Exploring the Cellulose/Xylan Specificity of the β-1,4-Glycanase Cex from *Cellulomonas fimi* through Crystallography and Mutation[†,‡]

Valerie Notenboom,[§] Camelia Birsan,[||] R. Antony J. Warren,[⊥] Stephen G. Withers,[||] and David R. Rose*[,§]

Protein Engineering Network of Centres of Excellence, Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, Toronto, and Departments of Chemistry and Microbiology, University of British Columbia, Vancouver, Canada

Received December 1, 1997; Revised Manuscript Received January 19, 1998

ABSTRACT: The retaining β-1,4-glycanase Cex from *Cellulomonas fimi*, a family 10 glycosyl hydrolase, hydrolyzes xylan 40-fold more efficiently than cellulose. To gain insight into the nature of its preference for xylan, we determined the crystal structure of the Cex catalytic domain (Cex-cd) trapped as its covalent 2-deoxy-2-fluoroxylobiosyl−enzyme intermediate to 1.9 Å resolution. Together with the crystal structure of unliganded Cex-cd [White, A., et al. (1994) *Biochemistry 33*, 12546−12552] and the previously determined crystal structure of the covalent 2-deoxy-2-fluorocellobiosyl−Cex-cd intermediate [White, A., et al. (1996) *Nat. Struct. Biol. 3*, 149−154], this structure provides a convincing rationale for the observed substrate specificity in Cex. Two active site residues, Gln87 and Trp281, are found to sterically hinder the binding of glucosides and must rearrange to accommodate these substrates. Such rearrangements are not necessary for the binding of xylobiosides. The importance of this observation was tested by examining the catalytic behavior of the enzyme with Gln87 mutated to Met. This mutation had no measurable effect on substrate affinity or turnover number relative to the wild type enzyme, indicating that the Met side chain could accommodate the glucoside moiety as effectively as the wild type Gln residue. Subsequent mutagenesis studies will address the role of entropic versus enthalpic contributions to binding by introducing side chains that might be more rigid in the unliganded enzyme.

The β-1,4-glycanase Cex[1] from the soil bacterium *Cellulomonas fimi* is capable of hydrolyzing both cellulose and xylan as well as a range of soluble aryl glycosides based upon glucose and xylose. The gene encoding Cex was cloned, expressed in *Escherichia coli*, and subsequently sequenced (*1*). When produced in *C. fimi*, Cex is glycosylated, but the recombinant form is not. Glycosylation has no apparent effect on catalytic activity; its main function appears to be protection against proteolysis (*2*). Cex is a 47 kDa protein comprising an N-terminal catalytic domain (35 kDa) and a C-terminal cellulose binding domain (12 kDa) held together by a linker peptide rich in proline and threonine residues. The domains in Cex retain their respective catalytic and cellulose-binding properties when separated by limited proteolysis (*3*, *4*).

Cex is a retaining β-glycanase (*5*). Hence, hydrolysis of substrates occurs with retention of anomeric configuration, by a double-displacement mechanism involving formation and hydrolysis of a covalent glycosyl−enzyme intermediate via oxocarbenium ion-like transition states, as proposed by Koshland (*6*). This mechanism implies the existence of two catalytic residues in the active site of Cex. One, the nucleophile, is present in the ionized form and stabilizes an oxocarbenium ion transition state, subsequently forming a glycosyl−enzyme intermediate; the other acts as the acid−base catalyst, protonating the glycosidic oxygen of the scissile bond. The nucleophilic residue was identified as Glu233 by trapping a covalent 2-deoxy-2-fluoroglucosyl−enzyme intermediate and then sequencing the purified glycopeptide isolated from proteolytic digests (*7*). The assignment was later confirmed by mutation−kinetic analysis (*8*). Site-directed mutation of Glu127 yielded mutants with kinetic parameters fully consistent with the role of acid−base catalyst for this residue (*9*). A detailed kinetic study of Cex (*10*) allowed delineation of the mechanism through identification of rate-limiting steps, as well as investigation of the transition-state structure for each step. Such mechanistic insights, derived from pre-steady- and steady-state kinetics, Brønsted relationships, kinetic isotope effect measurements, inactivation experiments, and pH studies, provide further

[†] This work is funded by the Protein Engineering Network of Centres of Excellence (PENCE).

[‡] Coordinates and structure factors for the structure described here have been deposited with the Brookhaven Protein Data Bank and will be accessible under code 2XYL.

* To whom correspondence should be addressed: Ontario Cancer Institute, 610 University Ave., Toronto, ON M5G 2M9, Canada. Phone: (416) 946-2970. Fax: (416) 946-6529. E-mail: drose@oci.utoronto.ca.

[§] University of Toronto.

[||] Department of Chemistry, University of British Columbia.

[⊥] Department of Microbiology, University of British Columbia.

[1] Abbreviations: Cex, *C. fimi* exoglycanase; CMC, carboxymethylcellulose; 2F-DNPX$_2$, 2″,4″-dinitrophenyl 2-deoxy-2-fluoro-β-D-xylobioside; 3,4-DNPC, 3″,4″-dinitrophenyl β-D-cellobioside; 3,4-DNPG, 3′,4′-dinitrophenyl β-D-glucoside; 3,4-DNPX, 3′,4′-dinitrophenyl β-D-xyloside; 3,4-DNPX$_2$, 3″,4″-dinitrophenyl β-D-xylobioside; PhC, phenyl β-D-cellobioside; PhX$_2$, phenyl β-D-xylobioside; PNPC, 4″-nitrophenyl β-D-cellobioside; PNPG, 4′-nitrophenyl β-D-glucoside; PNPX, 4′-nitrophenyl β-D-xyloside; PNPX$_2$, 4″-nitrophenyl β-D-xylobioside.

S0006-2960(97)02921-8 CCC: $15.00
Published on Web 03/13/1998

supporting evidence for the double-displacement mechanism of Cex and other retaining β-glycosidases.

