doi:10.1016/j.jmb.2007.05.082 J. Mol. Biol. (2007) 371, 1204–1218

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Crystal Structures of *Paenibacillus polymyxa* β-Glucosidase B Complexes Reveal the Molecular Basis of Substrate Specificity and Give New Insights into the Catalytic Machinery of Family I Glycosidases

Pablo Isorna[1], Julio Polaina[2], Lorena Latorre-García[2] Francisco Javier Cañada[3], Beatriz González[1] and Julia Sanz-Aparicio[1]*

[1]*Grupo de Cristalografía Macromolecular y Biología Estructural, Instituto de Química-Física "Rocasolano" CSIC, Serrano 119 28006-Madrid, Spain*

[2]*Instituto de Agroquímica y Tecnología de Alimentos CSIC, Apdo.73, 46100-Burjasot Valencia, Spain*

[3]*Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040-Madrid, Spain*

**Corresponding author*

Bacteria species involved in degradation of cellulosic substrates produce a variety of enzymes for processing related compounds along the hydrolytic pathway. *Paenibacillus polymyxa* encodes two homologous β-glucosidases, BglA and BglB, presenting different quaternary structures and substrate specificities. We previously reported the 3D-structure of BglA, which is highly specific against cellobiose. Here, we present structural analysis of BglB, a monomeric enzyme that acts as an exo-β-glucosidase hydrolyzing cellobiose and cellodextrins of higher degree of polymerization. The crystal structure of BglB shows that several polar residues narrow the active site pocket and contour additional subsites. The structure of the BglB–cellotetraose complex confirms these subsites, revealing the substrate-binding mode, and shows the oligosaccharide-enzyme recognition pattern in detail. Comparison between BglA and BglB crystal structures suggests that oligomerization in BglA can assist in fine-tuning the specificity of the active centre by modulating the loops surrounding the cavity. We have solved the crystal structure of BglB with bound thiocellobiose, a competitive inhibitor, which together with the BglB–cellotetraose complex delineate the general features of the aglycon site. The detailed characterization of the atomic interactions at the aglycon site show a recognition pattern common to all bacterial β-glucosidases, and presents some differences with the aglycon site in plant β-glycosidases essentially by means of a different orientation of the basal Trp. The crystal structures of of BglB with a covalently bound inhibitor (derived from 2-fluoroglucoside) and glucose (produced by hydrolysis of the substrate in the crystal), provide additional pictures of the binding events and the intermediates formed during the reaction. Altogether, this information can assist in the understanding of subtle differences of the enzyme mechanism and substrate recognition within this family of enzymes, and consequently it can help in the development of new enzymes with improved activity or specificity.

Keywords: β-glucosidase complexes; 3D structure; enzymatic mechanism

Abbreviations used: GH-1, family 1 glycoside hydrolases; ZmGlu1, β-glucosidase 1 from *Zea mays*; SbDhr, dhurrinase 1 from *Sorghum bicolour*; BglA, β-glucosidase A from *Paenibacillus polymyxa*; BglB, β-glucosidase B from *Paenibacillus polymyxa*; CAZy, carbohydrate active enzymes; 2FGlc, 2-deoxy-2-fluoroglycosides; 2F-DNPG, 2′,4′-dinitrophenyl 2-deoxy-2-fluoroglucoside; PNPG, *p*-nitrophenyl-β-D-glucopyranoside; Tcb, thiocellobiose; CTT, cellotetraose; dhurrin, (S)-4hydroxymandelonitrile-β-D-glucopyranoside; DimboaGlc, 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxacin-3(4H)-one-2-β-glucoside; SA, simulated annealing.

E-mail address of the corresponding author: xjulia@iqfr.csic.es

Introduction

Family 1 glycoside hydrolases (GH-1) include enzymes with different activities amply distributed among all sort of living organisms, which have in common the ability to hydrolyse β-glycosidic linkages of disaccharides, oligosaccharides or conjugated saccharides.[1] Among the members of this family are bacterial and fungal cellobiases, that play an important role in cellulolysis; phospho-β-galactosidases from lactic acid bacteria; plant enzymes involved in defence mechanisms against grazing herbivores; and intestinal lactase-phlorizin hydrolase from mammals.[2,3] The number of GH-1 enzymes whose three-dimensional structure has been solved has increased sharply in the last decade. Among the enzymes of known structure there are several from archaea, *Sulfolobus solfataricus*,[4] *Thermosphaera aggregans*,[5] *Pyrococcus horikoshii*;[6] from bacteria, *Lactococcus lactis*,[7] *Paenibacillus polymyxa*,[8] *Bacillus circulans*,[9] *Thermus nonproteolyticus*,[10] *Thermotoga maritima*;[11] from plants: *Trifolium repens*,[12] *Sinapsis alba*,[13] *Zea mays*,[14,15] *Sorghum bicolor*,[16] *Triticum aestivum*;[17] and the from the insect *Brevicoryne brassicae*.[18] Although different in quaternary structure, the monomeric form of all GH-1 enzymes is composed of a single-domain $(\beta/\alpha)_8$ barrel with a molecular mass of ~50 kDa. The hydrolysis of the β-glycosidic bond is carried out by a catalytic mechanism that retains the conformation of the anomeric carbon; two conserved glutamate residues acting as nucleophile and proton donor (see Figure 1).

Many structural features are common to the catalytic site of all the GH-1 enzymes, even if they are active with a large variety of substrates. On the other hand, even minor changes in the substrate structure may bring about large alterations of enzymatic behaviour. Consequently, the detailed characterization of the molecular basis of β-glycosidases substrate preference within the family should contribute to the understanding of the enzymatic specificity, a fundamental property of all biological systems, and essential for inhibitor design. The specificity of family GH-1 for the monosaccharide at the −1 subsite, the substrate non-reducing end (glycone), is known through the crystal structure of several complexes containing a ligand at this subsite.[8,11,13,19–21] These studies showed a highly conserved −1 subsite, with a common pattern in sugar recognition. Conversely, structural data for the +1, aglycone-binding site, is limited to two plant β-glucosidases from *Z. mays* (ZmGlu1) and *Sorghum bicolour* (SbDhr1) with ligands occupying this subsite.[16,22] These studies gave some clues about the molecular basis of the aglycone specificity which, however, were not confirmed by mutagenesis experiments.[20] Therefore, the available information is still insufficient to have a complete picture of the molecular basis of substrate specificity, as the aglycone-binding site is essential for defining substrate preference, and in the correct positioning of the substrate into the active site for the reaction to proceed. The characterization of the aglycone-binding site also demands localization of the additional subsites (+2, +3...) found in some enzymes.

The soil and plant-associated bacterium *P. polymyxa* (formerly *Bacillus polymyxa*) synthesizes two homologous β-glucosidases, BglA and BglB, belonging to GH-1.[23,24] Both enzymes are involved in the hydrolysis of cellodextrins but they show some structural and functional differences: BglA is a cellobiase with an unusual octameric quaternary structure.[2,8,25,26] In contrast, BglB is a monomeric enzyme that acts as an exo-β-glucosidase, hydrolyzing cellobiose and cellodextrins of higher degree of polymerization.[23,27,28] Here, we report on the crystallographic structure of BglB and its complexes, and describe the structural features that determine the functional properties of the enzyme.

Results

Crystal structure of BglB shows the architecture of the active site

Previous attempts to crystallize the native form of β-glucosidase B from *P. polymyxa* (BglB) expressed in *Escherichia coli* failed even though the protein was purified to apparent homogeneity. The fact that BglB is a monomeric protein with a molecular mass of about 52 kDa and an isoelectric point of about 5 makes this protein difficult to separate from other proteins of *E. coli* of very similar characteristics by

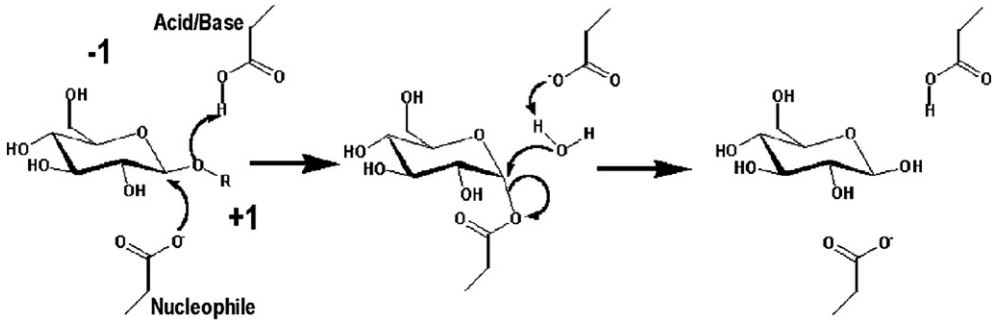

Figure 1. Scheme of the general glycosidase mechanism for retaining β-glycosidases, proceeding through a covalent intermediate.

