THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 275, No. 30, Issue of July 28, pp. 23020–23026, 2000
Printed in U.S.A.

Substrate Specificity in Glycoside Hydrolase Family 10

STRUCTURAL AND KINETIC ANALYSIS OF THE *STREPTOMYCES LIVIDANS* XYLANASE 10A*

Received for publication, January 7, 2000, and in revised form, March 24, 2000
Published, JBC Papers in Press, March 27, 2000, DOI 10.1074/jbc.M000129

Valérie Ducros‡§, Simon J. Charnock‡, Urszula Derewenda¶, Zygmunt S. Derewenda¶, Zbigniew Dauter‖, Claude Dupont, François Shareck**, Rolf Morosoli**, Dieter Kluepfel**, and Gideon J. Davies‡ ‡‡**

*From the ‡Department of Chemistry, Structural Biology Laboratory, University of York, Heslington, York YO10 5DD, United Kingdom, the ¶Department of Molecular Physiology and Biological Physics, University of Virginia Health Sciences Center, P. O. Box 10011, Charlottesville, Virginia 22906-0011, the ‖Brookhaven National Laboratory, Building 725A-X9, Upton, New York 11973, and the **Centre de Microbiologie et Biotechnologie, INRS-Institut Armand-Frappier, Laval, Québec H7V 1B7, Canada*

Endoxylanases are a group of enzymes that hydrolyze the β-1,4-linked xylose backbone of xylans. They are predominantly found in two discrete sequence families known as glycoside hydrolase families 10 and 11. The *Streptomyces lividans* xylanase Xyl10A is a family 10 enzyme, the native structure of which has previously been determined by x-ray crystallography at a 2.6 Å resolution (Derewenda, U., Swenson, L., Green, R., Wei, Y., Morosoli, R., Shareck, F., Kluepfel, D., and Derewenda, Z. S. (1994) *J. Biol. Chem.* 269, 20811–20814). Here, we report the native structure of Xyl10A refined at a resolution of 1.2 Å, which reveals many features such as the rare occurrence of a discretely disordered disulfide bond between residues Cys-168 and Cys-201. In order to investigate substrate binding and specificity in glycoside hydrolase family 10, the covalent xylobiosyl enzyme and the covalent cellobiosyl enzyme intermediates of Xyl10A were trapped through the use of appropriate 2-fluoroglycosides. The α-linked intermediate with the nucleophile, Glu-236, is in a $^4C_1$ chair conformation as previously observed in the family 10 enzyme Cex from *Cellulomonas fimi* (Notenboom, V., Birsan, C., Warren, R. A. J., Withers, S. G., and Rose, D. R. (1998) *Biochemistry* 37, 4751–4758). The different interactions of Xyl10A with the xylobiosyl and cellobiosyl moieties, notably conformational changes in the –2 and –1 subsites, together with the observed kinetics on a range of aryl glycosides, shed new light on substrate specificity in glycoside hydrolase family 10.

Endo-xylanases (EC 3.2.1.8) attack the internal β-xylosidic glycosidic linkages of the xylan backbone. Xylan, the major constituent of hemicellulose, is composed of a β-1,4-linked D-xylose backbone substituted at the 2′ and 3′ positions with L-arabinofuranose, D-glucuronic acid, and 4-*O*-methylglucuronic acid, and it displays a varying degree of acetylation, depending on the source. Xylanases have as yet unrealized potential in the pulp and paper industries for prebleaching and bio-pulping applications. They also find application in the food industries, including such roles as reducing the viscosity of poultry feeds and the pretreatment of silage grasses. The majority of xylanases fall into families 10 and 11 of the sequence-based classification of glycoside hydrolases (1–3). Structures of enzymes from glycoside hydrolase family 10 have been described for the Xyl10A (nomenclature according to Ref. 4) from *Streptomyces lividans* (5), the *Cellulomonas fimi* enzyme Cex (6), the *Pseudomonas cellulosa* Xyl10A (7), and the Xyl10A from *Clostridium thermocellum* (8), *Thermoascus aurantiacus* (9, 10), and *Penicillium simplicissimum* (11). All of these structures are of the intact catalytic core domains alone (glycoside hydrolases are frequently modular enzymes 12), although recently there have been preliminary reports of an intact family 10 xylanase structure with its xylan-binding domain.[1] Family 10 xylanases perform catalysis with net retention of configuration. The mechanism is a double displacement in which a covalent intermediate is formed and then hydrolyzed via oxocarbenium ion-like transition states, essentially as described by Koshland in 1953 (Ref. 13; for review, see Ref. 14). The covalent intermediate for these reactions has been trapped (15), and the three-dimensional analyses of the trapped covalent enzyme intermediates of the family 10 enzyme Cex with both 2-fluoro-substituted and natural ligands represent a significant advance in our knowledge of the catalytic mechanism of retaining glycoside hydrolases (16–18).

One interesting feature of family 10 xylanases is that, in addition to their xylanolytic activity, they display a range of activities against glucose-derived substrates such as cellulose. The *S. lividans* enzyme, *Sl*Xyl10A, is primarily a xylanase with little activity against glucose-based polymers. Its three-dimensional structure, at 2.6 Å, was the first family 10 enzyme structure to be reported (5). In this paper, we present its native structure refined at atomic (1.2 Å) resolution, together with the structure of both the trapped 2-fluoroxylobiosyl enzyme intermediate at 1.65 Å and the trapped cellobiosyl enzyme intermediate at 1.7 Å. Kinetic analysis of a series of aryl glycosides has also been performed, which reveal a preference for xylo-derived

* This work was funded in part by the Biotechnology and Biological Sciences Research Council (UK), European Union Contract BIO4-CT97-2303, the Natural Sciences and Engineering Research Council of Canada, and the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec. The Structural Biology Laboratory at York is supported by the Biotechnology and Biological Sciences Research Council. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1eow, 1eov, and 1eox) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ Recipient of a European Union fellowship for the Training and Mobility of Researchers.

‡‡ A Royal Society University Research Fellow. To whom correspondence should be addressed. Tel.: 44-1904-432596; Fax: 44-1904-410519; E-mail: davies@ysbl.york.ac.uk.

[1] Z. Fujimoto, personal communication.

substrates. Comparison of Xyl10A with the family 10 *C. fimi* enzyme Cex, which displays a more liberal substrate tolerance, gives insight into substrate specificity in this family of glycoside hydrolases.