The determination of the X-ray crystal structure of the Cex catalytic domain (Cex-cd) (*11*) provides a structural basis for explaining the wealth of biochemical information that is available for Cex and related glycosidases. Cex-cd folds into an eight-stranded parallel α/β-barrel, with an open cleft at the carboxyl-terminal end, proposed to be the active site. The two catalytic glutamates, Glu233 and Glu127, are located on either side of the cleft, being separated by a distance of 5.5 Å, consistent with a retaining mechanism. Such a separation is presumably optimal for the efficient formation of a glycosyl−enzyme intermediate on Glu233 of Cex-cd, while at the same time allowing Glu127 to protonate the aglycone in a concerted manner (*11*).

Diffusion of the mechanism-based inactivator 2,4-dinitrophenyl 2-deoxy-2-fluorocellobioside into crystals of Cex-cd led to the formation of a covalent 2-deoxy-2-fluorocellobiosyl−enzyme intermediate. The structure of this glycosyl−enzyme complex was determined by X-ray crystallography (*12*) and, together with the structural information on the catalytic domain alone, provided more insight into interactions at the active site that are crucial to the catalytic mechanism and substrate preference of Cex.

On the basis of amino acid sequence alignments, Cex has been assigned to family 10 of glycosyl hydrolases (*13*). Family 10 members are primarily categorized as endolytic β-1,4-xylanases, although modest activity against cellobiosides has been reported for several of them (*14*−*18*). In addition, low activity against carboxymethylcellulose (CMC) has been reported in some instances (*15*).

Clearly, glycosyl hydrolases of family 10 are largely xylanolytic but can to a varying degree hydrolyze cellulose as well, in contrast to the low-molecular weight family 11 enzymes which exclusively hydrolyze xylan. In the case of Cex, the k_{cat}/K_m values for the hydrolysis of aryl xylobiosides are 30−100 times higher than those for the corresponding aryl cellobiosides (*19*). Since xylan is a β-1,4-linked polymer of D-xylose, a saccharide unit similar to glucose but lacking the hydroxymethyl group on C-5, it seems likely that the presence of this group is somehow inhibitory to catalysis. Comparison of the complexed and uncomplexed Cex-cd structures reveals that upon formation of the 2-deoxy-2-fluorocellobiosyl−enzyme intermediate, the side chains of Gln87 and Trp281 reorient from their positions in the native enzyme to accommodate the C-5 hydroxymethyl groups of the distal and proximal glucosyl units, respectively (*12*). The aim of the study reported here is to assess whether this reorientation may contribute to the substrate specificity of Cex. The structure of the xylobiosyl−enzyme intermediate addresses whether this reorientation is also required in binding that substrate and whether there are other structural explanations for the specificity. An understanding of the atomic basis for substrate specificity could be achieved by attempting to modify the Cex active site through site-directed mutation to improve its specificity for xylan degradation.

MATERIALS AND METHODS

General Methods. Aryl glycoside substrates were synthesized as described previously (*10*, *20*−*22*) or, where available, were purchased from Sigma Chemical Co. All buffer reagents, unless otherwise specified, were also obtained from Sigma. Cex and CexQ87M were purified as described previously (*3*). The cellulose binding domain was removed from Cex by papain digestion, and the Cex-cd was purified as described previously (*20*).

Crystallization and Inhibitor Soak. The crystallization conditions for wild type Cex-cd are well established (*23*). Crystals appear overnight and generally grow to a maximum size of $1 \times 0.5 \times 0.5$ mm^3 within 3 days. Prior to addition of the substrate, the crystals were transferred to a larger volume of artificial mother liquor with a higher poly(ethylene glycol) concentration (from 10 to 15%) in which they were found to be most stable. 2,4-Dinitrophenyl 2-deoxy-2-fluoro-β-xylobioside (2F-DNPX_2) was added as a solid to the mother liquor surrounding the crystal prior to data collection. A soaking time of 48 h proved to be optimal in terms of both merging *R*-factors and substrate occupancy.

Data Collection and Processing. X-rays for diffraction data were generated from a CuKα target on a Rigaku RU200 generator operated at 6 kW (40 kV and 150 mA). Intensities were measured on a San Diego Multiwire Systems area detector, and data were indexed and reduced with the complementary software (*24*).

Refinement was performed by X-Plor version 3.851 (*25*). Seven percent of the data (1898 reflections) were set aside as an objective test set to monitor the refinement process and were not used during refinement except for map calculations where all data were included (*26*). The electron density map calculated using phases from only the protein model (unliganded Cex-cd, PDB entry code 2EXO) after 40 cycles of rigid body minimization, followed by simulated annealing from 3000 to 300 K, showed unambiguously the 2-deoxy-2-fluoroxylobioside moiety covalently bound to the catalytic nucleophile Glu233. Subsequent refinement involved several cycles of simulated annealing and positional refinement routines, after each of which electron density was calculated and carefully inspected. Ordered water molecules were included in the structure after every cycle of refinement based on $2|F_o| - |F_c|$ maps contoured at a map standard deviation of σ and $|F_o| - |F_c|$ density contoured at 3σ. The disaccharide was not included in the refinement process until the last cycle, in an effort to avoid biasing the electron density surrounding the sugar moiety. X-ray scattering factors for the fluorine atom were added. Standard X-plor carbohydrate parameter and topology files (param3.cho and toph3.cho, respectively) were adjusted to describe the substrate accurately, i.e., including the fluorine atom on the proximal xylose moiety and the covalent α-link between the nucleophile Glu233 and the substrate. Only half of the suggested energies were used for the sugar restraints, to allow for some flexibility during refinement and to increase the weight on experimental parameters. Dihedral and improper restraints were not imposed on the disaccharide. Individual *B*-factor refinement was applied after the *R*-free value showed no more convergence during energy minimization routines.