Table 1. Oligonucleotides used as primers for the amplification of the *bgl*B gene

Oligonucleotide	Sequence	Restriction site
JP27 (27mer)	5′-TGTGAGCTCTTGATTCCTGAAAGGACG-3′	SacI
AC25 (30mer)	5′-GGGTCTAGAATCCCCCTTTTCTATTTAAAA-3′	XbaI
JP434 (66mer)	5′-GTGAGCTCAGGAGGGATCATGCATCATCATCATCAT CATAGCGAGAATACCTTTATATTTCCTGCC-3′	SacI
JP435 (57mer)	5′-GGTCTAGATTATTAATGATGATGATGATGATGAAA CCCGTTCTTCGCCATCATTTGC-3′	XbaI

conventional procedures of gel-filtration and ion-exchange chromatography. Therefore, we added a histidine tag to the BglB molecule to facilitate its purification by affinity using a nickel-nitriloacetic acid (Ni-NTA) resin (for details, see Materials and Methods and Table 1). Finally, we were able to purify enough his-tagged BglB for crystallization and to solve the structure of the enzyme at 2.15 Å resolution. Experimental details and structure determination procedures are given in Materials and Methods and in Table 2. The crystals belong to the $P2_12_12_1$ space group, with one molecule in the asymmetric unit and 44% (v/v) solvent content within the cell. The structure of the monomeric BglB corresponds to the topology of a single $(\beta/\alpha)_8$ barrel, as expected for family 1 glycosidases (CAZy†). The barrel is closely packed in the face corresponding to the short loops linking different β/α repeats, whereas additional secondary structure elements are included within the β/α motifs. These long segments at the C-terminus of the β-strands (L1 to L8) define the active centre, the catalytic residues being located at the end of β4 (the acid/base Glu167) and β7 (the nucleophilic Glu356), respectively. This specific setting of the catalytic residues is a conserved feature in GH-1 enzymes and other β-glycosidase that altogether form the clan GH-A superfamily (CAZy).

Although the alignment of GH-1 enzymes shows sequence identity from 17–45% along its members, the tertiary structure is highly conserved, and the folding differences gather within the long loops at the C termini of the barrel β-strands (Figure 2(a)). Among them, L1, L4, L6 and L7 have been shown to be responsible for shaping the active site cavity by means of different loop length and amino acid composition (loops A–D).[8] In the case of BglB, they shape a rather narrow pocket (Figure 2(b)), as compared to the homologous BglA (Figure 2(c)). The different sequence compositions of these loops restrict the wide cavity found in BglA, and shape additional subsites in BglB.

The covalent intermediate is represented by the structure of 2-deoxy-2-fluoro-α-D-glucopyranosyl/BglB complex

The first crystallographic evidence proving that the enzymatic mechanism of retaining glycosidases proceeds *via* a covalent glycosyl–enzyme intermediate was obtained ten years ago.[29] This was achieved by using mechanism-based inhibitors such as 2-deoxy-2-fluoroglycosides (2FGlc) with a good leaving group by which a covalent intermediate species accumulates.[30] This kind of inhibitor has been used to obtain complexes from a number of different GH species, including one of the most intensively studied retaining glycosidases, the GH-22 hen egg-white lysozyme,[31] and have proven to be a valuable tool in delineating mechanism features within the different glycosidase families.

Crystals of native BglB were soaked overnight in a solution containing 2′,4′-dinitrophenyl 2-deoxy-2-fluoroglucoside (2F-DNPG). The X-ray crystallographic analysis showed electron density maps that clearly revealed the presence of a 2FGlc covalently linked to Glu356, thus representing the covalent intermediate. This moiety was modelled and further refined to the final statistics given in Table 2. C1 of the glucose ring is linked to Glu356 in an axial (α-anomeric) configuration. The pyranose ring does not show significant distortion from the chair conformation (Figure 3(a)). The sugar ring is stacked to Trp402, and a water molecule (located on the β-side of the pyranose ring at 3.5 Å from C1) is bridging the endocyclic O5 to the catalytic Glu167 through hydrogen bonds (Figure 4(a)). The 2F atom is stabilized by interaction to Asn166, and the hydrogen bonding pattern between sugar OH groups and the enzyme is conserved in the 2FGlc/BglA complex (PDB code 1E4I) and in other reported GH-1 covalent intermediates.[11,13,21] On the other hand, the only difference found between the free and complexed structures is the position of the Glu356 side-chain, which moves to link the sugar C1 (Figure 5(a)). Tyr298 also seems to be tilted in the complex, similar to that observed in the other GH-1 covalent complexes. Interestingly a role as acid catalyst in the deglycosidation step (the breakage of the covalent bond between the nucleophile and the sugar) has been associated with an equivalent tyrosine residue in the *E. coli* β-galactosidase, a GH-2 retaining glycosidase.[32] The short distance observed here, 2.6 Å, from Tyr298-OH to the fluoroglucose-Glu356 bridging oxygen, O1, would support this role. The additional interaction with the endocyclic O5, positioned at 3.4 Å, could reflect a transition-state stabilization function of this tyrosine residue, which has been proposed for an homologous tyrosine residue in the case of the GH-11 xylanase from *Bacillus agaradhaerens*.[33]

† http://afmb.cnrs-mrs.fr/CAZY/

Table 2. Data collection and refinement statistics

	1 BglB	2 2FGlc/BglB	3 Tcb/-BglB	4 Glc/BglB	4 CTT/BglB
A. *Data collection*					
Reactant		2F-DNPG	Thiocellobiose	Cellotetraose	Cellotetraose
Source	Rotating anode	Synchrotron	Rotating anode	Rotating anode	Rotating anode
Unit cell dimensions					
a (Å)	70.89	71.24	71.14	71.33	71.32
b (Å)	74.86	74.65	75.39	75.58	75.28
c (Å)	88.66	88.75	88.65	88.95	88.79
Limiting resolution (Å) (outer shell)	2.15 (2.25–2.15)	2.2 (2.32–2-2)	2.3 (2.40–2.31)	2.15 (2.23–2.15)	2.45 (2.58–2.45)
Unique reflections	28,210	24,641	21,447	25,919	17,936
R_{sym}	0.07 (0.48)	0.09 (0.38)	0.05 (0.45)	0.05 (0.29)	0.10 (0.49)
Completeness (%)	99.4 (99.9)	99.8 (100)	99.2 (95.1)	99.3 (97.4)	98.3 (89.2)
Mean multiplicity	6.0 (4.4)	10.4 (10.7)	5.4 (2.7)	4.5 (5.2)	4.4 (4.1)
Mean $I/\sigma I$	10.8 (2.0)	16 (3.4)	7.5 (1.9)	10.2 (2.5)	5.0 (1.8)
B. *Final refinement parameters*					
Protein atoms (non-H)	3620	3620	3620	3620	3620
Ligand atoms (non-H)		11	23	12	46
Solvent molecules	138	137	137	141	137
R-factor	0.21	0.19	0.21	0.20	0.21
R_{free}	0.25	0.25	0.26	0.25	0.27
r.m.s.d. from ideality					
Bond lengths (Å)	0.007	0.007	0.007	0.007	0.01
Bond angles (deg.)	1.3	1.3	1.3	1.3	1.6
Averaged *B*-factors (Å^2)					
Main chain (Å^2)	35.21	25.7	34.6	37.5	47.5
Side-chain (Å^2)	36.87	27.2	35.6	37.0	49.7
Ligand (Å^2)		34.4	56.2	33.6	54.5
PDB code	2O9P	2JIE	2O9R	2O9T	2Z1S

[a] $R_{sym} = \sum |I(h)i - <I(h)>| / \sum <I(h)>$.
[b] R-factor $= \sum (|F_{obs} - F_{calc}|) / \sum |F_{obs}|$ (R_{free} is equivalent to R-factor for a randomly selected 7% subset of reflections not used in structure refinement).

The structure of the thiocellobiose/BglB complex suggests the inhibition mechanism and the aglycone-binding site

Thio-oligosaccharides have been considered as promising non-covalent inhibitor candidates for structural biology studies among glycoside hydrolases. Such oligosaccharides, in which two or more carbohydrate residues are incorporated *via* S-glycosidic linkages, should conserve the global geometry of the natural substrate while being hydrolytically inert.[34] In fact, changes occur only at the glycosidic bond, the length of the S-glycosidic bond being greater (1.83 *versus* 1.41 Å) and the angle being smaller (97° *versus* 117°), and a more diverse conformational behaviour is observed in the torsional angle distribution around the aglyconic bond as compared to the respective values for O-glycoside analogues.[35]

By means of inhibition studies at different concentrations of a chromogenic substrate, *p*-nitrophenyl-β-D-glucopyranoside (PNPG) and inhibitor, we have confirmed that thiocellobiose (Tcb) behaved as a moderate competitive inhibitor of BglB ($K_i = 21$ mM) and compares favourably with the parent substrate, cellobiose ($K_m = 100$ mM), as expected for a substrate analogue-type inhibitor. Consequently, to obtain the complex structure with Tcb, native crystals were soaked for 20 min in a crystallization solution containing 20 mM inhibitor. The electron density maps allowed the clear identification of Tcb at the active site pocket, proving that the inhibitor has diffused into the active site, as shown in Figure 3(b). The conformation of thiocellobiose that fits better into the electron density envelope corresponds to the *syn* conformer, which is consistent with previous analysis of thiocellobiose bound to a GH-1 β-glucosidase from *Streptomyces* sp. by NMR, and is described to be the most populated in solution for free thiocellobiose.[35] Moreover, there seems to be no observable distortion from the $^4C_1$ ground-state conformation in any of the glucopyranosyl rings, and therefore both sugar rings have been modelled in the chair conformation, as observed earlier.[36,37] In other reported studies on glycoside hydrolase/inhibitor complexes, a substantial ring distortion has been observed at subsite −1.[38–42]

As Figure 4(b) shows, the inhibitor is recognized by the enzyme through an extensive net of hydrogen bonds. Remarkably, the catalytic Glu167 is located close to the glycoside S-atom in a position consistent with its role as proton donor, the two oxygen atoms of its carboxylate moiety being linked to S and O2. On the contrary, the nucleophile Glu356 is apparently located far from the scissile bond and interacts with O3. This O3 is also hydrogen bonded to a water molecule that, in turn, is located at the position corresponding to O3 in the previously described covalent intermediate. Other residues involved in recognition of the inhibitor at this subsite are Tyr298, which is hydrogen bonded to Glu356 and almost

stacked to the pyranose ring, and Trp402 linked to O4. This pattern is different from that found previously in the covalent 2FGlc/BglB intermediate, in which the glucopyranose ring adopts a different orientation and stacks to Trp402 (Figure 4(a)). Glu409, whose side-chain moves significantly from

