Structures of Oligosaccharide-Bound Forms of the Endoglucanase V from *Humicola insolens* at 1.9 Å Resolution[†,‡]

Gideon J. Davies,*[,§] Shirley P. Tolley,[§] Bernard Henrissat,[||] Carsten Hjort,[⊥] and Martin Schülein[⊥]

Department of Chemistry, University of York, Heslington, York YO1 5DD, England, Centre de Recherches sur les Macromolécules Végétales, CNRS, B.P. 53X, F-38041 Grenoble, France, and Novo-Nordisk a/s, Novo Allé, 2880 Bagsvaerd, Denmark

Received July 5, 1995; Revised Manuscript Received October 4, 1995[⊗]

ABSTRACT: Cellulose, a polymer of β-1,4-linked glucose residues, is the major polysaccharide component of plant cell walls and the most abundant biopolymer. The underlying mechanisms of the enzymatic degradation of cellulose are of increasing commercial and ecological significance. Endoglucanase V, from the cellulolytic soil hyphomycete *Humicola insolens*, is an endocellulase, the catalytic core of which consists of 210 amino acids and is known to hydrolyze the β-1,4 links with inversion of configuration at the anomeric carbon. The major products of cellulose hydrolysis are cellobiose and cellotriose. The crystal structures of the endoglucanase V (EGV) from *H. insolens*, in native, product (cellobiose), inactive mutant (D10N), and oligosaccharide-bound [$_{(D10N)}$−cellohexaose] forms, have been determined at resolutions of 1.9 Å or better. EGV consists of a six-stranded β-barrel domain with long interconnecting loops. A 40 Å groove exists along the surface of the enzyme, and this contains the catalytic residues, Asp 10 and Asp 121. The two catalytic aspartates sit to either side of the substrate binding groove in an ideal conformation for facilitating cleavage by inversion, their carboxyl groups being separated by approximately 8.5 Å. The complex between substrate and inactive mutant reveals excellent density for an oligosaccharide in six of the enzyme's seven substrate binding subsites. No sugar moiety, however, is seen bound to the −1 subsite at the point of cleavage. The geometry of the cleavage site suggets that the enzyme would favor the binding of sugars with an elongated glycosidic bond, as found in the transition state, as opposed to the binding of substrate. The oligosaccharide complexes reveal solvent water suitably placed for participation in a single displacement reaction as first suggested by Koshland in 1953 [Koshland, D. E. (1953) *Biol. Rev. 28*, 416−436]. A large conformational change takes place upon substrate binding. This "lid flipping" has the effect of increasing the hydrophobic environment of the catalytic proton donor, enclosing the active site at the point of cleavage, and bringing a third aspartate (Asp 114) in close proximity to the substrate. Site-directed mutagenesis of the catalytic residues has been used to confirm their significance in catalysis.

Cellulose is the major structural polysaccharide of plants and the most abundant macromolecule produced by living organisms. It is a polymer of β-1,4-linked glucopyranose residues, and the vast quantity of cellulolytic material gives its degradation a prime commercial significance. Cellulolytic turnover by living organisms has been estimated to be approximately 10^9 tons/year (Coughlan, 1985). The enzymatic degradation of cellulose is performed by a wide variety of plant, bacterial, and fungal species which possess endoglucanases, cellobiohydrolases, and β-glucosidases acting in a synergistic fashion to facilitate complete cleavage of the cellulose β-1,4 glycosidic bonds [for reviews, see Béguin and Aubert (1994), Henrissat (1994), and Wood (1992)]. Cellulases are represented in 10 of the 45 glycosylhydrolase families that have been classified (Henrissat, 1991; Henrissat & Bairoch, 1993). The structures of enzymes from five of the cellulase families have been published [Davies et al., 1993; Divne et al., 1994; Dominguez et al., 1995; Ducros et al., 1995; Juy et al., 1992; Rouvinen et al., 1990; Spezio et al., 1993; White et al., 1994; for review, Davies and Henrissat (1995)].

Cellulases usually consist of a catalytic core connected by a flexible linker region to a cellulose binding domain (Gilkes et al., 1991). In structural studies, intact multidomain cellulases have, thus far, proven inaccessible to protein crystallographic methods, and the majority of crystal structures are of the enzymes' catalytic cores. *Humicola insolens* is a soft-rot fungus that produces at least seven different cellulases (Schülein et al., 1993). The endoglucanse V (EGV)[1] from *H. insolens* is a protein of 284 amino acids consisting of a 210-residue catalytic core domain linked, *via* a heavily glycosylated serine and threonine rich linker region, to a cellulose binding domain (CBD). The CBD from the

[†] This work was funded in part by the Science and Engineering Research Council, the Biotechnology and Biological Sciences Research Council, Novo-Nordisk a/s, and the European Union (Contract BIO2-CT94-3018).

[‡] Coordinates for the structures described in this paper have been deposited with the Brookhaven Protein Data Bank (accession references: 1ENG, 2ENG, 3ENG, 4ENG).

* Corresponding author. Telephone: -44-1904-432596. Fax: -44-1904-410519. E-mail: davies@yorvic.york.ac.uk.

[§] University of York.

[||] CERMAV, Grenoble.

[⊥] Novo Nordisk a/s.

[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1995.

[1] Abbreviations: CBD, cellulose binding domain; CBH, cellobiohydrolase; CMC, carboxymethyl-substituted cellulose; EGV, endoglucanase V; HEWL, hen egg-white lysozyme; NAG, *N*-acetylglucosamine; NAM, *N*-acetylmuramic acid; PCR, polymerase chain reaction; PEG, polyethylene glycol.

0006-2960/95/0434-16210$09.00/0

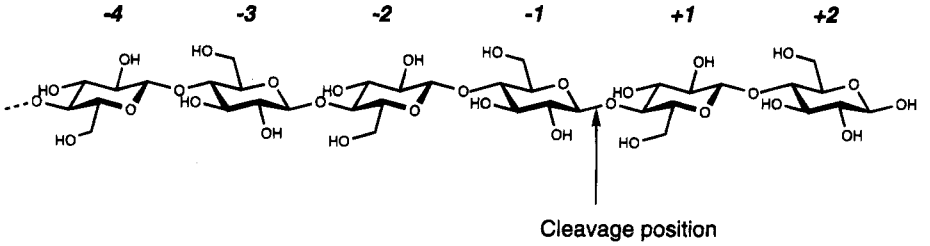

FIGURE 1: Cellulose polymer consisting of β-1,4-linked glucopyranosyl units. Endoglucanase V has six major sugar binding sites for substrate binding, labeled -4 to $+2$, and cleaves between the -1 and $+1$ subsites at the reducing end of the chain.

H. insolens EGV shows high sequence identity to the *Trichoderma reesei* CBH-I cellulose binding domain whose solution structure has been determined (Kraulis et al., 1989). We have determined the structure of the catalytic core of the EGV from *H. insolens* at a resolution of 1.5 Å by multiple isomorphous replacement techniques, utilizing the optimized anomalous scattering from the Lu and Hg derivatives (Davies et al., 1993, 1996). It is a family 45 enzyme (also known as cellulase family K), sharing sequence similarity with the endoglucanase B from the bacterium *Pseudomonas fluorescens* (Gilbert et al., 1990) and the fungal EGVs from *T. reesei* and *Fusarium oxysporum* (Saloheimo et al., 1994; Sheppard et al., 1994). The main structural feature of the enzyme is a six-stranded β-barrel domain with extended interconnecting loop regions. A long active site cleft separates the β-barrel domain from the remainder of the structure. This groove runs the full 42 Å across the surface of the enzyme. The fold is unrelated to any of the published glycosylhydrolase structures but is similar to barwin, an oligosaccharide-binding plant defense protein (Ludvigsen & Poulsen, 1992). The catalytic activity of barwin, if any, is not known.