Generation of Cex Gln87Met. The mutation Gln87Met was made in vitro by overlap extension (*27*), using pUC12-1.1*cex*(PTIS) (*28*). Synthetic oligonucleotides used for mutagenesis were prepared by the UBC Nucleic Acid and Protein Sequencing Unit (NAPS) with an Applied Biosystems 380A DNA synthesizer. The following primers were used, with underlining indicating where nucleotides were

changed: GTA TGG CAC TCG ATG CTG CCC GAC TGG and CCA GTC GGG CAG CAT CGA GTC CCA TAC. *E. coli* JM101 was transformed with the mutated plasmid. Transformants were selected on ampicillin, and the identity of the plasmid was checked therein by restriction digestion and sequencing of that part of the *cex* gene containing the mutation. The 500-base pair *Bam*HI−*Not*I fragment from an appropriate clone was used to replace the corresponding fragment in wild type pUC12-1.1cex(PTIS). After transformation of *E. coli* JM101 and selection on ampicillin, the plasmid was sequenced between the *Bam*HI and *Not*I sites to ensure that only the desired mutation was present. The plasmid, designated pCexQ187M, was used to transform various *E. coli* strains which were then tested for production of the mutant protein. *E. coli* Topp5 gave the best yield (data not shown) and was used for enzyme production.

Kinetic Analysis of the Gln87Met Mutant. The enzymic hydrolysis rates were measured by spectrophotometric detection of phenol release or by measuring the amount of reducing sugars liberated during hydrolysis of sugar polymers. Michaelis−Menten parameters for the enzymatic hydrolysis of all aryl glycoside substrates, with the exception of PhC and PhX_2, were determined using a continuous assay, by recording changes in absorbance using a UV/Vis Pye-Unicam 8700 spectrophotometer equipped with a temperature-controlled circulating water bath. Solutions of the appropriate substrate concentrations in 50 mM citrate or phosphate buffer (pH 5.5 or 7.0, respectively) with 1 mg of BSA mL^{-1} at 25 or 37 °C were preincubated within the spectrophotometer until thermally equilibrated; reactions were initiated by the addition of enzyme. Hydrolysis was continuously monitored at a wavelength at which there was a convenient absorbance difference between the initial glycoside and the phenol product as previously reported (*20*). The concentration of enzyme added was selected so that less than 10% of the total substrate was converted to product during analysis. Substrate concentrations employed ranged from 0.2 to 5 times the K_m value ultimately determined, wherever possible.

Hydrolysis rates for the substrates PhC and PhX_2 were determined at 37 °C using a stopped assay. Substrate at different concentrations in 50 mM phosphate buffer, with 1 mg of BSA mL^{-1} (pH 7.0) (190 μL), was prewarmed at 37 °C and the reaction initiated by the addition of 10 μL of enzyme. After an appropriate time, 0.6 mL of 2.0 M Na_3PO_4 (pH 12.1) was added to stop the reaction. The absorbance of the released phenolate at 288 nm was determined immediately, corrected for the spontaneous hydrolysis of substrate and the background absorbance of the enzyme ($\Delta\epsilon = 2.17$ mM^{-1} cm^{-1} for phenol at pH 12.1), and used to calculate rates. The program Grafit (*29*) was used to perform nonlinear regression analysis in the determination of K_m and V_{max}.

CM-cellulase and xylanase activities were measured by the *p*-hydroxybenzoic acid hydrazide (HBAH) method (*30*) which quantitates the production of reducing sugar. Initial rates of CM-cellulose hydrolysis were determined by incubating 100 μL of an enzyme solution with 900 μL of 2% CM-cellulose in 50 mM citrate buffer at pH 6.8 and 30 °C. Aliquots of 100 μL were removed at 0, 2, 4, 6, 8, and 10 min intervals and added to 1 mL of 5 mM NaOH in glass test tubes. After all samples were taken, 1 mL of HBAH reagent was added to each tube and the samples were heated in a boiling water bath for 12 min. When tubes had cooled to room temperature, A_{420nm} was measured and the amount of reducing sugar released was determined using a glucose standard curve generated under the same conditions. The activity is expressed as inverse seconds or nanomoles per milliliter of reducing sugars released per second per nanomoles per milliliter of enzyme. To ensure the validity of the assays for enzyme activity, concentrations of enzyme were chosen so that the production of reducing sugar in the assay fell within the linear response range of the standard curve. Appropriate controls for background reducing sugar and/or for the presence of chemical reagents that might interfere with the assay were always included with the assay procedures. Substrate for the xylanase assays was prepared by dissolving 6−10 g of birchwood xylan in 100−200 mL of distilled water at room temperature for 3−4 h with continuous stirring. Insoluble material was removed by centrifugation at 16000*g* in a GSA rotor for 20 min. The clear supernatant was lyophilized, and the dry, water-soluble product was used as substrate at a concentration of 0.2%; the release of reducing sugars was determined as described above.

RESULTS AND DISCUSSION

Covalent Cex-cd−2-Deoxy-2-fluoroxylobiosyl Intermediate Structure. Due to the very poor solubility of the slow substrate 2F-$DNPX_2$ at the concentration being sought (0.5 mM), soaking the Cex-cd crystal in a solution containing a significant amount of solid material proved to be necessary. The intent was that the substrate would partition into the crystal via the small amount in solution. Successful soaking was indicated by the appearance of yellow 2,4-dinitrophenolate in the solution overnight as a consequence of the reaction shown in Scheme 1, in which the covalent 2-deoxy-2-fluoroxylobiosyl−enzyme intermediate is formed. An additional indication of reaction occurring was the fact that the crystal surface appeared to be cracked a few hours after the addition of 2F-$DNPX_2$, but these cracks annealed over time.