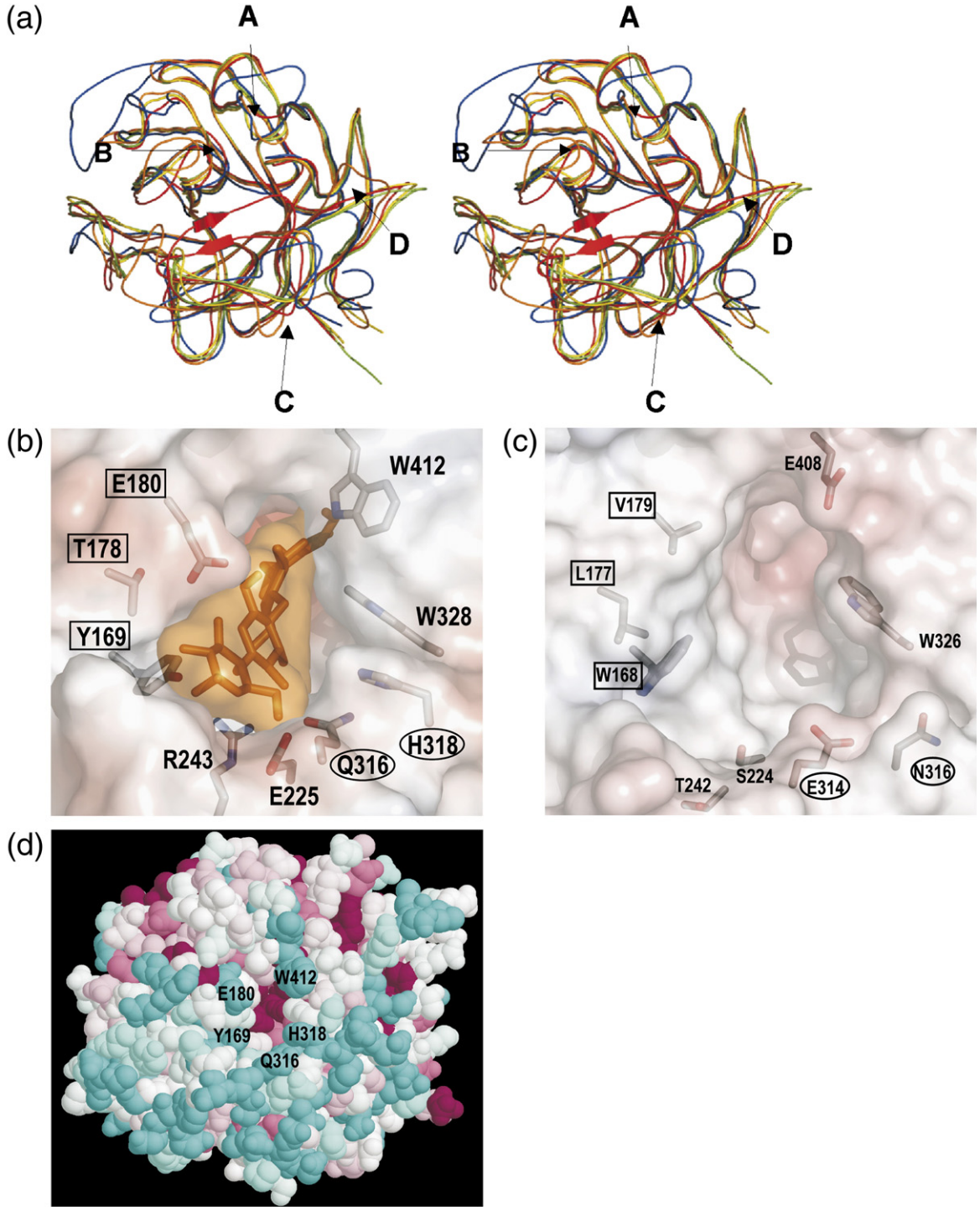

Figure 2. Folding and active site shape of family 1 glycoside hydrolases. (a) A stereoview of BglB C^{α} trace (yellow) in comparison with some representative enzymes: β-glycosidase from *S. solfataricus* (1GOW) in blue, β-glucosidase from *T. maritima* (1OD0) in green, 6-P-β-galactosidase from *L. lactis* (2PBG) in orange, and β-glucosidase P60.1 from *Z. mays* (1HXJ) in red. Loops A–D surrounding the active site are indicated with an arrow. (b) and (c) The molecular surface of the active site of BglB and BglA, respectively, showing the residues delineating the pocket as sticks. Residues from loops B (squar) and C (circle) are marked. A cellotetraose chain as observed in the CTT/BglB complex is shown at the BglB pocket in (b). (d) The conservation pattern for the GH-1 family, obtained using ConSurf.[58] BglB is presented as a space-filled model, coloured according to the conservation scores, from most variable (cyan) to most conserved (red). Highly variable residues protruding from the active site pocket are labelled.

the position in the native crystal, keeps the link to O4 and O6 already observed for the covalent complexes in both BglB and BglA,[8] but only through one of its O atoms.

The 3D structure of the enzyme in complex with the non-hydrolyzable S-glucosyl substrate analogue also provides information on binding at subsite +1. The inhibitor is recognized by the enzyme through a net of well-ordered water molecules that connect O1′ and O2′ to the side-chain of Tyr169, Gln316, Glu225 and Arg243 (Figure 4(b)). In addition, the conformation of the disaccharide is rather fixed through a link between both sugar rings occupying subsites −1 and +1, by means of a water molecule linking O3′ to O5 and O6, which is further connected to His181 through an additional water molecule. As seen in Figure 3(b), the glucopyranose ring at subsite +1 is stabilized in a hydrophobic pocket contoured by Trp328 indole at one face of the sugar ring, and Tyr169 and Leu174 side-chains at the other face. It is worth noting the lack of direct interactions between the sugar and the residues at subsite +1. All the polar interactions are mediated by water molecules, the residues closer to the sugar, i.e. Trp328, Leu174 and Tyr169, making essentially hydrophobic interactions.

Soaking BglB crystals with the substrate shows the substrate-binding site

Crystallographic studies of enzyme–substrate interaction have traditionally been done by genetic replacement of the catalytic residues, or by using substrate analogues through chemical modification of

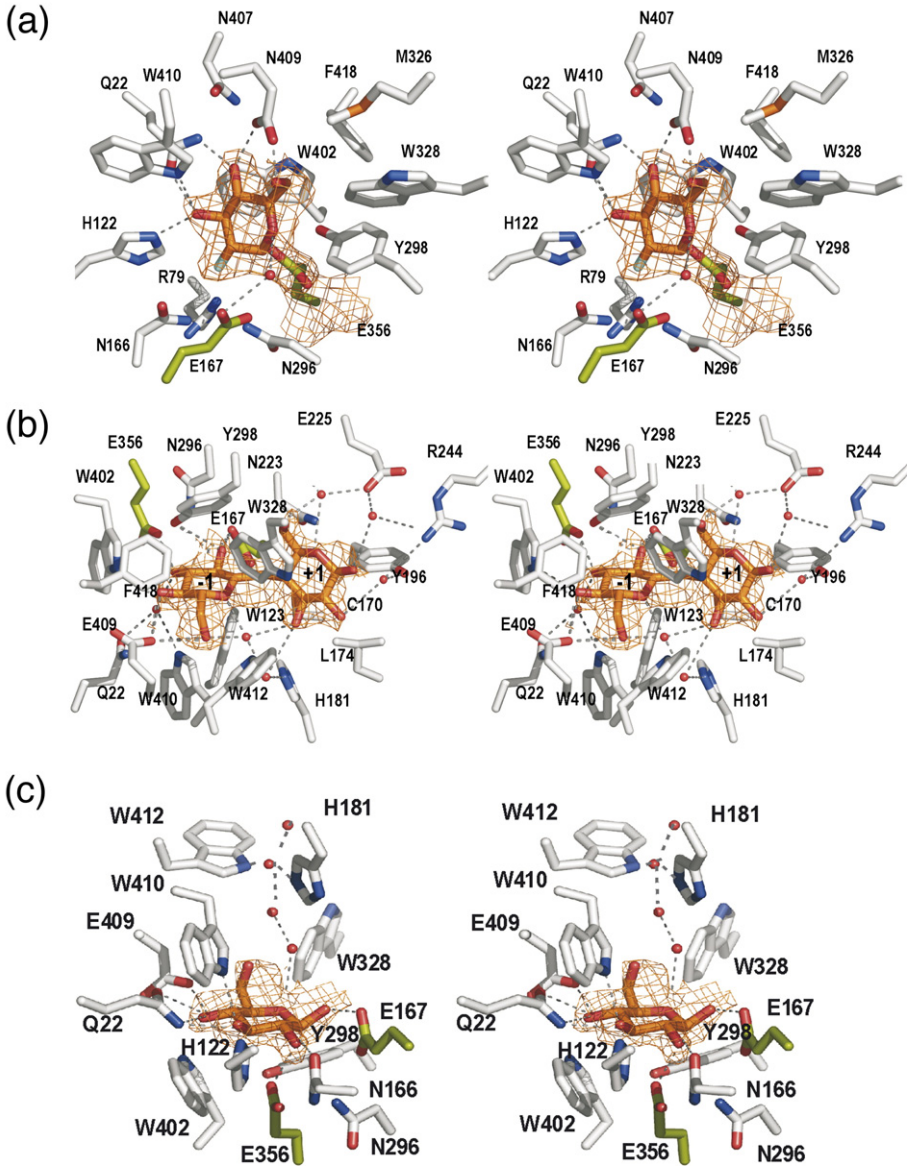

Figure 3. Stereoview of the active site in BglB complexes. Final 2*Fo*–*Fc* electron density for the refined ligands is contoured at 1σ for the (a) 2FGlc/BglB, (b) Tcb/BglB, (c) Glc/BglB and (d) CTT/BglB complexes. The catalytic glutamate residues are coloured in green, and polar interactions between the enzyme and the sugar are indicated by broken lines.

