Catalysis and specificity in enzymatic glycoside hydrolysis: a $^{2,5}B$ conformation for the glycosyl-enzyme intermediate revealed by the structure of the *Bacillus agaradhaerens* family 11 xylanase

Elisabetta Sabini[1], Gerlind Sulzenbacher[1,2], Miroslava Dauter[1], Zbigniew Dauter[1], Per Linå Jørgensen[3], Martin Schülein[3], Claude Dupont[4], Gideon J Davies[1] and Keith S Wilson[1]

Background: The enzymatic hydrolysis of glycosides involves the formation and subsequent breakdown of a covalent glycosyl-enzyme intermediate via oxocarbenium-ion-like transition states. The covalent intermediate may be trapped on-enzyme using 2-fluoro-substituted glycosides, which provide details of the intermediate conformation and noncovalent interactions between enzyme and oligosaccharide. Xylanases are important in industrial applications – in the pulp and paper industry, pretreating wood with xylanases decreases the amount of chlorine-containing chemicals used. Xylanases are structurally similar to cellulases but differ in their specificity for xylose-based, versus glucose-based, substrates.

Results: The structure of the family 11 xylanase, Xyl11, from *Bacillus agaradhaerens* has been solved using X-ray crystallography in both native and xylobiosyl-enzyme intermediate forms at 1.78 Å and 2.0 Å resolution, respectively. The covalent glycosyl-enzyme intermediate has been trapped using a 2-fluoro-2-deoxy substrate with a good leaving group. Unlike covalent intermediate structures for glycoside hydrolases from other families, the covalent glycosyl-enzyme intermediate in family 11 adopts an unusual $^{2,5}B$ conformation.

Conclusions: The $^{2,5}B$ conformation found for the α-linked xylobiosyl-enzyme intermediate of Xyl11, unlike the $^{4}C_1$ chair conformation observed for other systems, is consistent with the stereochemical constraints required of the oxocarbenium-ion-like transition state. Comparison of the Xyl11 covalent glycosyl-enzyme intermediate with the equivalent structure for the related family 12 endoglucanase, CelB, from *Streptomyces lividans* reveals the likely determinants for substrate specificity in this clan of glycoside hydrolases.

Addresses: [1]Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York, Y010 5DD, UK. [2]European Molecular Biology Laboratory, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany. [3]Novo-Nordisk A/S, Novo allé, Bagsvaerd, DK-2880, Denmark. [4]Centre de Microbiologie et Biotechnologie, INRS-Institut Armand-Frappier, Laval-des-Rapides, C.P. 100, Québec, Canada, H7N 4Z3.

Correspondence: Gideon J Davies
E-mail: davies@yorvic.york.ac.uk

Key words: boat conformation, enzyme specificity, transition state, xylanase

Received: **29 March 1999**
Accepted: **26 April 1999**

Published: **21 June 1999**

Chemistry & Biology July 1999, **6**:483–492
http://biomednet.com/elecref/1074552100600483

Introduction

Endoxylanases (EC 3.2.1.8) are found in five of the 71 glycoside hydrolase families (5, 10, 11, 43 and 54) [1,2]. They all perform hydrolysis using acid-base catalysis with cleavage of the β-1,4 bonds giving rise to either inversion or retention of anomeric configuration (reviewed in [3]). In 1953, Koshland [4] proposed the basic mechanistic schemes: inverting enzymes function using a single displacement reaction in which nucleophilic attack by water from below the sugar ring occurs simultaneously with leaving-group departure, whereas retaining enzymes operate using a double-displacement mechanism in which a covalent glycosyl-enzyme intermediate is formed and subsequently hydrolysed via oxocarbenium-ion-like transition states. Demonstration of the covalent glycosyl-enzyme intermediate in the reaction mechanism has come from the elegant use of 2-deoxy-2-fluoroglycosides [5] whose covalent intermediate becomes trapped on-enzyme, causing it to accumulate. Three-dimensional structures for the covalent glycosyl-enzyme intermediates of retaining β-glycoside hydrolases have been solved for enzymes from families 1, 5, 10 and 12 [6–11]. With the exception of enzymes from glycoside hydrolase family 11, all have revealed α-linked intermediates with an undistorted pyranose ring in the $^{4}C_1$ chair conformation.

The cell wall of terrestrial plants is a composite material in which cellulose, hemicellulose and lignin are in tight association. Xylans are amongst the most abundant hemicelluloses and constitute more than 30% of the dry weight of plants. In terrestrial plants, xylans are based on a β-1,4-linked-D-xylosyl backbone, whereas marine algae synthesise xylans that are based on a β-1,3-linked backbone (reviewed in [12]). The enzymatic hydrolysis of hemicellulose is of major importance in the pulp and paper industry: xylanase pretreatment in the bleaching process is effective in reducing the amount of environmentally toxic chlorine and chlorine-containing chemicals used. For application in

the bleaching process, an ideal xylanase should be highly thermostable and active at alkaline pH. Considerable effort has been expended into the study of structure–function relationships of these enzymes, which should aid their application in these processes.

The alkalophilic bacillus *Bacillus agaradhaerens* produces many glycoside hydrolases that are of interest because of their high catalytic activity at elevated pH. In this paper we describe the structure of the *B. agaradhaerens* xylanase from family 11 (hereafter Xyl11), at 1.78 Å resolution. The Xyl11 structure has the typical family 11 jelly-roll motif (also found in family 12 [13] whose members prefer glucose-derived substrates). Xyl11 performs catalysis with net retention of anomeric configuration. The covalent glycosyl-enzyme intermediate has been trapped using 2′,4′-dinitrophenyl 2-deoxy-2-fluoroxylobioside, and the structure determined at 2.0 Å resolution. The covalent xylobiosyl-enzyme intermediate adopts a $^{2,5}B$ boat conformation which, unlike the $^{4}C_1$ chair conformation, fulfils the stereochemical requirements for the oxocarbenium-ion-like transition state of the retaining mechanism. Comparison of the Xyl11 complex structure with that from the related family 12 cellulase CelB [11] reveals the likely basis for substrate specificity in this clan of enzymes.

Results and discussion

Structure of native Xyl11

Single cube-shaped crystals of Xyl11 grow in 7–14 days reaching a maximum size of 0.2 × 0.2 × 0.2 mm. Room temperature data for a monoclinic crystal form were collected on beamline X-31 at the EMBL Hamburg outstation. The space group is $P2_1$ with cell dimensions a = 37.9 Å, b = 79.7 Å, c = 74.7 Å and β = 102.4°. Data extend to 2.00 Å and are 95.3% complete with an R_{merge} of 6.2% (Table 1). Data for the orthorhombic form of Xyl11 were collected from single crystals at 120 K. The space group is $P2_12_12_1$ with cell dimensions a = 71.9 Å, b = 75.4 Å, c = 78.4 Å. Native data extended to 1.78 Å with an overall R_{merge} of 5.5% and are 99.1% complete (Table 1). In both crystal forms there are two molecules of Xyl11 in the asymmetric unit. All structures were refined with maximum likelihood techniques using REFMAC [14].