X-ray diffraction data were collected to 1.9 Å resolution at room temperature. A single crystal was sufficient for a complete data set, because of the resilience and high symmetry of the tetragonal crystals ($P4_12_12$). After three cycles of data reduction, the indexing parameters and merging residuals converged to yield the statistics shown in Table 1. The unit cell dimensions of the complexed crystal differed only slightly from those of the native ($\Delta a = \Delta b =$ 0.05 Å, $\Delta c = 0.28$ Å). Rigid body refinement indicated a slight rotation of the molecule in the unit cell compared to the packing of native Cex-cd. The structure and refinement statistics are listed in Table 2. The current model contains 2399 non-hydrogen protein atoms, 18 non-hydrogen saccharide atoms, and 123 ordered water molecules. Data between 10 and 1.9 Å resolution and with an F/σ of >2 refined to a crystallographic *R*-factor of 0.21 and an *R*-free of 0.26. Of the non-glycine residues, 91% have ϕ and ψ angles located in the most favored regions, 8% fall in the additional allowed regions, and the remaining residue Thr81 assumes the unusual cis conformation, as described previously (*10*).

The active site of Cex trapped as its 2-deoxy-2-fluoroxylobiosyl−enzyme intermediate is shown in Scheme 2,

Scheme 1

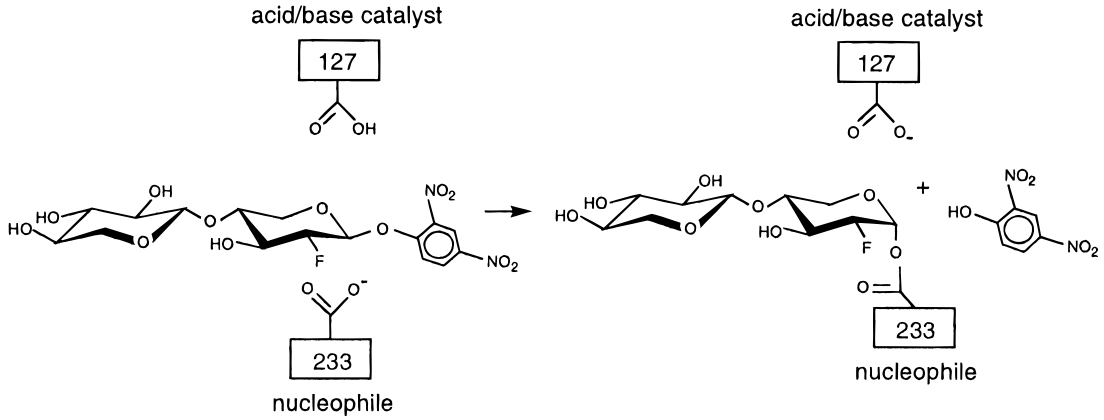


Table 1: Data Collection Statistics

resolution shell	all collected data to 1.90 Å	2.05−1.90 Å
number of reflections	197 626	25 539
number of unique reflections	25 857	5050
completeness (%)	99.94	99.90
I/σ	16.19	3.58
R_{sym}[a] (%)	8.03	27.03
cell dimensions (*a*, *b*, *c*) (Å)	88.18, 88.18, 81.29	

[a] $R_{sym} = \sum_h|(\langle I\rangle - I)|/\sum_h(\langle I\rangle)$, summed over all reflections, *h*, where $\langle I\rangle$ is the average of all equivalent measurements of that reflection.

Table 2: Structure and Refinement Statistics

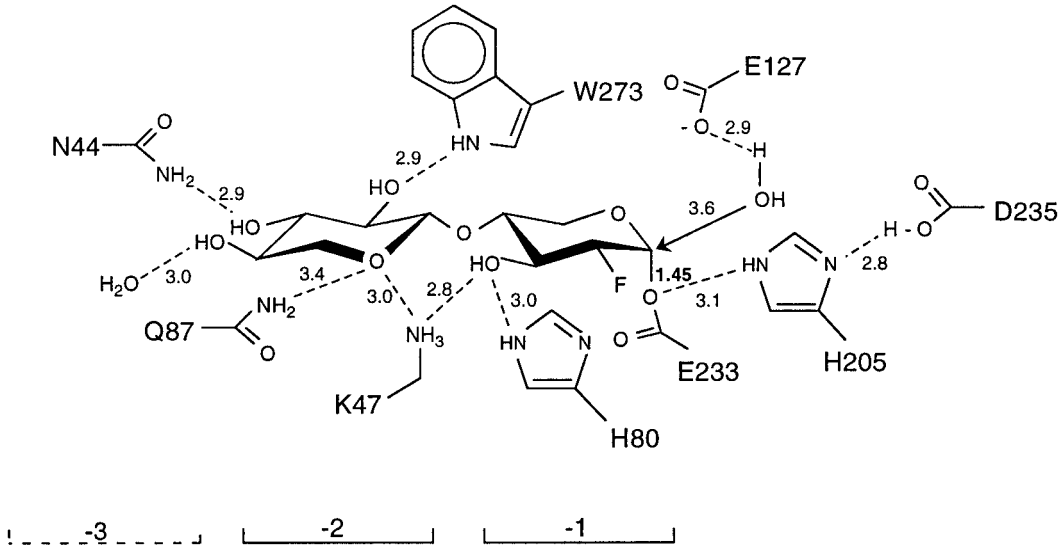
resolution range	10−1.9 Å
number of reflections ($F/\sigma > 2$)	25 480
R-factor[a]	0.212
R-free	0.260
number of protein atoms	2399
number of substrate atoms	18
number of solvent molecules	123
rms deviations from ideal	
bond lengths	0.005 Å
bond angles	1.23°
dihedrals	22.8°
impropers	1.03°
average *B*-factor	13.0 Å^2
average *B*-factor for substrate	6.64 Å^2
average *B*-factor for solvent molecules	23.20 Å^2

[a] *R*-factor $= \sum_h(|F_{obs}| - |F_{calc}|)/\sum_h|F_{obs}|$, summed over all reflections *h*.