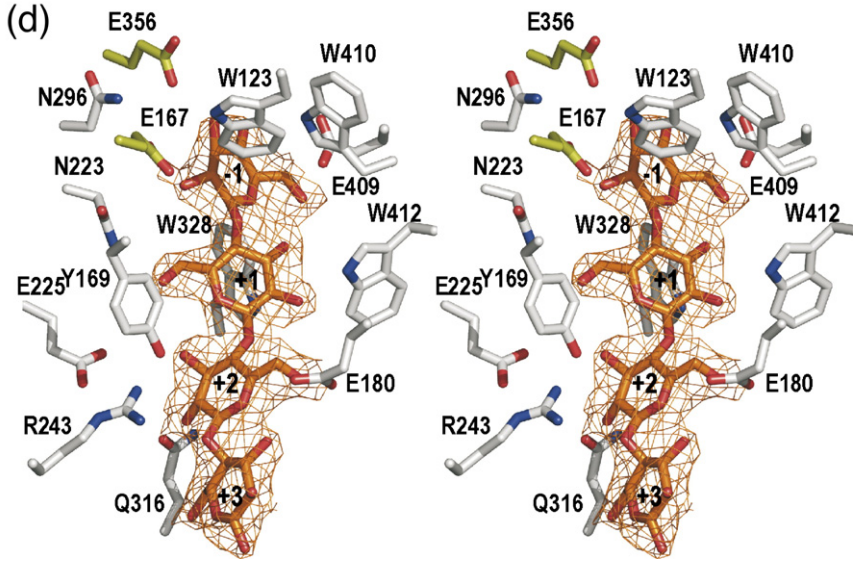

Figure 3 (*legend on previous page*)

the substrate, as is the case for the complexes described before. These two approaches are directed to slowing the hydrolysis reaction, therefore avoiding cleaveage and release of the substrates from the enzyme during the time necessary for a typical X-ray experiment. An alternative strategy has been proposed for the crystallographic analysis of glycosidases,[43] which exploits the potential use of sugars as a cryoprotectant.[44] The idea is to soak the crystal for a short time within a cryo-solution containing a high concentration of glucosides (20–25%), which should saturate the active site of all molecules and are subsequently inmobilized by flash-freezing in a nitrogen stream.

Several attempts to obtain complexes of BglB with cellotetraose have been made by soaking native crystals in saturated substrate solutions. A crystal was soaked for 15 min in this solution, and the crystallographic analysis revealed a clear density for a well ordered glucose molecule at the active site, as shown in Figure 3(c). In order to minimize the possible hydrolysis of the substrate, the soaking time was reduced to only 15 s; nevertheless, the crystallographic analysis revealed a weaker density but again consistent with a single glucose molecule (results not shown). Alternatively, a lower concentration of substrate was tried and a native crystal was soaked in a 5% (w/v) cellotetraose-containing solution for 1 min before being flash-frozen and exposed to X-ray radiation. Simulated annealing-omit maps calculated from these data showed a continuous density at the active site, revealing the presence of a sugar chain in the crystal. A cellotetraose moiety was built into the density and the complex was refined, giving the final electron density map shown in Figure 3(d), which reveals an extended conformation for the sugar chain and an essentially undistorted chair conformation for the glucose occupying the −1 subsite.

As shown in Figure 4(d), this glucose at subsite −1 is recognized by BglB through hydrogen bonding to the three glutamate residues also involved in Tcb binding. Thus, both oxygen atoms from the catalytic Glu167 are linked to O2, the nucleophile Glu356 is connected to O3, and Glu409 presents the bifurcated link to glycosides O4 and O6. On the other hand, and in agreement with what is observed in the complex with the thioanalogue Tcb, the glucose at subsite +1 does not make many polar interaction and is linked only to Trp412 and Asn223 through its O3′ and O6′, respectively. On the contrary, the glucose occupying subsite +2 makes an extensive net of hydrogen bonds to Glu180, Arg243 and Gln316, which were implicated in Tcb binding through water molecules occupying the same place as the substrate glucosyl at this subsite. Finally, it is remarkable that Glu180 and Gln316 define subsite +3 by accommodating the corresponding glucose unit through links to its proximal hydroxyl O3 and O6.

Structure of the complex glucose/BglB: inhibitor and/or product binding mode?

Soaking native crystals in saturated solutions of substrate led to the density shown in Figure 3(c). A glucose molecule in essentially a chair conformation is consistent with this density and refined to appropriate parameters. The sugar ring is occupying the active site in an orientation slightly tilted with respect to that found in the covalent BglB complex (Figure 5(a)), and presents the atomic interactions shown in Figure 4(c). The O3, O4 and O6 atoms are in the same positions as those found in the covalent complex, and keep similar hydrogen bonds to the enzyme with the only exception of Glu409, which moves away in this complex and still interacts with O4 but loses the link to O6. On the other hand, the positions of the oxygen at O1 and O2 enable them to interact strongly with the catalytic residues Glu167 and Glu356, respectively (see Figure 4(c)). The interaction of the acid catalyst

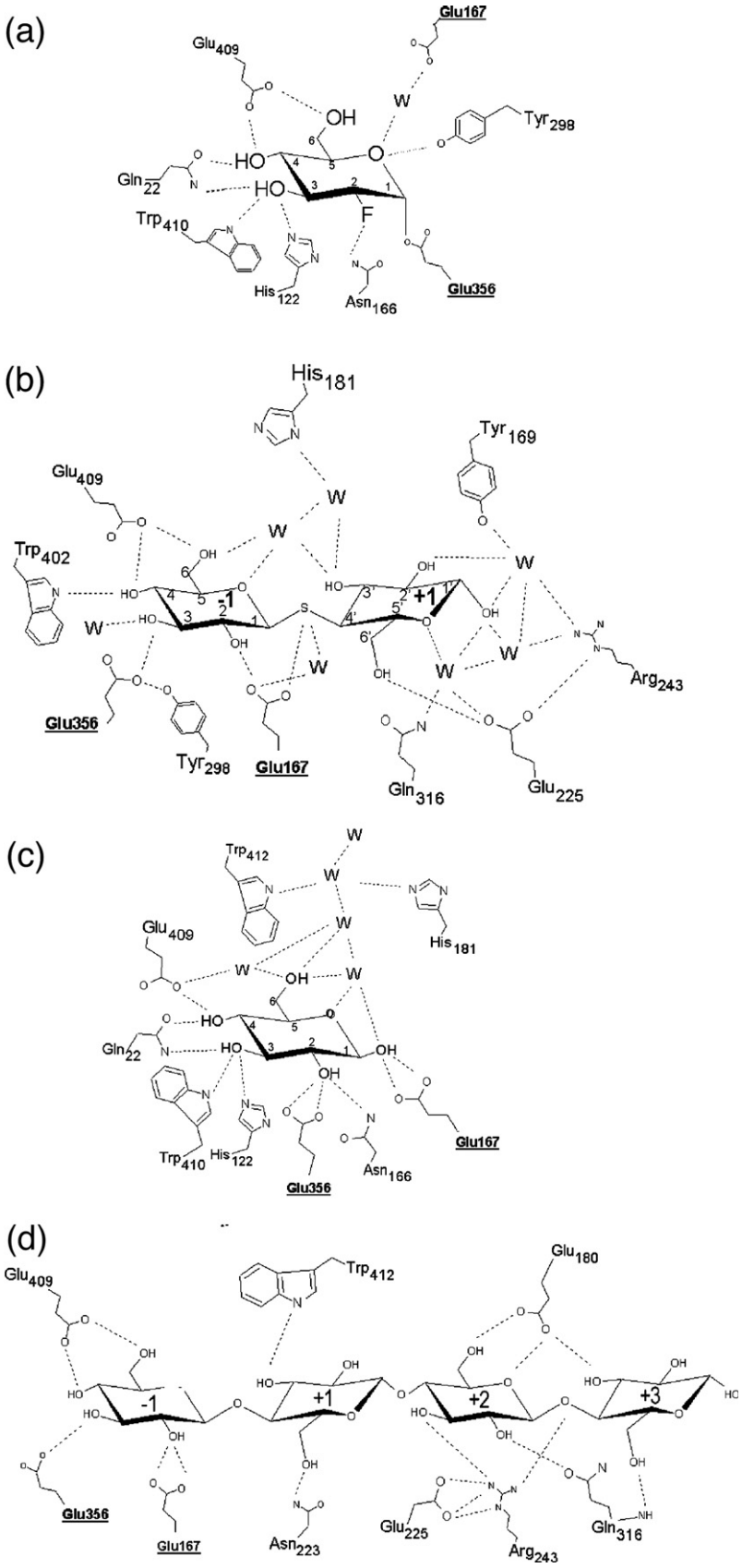

Figure 4. Bonding interactions of ligands in the active site of BglB. (a) 2-Deoxy-2-fluoro-α-D-glucosyl in the 2FGlc/BglB complex; (b) thiocellobiose in the Tcb/BglB complex; (c) glucose in the Glc/BglB complex; and (d) cellotetraose in the CTT/BglB complex. Only residues linked by hydrogen bond are shown in each case, and the catalytic glutamate residues are highlighted. Water molecules involved in ligand recognition are represented.