The orthorhombic form of Xyl11 was refined to give a final model structure with a crystallographic R value of 11.7% and a corresponding R_{free} of 17.7%. The two molecules of Xyl11 in the asymmetric unit have a packing density of 2.41 Å^3/Da and a solvent content of ~50% [15]. The final model consists of two molecules of 1664 protein atoms (residues 1–207), 611 solvent water molecules and a single xylose moiety unit in the –2 subsite of the enzyme. Xylose was not added at any point in the preparation so we must conclude that this 'contaminant' remains after the fermentation process. All the nonglycine residues have conformational angles in permitted regions of the Ramachandran plot [16] with none of these in 'generously allowed' regions as defined by PROCHECK [17]. Further structure quality indicators are given in Table 1. The amino-terminal glutamine residue is present as the modified cyclic pyroglutamate ring. The *B. agaradhaerens* family 11 xylanase, Xyl11, adopts the same overall fold

Table 1

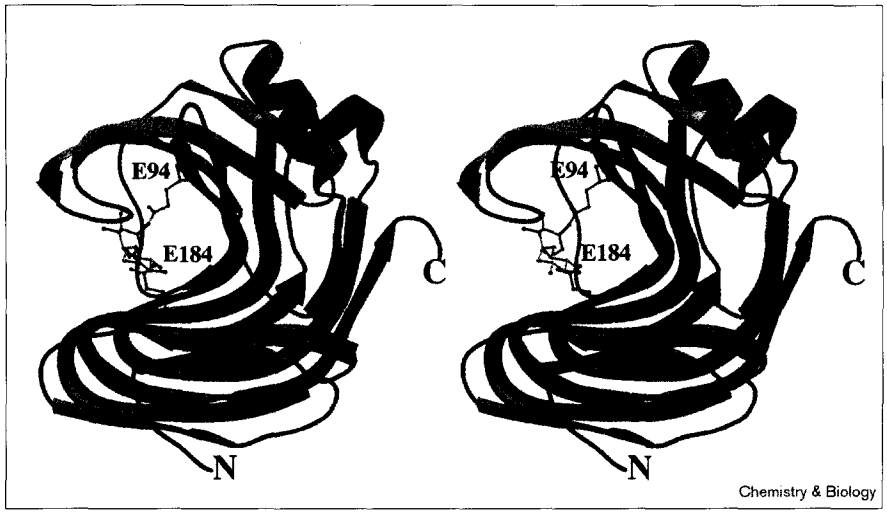
Data and refinement statistics for the *B. agaradhaerens* Xyl11.

	Native ($P2_1$)	Native ($P2_12_12_1$)	2-F-xylobiosyl intermediate
Data			
Resolution range (outer bin)	19.7–2.00	54.2–1.78	19.8–2.00
R_{merge} (%) (outer bin)	6.2 (24.1)	5.5 (16.2)	6.8 (15.4)
I/σI (outer shell)	13.4 (2.8)	27.4 (6.7)	22.5 (8.7)
Multiplicity (outer shell)	3.6 (3.3)	5.6 (3.2)	5.4 (5.6)
% Completeness (outer shell)	95.3 (78.7)	99.1 (98.3)	97.6 (82.5)
Refinement			
R factor (%)	16.2	11.7	14.2
R free (%)	22.6	17.7	18.9
No. protein atoms*	1661 (× 2 mols)	1664 (× 2 mols)	1641 (× 2 mols)
No. ligand atoms	–	10 (× 2 mols)	18 (× 2 mols)
No. solvent molecules	237	611	539
1-2 bond deviation (target; Å)	0.016 (0.020)	0.013	0.010
1-3 angle distance (target; Å)	0.039 (0.040)	0.031	0.029
Mean protein B-factor (Å^2)	25.6	16.5	16.4
Mean B mainchains A/B molecules (Å^2)	16/29	11/14	11/14
Mean B ligand (Å^2)	N/A	26.8/26.5	9.3/10.8
Mean B solvent (Å^2)	39.4	31.4	32.0

*This apparent discrepancy reflects the different modelling of alternative conformations.

Figure 1

Divergent stereo cartoon of the *B. agaradhaerens* Xyl11 covalent xylobiosyl-enzyme intermediate. This figure was drawn using BOBSCRIPT [48].



and topology to the other family 11 xylanase structures (for example [18–20]): a single domain polypeptide chain containing two β sheets, A and B, composed of five and eight β strands, respectively (Figure 1). In some family 11 xylanases, such as the *B. agaradhaerens* Xyl11 described here, the first β strand, A1, is absent. The two native structures of Xyl11, described here, are essentially the same with an overall root mean square (rms) deviation of 0.337 Å for mainchain atoms. The monoclinic form exhibits significantly higher disorder for the 'B' molecule, whose overall B value is 30 $Å^2$ compared to just 16 $Å^2$ for molecule A, presumably due to disorder within the crystal lattice. All further description refers to the orthorhombic crystal form, which does not display this disorder and which diffracts to higher resolution.

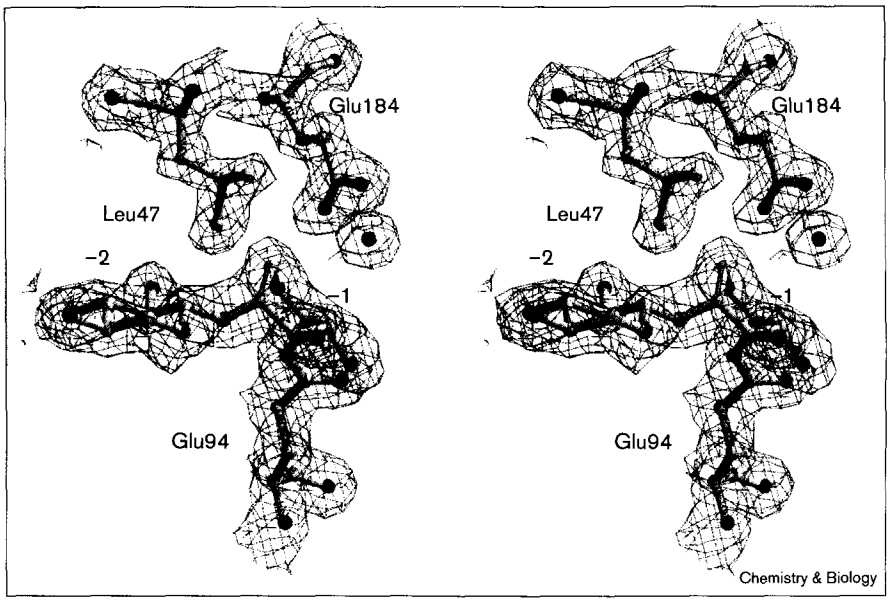
Structure of the covalent 2-F-xylobiosyl-enzyme intermediate