The structure of the native EGV revealed the presence of two completely conserved catalytic aspartates, either side of the substrate binding groove, with their carboxylate groups separated by 8−9 Å (Davies et al., 1993). This spatial separation is exactly as expected (McCarter & Withers, 1994) for EGV which has been shown to catalyze glycosyl group hydrolysis with inversion of configuration at the anomeric carbon (Schou et al., 1993). Studies have shown that EGV has six kinetically significant subsites for carbohydrate binding, hereafter referred to as -4 to $+2$ (Schou et al., 1993). It is able to cleave the β-1,4 glycosidic bond of the cellulose polymer between subsites -1 and $+1$ to yield either cellobiose or cellotriose as the product (Figure 1). The activity on carboxymethyl-substituted cellulose suggests that EGV functions as an endocellulase (EC 3.2.1.4), consistent with the observation that the active site is an open cleft as opposed to a tunnel (Schou et al., 1993; Schülein et al., 1993). EGV acts, therefore, on the disordered and amorphous regions of a cellulose fiber and has little or no catalytic activity on highly crystalline cellulose.

Acid-catalyzed glycosyl group hydrolysis, with inversion of configuration, requires a proton donor, whose function is to protonate the glycosidic oxygen, and a base, which activates the nucleophile through a deprotonation event. Our preliminary studies on the structure of a product (cellobiose) complex indicated likely roles for the two catalytic aspartates. Asp 121 is situated in a predominantly hydrophobic environment, and its carboxylate group makes a hydrogen bond to the O4 atom of the $+1$ subsite glucopyranosyl unit. This suggested that Asp 121 is the proton donor in the reaction, serving to protonate the glycosidic oxygen and assist aglycon departure. The carboxylate of Asp 10 is situated some 8−9 Å away. This allows both the substrate and the nucleophilic water molecule to bind between the two aspartates. Asp 10, therefore, acts as the catalytic base, activating a solvent water molecule for nucleophilic attack.

These proposals were made on the basis of our initial analyses of the cellobiose complex. In addition to a further analysis of the 1.5 Å native and 1.9 Å resolution EGV−cellobiose complex, here we describe the structure of an inactive mutant of one of the catalytic aspartates (D10N), both in an apo conformation and in a complex with the substrate cellohexaose [$EGV_{(D10N)}$−cellohexaose]. As well as providing information on oligosaccharide binding and the likely mechanism of catalysis of EGV, this ensemble of structures reveals a number of large conformational changes associated with a loop adjacent to the active site. This movement encloses the -1 and $+1$ subsites at the point of enzymatic cleavage, greatly increasing the hydrophobic nature of the environment adjacent to the catalytic proton donor, and brings a third aspartate residue over the -1 subsite. Kinetic analysis of site-directed mutants of these three catalytic aspartates confirms their probable roles in catalysis.

MATERIALS AND METHODS

Crystallization, Data Processing, and Structure Solution. Crystals of native EGV complexed with cellobiose, EGV mutant D10N, and EGV mutant D10N complexed with cellohexaose were grown essentially as described previously for the native enzyme (Davies et al., 1993). Crystals were grown by hanging-drop methods, from 10 mg mL^{-1} enzyme solutions buffered with 20 mM TRIS-HCl (pH 8.0) and using 10−22% (w/v) PEG 8000 as a precipitant. Cellobiose or cellohexaose was present in the crystallization conditions at a concentration of 5 mM. Crystals of the EGV D10N mutant complexed with cellohexaose grew only to a maximum width of 30 μm. This presented severe problems in crystal handling and mounting. In order to prevent crystal breakage, the single crystal used for data collection was mounted on a plateau within a 1 mm flattened glass capillary. The capillaries were flattened using a heated wire capillary flattener (Åkervall & Strandberg, 1971) kindly supplied by Professor Bror Strandberg at the University of Uppsala.

Data were collected on the RAXIS-II imaging plate system, using a Cu rotating anode source operating at 50 kV and 100 mA. All data were indexed manually and processed with the DENZO program (Z. Otwinowski, Yale University). Details of the data quality and completeness for the complexes and the D10N mutant are given in Table 1. All further computing used the CCP4 suite of programs

Table 1: Data Quality and Completeness for EGV (D10N) Mutant, the EGV−Cellobiose Complex, and the $EGV_{(D10N)}$−Cellohexaose Complexes

	D10N mutant		EGV−cellobiose		$EGV_{(D10N)}$−cellohexaose	
resolution (Å)	R_{merge}	completeness (multiplicity)	R_{merge}	completeness (multiplicity)	R_{merge}	completeness (multiplicity)
5.78	0.036	100.0 (3.5)	0.047	97.5 (3.5)	0.054	100.0 (3.6)
4.17	0.034	99.0 (3.5)	0.057	99.6 (3.6)	0.065	99.0 (3.6)
3.43	0.039	100.0 (3.6)	0.064	99.4 (3.6)	0.064	100.0 (3.6)
2.98	0.045	97.7 (3.6)	0.067	100.0 (3.7)	0.077	100.0 (3.6)
2.67	0.053	100.0 (3.6)	0.074	100.0 (3.6)	0.086	99.3 (3.5)
2.45	0.065	96.6 (3.6)	0.081	99.1 (3.6)	0.104	100.0 (3.5)
2.27	0.074	97.8 (3.7)	0.088	100.0 (3.5)	0.119	100.0 (3.5)
2.12	0.083	95.6 (3.7)	0.100	99.5 (3.5)	0.136	100.0 (3.5)
2.00	0.100	91.3 (3.3)	0.115	96.5 (3.1)	0.165	97.2 (3.0)
1.90	0.129	59.2 (3.0)	0.118	78.3 (2.3)	0.157	68.5 (1.4)
total	**0.050**	**91.4 (3.5)**	**0.071**	**96.3 (3.4)**	**0.085**	**95.1 (3.2)**

(Collaborative Computational Project, Number 4, 1994) unless otherwise stated. Molecular replacement was carried out using AMoRe (Navaza, 1994) with the native EGV structure as the search model. Extensive details of the refinement and analysis of the native EGV structure have been published elsewhere (Davies et al., 1996). The final native EGV model has a conventional crystallographic R factor of 0.105 with an R_{free} of 0.154 for all data between 8 and 1.5 Å resolution. Refinement was performed using standard procedures with the program PROLSQ (Hendrickson & Konnert, 1980), and manual rebuilding was performed with the program O (Jones et al., 1991). Restrained isotropic temperature factor refinement was considered appropriate at the resolution of 1.9 Å. Oligosaccharide atoms were not included in the refinement until refinement of the protein had reached convergence and the sugar density was clear and unambiguous. Stereochemical dictionaries for the refinement of the oligosaccharide moieties were calculated on the basis of the known structure of β-D-cellotetraose (Raymond et al., 1995). The final water model was completed using the ARP procedure (Lamzin & Wilson, 1993). Coordinates for the 1.5 Å native EGV structure, the 1.9 Å EGV−cellobiose structure, and the 1.9 Å $EGV_{(D10N)}$−cellohexaose structure have been deposited with the Brookhaven Protein Data Bank (Bernstein et al., 1977).