(Glu167) with the O1 is in accordance with its role as the protonating agent of the interglycosidic oxygen atom in the hydrolysis mechanism. On the other side, the interaction of O2 with the nucleophile residue (Glu356) must reflect the role of the substrate 2-OH in stabilizing the transition state, as has been proposed in some retaining β-glycosidases.[45] In fact, the pattern of polar interactions observed here is similar to that

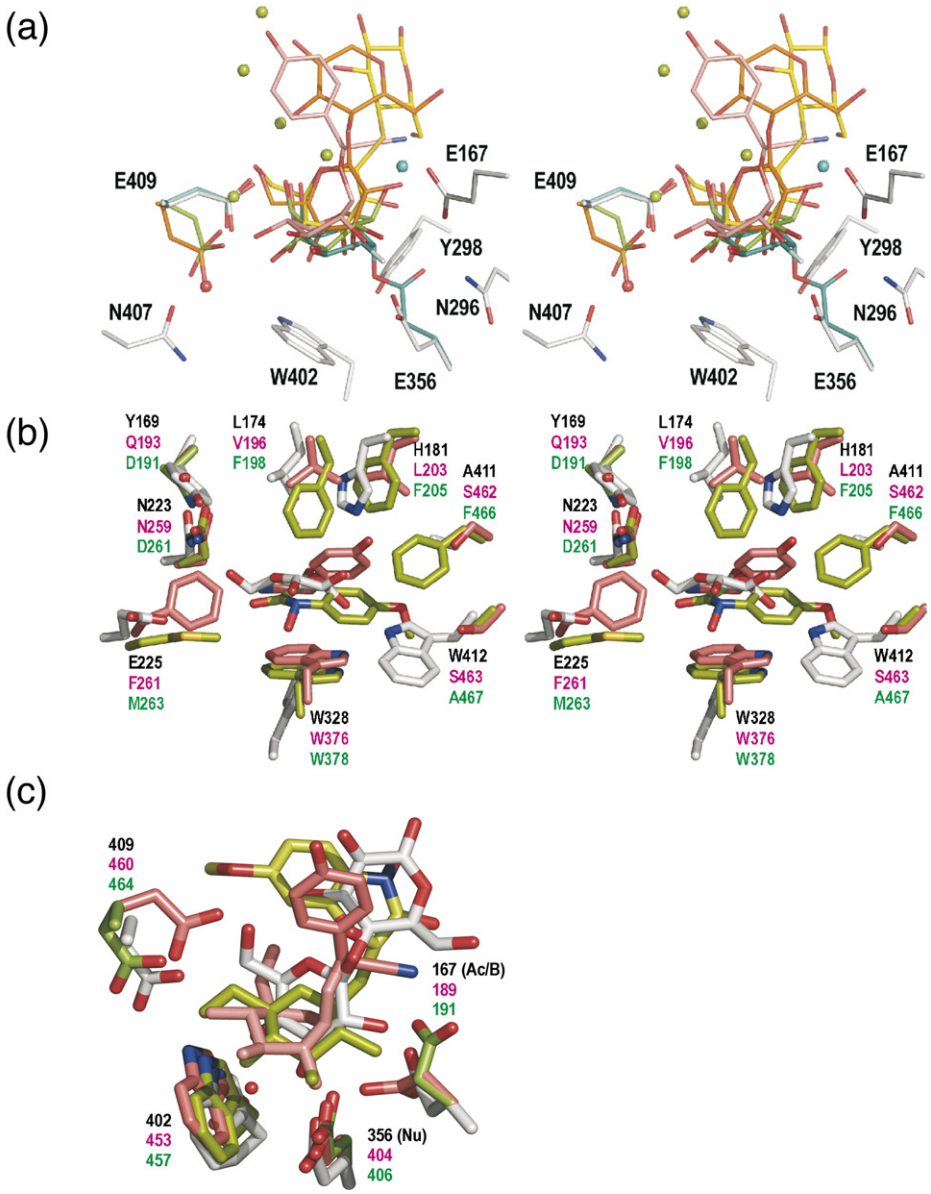

Figure 5. Comparison of BglB complexes with those from the maize and sorghum enzymes. (a) Superimposition of 2FGlc/BglB (cyan), Tcb/BglB (yellow), Glc/BglB (green) and CTT/BglB (orange) complexes onto native BglB (white). Only ligand moieties of the complexed structures are shown, with two exceptions: Glu356 in the covalent 2FGlc/BglB complex; and Glu409 presenting two positions, one observed in free BglB and the covalent complex, and the other common to Tcb/BglB, Glc/BglB, and CTT/BglB complexes. Water molecules found in the complexes are represented in the same colour as the corresponding ligands. The network of water molecules indicates a hydrophilic channel that may assist in providing the catalytic water molecule (in blue) and/or represent a product release pathway, as suggested by the superimposed aglycon moiety of dhurrin (pink), which lies in this channel. (b) A close-up view of the +1 site of the CTT/BglB complex (white) superimposed on the sorghum (SbDhr1) and maize (ZmGlu1) structures (salmon and green), both in complex with their natural substrates, dhurrin and DIMBOAGlc, respectively (PDB codes 1V03 and 1E56). All the reported GH-1 structures present the same orientation of the basal Trp328 observed in BglB, except those from plant species. Besides, all bacterial β-glucosidases present Leu/Val, His and Tyr/Trp, at the ceiling. Glu225 and Trp412 are unique to BglB. (c) The same three complexes described in (b), showing the corresponding ligands in the bound conformation, and the position of glucose at the −1 subsite. The glucose ring of DIMBOAGlc is rotated 60° with respect to that of dhurrin, while the ring is tilted in the CTT moiety. The −1 basal Trp and the conserved glutamate residues are shown.

found in the complex of the GH-1 β-glycosidase from *S. solfataricus*, with D-glucohydroximolactam,[21] which is proposed to mimic the oxocarbenium-ion-like transition state.

On the other hand, it is remarkable to note that in the structure of a barley β-D-glucan glucohydrolase, a retaining GH-3 enzyme that catalyses the removal of non-reducing glucosyl residues from β-D-glucans, a glucose molecule has been found tightly bound at the active site.[36] It is postulated that the glucose remains linked after the hydrolysis and is displaced only upon substrate binding.

Interestingly, the glucose makes polar interactions with the catalytic residues equivalent to those in the complex described here. Therefore, the glucose here might be indicating the product-binding mode, which is suggested also by its O1 atom being located very close to the catalytic water molecule found in the covalent Glc/BglB complex described above (Figure 5(a)). In any case, there is a net of well ordered water molecules keeping glucose in a rather fixed position and connecting O1, O5 and O6, to the solvent, as shown in Figures 3(c) and 4(c). Strikingly, when glucose was tested as an inhibitor, it behaved as very poor inhibitor and could not be assigned clearly to a pure competitive pattern. Only a lower limit, apparent $K_i > 50$ mM, could be estimated. Such a high inhibition constant does not correlate with a binding affinity necessary to justify the clear density shown by glucose in this Glc/BglB complex. The inhibitor capacity of glucose might be masked by its abilty to act as transglycosidation acceptor at high concentrations. Transglycosidation is a known secondary reaction for retaining glycosidases and it has been observed for BglB.[46]

Discussion

Bacteria species involved in the degradation of cellulosic substrates produce a variety of enzymes for processing related compounds along the hydrolytic pathway. This is the case of *P. polymyxa*, which encodes two homologous β-glucosidases, BglA and BglB, both presenting a polypeptide chain of 448 residues with 45% identity. The two enzymes present different quaternary structures, BglA being an octamer, whereas BglB exists as a monomer. Moreover, BglA is shown to be highly specific against cellobiose, while BglB is active also against oligosaccharides formed by more than two glucose units.[23] Despite the level of sequence homology between BglA and BglB (45%) being equivalent to that presented by either of them with respect to others members of GH-1, both enzymes show a highly conserved folding (438 of 448 residues superimpose their C^{α} atoms, giving an RMS of 0.84 Å) and remarkably, loops A–D responsible for active site shaping (loops L1, L4, L6 and L7 of the barrel) are in almost equivalent conformations. Therefore, the differential activity of BglB against longer substrates must be determined by the shape and the chemical nature of the residues located at the active site.

Thus, in contrast to the wide and open cavity previously reported in BglA, the active site of BglB shows many long polar residues that seem to narrow the pocket and contour additional subsites (Figure 2(b) and (c)). The BglB residues Tyr169, Thr178 and Glu180 from loop B, together with Glu225 and Arg243 from β5 and α5, respectively, clearly protrude into the cavity at a distal position from the catalytic residues and therefore restrict the pocket. The side-chains of Gln316 and His318, from loop C, point to the solvent and contribute to delineating the cavity. On the contrary, it is interesting to note that in BglA, the equivalent residues Glu314 and Asn316 are retracted, which may be determined by loop C being engaged in intermolecular contacts at the interface within each of the tight dimers of the BglA octamer. This suggests that oligomerization, in this case, can assist in fine-tuning the specificity of the active centre of BglA by modulating the loops surrounding the cavity. In addition, it is significant that sequence analysis reveals that some of these BglB residues, i.e. Glu225, Arg243 and Trp412. are very scarce or not found in the corresponding positions in the GH-1 family. Furthermore, Trp412, together with Tyr169, Glu180, Gln316 and His318 from loops B and C, present the greatest variability among the GH-1 enzymes (Figure 2(d)), which suggests that the great plasticity of these enzymes in processing such a different series of substrates is determined ultimately by the nature of a comparatively low number of residues.