B. agaradhaerens Xyl11 catalyses random cleavage of the β-1,4-glycosidic bond of the xylan backbone with net retention of anomeric configuration. The mechanism is widely believed to be a double-displacement reaction, essentially as proposed by Koshland [4]: a covalent glycosyl-enzyme intermediate is formed, and subsequently hydrolysed, via oxocarbenium-ion-like transition states. Such a mechanism requires a minimum of two catalytic carboxylates. One, the acid/base, functions initially as a Brønsted acid, protonating the glycosidic bond to assist leaving-group departure, whilst the second functions as a nucleophile and forms a covalent glycosyl-enzyme intermediate of inverted configuration. The catalytic intermediate is then hydrolysed by a water molecule, activated by deprotonation by the acid/base, which now functions as a Brønsted base. For family 11 xylanases, the acid/base is known and the nucleophile has been trapped using 2-fluoro (2-F)-substituted sugars [21]. The catalytic mechanism is believed to feature 'pK_a cycling' that regulates the pK_a of the acid/base along the reaction trajectory as has been demonstrated on the *Bacillus circulans* enzyme using ^{13}C nuclear magnetic resonance (NMR) spectroscopy [22].

For the *B. agaradhaerens* Xyl11, presented here, the two catalytic carboxylates are Glu94 (nucleophile) and Glu184 (acid/base), respectively. In order to trap the covalent xylobiosyl-enzyme intermediate, crystals of Xyl11 were bathed overnight in a solution containing 2′,4′-dinitrophenyl 2-deoxy-2-fluoro-β-xylobioside, which permits trapping and accumulation of the covalent 2-F-xylobiosyl-enzyme intermediate [21]. Data were collected from a single crystal at 120 K, to 2.00 Å resolution. They have an R_{merge} of 6.8% and are 97.6% complete (Table 1). As a result of the soaking, the cell dimensions are slightly perturbed from the native values to a = 71.91 b = 74.83 c = 78.35 Å. Observed electron density for the 2-F-xylobiosyl-enzyme intermediate is shown in Figure 2. The two catalytic glutamate residues are located on opposite sides of the active-site cleft. The C1 atom of the –1 subsite sugar (nomenclature as in [23]) is covalently linked to the OE2 atom of the nucleophile, Glu94, via an α-ester bond, as expected. Surprisingly, the –1 subsite sugar is found in a $^{2,5}B$ conformation, in marked contrast to all structures of retaining glycoside hydrolase covalent intermediates from other families. The implications of this observation are discussed below.

The 2-F-xylobiosyl-enzyme structure permits a description of the protein–ligand interactions in both the –2 and –1 subsites (Figure 3). The –2 subsite sugar is in the $^4C_1$ chair conformation, bound via aromatic stacking 'below' the aromatic ring plane of Trp19, a residue conserved throughout family 11 and family 12 xylanases. There are direct hydrogen bonds from the O3 and O2

Figure 2



Observed electron density for the covalent 2-F-xylobiosyl-enzyme intermediate. The electron-density map shown is a maximum-likelihood-weighted $2F_{obs}-F_c$ synthesis at a contour level of 0.4 electrons/Å^3. The nucleophile, Glu94, the catalytic acid, Glu184, Leu47 and a putative attacking water molecule are shown. The −1 subsite xylose ring adopts a $^{2,5}B$ conformation.

hydroxyl groups to Arg49 and from the O3 hydroxyl group to Glu17. The O2 hydroxyl group also makes a hydrogen bond to the Tyr85 hydroxyl group. The O4 atom interacts only with solvent water. The −1 subsite sugar is in the $^{2,5}B$ conformation with its anomeric carbon in a covalent α-ester linkage to the nucleophile Glu94 (OE2). Glu94 also makes a hydrogen bond from OE1 to the amide group of Gln143; interactions between the carbonyl oxygen of the nucleophile and a sidechain amide are important features of other unrelated glycoside hydrolases (for examples see [7,9,10]). The fluorine substituent at C-2 interacts with the NE1 from Arg129 and is 2.8 Å from the carbonyl oxygen (OE1) of the nucleophile. By analogy with other systems it seems likely that the OE1···O2 interaction plays an important role in transition-state stabilisation [10,24]. The O3 hydroxyl group interacts both with the mainchain carbonyl group of Pro133 and with the NH_2 group of Arg129. The catalytic acid/base Glu184 sits *syn* to the pyranoside endocyclic O5–C1 bond (as defined by Vasella and coworkers [25,26]) where it interacts with a solvent water molecule poised for nucleophilic attack at the anomeric centre. This water molecule lies 3.4 Å from the C1 of the sugar and the OE2-(Glu94)–C1–water angle is 168°, consistent with an 'in-line' nucleophilic attack. Our interpretation of the potential catalytic water is different to that described for the *B. circulans* enzyme [27]. In the case of the *B. agaradhaerens* enzyme we would predict that the 'catalytic' water interacts with Asn45 and the OE2 of the acid/base (Figure 3), and not with Tyr96 and OE1 of the acid/base as reported for the *B. circulans* enzyme. The two carboxylate oxygens of Glu184, OE2 and OE1, interact additionally with Asn45 and Tyr96, respectively. Close van der Waals' contacts are observed with Leu47. These might play a role both in maintaining the $^{2,5}B$ conformation and in the specificity for *xylo*-configured substrates.