Site-Directed Mutagenesis and Kinetics. The EGV cDNA gene had been cloned, on the basis of peptide sequence, using standard techniques. The gene was subcloned into pUC 19 (Yanisch-Perron et al., 1985) using PCR techniques, introducing a *Bam*HI site into the −10 position and a *Sal*I site just downstream of the TAG stop codon. This construct was named pCaHj 170. Site-directed mutants were made from this construct using the method of Nelson and Long (1989). Initially, a *Cla*I site was introduced into the EGV gene in position 533 using the following primer, 5′ AGTGCGATCGATTCCCCGACG 3′. The mutated EGV PCR fragment was digested with *Nco*I and *Xho*I and the resulting fragment used to replace the *Nco*I-*Xho*I portion (244−768) of the EGV gene in pCaHj 170. The resulting plasmid was named pCaHj 171.

The following primers were used to construct the D10N, D114N, and D121 mutants: D10N, 5′ CCGCTACTGGAACTGCTGCAAG 3′; D114N, 5′ CACTGGCGGTAATCTTGGCAGC 3′; and D121N, 5′ CAACCACTTCAATCTCAACATC 3′. The D10N mutation was subcloned from the PCR fragment as a *Bam*HI-*Nco*I fragment into pCaHj 171, and the D114N and D121N fragments were subcloned as *Nco*I-*Cla*I fragments. The resulting plasmids were sequenced to confirm the introduction of the mutation and as a control for PCR errors. These mutated EGV genes were then subcloned as *Bam*HI-*Sal*I fragments into the *Aspergillus oryzae* expression system (Christensen et al., 1988). Transformants were grown in flasks or 10 L tank fermentors using standard conditions.

Kinetic measurements were made on the full length EGV, although on small, soluble substrates, identical results are obtained with the catalytic core domain only. k_{cat} determinations were performed using both carboxymethylcellulose and reduced cellohexaose as substrates. Results were obtained using the coupled assay system as described by Schülein and co-workers (Schou et al., 1993), using 82 μM 2,6-dichoroindophenol and 40 nM cellobiose dehydrogenase and with reduced cellohexaose at concentrations between 0.005 and 2 mM. The data for steady-state kinetic determinations were analyzed using GRAFIT 3.0 (Leatherbarrow, R. J., Erihacus Software Ltd).

RESULTS

Structure of the Native Enzyme and the Inactive D10N Mutant

EGV consists of a six-stranded β-barrel domain with a number of disulfide-bonded interconnecting loops (Figure 2a). A long active site groove runs across the surface of the molecule, to either side of which sit the two catalytic aspartate residues 10 and 121. These two aspartates are completely conserved; indeed, they are the only completely conserved acidic residues, in all the sequences of family 45 enzymes (Figure 2b). In the native enzyme structure, Asp 121 exhibits static disorder with two distinct conformations with relative occupancies of 0.7 and 0.3. These two conformations may represent the two protonation states of this residue in the native enzyme conformation. There is no electron density for the loop from residue 112 to 117, and this is therefore presumed to be disordered in the native structure. Asp 121 sits in the center strand of a three-stranded mixed β-sheet that makes up one side of the active site cleft. Asp 10 sits on the opposite side of the cleft, its Cα and carboxylate group separated from the equivalent atoms of Asp 121 by 11.5 and 8.5 Å, respectively. A tyrosine residue, Tyr 8, makes up the floor of the active site cleft at this point. A solvent cavity protrudes into the structure adjacent to the hydroxyl function of Tyr 8, and this contains eight well-defined water molecules which may play

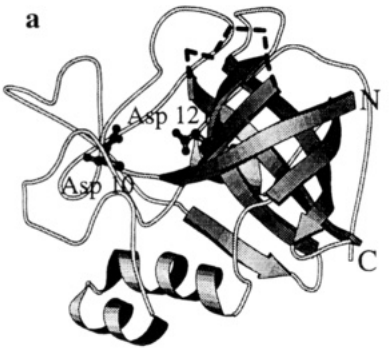

FIGURE 2: (a) Structure of the endoglucanase V from *Humicola insolens*. The two catalytic aspartate residues 10 and 121 sit to either side of the long active site cleft across the surface of the molecule. The figure was drawn with the MOLSCRIPT program (Kraulis, 1991). The dashed line indicates the position of residues 112−117 in the cellobiose complex. This loop is disordered in the native enzyme structure but ordered in both the complexes. (b) Sequence conservation around the two residues essential for catalysis in all the published family 45 sequences. The catalytic base, Asp 10, is marked with an asterisk and the proton donor (Asp 121) with a double dagger.

Table 2: Crystallographic Parameters for the EGV Crystal Forms

protein	space group	cell dimensions	resolution (Å)	R_{cryst}
native EGV[a]	$P2_1$	$a = 42.1, b = 51.7, c = 45.0, \beta = 109.4°$	1.48	0.105
EGV D10N	$P2_1$	$a = 41.9, b = 51.3, c = 44.6, \beta = 109.6°$	1.90	0.163
EGV−cellobiose	$P2_1$	$a = 41.8, b = 56.9, c = 37.4, \beta = 93.4°$	1.90	0.174
$EGV_{(D10N)}$−cellohexaose	$P2_1$	$a = 41.9, b = 51.9, c = 42.0, \beta = 97.7°$	1.90	0.170

[a] The data for the native EGV structure are from Davies et al. (1996).

a role in catalysis. Asp 121 sits in a predominantly hydrophobic environment. It is flanked, on one side, by the hydrophobic moieties of alanines 73 and 74 (also completely conserved), and the carboxylate group sits approximately 3.95 Å above the Cβ−Cγ atoms of Tyr 8. One of the carboxylate oxygens of Asp 121 hydrogen bonds to the hydroxyl group of Thr 6, which is, in turn, hydrogen-bonded to His 119. The other carboxylate oxygen makes a hydrogen bond to a single water molecule. The environment of Asp 121, together with its oligosaccharide interactions described below, leads us to believe that this residue is the catalytic proton donor. Asp 10 is in a more hydrophilic environment. The carboxylate group makes a number of interactions with solvent water molecules, each carboxylate oxygen making hydrogen bonds to two well-ordered solvent molecules. This environment, together with the sequence conservation and the spatial separation to Asp 121, suggests that Asp 10 functions as the catalytic base.

In order to try and obtain a complex of EGV with its substrate, an aspartate to asparagine mutant of the proposed catalytic base Asp 10 was constructed. This mutant is completely inactive in all kinetic studies; see below. In order to confirm the integrity of this mutant the structure of this mutant, was determined at 1.9 Å resolution. The details of the crystallographic parameters for this mutant and the oligosaccharide complexes are given in Table 2. The EGV D10N mutant crystallized with the same space group and cell dimensions as the native enzyme. The final crystallographic R factor is 0.163, with deviations from stereochemical ideality of 0.012 and 0.040 Å for bonds and angles, respectively. The D10N mutant, although completely inactive, has a structure almost isomorphous with that of the native enzyme, indicating that the lack of activity stems purely from the removal of the negative charge. There is a slight structural change from the native enzyme in that two waters bound to Asp 10 appear to be approximately 0.15 Å more distant from the carboxylate group. In all other respects, the structure is identical to that of the native EGV; the root mean square coordinate difference on main chain atoms between the two structures is 0.13 Å.