indicating the relevant interactions of the enzyme with the sugar moiety. The covalent nature of the glycosyl−enzyme species is clearly seen in Figure 1a, the contiguous electron density between the carboxyl side chain of Glu233 and the anomeric carbon of the sugar moiety being quite evident. The electron density displayed is that determined for the 2-deoxy-2-fluoroxylobiosyl−Cex-cd structure described in this paper. The glycosyl−enzyme bond between C-1 of the proximal sugar and the nucleophilic oxygen of Glu233 is 1.45 Å in length. Table 3 lists a selection of bond angles of the final refined conformation of the proximal (covalently linked) 2-deoxy-2-fluoroxylose unit, as compared to suggested parameters during refinement. The fluorine substituent at C-2 of the proximal saccharide is seen to be only slightly removed from its expected position in the relaxed $^4C_1$ conformation. The dihedral angle 3′-OH−C-3−C-2−2′-F measures 58°, as opposed to 66° for 3′-OH−C-3−C-2−2′-OH in the distal sugar. This slight dihedral "flattening" is also seen in the crystal structure of the cellobiosyl−Cex-cd covalent intermediate, in which the proximal 2′-hydroxyl group is retained, formed by a catalytically deficient mutant of the enzyme (V. Notenboom et al., unpublished experiments). Therefore, this is likely to be a characteristic of the covalent reaction intermediate rather than a consequence of the fluorine substitution. Interestingly, even larger distortion was observed in the 2-deoxy-2-fluorocellobiosyl−Cex-cd structure, with a dihedral angle of 21° being found (*11*). One consequence of, or perhaps reason for, this distortion is the apparent presence of a hydrogen bond between the fluorine at C-2 of the 2-fluorocellobiosyl moiety and the side chain amide nitrogen of Asn126, a highly conserved residue. No such strong interaction is seen in the structure of the 2-deoxy-2-fluoroxylobiosyl−Cex-cd, the separation of 3.3 Å being too great. A water molecule is found positioned between the acid−base catalyst Glu127 (2.9 Å) and C-1 (3.6 Å) of the covalently linked proximal xylose moiety in the latter structure, in an ideal position for attacking the anomeric carbon.

Apart from the slight changes in conformation of the sugar moiety, the significant differences in structure between the free Cex-cd, 2-deoxy-2-fluorocellobiosyl−Cex-cd, and 2-deoxy-2-fluoroxylobiosyl−Cex-cd are those seen within the enzyme active site, specifically in the positions of the side chains of Gln87 and Trp281. These changes are shown in Figure 1b which presents a superposition of the sugar moieties in the two covalent intermediates and of the positions of selected side chains for all three proteins. As can be seen, the indole side chain of Trp281 is displaced significantly from its position in the free enzyme upon formation of the 2-deoxy-2-fluorocellobiosyl intermediate, whereas no such change is seen upon formation of the 2-deoxy-2-fluoroxylobiosyl−enzyme intermediate. These conformational changes appear to be caused by repulsive steric interactions between the hydroxymethyl group at C-5 of the glucose of the proximal sugar and the tryptophan side chain. Such interactions do not occur when a xylose residue is present at this position. Similarly, there appear also to be repulsive, steric interactions between the hydroxymethyl group of the distal sugar and Gln87. In the native enzyme, the side chain of Gln87 is somewhat disordered, with side chain *B*-factors refined to an average of 47.5 Å^2; apparently, it is forced outward in the 2-deoxy-2-fluorocellobiosyl−Cex-cd intermediate. By contrast, the side chain of Gln87 is well

Scheme 2: Schematic Representation of Interactions in the Active Site Region of the Cex-cd−2F-Xylobioside Complex[a]



[a] Distances are measured in angstroms. The distance between 2′F and the carbonyl oxygen of E233 is 2.6 Å (not shown). The boxes below the diagram roughly indicate the area of the binding subsite. The −3 subsite is shown in dashed lines because it is not occupied by substrate.

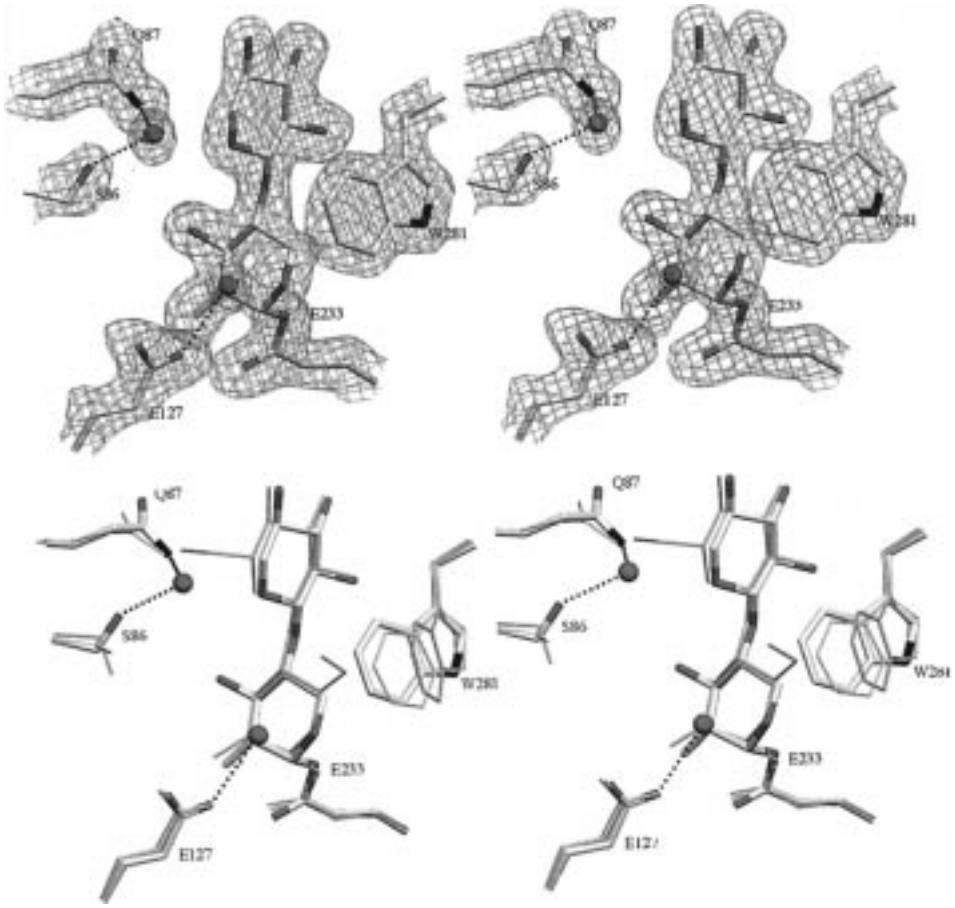
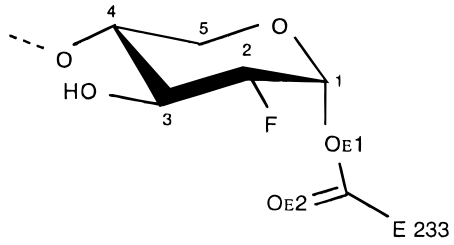