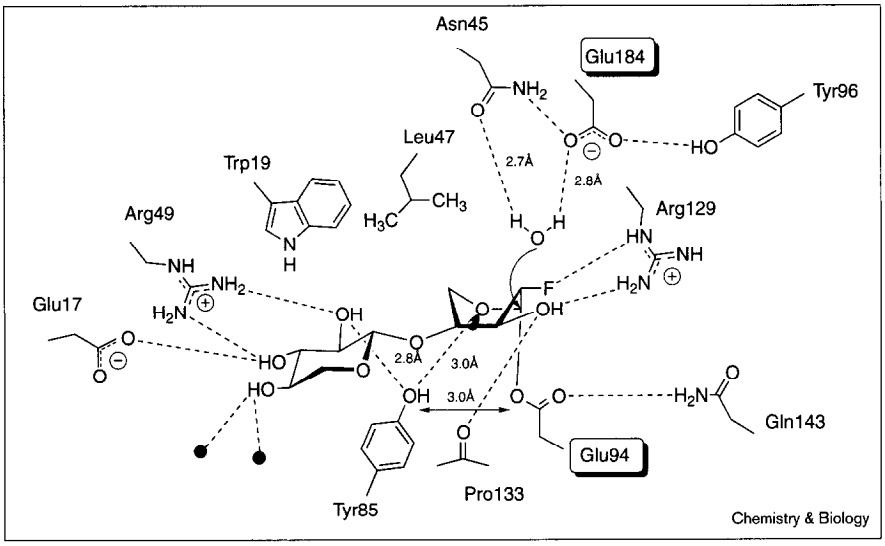
The 2-F-xylobiosyl-enzyme intermediate structure is very similar to the 'native' structure. The overall rms deviation between the native and the complex structure is 0.24 Å for all Cα atoms. The only significant difference occurs in the active site where Tyr186, which showed two distinct conformations in the native structure, adopts just one of these conformations in the complex. The second conformation in the native structure occupies what would be expected to be the position of the +1 subsite sugar in the Michaelis complex; indeed the hydroxyl group adopts a similar position to the putative attacking water observed in the intermediate.

Implications of the $^{2,5}B$ conformation for the catalytic mechanism

Structures of the covalent glycosyl-enzyme intermediates from other β-retaining glycoside hydrolase families determined, to date, have revealed a $^4C_1$ chair conformation for the −1 subsite sugar. This includes both *xylo*- and *gluco*-configured sugars [6–11]. In contrast, the −1 subsite sugar ring family 11 xylanases adopt an unusual $^{2,5}B$ conformation (a similar conformation to the Xyl11 structure described here has just been described for the equivalent covalent glycosyl-enzyme intermediate for the family 11 xylanase from *B. circulans* [27]). The formation and breakdown of the glycosyl-enzyme intermediate occurs via oxocarbenium-ion-like transition states. There is an absolute stereochemical requirement in such transition states for planarity of C5, O5, C1 and C2 as a result of the double-bond character across the O5–C1 bond [28]. The $^4C_1$ conformation does

Figure 3

Schematic representation of the interactions of Xyl11 with its 2-F-xylobiosyl-enzyme intermediate. Distances from the Tyr85 hydroxyl group are given and are as observed recently on the equivalent complex of the *B. circulans* enzyme [27].



not satisfy this criterion and ring distortion is required to fulfil the stereochemical obligations for glycosyl transfer. By contrast, the $^{2,5}B$ conformation found in the family 11 xylobiosyl-enzyme obeys these stereochemical requirements (Figure 4), which might assist the formation and breakdown of this covalent intermediate.

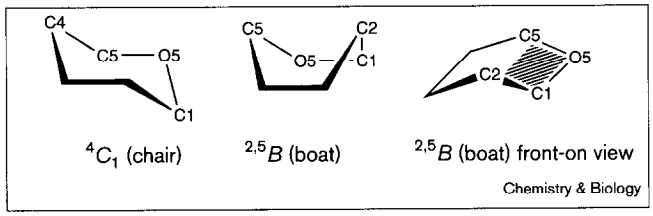
Two structural aspects could favour formation of the $^{2,5}B$ conformation. Tyr85 appears to make a geometrically sound hydrogen bond to one of the correctly orientated sp^3 lone pairs of the ring oxygen O5 (Figure 5). The ideal geometry for this interaction occurs only with the ring in a boat conformation — the lone pair of O5 points away from Tyr85 if the ring is modelled in a $^4C_1$ chair conformation. Tyr85 also 'interacts' with the nucleophilic oxygen OE2 of the Glu94 nucleophile (they are separated by 3.0 Å; Figure 3). For both of these interactions the tyrosine hydroxyl group must function as a hydrogen-bond donor, which is not simultaneously possible. The close proximity of the hydroxyl group of Tyr85 to the OE2 of the nucleophile might permit 'shuttling' of the hydrogen bond between O5 and OE2 during catalysis. Indeed, hydrogen-bond donation to O5, as indicated here, would presumably be detrimental at the oxocarbenium-ion-like transition state where positive-charge develops across the O5–C1 bond. It might be that, at this point, the hydrogen bond is shifted to the OE2 atom of the nucleophile to assist in cleavage of the C1–OE2 bond. The proposal that the $^{2,5}B$ conformation could exist for an oxocarbenium-ion-like transition state is not without precedent. β-Deuterium secondary kinetic-isotope effects measured on the unrelated yeast α-glucosidase system indicated that the pyranoside-ring C2–H bond lay perpendicular to the C1–O bond in the transition state. This could be rationalised only by a model in which a $^{2,5}B$ conformation was adopted in the bond-breaking transition state [29].

A boat conformation for the intermediate is unlikely to result merely from the easier ring interconversion of xylose compared to glucose. In the unrelated family 10 xylanases both *gluco-* and *xylo*-configured intermediates adopt the $^4C_1$ chair conformation [8–10]. The orientation of the distal –2 subsite sugar in family 10 xylanases is, however, markedly different to that observed in the family 11 enzymes. It is possible that stacking interactions and the environment of the sugar in the –2 subsite contribute to the –1 subsite conformation. Optimal stacking interactions occur only if the sugar in the –1 subsite is in the boat conformation. Furthermore, in Xyl11 Tyr85, described above, also interacts with O2 of the –2 subsite, presumably accepting a hydrogen bond from the sugar hydroxyl (Figures 4,5). This interaction would be lost were the –1 sugar to adopt a $^4C_1$ conformation.

Substrate specificity in glycoside hydrolase clan GH-C

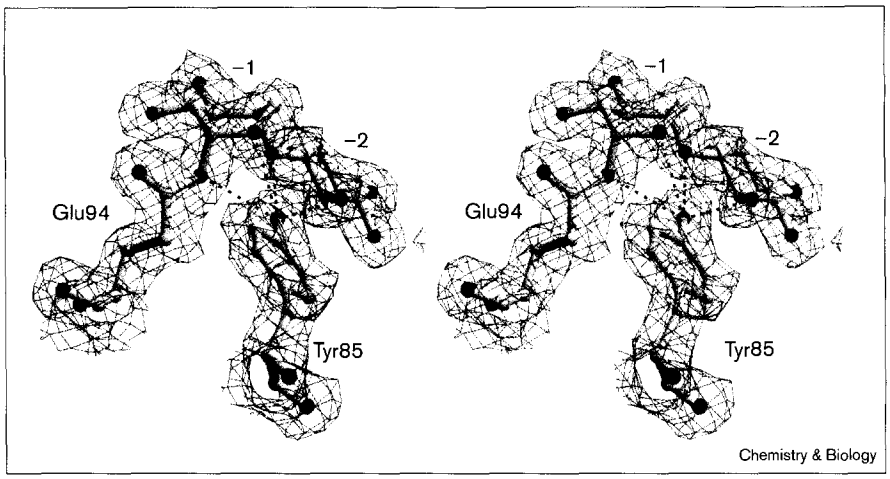
Family 11 xylanases, such as the *B. agaradhaerens* Xyl11, were predicted, on the basis of both conventional

Figure 4



Schematic drawing of the $^4C_1$ (chair) and $^{2,5}B$ (boat) conformations for the pyranoside ring. The $^{2,5}B$ conformation places C5, O5, C1 and C2 in-plane, fulfiling the stereochemical requirements demanded of an oxocarbenium ion, as can be seen from the 'front-on' view of the $^{2,5}B$ (boat).