MATERIALS AND METHODS

Enzyme Production and Purification—*S. lividans* strain IAF 19 was used for production of the truncated form of the Xyl10A (5). Seven-day-old cultures of *S. lividans* from Bennett-thiostreptron plates (19) were used as initial inoculum. The spores were scraped from the plates and inoculated into 12.5 ml of minimal M14 medium (20) and incubated for 18 h at 34 °C with agitation. Bacteria were recovered by centrifugation, used to inoculate 500 ml of the same medium, and allowed to grow for 69 h under the same conditions. Proteins were recovered from the culture supernatant by ultrafiltration on a 3-kDa cutoff membrane (Omega) and then dialyzed against 20 mM Tris-HCl buffer, pH 8.3, overnight. A 200-mg sample was adsorbed on a AP-2 15HR DEAE-HPLC column (Waters), and the proteins were eluted with a linear gradient to 1 M NaCl. The Xyl10A-containing fractions were pooled and concentrated on a 3-kDa cutoff membrane (Omega) using a pressurized stirred Amicon cell. Further purification to homogeneity was achieved by separation on a Superdex HR75 beaded column (3 × 60 cm) (Amersham Pharmacia Biotech) with 100 mM sodium citrate, pH 6.0, as the eluent. The purified Xyl10A-containing fractions were pooled, dialyzed, and freeze-dried.

Protein Analysis—Protein concentration was determined using bovine serum albumin as standard (Bio-Rad) (21). Determination of protein purity was assessed by SDS-polyacrylamide gel electrophoresis (22) and Western blot analysis with anti-Xyl10A antibodies (23). Analytical isoelectric focusing was carried out on PhastGel (Amersham Pharmacia Biotech) containing Pharmalyte carrier ampholytes from pH 3 to pH 9 using the automated PhastSystem (Amersham Pharmacia Biotech). The gels were silver-stained after the run according to the manufacturer's instructions.

Native Structure—Crystals were grown as described previously for the 2.6 Å analysis (5). Native data were collected, to 1.2 Å resolution, from a single crystal at room temperature at the EMBL Hamburg outstation, beamline X-31, at a wavelength of 0.87 Å. The crystal was mounted with the crystallographic b(b*) axis almost perpendicular to the spindle axis but offset sufficiently to prevent loss of data in the "blind region." Data were collected on a MAR 18 cm scanner, in three resolution sweeps to allow measurement of all intensities within the dynamic limitations of the detector. Data were processed and reduced with the DENZO and SCALEPACK programs (31, 32). Crystals belong to the space group $P2_12_12_1$ with the following cell dimensions: a = 70.25 Å, b = 46.93 Å, c = 86.39 Å, and $\alpha = \beta = \gamma = 90°$. There is a single molecule of Xyl10A in the asymmetric unit. Initial refinement used SHELXL (25). At a later stage, 5% of the observations were set aside for cross validation analysis (26) and were used to monitor various refinement strategies, such as geometric and temperature factor restraint values and the insertion of solvent water, and as the basis for the maximum likelihood refinement using the REFMAC program (24). Manual corrections of the model using the X-FIT routines of the program QUANTA (Molecular Simulations Inc.) were interspersed with cycles of least squares refinement using the maximum likelihood program REFMAC. Water molecules were added in an automated manner using ARP (27) and inspected manually prior to deposition. When isotropic refinement had converged (R_{cryst}, 0.14; R_{free}, 0.17), the model was extended to include calculated hydrogen scattering from "riding" hydrogens and refined with a restrained anisotropic treatment of the atomic displacement parameters as implemented in REFMAC (28).

Trapped Xylobiosyl Enzyme Intermediate—Prior to crystallization, the enzyme was washed with water on Centricon 10K membranes and concentrated to 30 mg ml^{-1}. The crystallization conditions described for the native enzyme (5) were modified in order to grow crystals under catalytically active conditions. Xyl10A in solution is active at pH 7.0, and the previously reported crystallization conditions were at pH 5.0. The protein was crystallized by the hanging drop method using a mother liquor composed of 0.1 M sodium HEPES buffer, pH 7.5, and 18% (w/v) polyethylene glycol 5000 (36% (v/v) of a 50% (w/v) stock solution), in the presence of 10% (v/v) isopropanol. The hanging drop consisted of 1 μl of protein solution together with 1 μl of mother liquor. Crystals grew over a period of 10–12 days to a maximum size of 0.1 × 0.4 × 0.4 mm. The soaking experiment was performed in a 10-μl drop of the reservoir solution in the presence of a small quantity of powdered 2′,4′-dinitrophenyl 2-deoxy-2-fluoro-β-xylobioside (a gift from Prof. S. G. Withers, University of British Columbia). Crystals were soaked for 12 h. Data collection was performed under cryogenic conditions. A single crystal was mounted in a rayon fiber loop and placed in a boiling nitrogen stream at 120 K. A cryoprotectant solution was made of 0.1 M sodium HEPES buffer, pH 7.5, 20% (w/v) polyethylene glycol 5000, 10% (v/v) isopropanol with the addition of glycerol to a final concentration of 15% (v/v). Data were collected using MAR Research image plate system together with a copper rotating anode and utilizing long, focusing, mirror optics (Yale/Molecular Structure Corp.). A total of 180° of data, to a resolution of 1.65 Å, were collected with an oscillation range of 1.0° per image. Data were processed and reduced using the DENZO and SCALEPACK programs (24). All further calculations used the CCP4 suite of programs unless otherwise stated. The crystals belong to the space group $P2_1$ with cell dimensions a = 42.21 Å, b = 81.06 Å, c = 131.62 Å and β = 102.79. There are two molecules in the asymmetric unit, giving a V_M of 2.7 $Å^3$/Da and a solvent content of approximately 44%. The native Patterson function, calculated at 4 Å resolution, revealed a peak with height approximately 62% of the origin peak height at position u = 0.50, v = 0.00, and w = 0.50 (not shown), which indicates that the two molecules lie in approximately the same orientation within the asymmetric unit.

The complexed structure was solved by molecular replacement using the coordinates of the native Xyl10A as a search model using the program AMoRE (29, 30). Five percent of the observations were immediately set aside for cross validation analysis (26) and were used to monitor various refinement strategies, such as geometric and temperature factor restraint values and the insertion of solvent water, and as the basis for the maximum likelihood refinement using the REFMAC program (24). The starting crystallographic R factor for the molecular replacement solution was 0.418 for all data between 15.0 and 1.65 Å. Manual corrections of the model using the X-FIT routines of the program QUANTA (Molecular Simulations Inc., San Diego, CA) were interspersed with cycles of least squares refinement using the maximum likelihood program REFMAC. Water molecules were added in an automated manner using ARP (27) and inspected manually prior to coordinate deposition.