FIGURE 1: (a, top) Stereodiagram of $2|F_o| - |F_c|$ electron density for the 2F-xylobiosyl−Cex-cd structure, contoured at 1.2σ, superposed on representative atoms of the final refined structure. The density was calculated from all data to 1.9 Å resolution and the final protein and water coordinates. The figure was drawn using the program Setor (*33*). (b, bottom) Stereorepresentation of a superposition of uncomplexed Cex-cd (*11*; PDB entry code 2EXO) shown in orange, 2F-cellobiosyl−Cex-cd (*12*; PDB entry code 1EXP) shown in green, and 2F-xylobiosyl−Cex-cd shown in bulky white, with oxygen atoms shown in red, nitrogens in blue, and the fluorine atom in yellow. Trp281, flanking the proximal sugar, does not need to reorient to accommodate the hydroxymethylene group as seen in the 2F-cellobiosyl−enzyme complex (a 31° rotation around its $C_\beta-C_\gamma$ bond). Gln87 is disordered in the 2F-cellobiosyl−enzyme structure but is stabilized in the 2F-xylobiosyl−enzyme complex through a water molecule and Ser86, which also adopts a different conformation in this interaction. The figure was drawn with Setor (*33*).

defined in the 2-deoxy-2-fluoroxylobiosyl−Cex-cd intermediate (average side chain *B*-factors of 11.8 Å^2) in a position previously occupied by the C-5 hydroxymethyl side chain of glucose; it appears to form a hydrogen bond with the

Table 3: Comparison of Ideal and Measured Angles



angle (deg)	imposed	observed
C-1−C-2−C-3	110.4	112.0
C-3−C-2−F	109.7	112.6
C-5−O-5−C-1	112.0	109.1
C-1−C-2−F	109.7	111.7
C-1−OE-1−C(E)	112.0	116.3
OE-1−C-1−C-2	109.7	111.3
OE-1−C(E)−OE-2	123.0	122.5
OE-1−C-1−O-5	112.1	109.7

endocyclic oxygen atom (O-5) of the distal sugar. Furthermore, Ser86 assumes a different side chain conformation in the 2-deoxy-2-fluoroxylobiosyl−Cex-cd intermediate, compared to that in both crystal structures described previously, now forming a hydrogen bond to a water molecule not seen in the 2-deoxy-2-fluorocellobiosyl−Cex-cd intermediate. In turn, this water molecule forms a fairly close interaction (2.9 Å) with Gln87, stabilizing its side chain even more (Figure 1a,b). When the structures are superimposed on other family 10 crystal structures found in the literature, the only differences in the active sites can be found at positions equivalent to Ser86 and Gln87 of Cex. For example, in XynZ from *Cellulomonas thermocellum* (*31*; PDB entry code 1XYZ), an Asn corresponds to Ser86 of Cex; and in XylA from *Pseudomonas fluorescens* ssp. *cellulosa* (*32*; PDB entry code 1CLX), a Pro and a Tyr correspond to Ser86 and Gln87, respectively, of Cex. It is possible that these positions, though located in the active site, are not as sensitive to evolutionary pressure as the strictly conserved residues with respect to their contribution to the catalytic mechanism. This could therefore result in a more promiscuous enzyme that recognizes both xylosides and glucosides.

It should be noted that differences in interactions during binding of the polymers xylan and cellulose at the more remote subsites could play a significant role in substrate specificity as well. In addition, the cellulose binding domain at the C-terminal end of Cex possibly affects hydrolysis rates of cellulose polymers. However, it appears that the specificity for xylan over cellulose substrate arises, at least in part, from repulsive steric interactions with Gln87 and Trp281 in the cellobiosyl−Cex-cd complex, as well as from the stabilization of Gln87 in the xylobiosyl−Cex-cd complex. Attempts to engineer an improved xylanase/cellulase ratio should logically focus upon these two residues initially. Our focus was directed in the first instance upon Gln87 because Trp281 cannot be replaced with a bulkier amino acid. It may be possible in the future to prevent displacement of Trp281 from the active site by introducing a bulkier residue behind Trp281, in the position occupied in the wild type enzyme by phenylalanine.

Our initial intention was to generate mutants in which Gln87 was replaced by Met, His, and Tyr. Met was chosen as a residue of slightly greater dimensions than Gln, and one with which hydrogen bonding possibilities are limited since it cannot donate a hydrogen bond. His was chosen as an amino acid with a bulkier side chain, but one which retains most of the hydrogen bonding possibilities, while Tyr is a larger residue with some polar character. Unfortunately, all attempts to obtain the His and Tyr mutants were unsuccessful; workable quantities of either protein could not be produced. Fortunately, sufficient quantities of the Met mutant were readily generated.