Figure 5

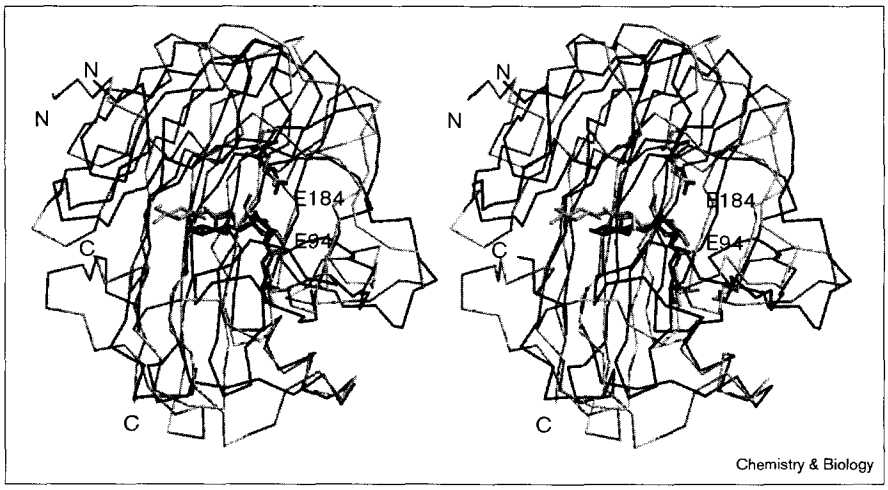


'Side-on' view of the covalent intermediate indicating the function of Tyr85. The electron-density map shown is a maximum-likelihood-weighted $2F_{obs}-F_c$ synthesis at a contour level of 0.4 electrons/Å^3.

sequence comparisons [30] and hydrophobic cluster analysis [31], to share a common three-dimensional fold and catalytic machinery with the celluases from family 12. This, in turn, led to their grouping as a clan of related glycoside hydrolases named clan GH-C [1]. These predictions were borne out by the three-dimensional structure determination of the cellulase CelB from *Streptomyces lividans* from family 12 [13]. Although the two enzyme families display low sequence identity (frequently below 20%), their three-dimensional structures share striking similarity and the active-site architecture is conserved to a high degree (Figure 6). Recently, the structure of the covalent cellotriosyl-enzyme intermediate for CelB has been determined at atomic resolution [11], which permitted description of the protein–ligand interactions through the –3 to –1 subsites, and revealed a $^4C_1$ conformation for the –1 subsite sugar in the glycosyl-enzyme intermediate. Comparison of the equivalent covalent intermediate complexes for the two families (Figure 7) allows us to examine the molecular basis for substrate specificity.

Family 11 xylanases are highly specific towards xylan. In contrast, the preferred substrate for family 12 cellulases is cellulose although they also retain significant activity on xylan. Xylan is a β-1,4-linked polymer of D-xylose and specificity presumably arises from discrimination against the C6–O6 primary substituent found on glucose. Further selection could also utilise the linear nature of a cello-oligosaccharide chain compared to the more helical structure of xylan [32]. In both families 11 and 12 the catalytic acid/base is structurally equivalent and is in a position for 'lateral' *syn*-protonation of the glycosidic bond. In family 11, the acid/base catalytic residue is packed between Leu47 and Tyr81 and hydrogen bonds to Asn45 and

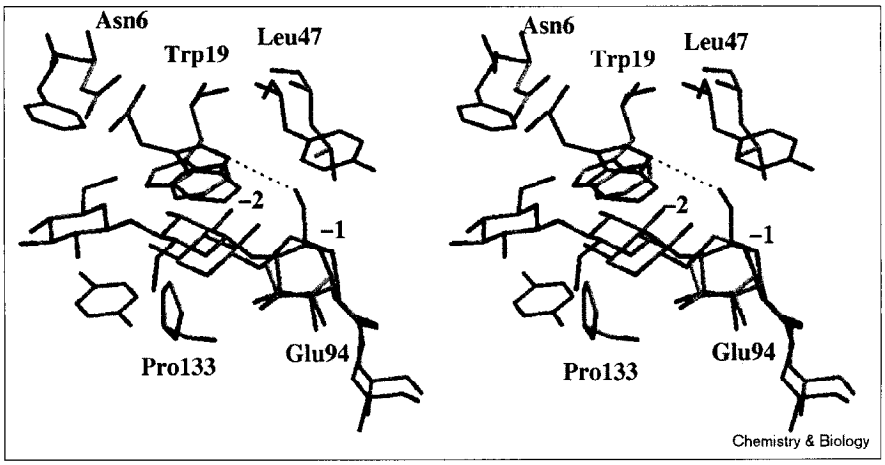
Figure 6



Divergent stereo overlap of the family 11 xylanase Xyl11 (this work, black) and the family 12 cellulase CelB (grey) [11,13]. The glycosyl-enzyme intermediate structures are shown.

Figure 7

Close-up of the –3, –2 and –1 subsites for the Xyl11 (yellow) and CelB (cyan) covalent intermediate complexes. The role of Leu47 and Pro133 in preventing the binding of glucose-based substrates to the family 11 structure can be seen. The –3 subsite of the family 12 enzyme possesses both tyrosine and phenylalanine residues (shown) that have no direct equivalent in family 11 although Asn6 could provide hydrogen bonding to a –3 subsite sugar.



Tyr96. This results in a perpendicular orientation of the carboxylate group with respect to the –1 subsite ring plane. The acid/base in family 12, however, is stabilised by just a single hydrogen bond (to Asn100), which leaves the carboxy plane parallel to the pyranose ring plane. This orientation allows Glu203 to interact both with the glycosidic oxygen and with the C6 hydroxyl group of glucose-derived substrates. In family 12, this C6 hydroxyl group hydrogen bonds to the NE1 of Trp24 which is involved in the –2 subsite stacking interaction. A structurally equivalent orientation of acid···C6-OH(–1)···Trp(–2) is also seen in the (*syn*-protonating) cellulases from family 7, such as the *Fusarium oxysporum* endoglucanase Cel7B [33,34].