Table 3: Geometry of the Glycosidic Linkages[a]

structure	ϕ (deg)	ψ (deg)	C1−O4′−C4′ (deg)	O5−O3′ (Å)
cellotetraose				
−4/−3	−96.6	−141.2	117.3	2.84
−3/−2	−93.7	−141.5	117.3	2.85
$EGV_{(D10N)}$−cellohexaose				
−4/−3	−88.0	−129.9	117.6	2.73
−3/−2	−58.9	−110.3	117.9	3.26
+1/+2	−65.9	−124.1	121.9	2.77
+2/+3	−91.1	−126.1	115.1	2.86
EGV−cellobiose				
+1/+2	−64.72	−112.4	119.4	2.99

[a] The values for the small molecule structure of cellotetraose are from Raymond et al. (1995).

Structure of the Product (Cellobiose) Complex

The structure of a complex of the active catalytic core of EGV cocrystallized with the product cellobiose was solved at 1.9 Å resolution. The structure has a crystallographic R factor, at 1.9 Å, of 0.160, with deviations from stereochemical ideality of 0.012 and 0.041 Å for bonds and angles, respectively. Cellobiose is found bound in the leaving group subsites +1 and +2. The sugar density is clear and unambiguous in both subsites, with the average B values for the two glucosyl units of 18 and 21 Å^2, respectively. Both the sugar rings are found in the full $^4C_1$ chair conformation, although deviations from the geometry of the "small-molecule" structures of cellodextrins are found (Gessler et al., 1994; Raymond et al., 1995). The geometry of the cellobiose, together with that observed for the oligosaccharide in the $EGV_{(D10N)}$−cellohexaose interactions described below, is given in Table 3. The O4 oxygen of the glucose unit in the +1 subsite makes a hydrogen bond to catalytic Asp 121 (distance, 2.6 Å). In a true substrate complex, the O4 atom

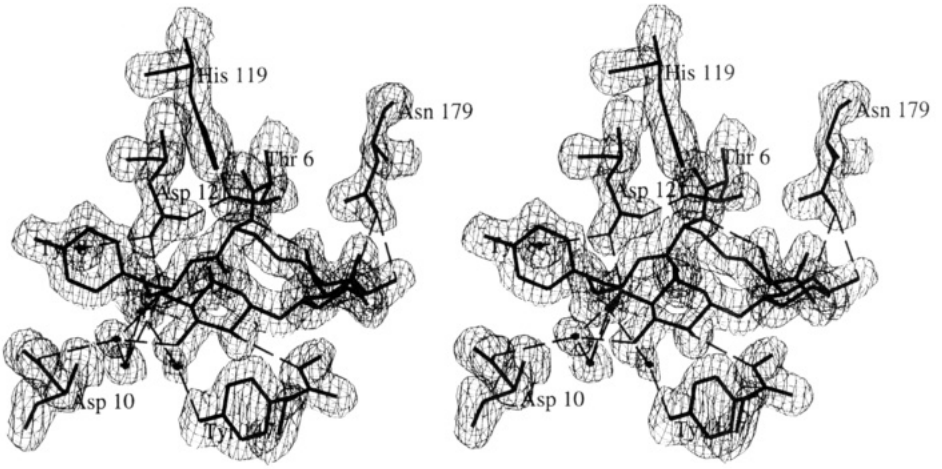

FIGURE 3: $2F_o - F_c$ density map, contoured at 0.44 eÅ^{-3}, for the cellobiose complex. Cellobiose is found bound in the "leaving-group" subsites +1 and +2.

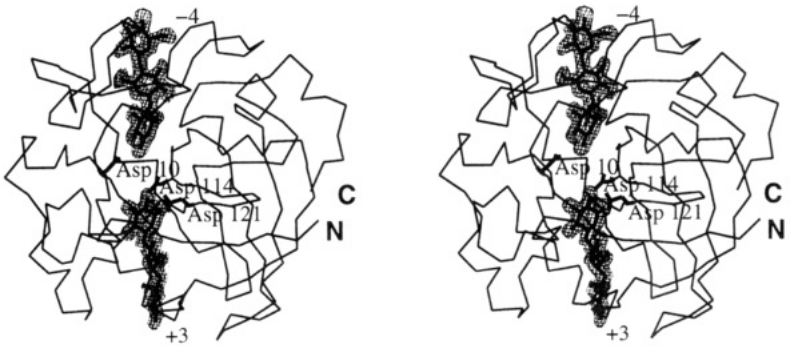

FIGURE 4: $2F_o - F_c$ density map, contoured at 0.44 eÅ^{-3} around the two cellotriose moieties observed in the EGV active site. Subsites −4 and +3 and aspartates 10, 114, and 121 are labeled for reference.

of the +1 subsite sugar would be the glycosidic oxygen linking the +1 subsite sugar to that in subsite −1. This is the atom which would become protonated by the catalytic proton donor. The proximity to Asp 121 suggests, therefore, that the Asp 121 residue functions as the catalytic proton donor in EGV. The O2 hydroxyl of the +1 subsite sugar hydrogen bonds to both the main chain carbonyl group of Gly 127 and the amide nitrogen of Gly 148. The O3 hydroxyl bonds to a water molecule and the carbonyl of Gly 128. The ring oxygen O5 makes a hydrogen bond to the O3 hydroxyl of the +2 subsite sugar. The O6 atom of the +1 subsite sugar makes a hydrogen bond to a solvent water molecule, which is linked *via* a second water molecule to the hydroxyl of Tyr 147. The tyrosine ring has moved some 2.4 Å from its position in the native structure to make this interaction. This is not achieved through a simple side chain rearrangement, but by a closing of the whole active site cleft, causing a reduction in its width of approximately 2 Å. In addition to this closing of the domains, the loop between strands V and VI (residues 112−117), disordered in the native conformation of the enzyme, becomes ordered. How the binding of cellobiose elicits an ordering of the loop is unclear. The only direct interaction between a loop amino acid and the cellobiose is *via* Asp 121. In turn, Asp 121 interacts *via* a complex hydrogen-bonding network involving Thr 6 through to the previously disordered side chain of His 119. The +2 subsite sugar makes fewer direct interactions with the protein. The O6 hydroxyl forms a hydrogen bond to the peptide amide group of Arg 7, and an exquisite interaction exists between the +2 subsite sugar and Asn 179. The side chain carbonyl makes a hydrogen bond to the O6 hydroxyl of the sugar, while the amide group donates hydrogen bonds to both the ring oxygen O5 and the free O1 hydroxyl. The cellobiose−enzyme interactions are shown, together with the appropriate density, in Figure 3.