Trapped Cellobiosyl Enzyme Intermediate—The protein was crystallized by the hanging drop method using a mother liquor composed of 0.1 M sodium HEPES buffer, pH 7.5, and 16% (w/v) polyethylene glycol 4000, in the presence of 10% (v/v) isopropanol. The hanging drop consisted of 1 μl of protein solution together with 1 μl of mother liquor. Crystals grew over a period of 10–30 days to a maximum size of 0.2 × 0.1 × 0.3 mm. The soaking experiment was performed in a 10-μl drop of the reservoir solution in the presence of a small quantity of powdered 2′,4′-dinitrophenyl 2-deoxy-2-fluoro-β-cellobioside (2F-$DNPG_2$), and data were collected essentially as described above for the xylobioside complex. A total of 180° of data, to a resolution of 1.70 Å, were collected with an oscillation range of 0.5° per image. Data were processed and reduced using the DENZO and SCALEPACK programs (31, 32). All further calculations used the CCP4 suite of programs unless otherwise stated. The crystals belong to the space group $P2_12_12_1$ with cell dimensions of a = 67.87 Å, b = 46.27 Å, c = 87.07 Å, and $\alpha = \beta = \gamma = 90°$. There is a single molecule of Xyl10A in the asymmetric unit. Because this cell is almost isomorphous with the original native cell, the same cross validation subset of reflections was used for refinement. Refinement was performed using the 1.2 Å native structure as the starting model with initial rigid body refinement at 6 Å, using AMoRE, to accommodate the cell shrinkage upon freezing and then continued as for the xylobioside complex.

Enzyme Assays on Aryl Glycosides—All reactions were performed in 50 mM sodium phosphate buffer, pH 7.0, at 37 °C with the addition of bovine serum albumin to a final concentration of 1mg ml^{-1}. The release of phenolic chromophores was monitored using a 500-μl quartz cuvette with a 10-mm path length at a wavelength of 400 nM. Data were recorded on an ATI Unicam UV/VIS *UV2* spectrophotometer and interpreted by nonlinear regression using the GRAFIT software. Wherever possible, a substrate concentration in the range of 0.2–2 times the K_m was used.

RESULTS AND DISCUSSION

Native Enzyme Structure at 1.2 Å—Data were collected from a single native enzyme crystal at room temperature at the EMBL Hamburg outstation to a resolution of 1.2 Å. The final data are 96% complete to 1.2 Å resolution with an overall R_{merge} ($\Sigma_{hkl}\ \Sigma_i|I_{hkli} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\ \Sigma_i\langle I_{hkl}\rangle$) of 0.051, a mean $I/\sigma(I)$ of 24, and a mean multiplicity of observations of 4.3 observations/reflection (Table I). The native structure has been

TABLE I
Refinement and structure quality statistics for the native and 2-F covalent-enzyme intermediate complexes of the S. lividans Xyl10A

	Native structure	Cellobiosyl enzyme intermediate	Xylobiosyl enzyme intermediate
Data quality			
Resolution of data (outer shell) (Å)	15–1.2 (1.22–1.20)	15–1.7 (1.76–1.70)	15–1.65 (1.71–1.65)
R_{merge} (outer shell)[a]	0.051 (0.35)	0.032 (0.121)	0.033 (0.097)
Mean I/σI (outer shell)	23.6 (3.4)	36.5 (9.4)	35.6 (12.9)
Completeness (outer shell) (%)	96 (91)	97 (82)	99 (94)
Multiplicity (outer shell)	4.3 (3.4)	4.3 (4.9)	3.5 (4.8)
Temperature (K)	293	120	120
Refinement			
PDB Code	1eow[b]	1eov[b]	1eox[b]
No. protein atoms/molecule	2385 (residues 1–302)	2385 (residues 1–302)	2410 (residues 1–309)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_1$
No. of protein molecules in asymmetric unit	1	1	2
No solvent waters	502	503	731
Resolution used in refinement (Å)	15–1.2	15–1.70	15–1.65
R_{cryst}	0.09	0.15	0.12
R_{free}	0.12	0.19	0.16
r.m.s. deviation 1–2 bonds (Å)	0.010	0.007	0.012
r.m.s. deviation 1–3 angles (Å)	0.022	0.022	0.028
r.m.s. deviation chiral volumes (Å^3)	0.098	0.093	0.123

[a] $R_{merge} = \Sigma_{hkl}\,\Sigma_i|I_{hkli} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\,\Sigma_i\,\langle I_{hkl}\rangle$.
[b] Protein Data Bank codes.

refined with anisotropic treatment of the atomic displacement parameters, initially using SHELXL (25) and later with a maximum likelihood treatment using REFMAC (24, 28). The final model structure consists of residues 1–302 and 506 solvent water molecules. This model has a crystallographic R value of 0.09, with a corresponding R_{free} of 0.12 for all observed data between 15 and 1.2 Å resolution. The deviations from stereochemical target values are 0.010 and 0.022 Å for the 1–2 and 1–3 bonding distances, respectively. Final refinement statistics are given in Table I. All of the nonglycine residues have conformational angles (φ and ψ) in permitted regions of the Ramachandran plot (33), with none of these in the "generously allowed" or "disallowed" regions as defined by PROCHECK (34).

The native structure of Xyl10A is essentially as reported previously (5). Xyl10A folds to form a standard $(\beta/\alpha)_8$ barrel with the two catalytic functions, the acid/base and the nucleophile on strands β-4 and β-7, as expected (3, 35–37) (Fig. 1). The structure refined at 1.2 Å resolution extends from residue 1 to residue 302 with no breaks or disorder in the main chain density. A number of residues show discrete double conformations, including what we believe to be a rare observation, a doubly configured disulfide between Cys-168 and Cys-201 (Fig. 2). The SG of Cys-201 is found in two equally occupied positions, both of which are involved in disulfide bond formation with the SG of Cys-168. We can exclude the possibility that the second conformation is the result of a partially free cysteine because it lies within S-S bonding distance of the adjacent SG and not at a van der Waals distance, as would be expected from a free Cys-SH. As was observed in the Cex structure (6), there is a nonprolyl *cis*-peptide between His-81 and Thr-82 (Fig. 3). This permits His-81 to hydrogen bond to the O(3) hydroxyl of the –1 subsite sugar, described in more detail below. The *S. lividans* Xyl10A overlaps well with other members of this sequence family. The native *S. lividans* Xyl10A gives r.m.s. overlaps of 0.85 Å (251 equivalent Cα atoms) with the *C. fimi* enzyme Cex (6), 0.98 Å (277 equivalent Cα atoms) with the Xyl10A from *C. thermocellum* (8), 1.0 Å (267 equivalent Cα atoms) for the *T. aurantiacus* xylanase (9), 1.1 Å (281 equivalent Cα atoms) for the xylanase from *P. simplicissimum* (11), and 1.3 Å (270 equivalent Cα atoms) with the *P. cellulosa* Xyl10A (7) (all overlaps calculated using LSQMAN in Ref. 38). In this sense, family 10 reveals itself to be much more structurally conserved than other glycoside hydrolase families, such as families 5 and 13, the members of which display a much greater diversity of sequence and structure.