In family 11 xylanases there is an invariant hydrophobic residue (leucine or valine) adjacent to the acid/base catalytic residue. This prevents accommodation of a glucose moiety in the –1 subsite on steric grounds alone. Moreover, this hydrophobic residue forces the invariant –2 subsite tryptophan (Trp24 in CelB2, Trp19 in Xyl11) into a position where its NE1 can no longer interact with a C6 hydroxyl group in the –1 subsite. This appears to confirm predictions made based on the family 12 structure alone [11] that in family 11 the hydrophobic group neighbouring the Brønsted acid/base plays a multifunctional role. It constrains the carboxylate of the acid/base to a vertical position, preventing interaction with a C6 hydroxyl group in the –1 subsite, and it favours binding of the more hydrophobic xylose-based polymer and precludes binding of cellulose substrates through a combination of steric hindrance and the depletion of hydrogen-bonding partners for the C6 hydroxyl group. Similar considerations apply with regard to substrate specificity in xylanase family 10. The *Cellulomonas fimi* enzyme CEX is capable of hydrolysing both cellulose and xylan but has a preference for xylose-based substrates. A tryptophan residue 'above' the –1 subsite would appear to disfavour binding of glucose-based substrates, but is sufficiently flexible to be displaced upon binding of cellobiose at a relatively low energetic cost. Cellobiose-based substrates are hydrolysed about 50–100-fold slower than xylobiose-based ones [9].

For the additional subsites, substrate discrimination is less clear. In the –2 subsite at the nonreducing end of the cleft, the 'thumb' of family 11 xylanases might play a further role in substrate specificity. This structural feature exhibits a high degree of sequence conservation. Pro133 is invariant, overlooks the cleft between subsites –1 and –2 and would seem to disfavour binding of a C5-hydroxymethyl substituent in the –2 subsite. The *S. lividans* CelB possesses a third nonreducing end binding site, designated –3, which possesses both Tyr66 and Phe8 which stack either side of the pyranoside ring. No equivalent interactions are possible in the family 11 xylanase, although Asn6 in this position could perhaps make hydrogen-bonding interactions to a putative sugar in the –3 subsite. Speculation on the substrate preferences for the aglycon sites (+1 to +3) is not possible based upon the covalent glycosyl-enzyme intermediate complexes.

Despite the explosion of three-dimensional structures for glycoside hydrolases in recent years (reviewed in [3]) substrate specificity amongst the related enzymes in not well understood. Although clan GH-C is a simple one that links a single family of xylanases (11) with a family containing cellulases (12), much larger clans exist that display an enormous range of specificities grafted onto a common fold and catalytic geometry [35]. The *B. agaradhaerens* Xyl11 structure, presented here, provides insights not only into substrate specificity, but also into how similar active sites can subtly change their geometry resulting in changed sugar-ring conformations and hence catalytic strategies. The enzyme appears to have evolved to proceed via a $^{2,5}B$ intermediate. Not only does this conformation conform to the

stereochemical constraints on the oxocarbenium-ion-like transition state, but it may also hinder the binding of glucose-based polymers.

Significance

Glycoside hydrolases are ubiquitous enzymes that utilise general acid/base chemistry to facilitate cleavage of glycosidic bonds. Xylanases are glycoside hydrolases that act on xylan which have found numerous applications in industrial processes and have great potential for replacing currently polluting technologies. The results described in this paper reveal the structure of the *Bacillus agaradhaerens* xylanase from glycoside hydrolase family 11 in both native and covalent-intermediate forms. The covalent glycosyl-enzyme intermediate, trapped using a 2-deoxy-2-fluoro β-D-xylobioside, displays an unusual $^{2,5}B$ conformation for the pyranoside ring at the catalytic centre. Such a ring conformation is different to the $^{4}C_1$ chair conformation seen for trapped intermediates on other families of glycoside hydrolases. This is important because such a conformation, unlike the $^{4}C_1$, obeys the stereochemical requirements for the oxocarbenium-ion transition state (C5, O1, C1 and C2 co-planar). This confirms that enzymatic glycoside hydrolysis could utilise different transition-state and intermediate structures, which opens up scope for the selective design of transition-state mimics.

Family 11 xylanases, which act on xylose-derived polysaccharides, form a 'clan' of related structures together with enzymes from family 12, which act on glucose-based polymers. Comparison of the respective intermediate complexes for enzymes from families 11 and 12 reveals the likely basis for substrate specificity in this clan. An invariant hydrophobic residue (leucine or valine) forms a strong van der Waals' interaction with the $^{2,5}B$-conformed pyranoside ring in the −1 subsite. This residue may also hinder the binding of glucose-based polymers for steric reasons, as well as preventing hydrogen-bonding interactions that occur in the related family 12 cellulase structures. Together, these results provide valuable insight into catalysis and substrate recognition in glycoside hydrolase clan GH-C.

Materials and methods

Cloning and expression of the B. agaradhaerans *family 11 xylanase gene*

Cells of *B. agaradhaerens* were harvested and genomic DNA isolated using standard procedures [36]. Genomic DNA was partially digested with restriction enzyme *Sau*3A, and size fractionated by electrophoresis on a 0.7% agarose gel. Fragments between 2 and 10 kb in size were isolated using electrophoresis onto DEAE-cellulose paper. Isolated DNA fragments were ligated to *Bam*HI-digested pSJ1678 plasmid DNA [37] and the ligation mixture was used to transform *E. coli* SJ2. The resulting gene library was screened on LB agar plates containing 0.5% AZCL-xylan (Megazyme) and 9 μg ml^{-1} chloramphenicol and incubated overnight at 37°C. A xylanase-positive colony, containing the *xyl11* gene in a 3.8 kb insert, was selected. The family 11 xylanase gene from *B. agaradhaerens* was cloned and expressed in a *Bacillus licheniformis* secretion/expression system [38]. The extracellular xylanase was purified using cation-exchange chromatography using S-Sepharose, at 278 K in 20 mM sodium acetate buffer at pH 5. The enzyme (pI 8.8) was eluted using a gradient of 10 l 0.5 M NaCl. The xylanase pool was concentrated and desalted on a Sephadex 75 column in 50 mM sodium acetate pH 6.0. The purified enzyme has an extinction coefficient of 56590 M^{-1}cm^{-1} and a molecular weight (M_r) of 22.5 kDa.

