & Withers, S.G. (1994). Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* **4**, 885–892.
7. Blake, C.C.F., Koenig, D.F., Mair, G.A., North, A.C.T., Phillips, D.C. & Sarma, V.R. (1965). Structure of hen egg white lysozyme. A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **206**, 757–763.
8. Matthews, B.W. & Remington, S.J. (1974). The three-dimensional structure of lysozyme from bacteriophage T4. *Proc. Natl. Acad. Sci. USA* **71**, 4178–4182.
9. Burmeister, W., Henrissat, B., Bosso, C., Cusack, S. & Ruigrok, R. (1993). Influenza B neuraminidase can synthesize its own inhibitor. *Structure* **1**, 19–26.
10. Crennel, S.J., Garman, E.F., Laver, W.G., Vimr, E.R. & Taylor, G.L. (1993). Crystal structure of a bacterial sialidase (from *Salmonella typhimurium* LT2) shows the same fold as an influenza virus neuraminidase. *Proc. Natl. Acad. Sci. USA* **90**, 9852–9856.
11. Lowe, G. & Sheppard, G. (1968). Acetamido-group participation in lysozyme catalysis. *J. Chem. Soc., Chem. Commun.* 529–530.
12. Capon, B. (1969). Mechanisms in carbohydrate chemistry. *Chem. Rev.* **69**, 407–498.
13. Piszkiewicz, D. & Bruice, T.C. (1968). Glycoside hydrolysis. II. Intramolecular carboxyl and acetamido group catalysis in β-glycoside hydrolysis. *J. Am. Chem. Soc.* **90**, 2156–2163.
14. Thunnissen, A.-M., *et al.*, & Dijkstra, B.W. (1994). Doughnut-shaped structure of a bacterial muramidase revealed by X-ray crystallography. *Nature* **367**, 750–753.
15. Weaver, L.H., Grütter, M.G. & Matthews, B.W. (1995). The refined structures of goose lysozyme and its complex with a bound trisaccharide show that the "goose-type" lysozymes lack a catalytic aspartate residue. *J. Mol. Biol.* **245**, 54–68.
16. Terwisscha van Scheltinga, A.C., Kalk, K.H., Beintema, J.J. & Dijkstra, B.W. (1994). Crystal structure of hevamine, a plant defence protein with chitinase and lysozyme activity, and its complex with an inhibitor. *Structure* **2**, 1181–1189.
17. Strynadka, N.C.J. & James, M.N.G. (1991). Lysozyme revisited: crystallographic evidence for distortion of an *N*-acetylmuramic acid residue bound in site D. *J. Mol. Biol.* **220**, 401–424.
18. Kuroki, R., Weaver, L.H. & Matthews, B.W. (1993). A covalent enzyme–substrate intermediate with saccharide distortion in a mutant T4 lysozyme. *Science* **262**, 2030–2033.
19. Tanaka, Y., Tao, W., Blanchard, J.S. & Hehre, E.J. (1994). Transition state structures for the hydrolysis of α-D-glucopyranosyl fluoride by retaining and inverting reactions of glycosylases. *J. Biol. Chem.* **269**, 32306–32312.
20. Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarity. *Biochem. J.* **280**, 309–316.
21. Henrissat, B. & Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **293**, 781–788.
22. Törrönen, A., Kubicek, C.P. & Henrissat, B. (1993). Amino acid sequence similarities between low molecular weight endo-1,4-β-xylanases and family H cellulases revealed by clustering analysis. *FEBS Lett.* **321**, 135–139.
23. Divne, C., *et al.*, & Jones, T.A. (1994). The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science* **265**, 524–528.
24. Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.-P. & Davies, G. (1995). Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc. Natl. Acad. Sci. USA* **92**, 7090–7094.
25. Jenkins, J., Lo Leggio, L., Harris, G. & Pickersgill, R. (1995). β-glucosidase, β-galactosidase, family A cellulases, family F xylanases and two barley glycanases form a superfamily with 8-fold β/α architecture and with two conserved glutamates near the carboxy-terminal ends of β-strands four and seven. *FEBS Lett.* **362**, 281–285.
26. Holm, L. & Sander, C. (1994). Structural similarity of plant chitinase and lysozymes from animals and phage. *FEBS Lett.* **340**, 129–132.
27. Fukamizo, T., Koga, D. & Goto, S. (1995). Comparative biochemistry of chitinases — anomeric form of the reaction products. *Biosci. Biotechnol. Biochem.* **59**, 311–313.
28. Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K.C. & Jones, T.A. (1990). Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science* **249**, 380–386.
29. Orengo, C., Jones, D.T. & Thornton, J.M. (1994). Protein superfamilies and domain superfolds. *Nature* **372**, 631–634.
30. Chothia, C. (1992). One thousand families for the molecular biologist. *Nature* **357**, 543–544.
31. Barrett, T., Suresh, C.G., Tolley, S.P., Dodson, E.J. & Hughes, M.A. (1995). The crystal structure of a cyanogenic β-glucosidase from white clover, a family 1 glycosyl hydrolase. *Structure* **3**, 951–960.
32. Jacobson, R.H., Zhang, X.-J., DuBose, R.F. & Matthews, B.W. (1994). Three-dimensional structure of β-galactosidase from *E. coli*. *Nature* **369**, 761–766.
33. Ducros, V., *et al.*, & Haser, R. (1995). Crystal structure of the catalytic domain of a bacterial cellulase belonging to family 5. *Structure* **3**, 939–949.
34. Spezio, M., Wilson, D.B. & Karplus, P.A. (1993). Crystal structure of the catalytic domain of a thermophilic endocellulase. *Biochemistry* **32**, 9906–9916.
35. Divne, C., *et al.*, & Jones, T.A. (1994). The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science* **265**, 524–528.
36. Juy, M., *et al.*, & Aubert, J.-P. (1992). Three-dimensional structure of a thermostable bacterial cellulase. *Nature* **357**, 89–91.
37. Derewenda, U., *et al.*, & Derewenda, Z.S. (1994). Crystal structure, at 2.6 Å resolution, of the *Streptomyces lividans* xylanase A, a member of the F family of β-1,4-D-glycanases. *J. Biol. Chem.* **269**, 20811–20814.
38. Wakarchuk, W.W., Campbell, R.L., Sung, W., Davoodi, J. & Yaguchi, M. (1994). Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. *Protein Sci.* **3**, 467–475.
39. Boel, E., *et al.*, & Woldike, H.F. (1990). Calcium binding in α-amylases: an x-ray diffraction study at 2.1-Å resolution of two enzymes from *Aspergillus*. *Biochemistry* **29**, 6244–6249.
40. Mikami, B., *et al.*, & Sacchettini, J.C. (1993). The 2.0 Å resolution structure of soybean β-amylase complexed with α-cyclodextrin. *Biochemistry* **32**, 6836–6845.
41. Aleshin, A., Golubev, A., Firsov, L.M. & Honzatko, R.B. (1992). Crystal structure of glucoamylase from *Aspergillus awamori* var. *X100* to 2.2 Å resolution. *J. Biol. Chem.* **267**, 19291–19298.

42. Keitel, T., Simon, O., Borriss, R. & Heinemann, U. (1993). Molecular and active-site structure of a *Bacillus* 1,3-1,4-β-glucanase. *Proc. Natl. Acad. Sci. USA* **90**, 5287–5291.
43. Varghese, J.N., Garrett, T.P.J., Colman, P.M., Chen, L., Høj, P.B. & Fincher, G.B. (1994). Three-dimensional structure of two plant β-glucan endohydrolases with distinct substrate specificities. *Proc. Natl. Acad. Sci. USA* **91**, 2785–2789.
44. Perrakis, A., *et al.*, & Vorgias, C.E. (1994). Crystal structure of a bacterial chitinase at 2.3 Å resolution. *Structure* **2**, 1169–1180.
45. Hart, P.J., Monzingo, A.F., Ready, M.P., Ernst, S.R. & Robertus, J.D. (1993). Crystal structure of an endochitinase from *Hordeum vulgare* L. seeds. *J. Mol. Biol.* **229**, 189–193.
46. Weaver, L.H. & Matthews, B.W. (1987). Structure of bacteriophage T4 lysozyme refined at 1.7 Å resolution. *J. Mol. Biol.* **193**, 189–200.
47. Varghese, J.N., Laver, W.G. & Colman, P.M. (1983). Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature* **303**, 35–40.
48. Davies, G.J., *et al.*, & Schülein, M. (1993). Structure and function of endoglucanase V. *Nature* **365**, 362–364.
49. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950.

Gideon Davies, Department of Chemistry, University of York, Heslington, York, YO1 5DD, UK and Bernard Henrissat, Centre de Recherches sur les Macromolécules Végétales, CNRS, BP 53, F-38041 Grenoble Cedex 9, France.