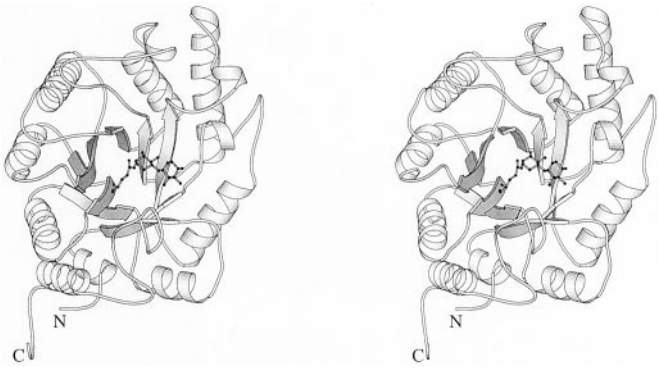

FIG. 1. **Divergent stereo ribbon diagram for the trapped xylobiosyl enzyme intermediate complex of Xyl10A.** The xylobioside moiety and the nucleophile Glu-236 are shown by a ball-and-stick representation. This figure was drawn with the MOLSCRIPT program (55).

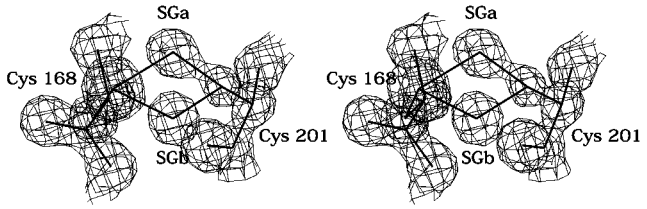

FIG. 2. **Observed electron density for the doubly conformed disulfide bridge between Cys-168 and Cys-201.** The 1.2-Å map is a maximum likelihood/σ_A weighted $2F_o - F_c$ synthesis at a contour level of 1.0 electron/Å^3 and is shown in divergent stereo.

Structure of the Trapped 2-F Xylobiosyl Enzyme Intermediate at 1.65 Å—A summary of the data quality and completeness is given in Table I. The data consist of 242585 observations of 66881 unique reflections. The final data are 99.3% complete to 1.65 Å resolution with an overall R_{merge} ($\Sigma_{hkl}\,\Sigma_i|I_{hkli} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\,\Sigma_i\langle I_{hkl}\rangle$) of 0.033, a mean I/$\sigma$(I) of 36, and a mean multiplicity of observations of 3.5 observations/reflection. The final model, featuring residues 1–309, has a crystallographic R value of 0.126, with a corresponding R_{free} of 0.162 for observed data between 20 and 1.65 Å resolution. This model has deviations from stereochemical target values of 0.012 and 0.028 Å (corre-

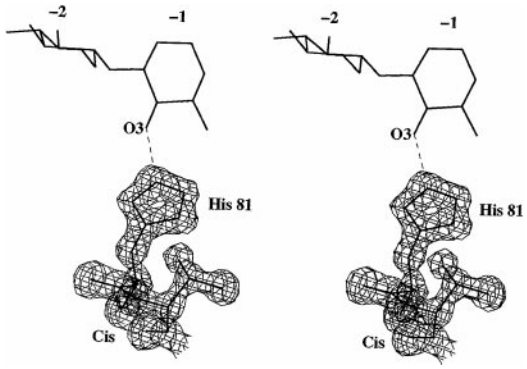

FIG. 3. **Observed electron density for the nonprolyl *cis*-peptide between Cys-168 and Cys-201 that permits His-81 to interact with substrate.** The 1.2-Å map is a maximum likelihood/σ_A weighted $2F_o - F_c$ synthesis at a contour level of 1.0 electron/Å^3 and is shown in divergent stereo.

sponding to approximately 1.1°), for 1–2 and 1–3 bonds, respectively. Final refinement statistics for the 1.65 Å complex structure are given in Table I. As with the native structure, all of the nonglycine residues have conformational angles (φ and ψ) in permitted regions of the Ramachandran plot (33), with none of these in the generously allowed or disallowed regions as defined by PROCHECK (34).

The structure of the trapped xylobiosyl enzyme intermediate, in space group $P2_1$, reveals more extended density for the C-terminal tail than was observed in the native enzyme structure, in which just 302 amino acids could be successfully modeled. Mass spectrometry data (not shown) indicate that the crystallized material in both cases is a 313-amino acid species. We assume that the remaining C-terminal amino acids are disordered in-crystal, the additional residues seen in the complex structure presumably resulting from slight changes in crystal-packing environment and not a lower degree of disorder due to crystal freezing. The structure for the frozen cellobiosyl enzyme intermediate, in the same space group ($P2_12_12_1$) as the room temperature native structure, is also only visible to residue 302. Other than this, no conformational changes are revealed upon formation of the intermediate. The native and xylobiosyl enzyme intermediate structures overlap with an r.m.s. deviation of 0.45 Å for all 302 equivalent Cα atoms.

Xyl10A is a retaining glycoside hydrolase. The mechanism features the formation, and subsequent breakdown, of a covalent glycosyl enzyme intermediate flanked by oxocarbenium ion-like transition states (for reviews, see Refs. 14, 39, and 40) (Fig. 4). The 2-fluoro-xylobiosyl enzyme intermediate was trapped through the use of a 2′,4′-dinitrophenyl 2-deoxy-2-fluoro-β-xylobioside, the basis of a trapping experiment using 2-fluoroglycosides having been described previously (see, for example, Refs. 15, 17, and 41). The 2-F xylobiosyl enzyme intermediate for Xyl10A forms an α-ester linkage to the catalytic nucleophile, Glu-236, as expected (Fig. 5). The –1 subsite sugar ring is in the $^4C_1$ chair conformation, as described previously for family 10 xylanase from *C. fimi* Cex (17, 18) but markedly different form the $^{2,5}B$ boat conformation observed for the xylanases from family 11 (42, 43). In contrast to many other systems studied, but consistent with other family 10 structures, the carbonyl oxygen of the nucleophile makes a close interaction (2.7 Å) with the fluorine atom at C2. In the equivalent family 1, 5, and 12 complexes, there appears to be more conformational freedom for the nucleophile, which is able to rotate to avoid this unfavorable interaction (see, for example, Refs. 41, 44, and 45). The catalytic acid/base lies *anti* to the endocyclic O(5)-C(1) bond (as defined by Heightman and Vasella (46)). In the trapped covalent intermediate, it makes a 2.7