Crystallisation, data collection and processing

Two different crystallisation conditions were used for Xyl11 giving rise to the monoclinic and orthorhombic crystal forms, respectively. Xyl11 protein, prepared as above, was washed and concentrated to 10 mg ml^{-1} on Filtron 10 K membranes. Crystals of the monoclinic form were grown using the hanging-drop method from 0.1 M citrate buffer pH 5.06, 20% isopropanol with 20% PEG 4 K as precipitant. Crystals of the orthorhombic form were also obtained by hanging-drop vapour diffusion. Hanging drops (3 μl), consisting of 2 μl protein (10 mg ml^{-1} in 100 mM sodium acetate pH 6.0) + 1 μl reservoir solution (100 mM MES pH 6.5, 30% ammonium sulphate), were equilibrated at room temperature over 1 ml of reservoir solution for 7–14 days.

Preliminary data for Xyl11 were collected from a single crystal of the monoclinic form at the EMBL Hamburg Outstation beamline X-31. Data collection was performed at room temperature from a single crystal mounted in a glass capillary. Crystals of the orthorhombic form of Xyl11 were transferred, stepwise, to a cryoprotectant stablising solution consisting of 30% ammonium sulphate, 100 mM MES pH 6.5 and 10, 20 or 30% (v/v) glycerol. Crystals were mounted in a rayon fibre loop and flash-frozen in a boiling N_2 stream at 120 K. X-ray data were measured using an RAXIS II image-plate, with CuKα radiation ($\lambda = 1.5418$ Å) from a rotating-anode generator, with long-focusing mirror optics, operating at 50 kV and 100 mA. In order to obtain the 2-fluoroxylobiosyl-enzyme intermediate, crystals were harvested into the stabilising solution, and solid 2′,4′-dinitrophenyl 2-deoxy-2-fluoro-β-xylobioside added. The low solubility of this compound precludes conventional soaking strategies. Following overnight reaction, a single crystal was prepared for data collection at 120 K as described above. All data were processed and reduced with the HKL suite [39] and all subsequent computing used the CCP4 suite [40] unless otherwise stated.

Structure solution and refinement

The structure from the monoclinic crystal form of Xyl11 was solved by molecular replacement using the *Bacillus pumilus* family 11 xylanase (G.S., Z.D., V. Lamzin, and K.S.W., unpublished observations and [41]) as a search model with the program AMoRe [42,43]. Two significant solutions were identified and manual corrections to the model carried out using QUANTA (Molecular Simulations Inc. San Diego, USA). A small percentage (5%) of the reflections were set aside for cross-validation analysis [44] and the structure was refined using maximum-likelihood methods (REFMAC [14]). A partially refined model from this crystal form was used to solve the structure of the orthorhombic form of Xyl11 by molecular replacement again with AMoRe. Data in the resolution range 20 to 2.5 Å were used together with an outer radius of Patterson intergration of 18 Å. The model was refined using all observed data (20 to 1.78 Å). A two Gaussian bulk solvent correction was applied according to Babinet's principle [14]. Tight noncrystallographic symmetry restraints were imposed initially and then slowly released as judged appropriate from the behaviour of the free R subset of reflections. Indeed, the model was finally refined without imposition of NCS restraints. The behaviour of R_{free} also indicated that both the contribution of scattering from 'riding' hydrogen atoms and an anisotropic model for the atomic displacement parameters of the heavy atoms was appropriate. The isotropic refinement converged with values for R and R_{free} of 14.8 and 18.5, respectively. Anisotropic refinement within REFMAC, applying 'similarity', 'sphericity' and 'rigid-bond' restraints [45], resulted in drops of 3% for the working R value and 0.9% for R_{free}. Water molecules were inserted automatically using ARP [46] and

then checked manually prior to coordinate deposition. All water molecules display B values less than 60 Å^2.

The complex structure was refined essentially as described for the native structure above, with the exception that anisotropic modelling of the atomic displacement parameters was not deemed appropriate by the behaviour of R_{free}. A stereochemical dictionary for the ligand was prepared from a model structure generated in QUANTA. The ligand was refined without conformational torsion-angle restraints in order to prevent bias towards a particular ring conformation.

Accession numbers

Coordinates and observed structure-factor amplitudes for the structures described in this paper have been deposited with the PDB [47] with codes 1QH7 and 1QH6.

Acknowledgements

This work was funded, in part, by the Biotechnology and Biological Sciences Research Council, the European Union (EC-BIOTECH 003R04292), Novo-Nordisk A/S and the University of York. G.J.D. is a Royal Society University Research Fellow. The authors would like to thank Steve Withers for discussing unpublished results and provision of 2′4′ dinitrophenyl 2-fluoro-2-deoxy β-D-xylobioside and the staff of the EMBL Hamburg outstation for the provision of data collection facilities.