Table 4: Kinetic Parameters for the Hydrolysis of β-Glycosides by Cex and the Gln87Met Mutant at 37 °C and pH 7.0

substrate	enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})
PNPG	Cex wt	0.024	8.3	2.9×10^{-3}
	Gln87Met	0.024	7.4	3.2×10^{-3}
3,4-DNPG	Cex wt	2.9	6.5	0.45
	Gln87Met	1.2	6.0	0.20
PNPC	Cex wt	15.8	0.60	26.3
	Gln87Met	9.2	0.34	27.1
3,4-DNPC	Cex wt	9.7	0.14	69
	Gln87Met	8.3	0.065	128
PhC	Cex wt	−[b]	−	1.0×10^{-4}
	Gln87Met	−	−	2.3×10^{-4}
PNPX	Cex wt	2.6	20	0.13
	Gln87Met	1.2	67	0.018
3,4-DNPX	Cex wt	22	7.9	2.8
	Gln87Met	9.0	7.2	1.2
$PNPX_2$	Cex wt	39.8	0.018	2200
	Gln87Met	57.9	0.12	482
$3,4\text{-}DNPX_2$	Cex wt	22	0.012	1840
	Gln87Met	38.2	0.038	1005
PhX_2	Cex wt	10.4	8.6	1.2
	Gln87Met	18.7	21.2	0.9
2% CMC	Cex wt	0.052[a]	−	−
	Gln87Met	0.025[a]	−	−
1% xylan	Cex wt	0.47[a]	−	−
	Gln87Met	0.40[a]	−	

[a] Specific activity expressed as inverse seconds or nanomoles per milliliter of reducing sugars liberated per second per nanomoles per milliliter of protein. [b] −, parameter not determined.

Catalytic Properties of the Gln87Met Mutant. The Gln87Met mutant behaved in a manner virtually identical to that of the wild type enzyme during purification; both enzymes were desorbed from the cellulose affinity column by approximately the same volume of eluant, and both ran as a single bands with an M_r of 47 kDa on SDS−PAGE (>95% purity by inspection). The yield of purified enzyme ranged from 15 to 25 mg per liter of liquid culture. Mass spectrometric analysis confirmed the molecular mass of the mutant protein as 47 120 Da, exactly as predicted. Correct folding was indicated by the CD spectrum and the thermal stability of the mutant enzyme being virtually identical to those of the wild type enzyme.

Values of k_{cat} and K_m for the mutant and wild type enzymes were determined on a range of glucoside, cellobioside, xyloside, and xylobioside substrates (Table 4). Substitution of Gln87 by Met in the binding site of Cex did not alter the K_m and k_{cat} values for the two series of substrates significantly. This is best expressed in the form of the linear free energy relationship in which values of $\log(k_{cat}/K_m)$, the specificity constant, for each substrate with the Gln87Met mutant, are plotted against the same parameter for the wild type enzyme (Figure 2). The excellent correlation coefficient ($\rho = 0.974$) and the slope of unity ($r = 1.03$) confirm that

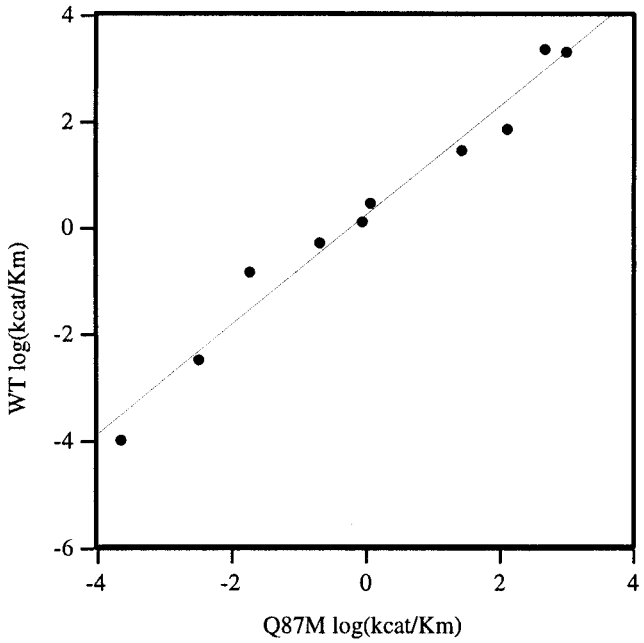


FIGURE 2: Plot of $\log(k_{cat}/K_m)$ for hydrolysis of a series of aryl glycosides by the Gln87Met mutant of Cex versus the same parameter for the wild type enzyme.

the two enzymes have essentially identical specificities.

The small variations observed are most likely due to perturbations in the local environment within the protein. Thus, in contrast to what had been hoped, the Gln87Met mutant did not show a decreased affinity for cello−oligosaccharide substrates. Presumably, the methionine residue is still sufficiently flexible to undergo a rearrangement similar to that observed for Gln87 upon binding the gluco substrates in the active site. It is possible that the hydrogen bonding interaction between Gln87 and the endocyclic oxygen of the distal sugar either does not play a very important role in specificity or is compensated for in some way. It is likely, therefore, that the binding interactions with the C-5 hydroxymethylene groups in the active site are more complex than initially thought, and that a single point mutation in the catalytic domain of Cex may be insufficient to produce the desired changes in the specificity of the enzyme. It is significant in this context that the Gln87Ser mutant of Cex, constructed in a manner similar to that of the Gln87Met mutant, behaves kinetically like the latter with the same substrates (A. MacLeod, personal communication).