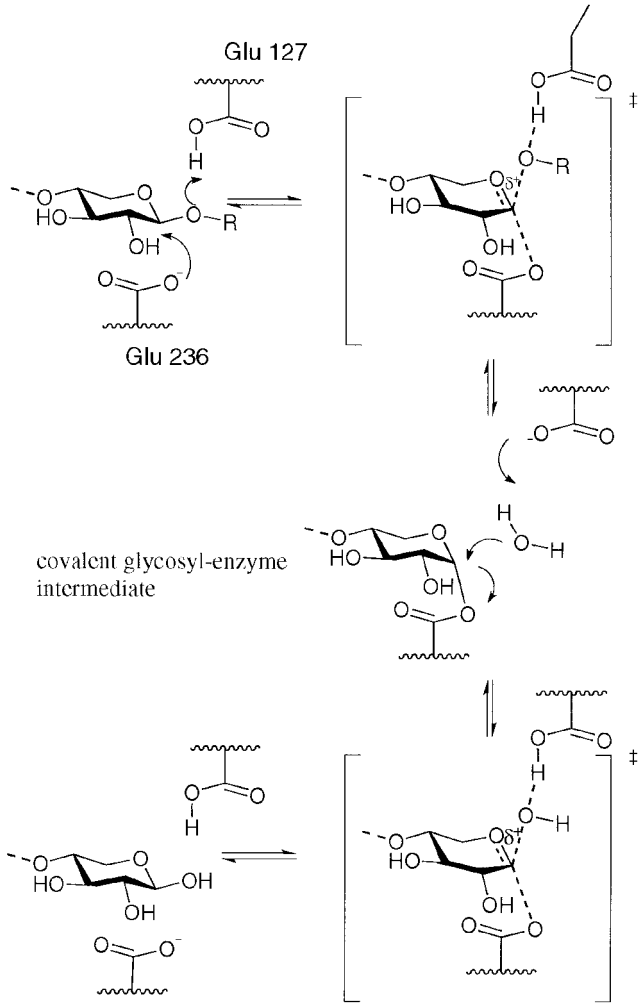

FIG. 4. **Double displacement reaction mechanism as applied to the *S. lividans* Xyl10A.**

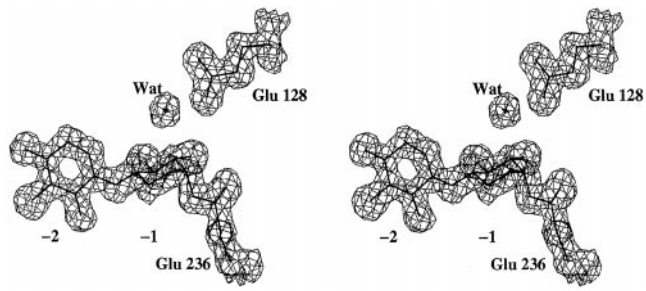

FIG. 5. **Observed electron density for the xylobiosyl enzyme intermediate.** The 1.65 Å map is a maximum likelihood/σ_A weighted $2F_o - F_c$ synthesis at a contour level of 0.47 electrons/Å^3 and is shown in divergent stereo.

Å hydrogen bond to a solvent water molecule poised to make a nucleophilic attack at C1 (Figs. 5 and 6). In the –2 subsite, the O(4) hydroxyl interact only with solvent water molecules, whereas O(3) makes a 3.05 Å hydrogen bond to Asn-45. The O(2) hydroxyl interacts both with the carboxylate of Glu-44 and with the indole NH of Trp-266. A lysine, Lys-48, appears to be donating a hydrogen bond to the ring O(5) oxygen. In the –1 subsite, Trp-274 forms a hydrophobic lid to the binding site stacking over the –1 subsite xylose ring, whereas hydrogen bonds to O(3) are provided by Lys-48 and His-81. The 2-position fluorine makes hydrogen bonds with His-81, Asn-127, and the side chain carbonyl oxygen of the nucleophile Glu-236, described above.

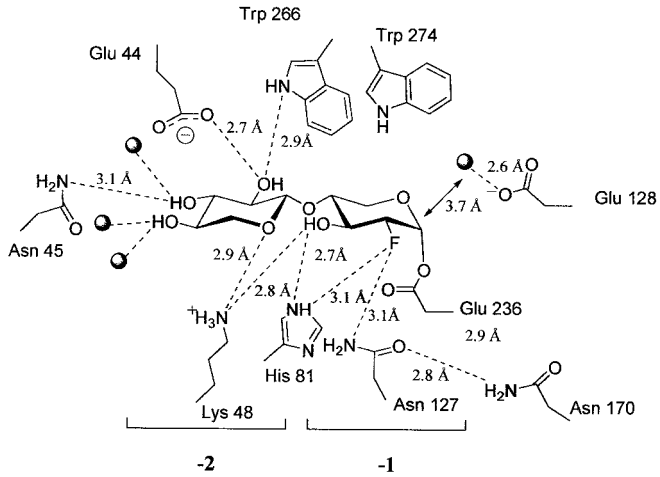

FIG. 6. **Schematic diagram of the protein-ligand interactions for the 2F-xylobiosyl enzyme intermediate of Xyl10A.** Distances less than 3.2 Å are indicated.

Noncovalent interactions of the hydroxyl substituent at C(2) are known to make a particularly important contribution to transition state stabilization (16, 47). The major factor may be the interaction between the carbonyl of the nucleophile with the O(2) hydroxyl at the transition state (18). Of the other O(2) interactions, site-directed mutagenesis of the *S. lividans* enzyme indicates a catalytic role, as opposed to substrate binding role, for Asn-127 (48). Structural work on the *C. fimi* enzyme also points to a catalytic role for this residue, in which it functions in the correct positioning of the nucleophile through a hydrogen bond network with a second Asn, equivalent to Asn-170 in Xyl10A (Fig. 6). These proposals are supported by the structures of the related family 5 mannanase from *Thermomonospora fusca* and *Trichoderma reesei* (49, 50). These enzymes recognize the C(2) epimer of glucose, yet they retain an equivalent Asn at this position. The role of His-81 has been confirmed by site-directed mutagenesis. Mutation of His-81 of Xyl10A to serine, reduces k_{cat} on polymeric substrates by a factor of 20 (51).