Structure of the EGV$_{(D10N)}$−Cellohexaose Complex

The inactive D10N mutant was used to prepare a substrate complex in order to allow a more complete description of the enzyme−substrate interactions. The structure was solved by molecular replacement at 1.9 Å resolution. The crystallographic *R* factor is 0.17 for all data between 10 and 1.9 Å, with deviations from stereochemical ideality of 0.011 and 0.042 Å for bonds and angles, respectively. Even in the initial molecular replacement map, prior to refinement, the electron density for sugars in sites −4, −3, −2, +1, +2, and +3 was clear and unambiguous. All atoms of the glucose units are clearly visible (Figure 4). This is reflected in the temperature factors for the sugars after refinement. The average *B* values for the glucose units in the −4 to +3 subsites are 25, 9, 8, 10, 15, and 42 Å^2, respectively, indicating that the mobility of the sugars increases toward the peripheries of the active site cleft. All the glucopyranosyl rings are found in the full $^4C_1$ chair conformation. No sugar density is observed in the −1 subsite. A striking feature of the interactions along the active site cleft is the absence of stacking interactions with aromatic residues. These have been observed for both sugar binding proteins in general and many of the previously determined cellulase structures: the CBH-II from *T. reesei* (Rouvinen et al., 1990), endoglucanase E2 from *Thermanospora fusca* (Spezio et al., 1993), CelD

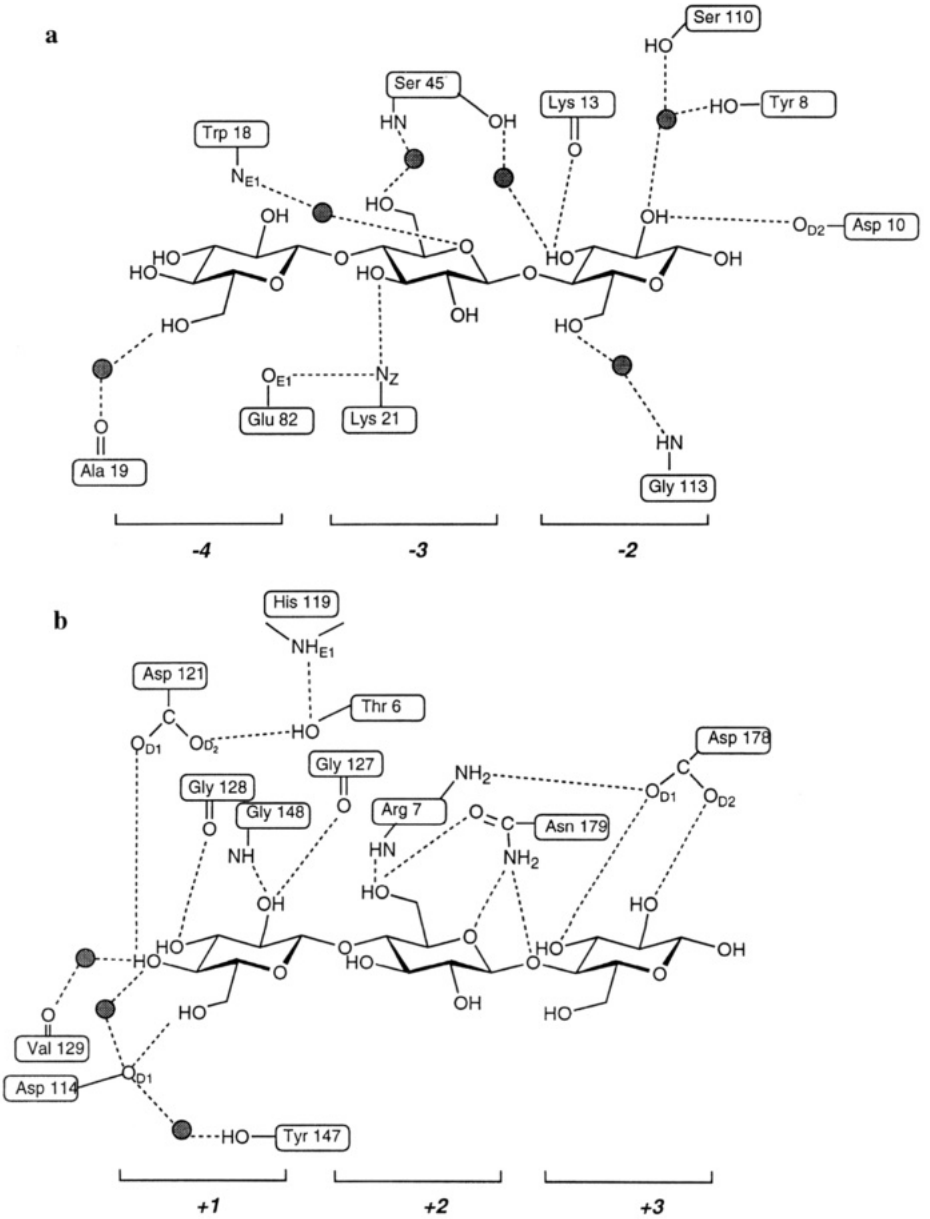

FIGURE 5: Schematic diagram of the enzyme−oligosaccharide hydrogen bond network in the $EGV_{(D10N)}$−cellohexaose complex in the (a) −4, −3, and −2 subsites and (b) the +1, +2, and +3 subsites. Water molecules are shown as shaded spheres.

from *Clostridium thermocellum* (Juy et al., 1992), CBH-I from *T. reesei* (Divne et al., 1994), and the endoglucanse I from both *H. insolens* and *F. oxysporum* (G. J. Davies and M. Schülein, unpublished observations). With the solitary exception of the −4 subsite, no aromatic stacking interactions are observed in the EGV active site. EGV has an absolute specificity for cellulose, as opposed to xylan. The interactions of the O6 hydroxyls, which are absent in xylose, are consistent with this observation. The O6 hydroxyl groups, in all the subsites, apart from the peripheral +3 subsite, make interactions with the protein molecule. Those in subsites +1 and +2 interact *via* direct hydrogen bonds to the protein and those in the −4, −3, and −2 subsites through solvent-mediated hydrogen bonds. Figure 5 shows the sugar-protein interactions and a list of the hydrogen bonds between sugar and protein, both direct and mediated by solvent, is given in Table 4. The observation that seven sugar units fit in the active site, with the +3 subsite sugar showing high mobility, is consistent with the kinetic observations that EGV has six kinetically distinct subsites for carbohydrate binding.

(a) Subsite −4. As mentioned above, subsite −4 is the only glucosyl binding site with the "classical" stacking above a tryptophan residue as has been observed in many sugar binding proteins. The hydrophobic α-face of the sugar sits approximately 3.5 Å above the indole ring of Trp 18. There are no other direct interactions with the enzyme besides a water-mediated hydrogen bond between the O6 hydroxyl and the carbonyl oxygen of Ala 19.

(b) Subsite −3. There are direct sugar−protein hydrogen bonds between the O3 hydroxyl and the NZ of Lys 21. Lys 21 is anchored in this position by a salt link to Glu 82. There are no other direct interactions with the enzyme. The O6 hydroxyl makes a water-mediated hydrogen bond with the amide group of Ser 45, and the O5 oxygen makes a hydrogen bond with the indole NE1 of Trp 18.

(c) Subsite −2. Again, the −2 subsite sugar makes remarkably few direct interactions with the protein. The O3 hydroxyl makes a direct hydrogen bond to the carbonyl oxygen of Lys 13, and the O2 with Asp 10, but all other −2 subsite interactions are mediated by solvent water molecules.