Detailed analysis of the relevant atomic interactions within each subsite comes from the crystal structure of the BglB complexes, the complex of BglB with the inhibitor Tcb, and that with the substrate CTT, providing the molecular basis for the specific substrate preference of BglB. First, the absence of direct polar interactions observed in the complex Tcb/BglB at the aglycone, the glucose of the inhibitor being linked to the enzyme through a network of water molecules hydrogen bonded to distal residues of BglB (Figure 4(b)). Second, the CTT/BglB complex reveals a tighter binding of the glucose placed at subsite +2, which is recognized by the enzyme through an extensive network of hydrogen bonds to Glu180, Arg243 and Gln316 (Figure 4(d)). In contrast, the glucosyl residues placed at subsite +1 and +3 are constrained less tightly, as deduced by the fewer polar interactions made by their OH groups. Taken together, these observations underline the structural requirements causing the observed activity of BglB against longer oligosaccharides than BglA, three being preferred over four glucose units.

It is worth comparing the aglycone (+1) binding site of the CTT/BglB complex with the previously reported GH-1 β-glycosidases from plants, complexed with their natural substrates (Figure 5(b)), SbDhr1-dhurrin (PDB code 1V03) and ZmGlu1-DIMBOAGlc (PDB code 1E56). In both plant enzymes, the aglycone site is described as basically formed by two walls that embrace the aglycone,[16,22] the basal platform being shaped by a highly conserved Trp (Trp376 in SbDhr1, and Trp378 in ZmGlu1). The differences in specificity between both enzymes are attributed to the different "ceiling", and in ZmGlu1, three non conserved Phe residues make a hydrophobic pocket able to accommodate different aromatic substrate analogues. On the contrary, in SbDhr1, two of these Phe residues are substituted by less bulky hydrophobic residues, and two polar interactions established by Ser462 and Asn259 restrict the position of the aglycone, for which the enzyme

is highly specific. It is interesting to highlight that the conformation adopted by the basal Trp side-chain in both plant enzymes, and in the reported β-glucosidase from *T. repens* (PDB code 1CBG), is unique and different from that found in BglB Trp328, and the other GH-1 enzymes whose crystal structure is known. This might determine a particular orientation of the aglycone that is possibly common to plant enzymes, presenting highly variable aglycone moieties.

On the other hand, Leu/Val, His and Ala are present in all bacterial β-glucosidases at positions ascribed to the ceiling, i.e. Leu174, H181 and Ala411 in BglB, and the same pattern is observed in bacterial phospho-β-galactosidases, where Gly, Phe and Ser as conserved residues. Consequently, these residues shape a common environment for the aglycone within each of these groups of enzymes, while the greater variability found in eukaryote enzymes must reflect their implications in determining the different specificities. Thus, we think that the complex CTT/BglB described here delineates the general features of the +1 site in bacterial β-glucosidases. As a common pattern, the aglycone must be located in a hydrophobic environment, enclosed by the basal Trp328, and Leu174, His181 and Tyr169 (most frequently a Trp), in the other side of the ring (Figure 5(b)). The different substrate preference among these enzymes should be determined by the variable residues. In the case of BglB, Glu225 and Trp412 are unique in this position and must be essential in positioning the glucose at +1, as mentioned above.

In the reported crystal structures of ZmGlu1 complexed with the substrate and analogues, the sugar ring at the −1 site was always poorly defined, the electron density being very weak.[14,22] Therefore, it was not possible to define a specific conformation of the glucopyranoside ring at the −1 site, a fact that was attributed to the active site being able to recognize both the ground-state chair conformation, and the postulated distorted skew-boat conformation. On the contrary, in the structure of SbDhr1 in complex with its natural substrate dhurrin, the electron density was well defined and clearly showed the substrate bound in the distorted conformation that represents the Michaelis complex.[16] In our case, the electron density clearly shows the positions and conformation of Tcb and CTT at the active site (Figure 3(b) and (d)). Although the possibility of small distortions ca not be excluded at the present resolution (2.3 Å and 2.45 Å, respectively), it appears that both oligosaccharides are bound at the active site showing the glycone moiety essentially in a chair conformation, as observed in other retaining glycosidases.[36,37] This is in contrast with that observed in several exo and endo-acting β-D-glycosidases. including glycosyl ligands significantly distorted at subsite −1.[38–43] It was proposed recently that structurally related enzymes do not necessarily adopt the same catalytic pathways and, in general, exo-acting enzymes in which the glycone sugar is not tethered to an extended oligosaccharide chain, are able to explore larger conformational space than endo-acting enzymes whose subsite −1 glycone is necessarily restrained by the oligosaccharide chain to which it is linked.[47] Consequently, structurally related enzymes working on chemically different sugars may harness alternative agendas that reflect the chemistry of their substrate. Furthermore, trapped Michaelis complexes with axial leaving groups have been reported showing different conformations at the −1 site ring.[42,48] Thus, it is difficult to postulate any specific conformation for an activated enzyme–substrate complex, but such a distortion would indeed bring the substrate C1 atom closer to Glu356, therefore enabling the nucleophilic attack. Details of how this process may come about is unknown, but Glu409 is likely to play an essential role, as discussed below. On the other hand, the inhibitory capacity of thiocellobiose might be related to some impairment in the course of this conformational pathway along the enzymatic hydrolysis, owing to the different chemical nature of the sulphur-glycoside bond. Finally, we suggest that this activated complex would preserve essentially the same interaction pattern, as the residues involved in substrate recognition present long, flexible side-chains, able to accommodate slight changes in the sugar chain.

As suggested above, an aspect worthy of comment concerns the position of Glu409. This glutamate residue is highly conserved among GH-1 β-glycosidases and presents a bifurcated hydrogen bond to O4 and O6, observed in all reported GH-1 complexes.[49] The conformational freedom characteristic of each enzyme allows this glutamate side-chain to adapt its position to an axial O4, therefore being able to recognize also galacto-configured substrates in some enzymes.[8] Another role ascribed to this glutamate residue has been given in the maize enzyme, as guiding the movement of the glucosyl moiety into the correct distorted position required for the nucleophilic attack to occur.[16] Interestingly, Glu409 presents two states among the crystal structures of BglB described here, as shown in Figure 5(a). In the free enzyme and the covalent intermediate, its position is equivalent to that found in the Michaelis complex from sorghum, and the same is observed in the reported complexes of *T. maritima*,[11] and *S. solfataricus*[21] with transition-state analogues. On the other hand, the Glc/BglB, Tcb/BglB and CTT/BglB complexes are similar to ZmGlu1-DimboaGlc,[16] which is assumed to represent the substrate-bound mode. This highlights the possible role of the glutamate residue as a switch between the intermediate, tighter bound, species while the other state of the glutamate involves the substrate and possibly the product binding mode. In this context, the superposition of the BglB complexes shown in Figure 5(a) also shows a net of well-ordered molecules along the different complexes that define a hydrophilic channel open to the solvent, probably involved in providing the catalytic water molecule. Additional comparison with the SbDhr1–dhurrin complex[16] locates the aglycone moiety of dhurrin in

this channel, suggesting a possible product release pathway. Considering the hydrogen bonding pattern of the water molecules network (Figures 3 and 4), the side-chains of His181 and Trp412 might be connected to this role.

In conclusion, the three-dimensional structure of BglB complexes reported here provides the molecular basis of the different functionalities found in BglA and BglB, two very closely related enzymes, and revals new details of the machinery involved in modulating the specificity along the GH-1 enzymes. Moreover, these complexes illustrate different stages through the reaction course, and they could assist in the mechanistic interpretation of enzyme inhibition and in the development of new enzymes with improved activity or specificity.

Materials and Methods

Expression and purification of BglB

The *bgl*B gene was amplified by PCR in reactions that added a series of six histidine residues to either the amino or the carboxy-terminal residue of the protein, or to both. The primers used in these reactions are given in Table 1. Primers JP434 (direct) and JP435 (reverse) contain suitable restriction sites to facilitate the cloning of the amplified gene and six histidine-encoding codons (CAT in the coding strand and ATG in the complementary). Following a SacI site, JP434 contains a Shine-Dalgarno (ribosome-binding) sequence (AGGAGGGACT) followed by an initiation codon (ATG) and 27 nt corresponding to nucleotides 3–30 of the *bgl*B gene sequence. The elements of JP435 are a XbaI site followed by two stop codons (TTA), six ATG codons that correspond to six histidine codons in the complementary strand, and 25 nt corresponding to the final ones of the open reading frame of the gene. The two other primers used, JP27 (direct) and AC25 (reverse), are standard primers containing a suitable restriction site followed by a series of nucleotides corresponding to the beginning or the end of the gene-coding sequence. Four amplification reactions were carried out using each one of the following combinations of primer: JP27+AC25; JP27+JP435; JP434+AC25 and JP434+JP435. A plasmid with the *bgl*B gene[24] was used as the template. The four amplified DNA fragments were digested with SacI and XbaI and ligated to plasmid ΔpUC18[26] opened with the same enzymes. The ligation mixtures were used to transform *E. coli* XL1 Blue (Stratagene). The screening of β-glucosidase activity in the transformants[23] revealed that the clones expressing the construction obtained by amplification with primers JP434+AC25 (His tag at the amino-terminal end of the protein) showed activity similar to those expressing the control construction JP27+AC25 (without the His tag). Therefore, this construction with six His residues at the N terminus of BglB was used for large-scale production of the enzyme, which was subsequently purified by affinity to nickel using a nickel-nitriloacetic acid (Ni-NTA) resin (Qiagen).