Substrate Specificity in Family 10 Xylanases—Enzymes from glycoside hydrolase family 10 are essentially xylanolytic but show a range of catalytic activities against glucose-derived substrates. The *C. fimi* enzyme Cex is frequently described as multifunctional because its activity on aryl glycosides favor xylose-derived substrates over glucose-derived ones by a factor of approximately 40 (17). On highly polymeric or natural substrates, however, Cex is a substantially better xylanase than a cellulase. The latter comparison is confused by the significant ambiguity in the reported activity on xylan with a 1000-fold difference between birchwood xylan (17) and xylan from oat spelts (52): 0.47 s^{-1} *versus* 422 s^{-1}. One possibility is that the different degree of acetylation renders birchwood xylan a substantially poorer substrate. The noncovalent interactions of the 2-F-xylobiosyl enzyme intermediate of Xyl10A are essentially identical to that described previously for the *C. fimi* enzyme (17, 18). The fact that the interactions appear to be almost identical, yet the kinetics on glucose and xylose-derived substrates appear dissimilar, is intriguing.

Kinetic data were therefore collected for both Cex and Xyl10A using a range of paranitrophenol-based glycosides: PNPX,[2] $PNPX_2$, PNPG, and $PNPG_2$ (Table II). The *S. lividans* enzyme displays values typical for a family 10 xylanase, such as the *P. cellulosa* enzyme, the values for which are included for comparison (52). Compared with the *C. fimi* enzyme, Xyl10A displays a much lower activity on all aryl glycosides. Even on the best substrate tested, PNP-xylobiose, the *S. lividans* displays a k_{cat}/K_m some 38-fold lower than the *C. fimi* enzyme. Others have pointed out that these results may reflect the ability of Cex to use aryl-glycosides as substrates more efficiently than other family 10 enzymes, because the activity of Cex on natural substrates, such as xylotriose, is very similar to other family 10 xylanases (52, 53). In the accompanying paper by Andrews *et al.* (53), kinetics on 1-F-substituted substrates indicate that it is the preferential ability of Cex to utilize binding energy from aromatic aglycons that gives rise to these differences. The data on aryl xylo and glucosides reveal that, in addition to being a worse catalyst for these substrates, Xyl10A also shows less tolerance for glucose-based substrates compared with Cex. The ratio of k_{cat}/K_m (Cex_{PNPX2}) to k_{cat}/K_m (Cex_{PNPG2}) shows that Cex hydrolyses aryl xylobiosides approximately 140 times more efficiently than the corresponding cellobiosides. For Xyl10A, the equivalent ratio is over 714. This indicates that Xyl10A is approximately 5 times more specific for aryl xylobiosides than for any cellobiosides, with a value of approximately 10 achieved with the respective monosaccharide derivatives.

Notenboom *et al.* (17) have proposed that two structural features are responsible for the ability of Cex to accommodate glucose-derived substrates in the –1 and –2 subsites. In the –1 subsite, Trp-281 (equivalent to Trp-274 in Xyl10A) moves to permit binding of the C(6)-OH group of glucose. In the –2 subsite, Gln-87 (equivalent to Gln-88 in Xyl10A) also moves, becoming disordered beyond C_β, to accommodate the exocyclic C(6)-OH substituent in this subsite (17). Because Xyl10A has interactions identical to those of Cex in these subsites, described above, it seems likely that its lower tolerance for glucose-derived substrates results from an impaired ability to undergo the required conformational changes. Trp-274 over the –1 subsite is free to move in the Cex enzyme; in Xyl10A, however, this tryptophan residue stacks below the bulky side chain of Arg-275, with which it makes close van der Waals interactions. Any conformational rearrangement would perhaps also involve a concerted motion of Arg-275 at a consequently greater energetic penalty.

Structure of the Trapped 2-F Cellobiosyl Enzyme Intermediate at 1.70 Å—In order to investigate this hypothesis, we have determined the structure of the 2-F cellobiosyl enzyme intermediate for the *S. lividans* Xyl10A at 1.7 Å resolution. A summary of the data and model quality is given in Table I. The final data are 97% complete to 1.70 Å resolution with an overall R_{merge} ($\Sigma_{hkl}\,\Sigma_i|I_{hkli} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\,\Sigma_i\,\langle I_{hkl}\rangle$) of 0.032, a mean $I/\sigma(I)$ of 36.5, and a mean multiplicity of observations of 4.3 observations/reflection. The final model, featuring residues 1–302, has a crystallographic R value of 0.15, with a corresponding R_{free} of 0.19 for all observed data between 15 and 1.7 Å resolution. This model has deviations from stereochemical target values of 0.007 and 0.022 Å (corresponding to approximately 1.1°), for 1–2 and 1–3 bonds, respectively.

The trapping experiment, with the 2-F cellobioside, was again performed with crystals grown at pH 7.0. The cellobiosyl enzyme intermediate binds a covalent α-ester linkage to the catalytic nucleophile, Glu-236, as expected (Fig. 7). The –1 subsite ring is in $^4C_1$ chair conformation, but there is no distortion away from sp^3 geometry for C(2), as was observed for the equivalent complex of the *C. fimi* enzyme (16). In order to accommodate the C(6)-OH group of glucose, the active site undergoes conformational changes, as predicted. In the –2 subsite, Gln-88 moves to permit binding of the C(6)-OH substitu-

[2] The abbreviations used are: PNPX, 4-nitrophenyl β-D-xyloside; $PNPX_2$, 4-nitrophenyl β-D-xylobioside; PNPG, 4-nitrophenyl β-D-glucoside; $PNPG_2$, 4-nitrophenyl β-D-cellobioside.

TABLE II
Kinetic parameters for the hydrolysis of β-glycosides by S. lividans Xyl10A (this study), P. cellulosa Xyl10A (from Ref. 52) and C. fimi Cex (this study plus Ref. 17)

Substrate	Enzyme	k_{cat}	K_M	k_{cat}/K_M
		s^{-1}	mM	$s^{-1}\ mM^{-1}$
PNPG	Xyl10A, *S. lividans*[a]	0.011	183	6.9×10^{-4}
	Xyl10A, *P. cellulosa*[b]	ND[c]	ND	ND
	Cex, *C. fimi*[a]	0.096 (0.024)	15.9 (8.3)	6.0×10^{-3} (2.9×10^{-3})
$PNPG_2$	Xyl10A, *S. lividans*[a]	3.6	50.5	0.071
	Xyl10A, *P. cellulosa*[b]	2.6	50	0.052
	Cex, *C. fimi*[a]	14.5 (15.8)	1.05 (0.60)	13.8 (26.3)
PNPX	Xyl10A, *S. lividans*[a]	0.32	375	8.1×10^{-4}
	Xyl10A, *P. cellulosa*[b]	0.15	308	4.8×10^{-4}
	Cex, *C. fimi*[a]	1.9 (2.6)	20.8 (20)	0.09 (0.13)
$PNPX_2$	Xyl10A, *S. lividans*[a]	29.4	0.58	50.7
	Xyl10A, *P. cellulosa*[b]	86.4	0.57	151.6
	Cex, *C. fimi*[a]	37.1 (39.8)	0.019 (0.018)	1952 (2200)

[a] This study (data for Cex according to Ref. 17 are shown in parentheses for comparison).
[b] Data from Ref. 52.
[c] ND, not detectable.