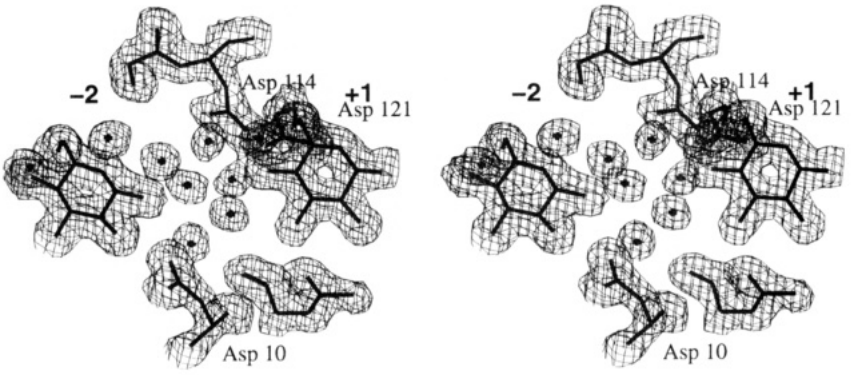

FIGURE 6: $2F_o - F_c$ density map, contoured at 0.44 eA^{-3}, for the EGV D10N mutant−cellohexaose complex in the vicinity of the active site. The ordered water molecules in the −1 subsite are clearly visible as the density for Asp 114 which has moved over 12 Å compared to its position in the cellobiose complex.

Table 4: Summary of the Hydrogen Bonds (≤3.1 Å) in the EGV$_{(D10N)}$−Cellohexaose Complex

direct hydrogen bond to protein				water-mediated hydrogen bond				
subsite	substrate atom	*d* (Å)	protein atom	substrate atom	*d* (Å)	water molecule	*d* (Å)	protein atom
−4				O6	2.8	9OW3	3.0	Ala 19 O
−3	O3	3.0	Lys 21 NZ	O5	2.9	8OW8	2.6	Trp 18 NE1
				O6	2.8	9OW1	2.8	Ser 45 N
−2	O2	3.1	Asp 10 OD2	O2	2.8	8OW6	2.8	Tyr 8 OH
	O3	2.7	Lys 13 O	O2	2.8	8OW6	2.6	Ser 110 OG
				O3	2.7	9OW2	2.7	Ser 45 OG
				O6	2.8	10OW1	2.9	Gly 113 N
+1	O2	2.9	Gly 127 O	O4	2.7	11OW1	3.1	Val 129 O
	O2	2.9	Gly 148 N	O4	3.1	11OW3	2.5	Asp 114 OD1
	O3	2.8	Gly 128 O					
	O4	2.6	Asp 121 OD1					
	O6	2.5	Asp 114 OD1					
+2	O5	3.0	Asn 179 ND2					
	O6	2.4	Asn 179 OD1					
	O6	3.1	Arg 7 N					
+3	O2	2.3	Asp 179 OD2					
	O3	2.9	Asp 178 OD1					
	O4	3.1	Asn 179 ND2					

The O2 hydroxyl interacts with the hydroxyl of Tyr 8 and the hydroxyl function of Ser 110 *via* a single solvent water molecule. The free O1 hydroxyl interacts with a network of water molecules, none of which could be said to make good, direct, hydrogen bonds to the enzyme. The interactions between the O4, O5, and O6 atoms are of particular importance since they seem to be responsible for the conformational changes which are discussed more fully below. The O4 and O5 atoms make weak solvent-mediated interactions with the OD2 of Asp 114. The O6 atom also makes a critical solvent-mediated interaction with the peptide amide group of Gly 113. These latter interactions are probably those which contribute most to the conformational change since Gly 113 seems to act as the "hinge" for this movement.

(d) Subsite −1. Subsite −1 is not occupied by oligosaccharide in this complex. Instead, the subsite is occupied by a number of discrete water molecules (Figure 6). While it is a formal possibility that the −1 subsite sugar is disordered, this seems extremely unlikely. The water molecules in its place are extremely well-defined, leaving no room for a sugar ring, and the flanking atoms of the adjacent −2 and +1 subsite sugars are well-ordered: the O1 atom of sugar −2 with a *B* value of 8.6 Å^2 and the O4 atom of the +1 subsite sugar with a B value of 9.7 Å^2. Another possibility is that the cellohexaose has been cleaved by the D10N mutant in the high protein concentration of the crystal/crystallization conditions, and the cleaved sugars (presumably cellotriose) have migrated within the sugar binding sites to leave site −1 unoccupied. Kinetic studies have shown that the D10N mutant is totally inactive in all assays, and this possibility seems unlikely. In addition, crystallographic studies of cellotriose and cellotetraose complexes of EGV (data not shown) indicate that the crystals have different cell dimensions and show very poor disordered density in the active site, which could not be modeled as oligosaccharide. It seems more probable that binding to the −1 (cleavage) subsite is energetically unfavorable for substrate; this is discussed in more detail below. What we observe is two cellohexaose molecules. The first has three glucose units disordered in the solvent and then three units entering the −4, −3, and −2 subsites. A second cellohexaose molecule binds in the +1, +2, and +3 subsites, with the last three units again being disordered in the solvent. The gradual change in *B* values of the sugars, which increase toward the entrance and exit to the active site cleft, supports this proposal. Indeed, electron density which cannot satisfactorily be modeled as solvent exists adjacent to the −4 and +3 subsite sugars; it is too diffuse however, to be modeled as ordered oligosaccharide either. A conformational change,

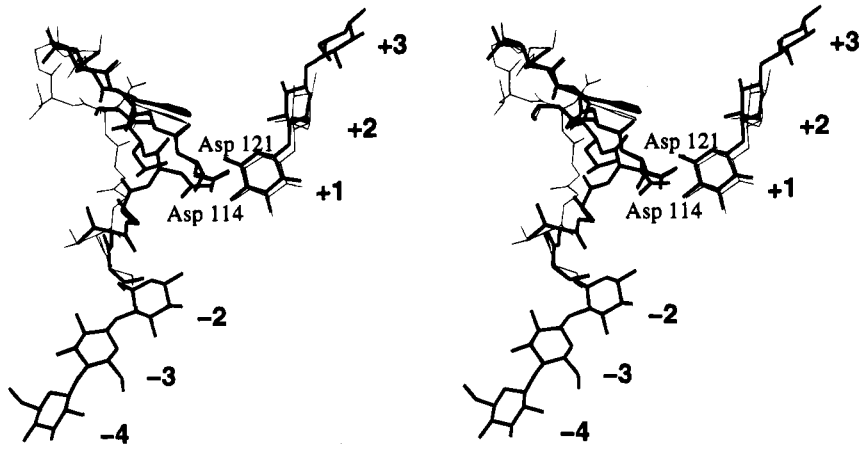

FIGURE 7: Protein atoms involved in the loop movement. Faint lines represent the atomic positions in the cellobiose complex and thicker lines those in the $EGV_{(D10N)}$−cellohexaose complex. Labels are shown for the positions of Asp 114 and Asp 121 as they are found in the cellohexaose complex.

involving the disordered loop, between residues 112 and 117, has resulted in Asp 114 moving into a position over the −1 subsite.