Enzyme inhibition assays

The enzymatic activity of β-glucosidase BglB was followed using PNPG as substrate and measuring the released *p*-nitrophenol by UV absorbance at 400 nm. The enzyme (3.5 μg/ml) was incubated with the substrate (five different concentrations, from 0 to 7 mM, for the inhibition studies) in 50 μl of buffer (50 mM sodium phosphate, pH 7) and in the presence of several concentrations of thiocellobiose (0, 5 mM, 10 mM, 20 mM and 40 mM) or glucose (0, 10 mM, 20 mM and 40 mM). After incubation at 37 °C for 10 min, the samples were diluted with 1 M sodium carbonate and the absorbance at 400 nm was measured. Double reciprocal Lineweaver-Burk plots ($1/V$ *versus* $1/S$) were used to extract apparent K_m values and to calculate competitive inhibition constants K_i.

Crystallization and data collection

Equal amounts of protein solution (10 mg/ml of BglB in 10 mM Tris–HCl, pH 8) and precipitant buffer (100 mM Bis-Tris (pH 6.5), 30% (w/v) PEG 550 MME, 0.05 M $CaCl_2$) were mixed and vapour equilibrated against a reservoir containing the latter solution. Crystallization experiments were carried out at room temperature. Crystals grew in two to three weeks to of maximum size 0.3 mm×0.2 mm×0.2 mm, belong to space group $I2_12_12_1$, with one molecule per asymmetric unit and 44% (v/v) solvent content in the cell. For data collection, native crystals were transferred to cryoprotectant solutions containing 40% (w/v) PEG 550 MME before being cooled to 120 K in a stream of nitrogen (Oxford Instruments). An X-ray radiation (Cu, Kα) data set was collected using an in-house Bruker-Nonius rotating anode generator, operating at 40 kV and 100 mA, equipped with Osmic focusing mirrors and Kappa2000 CCD detector. A total of 245 data frames were collected using 0.5° oscillations and 5 min exposure times. The diffraction data were analyzed with HKL-2000 package to yield integrated intensities.[50] A summary of data collection and data reduction statistics is shown in Table 2.

Structure solution and refinement

The structure of BglB was solved by the molecular replacement method. Initial phases were obtained with MOLREP using the atomic coordinates of β-glucosidase A from *P. polymyxa* as the search model (PDB code 1E4I).[51] A clear solution was found that after rigid-body fitting produced an *R*-factor of 0.49 with a correlation coefficient of 0.47 for data up to 3 Å. Crystallographic refinement was performed by using CNS with flat bulk-solvent correction, including low-resolution data to 50 Å, and using maximum likelihood target features.[52] A subset of 7% randomly selected structure-factor amplitudes were excluded from automated refinement to compute a free *R*-factor throughout the refinement. Several rounds of simulated annealing (SA)-refinement were alternated with rebuilding by the program O.[53] At the last stage, water molecules were included combined with more rounds of positional and individual restrained *B*-factor refinement, which led to a final *R*-factor of 0.21 (R_{free}=0.25) for all data set up to 2.15 Å resolution. Final refinement parameters are reported in Table 2. Stereochemistry of the model was checked by using PROCHECK[54] and the Figures were generated with PYMOL‡.

‡ http://pymol.sourceforge.net/

Crystal structure of the complexes

2F-DNPG, thiocellobiose, and cellotetraose used in the soaking experiments were purchased from SIGMA.

2-Deoxy-2-fluoro-α-D-glucopyranoside/BglB complex

A crystal was transferred into a fresh solution of 6 mM 2F-DNPG and 35% (w/v) PEG 550 MME in 100 mM Bis-Tris (pH 6.5). After soaking overnight, the crystal was flash-frozen to 100 K. A full data set was collected at the ESRF using BM16 beam line, up to 2.2 Å. Analysis of the diffraction data was performed using MOSFLM[55] merged with the CCP4 suite,[56] and the atomic structure was refined by CNS.[52] SA-omit maps were calculated at the active-site cavity showing the presence of the molecule 2FGlc linked to the nucleophile, and this was manually built into the model and included in the refinement. Final statistics are given in Table 2.

Thiocellobiose/BglB complex

A crystal was soaked in 100 mM Bis-Tris, pH 6.5, 35% (w/v) PEG 550 MME, 50 mM $CaCl_2$, 10 mM thiocellobiose for 20 min before being flash-frozen to 120 K in a nitrogen stream. A full data set was collected in a Bruker-Nonius rotating anode generator, equipped Kappa2000 CCD detector. A total of 254 data frames were collected using 0.7° oscillations and 10 min exposure times. The diffraction data were indexed, integrated and scaled using HKL-2000 program suite.[50] Several SA-omit maps were calculated using different radius spheres around catalytic residues, which unequivocally revealed the presence of an inhibitor molecule at the active site. Thiocellobiose was included in the model and refined with CNS.[52]

Glucose/BglB complex

A crystal was transferred to a crystallization solution (100 mM Bis-Tris (pH 6.5), 30% PEG 550 MME, 0.05 M $CaCl_2$) saturated with cellotetraose, and after 15 min was flash-frozen to 120 K. A full data set to 2.15 Å resolution was collected on a 345 MAR Research Imaging Plate detector, using an ENRAF-NONIUS rotating anode generator. Data were processed using the MOSFILM[55] program package and merged with the CCP4 suite.[56] SA-omit maps were calculated to find the active-site occupation, and the structure was refined by CNS.[52] A summary of data collection and refinement statistics is given in Table 2.

Cellotetraose/BglB complex

A crystal was transferred to a stabilizing solution (100 mM Bis-Tris (pH 6.5), 35% PEG 550 MME, 50 mM $CaCl_2$, 5% (w/v) cellotetraose) and after 1 min was flash-frozen to 120 K. A full data set to 2.45 Å resolution was collected on a 345 MAR Research Imaging Plate detector, using an ENRAF-NONIUS rotating anode generator. Data were processed using the MOSFILM[55] program package and merged with the CCP4 suite.[56] SA-omit maps were calculated to find the active-site occupation, and the structure was refined by CNS.[52] and Refmac.[57] A summary of data collection and refinement statistics is given in Table 2.

Acknowledgements

This work has been supported by grants BIO2004-03773-C04-04, BIO2004-03773-C04-02 and CTQ2006-10874-C02-01 from the Ministerio de Educación y Ciencia, Spain.

References

1. Coutinho, P. M. & Henrissat, B. (1999). Carbohydrate-active enzymes: an integrated database approach. In *Recent Advances in Carbohydrate Bioengineering* (Gilbert, H. J., Davies, G., Henrissat, B. & Svensson, B., eds), pp. 3–12, The Royal Society of Chemistry, Cambridge, UK.
2. Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Gonzalez, B., Lopez-Camacho, C. & Polaina, J. (1998). Structural basis of increased resistance to thermal denaturation induced by single amino acid substitution in the sequence of β-glucosidase A from *Bacillus polymyxa*. *Proteins: Struct. Funct. Genet.* **33**, 567–576.
3. Bhatia, Y., Mishra, S. & Bisaria, V. S. (2002). Microbial β-glucosidases: cloning, properties, and applications. *Crit. Rev. Biotechnol.* **22**, 375–407.
4. Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. & Pearl, L. H. (1997). Crystal structure of the β-glycosidase from the hyperthermophilic archeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability. *J. Mol. Biol.* **271**, 789–802.
5. Chi, Y. I., Martinez-Cruz, L. A., Jancarik, J., Swanson, R. V., Robertson, D. E. & Kim, S. H. (1999). Crystal structure of the β-glycosidase from the hyperthermophile *Thermosphaera aggregans*: insights into its activity and thermostability. *FEBS Letters*, **445**, 375–383.
6. Akiba, T., Nishio, M., Matsui, I. & Harata, K. (2004). X-ray structure of a membrane-bound β-glycosidase from the hyperthermophilic archaeon *Pyrococcus horikoshii*. *Proteins: Struct. Funct. Genet.* **57**, 422–431.
7. Wiesmann, C., Beste, G., Hengstenberg, W. & Schulz, G. E. (1995). The three-dimensional structure of 6-phospho-β-galactosidase from *Lactococcus lactis*. *Structure*, **3**, 961–968.
8. Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Lequerica, J. L. & Polaina, J. (1998). Crystal structure of β-glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family 1 glycosyl hydrolases. *J. Mol. Biol.* **275**, 491–502.
9. Hakulinen, N., Paavilainen, S., Korpela, T. & Rouvinen, J. (2000). The crystal structure of β-glucosidase from *Bacillus circulans* sp. alkalophilus: ability to form long polymeric assemblies. *J. Struct. Biol.* **129**, 69–79.
10. Wang, X., He, X., Yang, S., An, X., Chang, W. & Liang, D. (2003). Structural basis for thermostability of β-glycosidase from the thermophilic Eubacterium *Thermus nonproteolyticus* HG102. *J. Bacteriol.* **185**, 4248–4255.
11. Zechel, D. L., Boraston, A. B., Gloster, T., Boraston, C. M., MacDonald, J. M., Tilbrook, D. M. *et al.* (2003). Iminosugar glycosidase inhibitors: structural and thermodynamic dissection of the binding of isofagomine and 1-deoxynojirimycin to β-glucosidases. *J. Am. Chem. Soc.* **125**, 14313–14323.
12. 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.

13. Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S. & Henrissat, B. (1997). The crystal structures of *Sinapis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. *Structure*, **5**, 663–675.
14. Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W. P., Bevan, D. R., Henrissat, B. & Esen, A. (2001). Crystal structure of a monocotyledon (maize ZmGlu1) β-glucosidase and a model of its complex with p-nitrophenyl β-D-thioglucoside. *Biochem. J.* **354**, 37–46.
15. Zouhar, J., Vevodova, J., Marek, J., Damborsky, J., Su, X.-D. & Brzobohaty, B. (2001). Insights into the functional architecture of the catalytic center of a maize β-glucosidase Zm-p60.1. *Plant Physiol.* **127**, 973–985.
16. Verdoucq, L., Moriniere, J., Bevan, D. R., Esen, A., Vasella, A., Henrissat, B. & Czjzek, M. (2004). Structural determinants of substrate specificity in family 1 β-glucosidases: Novel insights from the crystal structure of sorghum Dhurrinase-1, a plant β-glucosidase with strict specificity, in complex with its natural substrate. *J. Biol. Chem.* **279**, 31796–31803.
17. Sue, M., Yamazaki, K., Yajima, S., Nomura, T., Matsukawa, T., Iwamura, H. & Miyamoto, T. (2006). Molecular and Structural Characterization of Hexameric β-D-Glucosidases in Wheat and Rye. *Plant Physiol.* **141**, 1237–1247.
18. Husebye, H., Arzt, S., Burmeister, W. P., Haertel, F. V., Brandt, A., Rossiter, J. T. & Bones, A. M. (2005). Crystal structure at 1.1 Å resolution of an insect myrosinase from *Brevicoryne brassicae* shows its close relationship to β-glucosidases. *Insect Biochem. Mol. Biol.* **35**, 1311–1320.
19. Wiesmann, C., Hengstenberg, W. & Schulz, G. E. (1997). Crystal structures and mechanism of 6-phospho-β-galactosidase from *Lactococcus lactis*. *J. Mol. Biol.* **269**, 851–860.
20. Verdoucq, L., Czjzek, M., Moriniere, J., Bevan, D. R. & Esen, A. (2003). Mutational and structural análisis of aglycone specificity in maize and sorgun β-glucosidases. *J. Biol. Chem.* **278**, 25055–25062.
21. Gloster, T. M., Roberts, S., Ducros, V., Perugino, G., Rossi, M., Hoos, R. *et al.* (2004). Structural studies of the β-glycosidase from *Sulfolobus solfataricus* in complex with covalently and noncovalently bound inhibitors. *Biochemistry*, **43**, 6101–6109.
22. Czjzek, M., Cicek, M., Zamboni, V., Bevan, D. R., Henrissat, H. & Esen, A. (2000). The mechanism of substrate (aglycone) specificity in β-glucosidases is revealed by crystal structures of mutant maize β-glucosidase-Dimboa, -DimboaGlc, and –dhurrin complexes. *Proc. Natl Acad. Sci. USA*, **97**, 13555–13560.
23. Gonzalez-Candelas, L., Aristoy, M. C., Polaina, J. & Flors, A. (1989). Cloning and characterization of two genes from *Bacillus polymyxa* expressing β-glucosidase activity in *Escherichia coli*. *Appl Environ Microbiol.* **55**, 3173–3177.
24. Gonzalez-Candelas, L., Ramon, D. & Polaina, J. (1990). Sequences and homology analysis of two genes encoding β-glucosidases from *Bacillus polymyxa*. *Gene*, **30**, 31–38.
25. Lopez-Camacho, C., Salgado, J., Lequerica, J. L., Madarro, A., Ballestar, E., Franco, L. & Polaina, J. (1996). Amino acid substitutions enhancing thermostability of *Bacillus polymyxa* β-glucosidase A. *Biochem. J.* **15**, 833–838.
26. Gonzalez-Blasco, G., Sanz-Aparicio, J., Gonzalez, B., Hermoso, J. A. & Polaina, J. (2000). Directed evolution of β-glucosidase A from *Paenibacillus polymyxa* to thermal resistance. *J. Biol. Chem.* **275**, 13708–13712.
27. Arrizubieta, M. J. & Polaina, J. (2000). Increased thermal resistance and modification of the catalytic properties of a β-glucosidase by random mutagenesis and *in vitro* recombination. *J. Biol. Chem.* **275**, 28843–28848.
28. Zubillaga, R. A., Garcia-Hernandez, E., Camarillo-Cadena, M., Leon, M. & Polaina, J. (2006). Effect of a new ionic pair on the unfolding activation barrier of β-glucosidase B. *Protein Pept Letters*, **13**, 113–118.
29. White, A., Tull, D., Johns, K., Withers, S. G. & Rose, D. (1996). Crystallographic observation of a covalent catalytic intermediate in a β-glycosidase. *Nature Struct. Biol.* **3**, 149–152.
30. 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.
31. Vocadlo, D. J., Davies, G. J., Laine, R. & Withers, S. G. (2001). Catalysis by hen egg-white lysozyme proceeds *via* a covalent intermediate. *Nature*, **412**, 835–838.
32. Roth, N. J., Penner, R. N. & Huber, R. E. (2003). β-Galactosidase (*E. coli*) with double substitutions shows that Tyr-503 acts independently of Glu461 but cooperatively with Glu537. *J. Protein Chem.* **22**, 663–668.
33. Rye, C. S. & Withers, S. G. (2000). Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* **4**, 573–580.
34. Driguez, H. (2001). Thiooligoscharides as tools for structural biology. *ChemBiochem*, **2**, 311–318.
35. Montero, E., Vallmitjana, M., Pérez-Pons, J. A., Querol, E., Jiménez-Barbero, J. & Cañada, F. J. (1998). NMR studies of the conformation of thiocellobiose bound to a β-glucosidase from *Streptomyces* sp. *FEBS Letters*, **421**, 243–248.
36. Hrmova, M., Varghese, J. N., de Gori, R., Smith, B. J., Driguez, H. & Fincher, G. B. (2001). Catalytic mechanisms and reaction intermediates along the hydrolytic pathway of a plant β-D-glucan glucohydrolase. *Structure*, **9**, 1005–1016.
37. Hrmova, M., DeGori, R., Smith, B. J., Fairweather, J. K., Driguez, H., Varghese, J. N. & Fincher, G. B. (2002). Structural basis for broad substrate specificity in higher plant β-D-glucan glucohidrolases. *Plant Cell*, **14**, 1033–1052.
38. Sulzenbacher, G., Driguez, H., Henrissat, B., Schülein, M. & Davies, G. J. (1996). Structure of the *Fusarium oxysporum* endoglucanase I with a nonhydrolyable substrate analogue: substrate distortion gives rise to the preferred axial orientation for the leaving group. *Biochemistry*, **35**, 15280–15287.
39. Davies, G. J., Withers, S. G. *et al.* (1998). Snapshots along an enzymatic reaction coordinate: analysis of a retaining β-glycoside hydrolase. *Biochemistry*, **37**, 11707–11713.
40. Zou, J. Y., Jones, T. A. *et al.* (1999). Crystallographic evidence for substrate distortion and protein conformational changes during catalysis in cellobiohydrolases Cel6A *Trichoderma reesei*. *Structure*, **7**, 1035–1045.
41. Garcia-Herrero, A., Montero, E., Muñoz, J. L., Espinosa, F. F., Vián, A., García, J. L. *et al.* (2002). Conformational selection of glycomimetics at enzyme catalytic sites: experimental demonstration of the binding of distinct high-energy distorted conformations of C-, S-, and O-glycosides by *E. coli* β-galactosidases. *J. Am. Chem. Soc.* **124**, 4804–4810.
42. Vasella, A., Davies, G. J. & Böhm, M. (2002). Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* **6**, 619–629.

43. Schmidt, A., Gübitz, G. M. & Kratky, C. (1999). Xylan binding subsite mapping in the xylanase from *Penicillium simplicissimum* using xylooligosaccharides as cryo-protectant. *Biochemistry*, **38**, 2403–2412.
44. Garman, E. F. & Schneider, T. R. (1997). Macromolecular cryocrystallography. *J. Appl. Crystallog.* **30**, 211–237.
45. Withers, S. G. (2001). Mechanisms of glycosyl transferases and hydrolases. *Carbohydr. Pol.* **44**, 325–337.
46. Montero, E. (1998). Intestinal lactase and other glycosidases: hydrolysis and transglycosidation reactions and NMR studies of enzyme-substrate interactions. PhD thesis, Universidad Complutense de Madrid.
47. Money, V. A., Smith, N. L., Scaffidi, A., Stick, R. S., Gilbert, H. J. & Davies, G. J. (2006). Substrate distortion by a lichenase highlights the different conformational itineraries harnessed by related glycoside hydrolases. *Angew. Chem.* **118**, 5260–5264.
48. Varrot, A., Frandsen, T. P., Driquez, H. & Davies, G. J. (2002). Structure of the *Humicola insolens* cellobiohydrolase Cel6A D416A mutant in complex with a non-hydrolysable substrate analogue, methyl cellobiosyl-4-thio-β-cellobiose, at 1.9 Å. *Acta Crystallog. sect. D*, **58**, 2201–2204.
49. Marana, S. R. (2006). Molecular basis of substrate specificity in family 1 glycoside hydrolases. *IUBMB Life*, **58**, 63–73.
50. Otwinowski, Z. & Minor, W. (1997). A processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
51. Vagin, A. & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Crystallog.* **30**, 1022–1025.
52. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
53. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein methods in electron density maps and the location of errors in these models. *Acta Crystallog. sect A*, **47**, 110–119.
54. Laskowsky, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
55. Leslie, A. G. W. (1992). Molecular data processing. In *Crystallographic Computing 5: From Chemistry to Biology* (Moras, D., Podjarny, A. D. & Thierri, J. C., eds), pp. 39–50, Oxford University Press, Oxford, UK.
56. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
57. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
58. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T. & Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucl. Acids Res.* **33**, 299–302.

Edited by M. Guss

(Received 20 April 2007; received in revised form 22 May 2007; accepted 25 May 2007)
Available online 2 June 2007