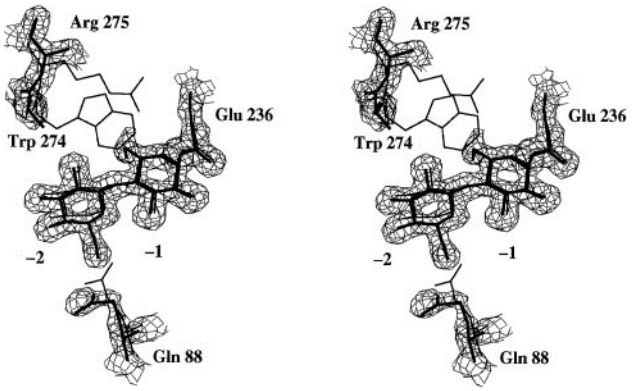

FIG. 7. **Observed electron density for the cellobiosyl enzyme intermediate.** The 1.7-Å map is a maximum likelihood/σ_A weighted $2F_o - F_c$ synthesis at a contour level of 0.40 electrons/Å^3 and is shown in divergent stereo. The cellobiosyl enzyme coordinates are shown in *boldface,* and the xylobiosyl enzyme intermediate coordinates are overlaid in *faint lines*.

ent. It does not become disordered, as was observed with the Cex enzyme (16, 17), but instead, the side chain rotates approximately 120 degrees around Ç$_2$, displacing the amide moiety by approximately 4 Å. The NZ of Lys-48 makes a small movement of 0.5 Å, so that in the cellobiosyl enzyme intermediate, it interacts with the C(6)-OH of the –2 subsite sugar, rather than the O(3) hydroxyl in –1, as was observed in the xylobiosyl enzyme intermediate.

The greatest conformational changes take place in the –1 subsite and give insight into the difference in the apparent specificities between Cex and Xyl10A. In Cex, a C(6)-OH substituent was simply accommodated by a small rotation of the Trp side chain around Ç$_1$. In Xyl10A, the equivalent tryptophan, Trp-274, stacks against the side chain of Arg-275. In the Xyl10A cellobiosyl enzyme intermediate, both Trp-274 and Arg-275 become completely disordered. No electron density is visible beyond CB for Arg-275, and even the CB density is poor for Trp-274. This side chain disorder is matched by main chain movements of 0.3–0.7 Å for the region from Ser-273 to Ser-276 (Fig. 7). This does seem to confirm that there is a greater structural and energetic penalty for glucose binding to the –1 subsite of Xyl10A than there is in Cex. Similar proposals have been made for the xylan specificity of the Xyl10A from *P. cellulosa*, in which Leu-314 similarly blocks movement of the –1 subsite tryptophan in a way similar to the role of Arg-275 in the *S. lividans* enzyme. These proposals have recently been tested by site-directed mutagenesis, and indeed, the L314A mutation improves the relative activity of the *P. cellulosa* enzyme for glucose over xylose-derived substrates by a factor of almost 6500 (53).

Conclusions—Family 10 xylanases display a range of tolerance for glucose-derived substrates, such as aryl cellobiosides. Rose and co-workers (16, 17) demonstrated that for the *C. fimi* enzyme, Cex discrimination resides, at least in part, in the –2 and –1 subsites, components of which must move in order to accommodate the additional hydroxymethyl group of glucose. For the *S. lividans* XylA, described here, a similar mechanism applies, but movement of the equivalent tryptophan in the –1 subsite is sterically hindered by a stacking interaction with the adjacent arginine. Both residues move, in concert with a main chain displacement, to allow formation of the cellobiosyl enzyme intermediate. Such an upheaval leaves both residues disordered and contributing little to the structural integrity of the binding site, all of which is reflected in a lower tolerance for glucose-based substrates. Family 11 xylanases are structurally unrelated to the family 10 enzymes but strongly linked to the cellulases from family 12 in glycoside hydrolase clan GH-C (45, 54). Similar steric considerations also apply for the specificity of these enzymes. The three-dimensional structures of family 11 and 12 enzymes and their catalytic machinery are conserved; specificity for xylose over glucose derived substrates again resides, to a large extent, in the –1 subsite. In the xylanolytic members of the clan, an invariant hydrophobic residue both prevents access to glucose-based substrates and also assists in favoring an unusual $^{2,5}B$ (boat) ring conformation that is less accessible to a glucopyranose ring (42, 43). The corresponding cellulases from family 12 bind the intermediate in a $^4C_1$ (chair) conformation. In the cellulases, the environment both permits the C(6)-OH substituent and provides hydrogen bond partners for the O(6) hydroxyl through subtle reorientation of an invariant tryptophan residue over the –2 subsite.

Many of the 77 sequence-based glycoside hydrolase families contain enzymes that display broad substrate specificity. The mechanisms whereby family members discriminate between different substrates is not widely understood. In recent years, the development of appropriate substrates to trap stable enzyme-bound species has allowed investigation of closely related enzymes that display different specificities. In the accompanying paper, Gilbert and co-workers (53) demonstrate that such structural and kinetic information may be used to guide a protein engineering program to change the substrate specificity of these enzymes. The present work opens up possibilities for the tailoring of glycoside hydrolase specificity for novel applications.

Acknowledgments—We are grateful to Tony Warren and Annabelle Varrot for supplying S. J. C. with Cex and to Prof. Harry Gilbert for useful discussions.