(e) Subsites +1, +2, and +3. The two sugars in subsites +1 and +2 occupy positions almost identical to those of the sugars in these subsites in the cellobiose complex discussed above. Again the +1 sugar O4 atom makes a 2.6 Å hydrogen bond to Asp 121, confirming its likely role as the proton donor in the reaction. The only difference in the binding compared to cellobiose comes from the direct interaction of Asp 114 with the O6 position of the +1 subsite sugar and a solvent-mediated interaction with the O4 group. These extra interactions are made in place of hydrogen bonds to water molecules which these groups make in the cellobiose complex. This direct interaction between the OD1 atom of Asp 114 and the O6 of sugar +1 is again likely to be a crucial interaction involved in the absolute specificity of the enzyme for cellulose as opposed to xylan. Subsite +3 is the least well-defined of the sugar binding sites having an average *B* value for the sugar atoms of 42 $Å^2$. The sugar sits on the periphery of the enzyme surface, with the only interaction being between Asp 178 with both the O3 and O2 positions of the sugar. All the other potential hydrogen bond donors and acceptors of the +1, +2, and +3 subsite sugars make hydrogen bonds to solvent water molecules, which are not linked directly to the enzyme. Presumably, the paucity of interactions of the leaving group subsite sugars is energetically sensible in allowing product release and preventing cellobiose from acting as an inhibitor for the reaction.

Conformational Change

As mentioned above, the loop from residues 112 to 117 is disordered in the native EGV structure. The sequence of this loop is ^{112}Gly-Gly-Asp-Leu-Gly-Ser^{117}. Glycine rich loops such as these are indicative of high mobility, and indeed, such a loop plays a dynamic role in catalysis by pancreatic α-amylase (Qian et al., 1994). In EGV, this loop occurs between strands V and VI and corresponds to a loop which is disordered in the solution structure of the only protein of related structure, barwin (Ludvigsen & Poulsen, 1992). When complexed with cellobiose, this loop becomes ordered but makes few interactions with the cellobiose itself. In this conformation, residues Asp 114 and Leu 115 are both solvent-exposed. The structure of the D10N mutant−cellohexaose complex, however, reveals a significantly different conformation for this loop (Figure 7). In this case, the interactions with the sugars are defined better. The magnitude of the loop movement is substantial (Figure 8). The main chain atoms of residues Asp 114 and Leu 115 have moved approximately 7 Å, and the apex of the side chain of Leu 115 has "flipped" in moving almost 15 Å. The results of this are twofold. The first is to position Leu 115 adjacent to Asp 121, making the environment of the acid group considerably more hydrophobic, and presumably assisting in keeping the pK_a of Asp 121 suitably raised, and the second is that Asp 114 becomes positioned immediately above the vacant −1 subsite and interacts with the O6 hydroxyl of the +1 subsite sugar. The combined effect of this "lid flipping" is to form a miniature tunnel over the cleavage site. In addition to the many beneficial interactions that this movement bestows, it may also serve to prevent access of potential nucleophiles other than solvent water to the active site. The basis of this conformational change seems to stem entirely from the interactions of the −2 subsite sugar. Two interactions may be critical. The ring oxygen O5 of the −2 subsite sugar makes a water-mediated hydrogen bond to the peptide amide group of Gly 113. This residue is completely disordered in the native enzyme structure and has been displaced some 3.5 Å from its observed position in the cellobiose complex in order to make this interaction. A second rearrangement takes place in allowing the hydroxyl of Ser 45 to interact, again *via* solvent, with the O3 hydroxyl of the −2 sugar. In addition to these interactions, there are poor solvent-mediated hydrogen bonds between the O4 and O5 atoms with Asp 114.

Site-Directed Mutagenesis

Site-directed mutagenesis of the three aspartates implicated in catalysis was performed. All three asparates were mutated to asparagine. The results of these mutations on the catalytic activity of EGV, measured by the hydrolysis of reduced cellohexaose, are shown in Table 5. Mutation of either Asp 10 or Asp 121 causes total inactivation of the enzyme within the limits of accuracy of the assay. This is consistent with the absolute sequence conservation of these residues and their fundamental roles in catalysis. Mutation of Asp 114 to Asn reduces the k_{cat}/K_m on reduced cellohexaose by a factor of

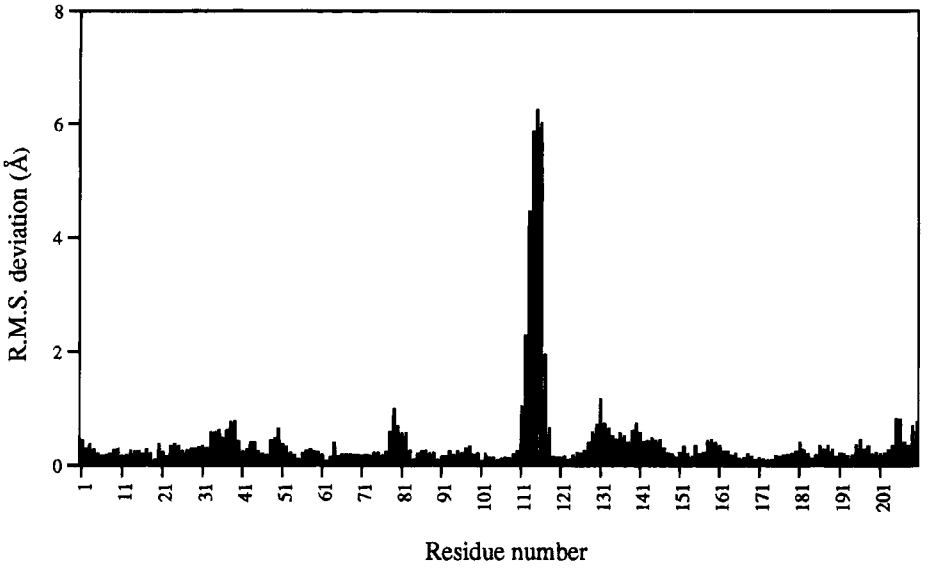

FIGURE 8: Root mean square deviations in main chain positions between the EGV−cellobiose complex and the $EGV_{(D10N)}$−cellohexaose complex. The loop from 112 to 117 is disordered in the native EGV structure.

Table 5: Steady State Kinetic Parameters for the Hydrolysis of Reduced Cellohexaoase by Native EGV and the Site-Directed Mutants of the Catalytic Aspartates

protein	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($s^{-1}\,\mu M^{-1}$)
EGV native	17.7	139	1.27×10^{-1}
D10N	<0.0001	–	–
D121N	<0.0001	–	–
D114N	0.11	158	6.96×10^{-4}

160. Again this is consistent with a significant, but more minor, role for Asp 114. This mutation also reduces the enzymatic activity on soluble polymeric substrates such as CMC by a factor of approximately 20 (data not shown), but these data are not straightforward to interpret as the activity determinations are only possible at a concentration of substrate below the K_m. The change in k_{cat}/K_m for the D114N mutant with reduced cellohexaose as substrate comes mainly from a reduction in k_{cat}, suggesting that the role of this residue is catalytic as opposed to binding. The function of Asp 114 is perhaps in transition state stabilization or in the regeneration of the correct protonation states of the catalytic residues. An aspartate in an equivalent position in this loop (in a Gly-X-Asp hydrophobic motif) is found in the other EGV sequences apart from the *T. reesei* enzyme. The *T. reesei* sequence, which is significantly different from the *H. insolens*, *F. oxysporum*, and *P. fluorescens* sequences in a great many areas, has no equivalent residue in this region of the sequence.