References

1. Henrissat, B. & Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **316**, 695-696.
2. Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem. J.* **280**, 309-316.
3. Davies, G., Sinnott, M.L. & Withers, S.G. (1997). Glycosyl transfer. In *Comprehensive Biological Catalysis* Vol. 1. (Sinnott, M.L. ed.) pp. 119-209. Academic Press, London.
4. Koshland, D.E. (1953). Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev.* **28**, 416-436.
5. Withers, S.G., Street, I.P., Bird, P. & Dolphin, D.H. (1987). 2-deoxy-2-fluoroglucosides: a novel class of mechanism-based glucosidase inhibitors. *J. Am. Chem. Soc.* **109**, 7530-7531.
6. Burmeister, W.P., Cottaz, S., Driguez, H., Palmieri, S. & Henrissat, B. (1997). The crystal structures of *Sinapis alba* myrosinase and of a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinary of an *S*-glycosidase. *Structure* **5**, 663-675.
7. Davies, G.J., *et al.*, & Withers, S.G.(1998). Snapshots along an enzymatic reaction coordinate: analysis of a retaining β-glycoside hydrolase. *Biochemistry* **37**, 11707-11713.
8. White, A., Tull, D., Johns, K., Withers, S.G. & Rose, D.R. (1996). Crystallographic observation of a covalent catalytic intermediate in a β-glycosidase. *Nat. Struct. Biol.* **3**, 149-154.
9. Notenboom, V., Birsan, C., Warren, R.A.J., Withers, S.G. & Rose, D.R. (1998). Exploring the cellulose/xylan specificty of the β-1,4-glycanase Cex from *Cellulomonas fimi* through crystallography and mutation. *Biochemistry* **37**, 4751-4758.
10. Notenboom, V., *et al.*, & Withers, S.G.(1998). Insights into transition state stabilisation of the β-1,4 glycosidase Cex by covalent intermediate accumulation in active site mutants. *Nat. Struct. Biol.* **5**, 812-818.
11. Sulzenbacher, G., *et al.*, & Davies, G. (1999). The crystal structure of the 2-fluorocellotriosyl-complex of the *Streptomyces lividans* endoglucanase, CelB2, at 1.2 Å resolution. *Biochemistry* **38**, 4826-4833.
12. Visser, J., Beldman, G., Someran, M.A.K.-v. & Voragen, A.G.J. (1992). Xylans and xylanases. In *Progress in Biotechnology* Vol. 7. Elsevier, Amsterdam.
13. Sulzenbacher, G., Shareck, F., Morosoli, R., Dupont, C. & Davies, G.J. (1997). The *Streptomyces lividans* family 12 endoglucanase: construction of the catalytic core, expression and X-ray structure at 1.7 Å resolution. *Biochemistry* **36**, 16032-16039.
14. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum likelihood method. *Acta Crystallogr. D* **53**, 240-255.
15. Matthews, B.W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497.
16. Ramachandran, G.N., Ramakrishnan, C. & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95-99.
17. Laskowski, R.A., McArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 282-291.
18. Campbell, R.L. *et al.* (1993). A comparison of the structures of the 20 kDa xylanases from *Trichoderma harzianum* and *Bacillus circulans*. In *Proceedings of the Second TRICEL Symposium on* Trichoderma reesei *Cellulases and Other Hydrolases* Vol. 8. (Suominen, P. & Reinikainen, T. eds). pp. 63-72, Foundation for Biotechnical and Industrial Fermentation Research, Espoo.
19. Törrönen, A., Harkki, A. & Rouvinen, J. (1994). Three-dimensional structure of endo-1,4-β-xylanase II from *Trichoderma reesei*: two conformational states in the active site. *EMBO J.* **13**, 2493-2501.
20. Gruber, K., *et al.*, & Kratky, C. (1998). Thermophilic xylanase from *Thermomyces langinosus*: high resolution X-ray structure and modelling studies. *Biochemistry* **37**, 13475-13485.
21. Miao, S., Ziser, L., Aebersold, R. & Withers, S.G. (1994). Identification of glutamic acid 78 as the active site nucleophile in *Bacillus subtilis* xylanase using electrospray tandem mass spectometry. *Biochemistry* **33**, 7027-7032.
22. McIntosh, L.P., *et al.*, & Withers, S.G. (1996). The pK_a of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a ^{13}C-NMR study of *Bacillus circulans* xylanase. *Biochemistry* **35**, 9958 9966.
23. Davies, G.J., Wilson, K.S. & Henrissat, B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557-559.
24. Namchuk, M.N. & Withers, S.G. (1995). Mechanism of *Agrobacterium* β-glucosidase: kinetic analysis of the role of noncovalent enzyme/substrate interactions. *Biochemistry* **34**, 16194-16202.
25. Heightman, T.D. & Vasella, A.T. (1999). Recent insights into inhibition, structure and mechanism of configuration-retaining glycosidases. Angew. *Chem. Int Ed.* **38**, 750-770.
26. Varrot, A., Schulein, M., Pipelier, M., Vasella, A. & Davies, G.J. (1999). Lateral protonation of a glycosidase inhibitor: structure of the *Bacillus agaradhaerens* Cel5A in complex with a *cellobio*-derived imidazole at 0.97 Å resolution. *J. Am. Chem. Soc.* **121**, 2621-2622.
27. Sidhu, G. *et al.*, & Breyer, G.D. (1999). Sugar ring distortion in the glycosyl-enzyme intermediate of a family G/11 xylanase. *Biochemistry* **38**, 5346-5354.
28. Sinnott, M.L. (1990). Catalytic mechanisms of enzymic glycosyl transfer. *Chem. Rev.* **90**, 1171-1202.
29. Hosie, L. & Sinnott, M.L. (1985). Effects of deuterium substitution alpha and beta to the reaction centre, 18O substitution in the leaving group, and aglycone acidity on hydrolyses of aryl glucosides and glucosyl pyridinium ions by yeast alpha-glucosidase. A probable failure of the antiperiplanar-lone-pair hypothesis in glycosidase catalysis. *Biochem. J.* **226**, 437-446.
30. Saarilahti, H.T., Henrissat, B. & Pavla, E.T. (1990). CelS: a novel endoglucanase identified from *Erwinia carotovora* subsp. *carotovora* *Gene* **90**, 9-14.
31. 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.
32. Atkins, E.D.T. (1992). Three-dimensional structure, interactions and properties of xylans. In *Xylanas and Xylanases* Vol. 7. (Visser, J., Beldman, G., Someren, M.A.K.v. & Voragen, A.G.J., eds) pp. 39-50. Elsevier, Amsterdam.
33. Sulzenbacher, G., Driguez, H., Henrissat, B., Schülein, M. & Davies, G.J. (1996). Structure of the *Fusarium oxysporum* endoglucanase I with a nonhydrolysable substrate analogue: substrate distortion gives rise to a pseudo-axial orientation for the leaving group. *Biochemistry* **35**, 15280-15287.
34. Sulzenbacher, G., Schülein, M. & Davies, G.J. (1997). Structures of the endoglucanase I from Fusarium oxysporum: native, cellobiose and 3,4-epoxybutyl β-D-cellobioside inhibited forms at 2.3 Å resolution. *Biochemistry* **36**, 5902-5911.
35. Henrissat, B., *et al.*, & Davies, G. (1995). Conserved catalytic machinary and the prediction of a common fold for several families of glycosyl hydrolases. *Proc. Natl Acad. Sci. USA* **92**, 7090-7094.
36. Pitcher, D.G., Saunders, N.A. & Owen, R.J. (1989). Rapid extraction of bacterial genomic DNA with guanidinium thiocyanate. *Lett. Appl. Microbiol.* **8**, 151-156.
37. Jørgensen, S. (1994). An amolytic enzyme, in particular a pyrococcus alpha amylase. World patent WO 94/19454, 1 September 1994.
38. Jørgensen, S., Jørgensen, P.L. & Diderichsen, B.K. (1991). Stable integration of DNA in bacterial genomes. World patent WO 91/09129, 27 June 1991.

39. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology: Macromolecular Crystallography, part A* Vol. 276. (C W Carter, J. & Sweet, R.M., eds) pp. 307-326. Academic Press, London & New York.
40. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760-763.
41. Sulzenbacher, G.D. (1999). Structural studies of cellulases and xylanases [PhD Thesis]. University of York, York, UK.
42. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157-163.
43. Navaza, J. & Saludijan, P. (1997). AMoRe: an automated molecular replacement program package. *Methods Enzymol.* **276**, 581-594.
44. Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.
45. Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S. & Dodson, E.J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D* **55**, 247-255.
46. Lamzin, V.S. & Wilson, K.S. (1993). Automated refinement of protein models. *Acta Crystallogr. D* **49**, 129-147.
47. Bernstein, F.C., *et al.*, & Tasumi, M. (1977). The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535-542.
48. Esnouf, R.M. (1997). An extensively modified version of MolScript that includes greatly enhanced colouring capabilities. *J. Mol. Graphics* **15**, 133-138.

Because *Chemistry & Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cmb – for further information, see the explanation on the contents pages.