REFERENCES

1. Henrissat, B. (1991) *Biochem. J.* **280,** 309–316
2. Henrissat, B., and Bairoch, A. (1993) *Biochem. J.* **293,** 781–788
3. Henrissat, B., and Bairoch, A. (1996) *Biochem. J* **316,** 695–696
4. Henrissat, B., Teeri, T. T., and Warren, R. A. (1998) *FEBS Lett.* **425,** 352–4
5. Derewenda, U., Swenson, L., Green, R., Wei, Y., Morosoli, R., Shareck, F., Kluepfel, D., and Derewenda, Z. S. (1994) *J. Biol. Chem.* **269,** 20811–20814
6. White, A., Withers, S. G., Gilkes, N. R., and Rose, D. R. (1994) *Biochemistry* **33,** 12546–12552
7. Harris, G. W., Jenkins, J. J., Connerton, I., Cummings, N., Leggio, L. L., Scott, M., Hazlewood, G. P., Laurie, J. I., Gilbert, H. J., and Pickersgill, R. W. (1994) *Structure* **2,** 1107–1116
8. Dominguez, R., Souchon, H., Spinelli, S., Dauter, Z., Wilson, K. S., Chauvaux, S., Béguin, P., and Alzari, P. M. (1995) *Nature Struct. Biol.* **2,** 569–576
9. Lo Leggio, L., Kalogiannis, S., Bhat, M. K., and Pickersgill, R. W. (1999) *Proteins* **36,** 295–306
10. Natesh, R., Bhanumoorthy, P., Vithayathil, P. J., Sekar, K., Ramakumar, S., and Viswamitra, M. A. (1999) *J. Mol. Biol.* **288,** 999–1012
11. Schmidt, A., Schlacher, A., Steiner, W., Schwab, H., and Kratky, C. (1998) *Protein Sci.* **7,** 2081–2088
12. Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. J. (1991) *Microbiol. Rev.* **55,** 303–315
13. Koshland, D. E. (1953) *Biol. Rev.* **28,** 416–436
14. Davies, G., Sinnott, M. L., and Withers, S. G. (1997) in *Comprehensive Biological Catalysis* (Sinnott, M. L., ed) Vol. 1, pp. 119–209, Academic Press, London
15. Withers, S. G., Street, I. P., Bird, P., and Dolphin, D. H. (1987) *J. Am. Chem. Soc.* **109,** 7530–7531
16. White, A., Tull, D., Johns, K., Withers, S. G., and Rose, D. R. (1996) *Nat. Struct. Biol.* **3,** 149–154
17. Notenboom, V., Birsan, C., Warren, R. A. J., Withers, S. G., and Rose, D. R. (1998) *Biochemistry* **37,** 4751–4758
18. Notenboom, V., Birsan, C., Nitz, M., Rose, D. R., Warren, R. A. J., and Withers, S. G. (1998) *Nat. Struct. Biol.* **5,** 812–818
19. Jones, K. J. (1949) *J. Bacteriol.* **57,** 141–145
20. Mondou, F., Shareck, F., Morosoli, R., and Kluepfel, D. (1986) *Gene* **49,** 323–329
21. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) *J. Biol. Chem.* **193,** 265–275
22. Laemmli, U. K. (1970) *Nature* **227,** 680–685
23. Towbin, H., Staehelin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76,** 4350–4354
24. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53,** 240–255
25. Sheldrick, G. M., and Schneider, T. R. (1997) *Methods Enzymol.* **277,** 319–343
26. Brünger, A. T. (1992) *Nature* **355,** 472–475
27. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **49,** 129–147
28. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 247–255
29. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50,** 157–163
30. Navaza, J., and Saludijan, P. (1997) *Methods Enzymol.* **276,** 581–594
31. Otwinowski, Z. (1993) in *Data Collection and Processing: Proceedings of the CCP4 Study Weekend* (Sawyer, L., Isaacs, N., and Bailey, S., eds) pp. 56–62, Science and Engineering Research Council, Daresbury, United Kingdom
32. Otwinowski, Z., and Minor, W. (1997) in *Methods in Enzymology: Macromolecular Crystallography, Part A* (Carter, C. W., Jr., and Sweet, R. M., eds) Vol. 276, pp. 307–326, Academic Press, London
33. Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963) *J. Mol. Biol.* **7,** 95–99
34. Laskowski, R. A., McArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 282–291
35. Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.-P., and Davies, G. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 7090–7094
36. Jenkins, J., Leggio, L. L., Harris, G., and Pickersgill, R. (1995) *FEBS Lett.* **362,** 281–285
37. Henrissat, B., and Davies, G. J. (1997) *Curr. Opin. Struct. Biol.* **7,** 637–644
38. Kleywegt, G. J., and Jones, T. A. (1994) *ESF/CCP4 Newsletter* **31,** 9–14
39. Sinnott, M. L. (1990) *Chem. Rev.* **90,** 1171–1202
40. McCarter, J. D., and Withers, S. G. (1994) *Curr. Opin. Struct. Biol.* **4,** 885–892
41. Davies, G. J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A. M., Schülein, M., and Withers, S. G. (1998) *Biochemistry* **37,** 11707–11713
42. Sidhu, G., Withers, S. G., Nguyen, N. T., McIntosh, L. P., Ziser, L., and Brayer, G. D. (1999) *Biochemistry* **38,** 5346–5354
43. Sabini, E., Sulzenbacher, G., Dauter, M., Dauter, Z., Jørgensen, P. L., Schülein, M., Dupont, C., Davies, G. J., and Wilson, K. S. (1999) *Chem. Biol.* **6,** 483–492
44. Burmeister, W. P., Cottaz, S., Driguez, H., Palmieri, S., and Henrissat, B. (1997) *Structure* **5,** 663–675
45. Sulzenbacher, G., Withers, S., Dupont, C., and Davies, G. (1999) *Biochemistry* **38,** 4826–4833
46. Heightman, T. D., and Vasella, A. T. (1999) *Angew. Chem. Int. Ed.* **38,** 750–770
47. Namchuk, M. N., and Withers, S. G. (1995) *Biochemistry* **34,** 16194–16202
48. Roberge, M., Dupont, C., Morosoli, R., Shareck, F., and Kluepfel, D. (1997) *Protein Eng.* **10,** 399–403
49. Hilge, M., Gloor, S. M., Rypniewski, W., Sauer, O., Heightman, T. D., Zimmermann, W., Winterhalter, K., and Piontek, K. (1998) *Structure* **6,** 1433–44
50. Sabini, E., Schubert, H., Murshudov, G., Wilson, K. S., Siika-Aho, M., and Penttilä, M. (2000) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56,** 3–13
51. Roberge, M., Shareck, F., Morosoli, R., Kluepfel, D., and Dupont, C. (1997) *Biochemistry* **36,** 7769–7775
52. Charnock, S. J., Spurway, T. D., Xie, H., Beylot, M. H., Virden, R., Warren, R. A., Hazlewood, G. P., and Gilbert, H. J. (1998) *J. Biol. Chem.* **273,** 32187–32199
53. Andrews, S., Charnock, S. J., Lakey, J. H., Davies, G. J., Claeyssens, M., Nerinckx, W., Underwood, M., Sinnott, M. L., Warren, A. J., and Gilbert, H. J. (2000) *J. Biol. Chem.* **275,** 23027–23033
54. Sulzenbacher, G., Shareck, F., Morosoli, R., Dupont, C., and Davies, G. J. (1997) *Biochemistry* **36,** 16032–16039
55. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950