In conclusion, the crystallographic analysis of the Cex-cd−2F-xylobiosyl covalent intermediate has provided a possible structural rationale for the preference of Cex for xylan over cellulose. The C-5 substituent on the glucose rings causes the displacement of two side chains relative to the unliganded enzyme, Gln87 and Trp281, a displacement that is not necessary for the binding of xylobiose. An additional component may be the reduced strain of the sugar ring in the xylobiosyl−enzyme covalent intermediate, as compared to that of the cellobiosyl complex. A first attempt to test these conclusions with a Gln87 to Met mutant did not succeed in altering the cellulose/xylan specificity. A Met side chain presumably retains enough flexibility to accommodate a glucose moiety as well as xylose in the −2 subsite. Alternative strategies could include the incorporation of a bulkier side chain in this position or the substitution by a smaller side chain, to test for the opposite result, an improvement in cellulase activity. As the Gln side chain is disordered in both the wild type and cellobiose intermediate, but is observed to be well ordered in the xylobiose complex, there may be an entropic cost to binding that is actually greater for xylan binding than for cellulose binding. Therefore, the effects of various side chains at position 87 may be revealing in terms of the entropic versus the enthalpic contributions to binding. Structure-based studies along these lines are underway.

ACKNOWLEDGMENT

We thank Dr. André White for his enthusiastic support and help.

REFERENCES

1. O'Neill, G., Goh, S. H., Warren, R. A. J., Kilburn, D. G., and Miller, R. C., Jr. (1986) *Gene 44*, 325.
2. Langsford, M. L., Gilkes, N. R., Singh, B., Moser, B., Miller, R. C., Jr., Warren, R. A. J., and Kilburn, D. G. (1987) *FEBS Lett. 225*, 163.
3. Gilkes, N. R., Warren, R. A. J., Miller, R. C., Jr., and Kilburn, D. G. (1988) *J. Biol. Chem. 263*, 10401.
4. Gilkes, N. R., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. J. (1989) *J. Biol. Chem. 264*, 17802.
5. Withers, S. G., Dombroski, D., Berven, L. A., Kilburn, D. G., Miller, R. C., Jr., Warren, R. A. J., and Withers, S. G. (1986) *Biochem. Biophys. Res. Commun. 139*, 487.
6. Koshland, D. E. (1953) *Biol. Rev. 28*, 416.
7. Tull, D., Withers, S. G., Gilkes, N. R., Kilburn, D. G., Warren, R. A. J., and Aebersold, R. (1991) *J. Biol. Chem. 266*, 15621.
8. MacLeod, A. M., Tull, D., Warren, R. A. J., and Withers, S. G. (1996) *Biochemistry 35*, 13165.
9. MacLeod, A. M., Lindhorst, T., Withers, S. G., and Warren, R. A. J. (1994) *Biochemistry 33*, 6571.
10. Tull, D., and Withers, S. G. (1994) *Biochemistry 33*, 6363.
11. White, A., Withers, S. G., Gilkes, N. R., and Rose, D. R. (1994) *Biochemistry 33*, 12546.
12. White, A., Tull, D., Johns, K., Withers, S. G., and Rose, D. R. (1996) *Nat. Struct. Biol. 3*, 149.
13. Henrissat, B., and Bairoch, A. (1993) *Biochem. J. 293*, 781.
14. Grepinet, O., Chebrou, M.-C., and Beguin, P. (1988) *J. Bacteriol. 170*, 4582.
15. Luthi, E., Love, D. R., McAnulty, J., Wallace, C., Caughey, P. A., Saul, D., and Berguist, P. L. (1990) *Appl. Environ. Microbiol. 57*, 694.
16. Lin, L.-L., and Thompson, J. A. (1991) *Mol. Gen. Genet. 228*, 55.
17. Shareck, F., Roy, C., Yaguchi, M., Morosoli, R., and Kluepfel, D. (1991) *Gene 107*, 75.
18. Haas, H., Herfurth, E., Stoffler, G., and Redl, B. (1992) *Biochim. Biophys. Acta 1117*, 279.
19. Gilkes, N. R., Claeyssens, M., Aebersold, R., Henrissat, B., Meinke, A., Morrison, H. D., Kilburn, D. G., Warren, R. A. J., and Miller, R. C., Jr. (1991) *Eur. J. Biochem. 202*, 367.
20. Ziser, L., Setyawati, I., and Withers, S. G. (1995) *Carbohydr. Res. 274*, 137.
21. Kempton, J. B., and Withers, S. G. (1992) *Biochemistry 31*, 9961.
22. . Ziser, L., and Withers, S. G. (1994) *Carbohydr. Res. 265*, 9.
23. Bedarkar, S., Gilkes, N. R., Kilburn, D. G., Kwan, E., Rose, D. R., Miller, R. C., Jr., Warren, R. A. J., and Withers, S. G. (1992) *J. Mol. Biol. 228*, 693.
24. Howard, A. J., Neilsen, C., and Xuong, N. H. (1985) *Methods Enzymol. 114*, 452.
25. Brünger, A. T., Kuriyan, J., and Karplus, M. (1987) *Science 235*, 458.
26. Kleywegt, G. J., and Brünger, A. T. (1996) *Structure 4*, 897.

27. Higuchi, R. (1990) in *PCR protocols: A guide to methods and applications* (Innis, M. A., Gelford, D. H., Sninsky, J. J., and White, T. J., Eds.) p 117, Academic Press, San Diego.
28. O'Neill, G. P., Kilburn, D. G., Warren, R. A. J., and Miller, R. C., Jr. (1986) *Appl. Environ. Microbiol. 52*, 737.
29. Leatherbarrow, R. J. (1990) *GraFit Version 2.0*, Erithacus Software Ltd., Staines, U.K.
30. Miller, G. L., Blum, R., Glennon, W. E., and Burton, A. L. (1960) *Anal. Biochem. 1*, 127.
31. Dominguez, R., Souchon, H., Spinelli, S., Dauter, Z., Wilson, K. S., Chauvaux, S., Beguin, P., and Alzari, P. M. (1995) *Nat. Struct. Biol. 2*, 569.
32. Harris, G. W., Jenkins, J. A., Connerton, I., Cummings, N., Lo Leggio, L., Scott, M., Hazelwood, G. P., Laurie, J. I., Gilbert, H. J., and Pickersgill, R. W. (1994) *Structure 2*, 1107.
33. Evans, S. V. (1993) *J. Mol. Graphics 11* (2), 134.

BI9729211