Recent work by Withers and co-workers on the inverting cellulase from family 6, CenA, has also indicated the importance of a third catalytic group in catalysis by this enzyme (Damude et al., 1995). In this structure, mutation to the catalytic proton donor or the base again indicates the crucial role of these groups in catalysis. Mutation of another residue, Asp 287, has a significant but not critical effect on the k_{cat} for the reaction. This effect is much larger on substrates requiring significant acid catalysis as opposed to substrates with an aglycon of low pK_a. The authors conclude that the role of this third residue may well be to assist in keeping a suitable environment in which the pK_a of the catalytic proton donor is elevated. This is indicated by the fact that these residues are separated only by 3.8 Å. The position of Asp 114 in the EGV complex, separated by 5.6 Å from the proton donor Asp 121, may be consistent with a similar role for Asp 114. Indeed, in the native EGV structure, Asp 121 is found in two discrete conformations, perhaps indicative of two protonation states for this residue in the native state. It is possible that the conformational change, which, in addition to bringing Asp 114 into reasonable proximity to Asp 121, also increases the hydrophobic environment of the catalytic proton donor, is fundamentally responsible for achieving the correct pK_a for Asp 121.

DISCUSSION

Mechanism of Catalysis by EGV. No sugar is seen bound in the −1 subsite, which is instead occupied by a number of discrete water molecules. There is an 8.5 Å separation between the C1 atom of the sugar in the −2 subsite and the C4 atom of the +1 subsite. This is 1 Å longer than is observed for the equivalent distance in either the crystal structures of cellodextrins or the cellotriose moieties observed in the ABC and EFG subsites of EGV. Neither a $^4C_1$ chair hexopyranose, with regular glycosidic bond lengths and angles, nor a distorted sugar ring such as is found in the HEWL NAM-NAG-NAM complex (Strynadka & James, 1991) spans this distance satisfactorily. One possible explanation for this apparent anomaly is that the observed sugar binding constitutes some form of artifact. The 1.9 Å electron density maps for both the $EGV_{(D10N)}$−cellohexaose and EGV−cellobiose complexes are well-defined and unambiguous. The low-temperature factors found for the modeled sugars are characteristic of precise positioning. The sugars found in the +1 and +2 subsites in the crystallographically independent $EGV_{(D10N)}$−cellohexaose complex and EGV-cellobiose complex occupy nearly identical positions. All these factors are indicative of an authentic sugar model. Perhaps an alternative explanation lies in the fact that the enzyme has evolved not to optimally bind to the substrate but, instead, to interact more favorably with the transition state. Acid-catalyzed hydrolysis of glycosidic bonds does indeed involve a substantial increase in the length of the scissile bond upon protonation, with values of up to 0.8 Å having been proposed (Tanaka et al., 1994). Changes

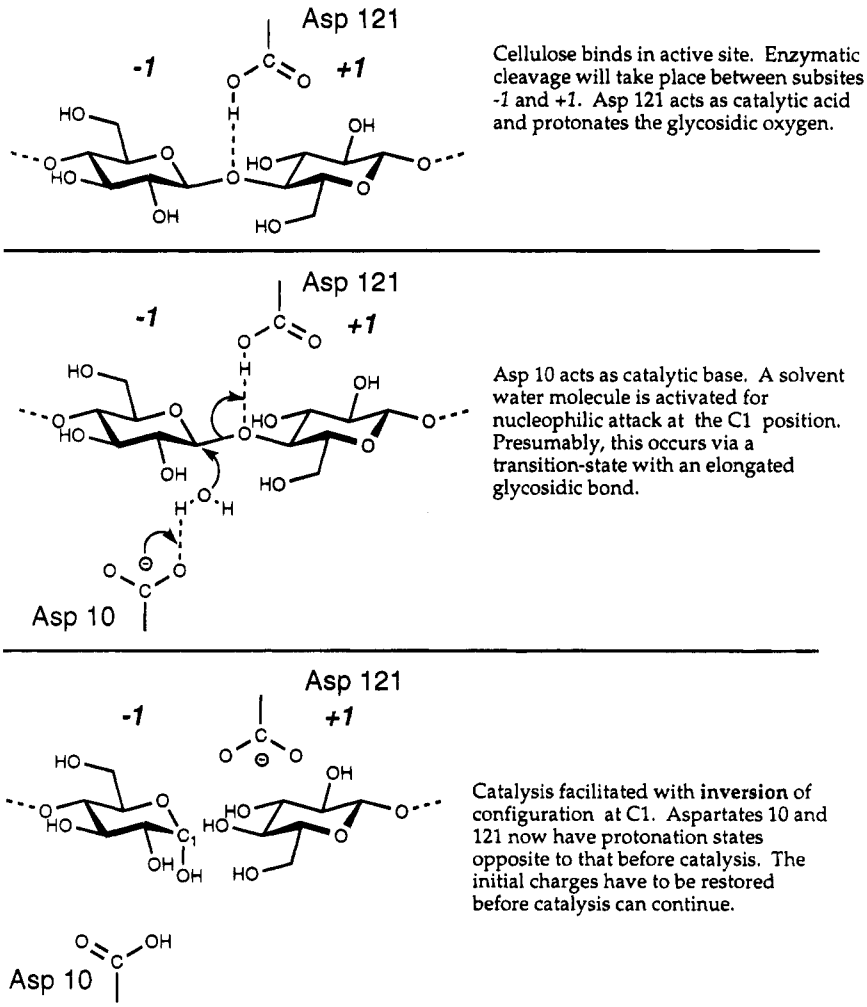

FIGURE 9: Likely mechanism for catalysis by EGV, leading to inversion at the anomeric carbon.

of this magnitude are quite consistent with the additional length of the −1 subsite. If this bond lengthening in the transition state is favored by the enzyme itself, then it is not surprising that, in the inactive $_{D10N}$−cellohexoase structure, binding of substrate across the −2 to +1 subsites is energetically unfavorable and therefore not observed. Comparison of the geometry of the oligosaccharide bound to EGV with that of the known cellotetraose structures reveals small but significant differences (Table 2). In general, the geometry of the sugars deviates more from the known cellotetraose geometry the closer the sugars are to the −1 subsite. This is particularly reflected in the ϕ/ψ angles, which deviate by up to 40° between the −3 and −2 subsite sugars, and in the 4.6° difference in the glycosidic bond angle for the +1 to +2 linkage. The absence of a sugar in the −1 subsite means that it is not possible to speculate how EGV breaks the O5−O3′ hydrogen bond between the −1 and +1 subsite sugars that would be present in the natural substrate.

It is possible to model a feasible orientation for the −1 subsite sugar. This is based on a continuation of the natural twist of the sugars through the −4, −3, and −2 subsites, although it must be stated that such a simple model would be difficult to rationalize with the orientation of the +1 subsite sugar, for the reasons stated above. In this orientation, the α-face of the −1 subsite-modeled sugar points toward Asp 10, with the glycosidic oxygen being protonated by Asp 121. Between Asp 10 and the C1 position, there is room for a single water molecule. This, together with the absolute sequence conservation of these residues and their respective environments, suggests that catalysis by EGV probably involves the activation of a water molecule for nucleophilic attack on C1, by Asp 10. This orientation of catalytic groups is as predicted for inverting enzymes by Koshland (1953). This would suggest a simple single displacement, S_N2, model for catalysis by EGV (Figure 9). Others have pointed out the likely similarity in the transition states of both inverting and retaining glycosylhydrolases (Tanaka et al., 1994). The inverting mechanism of EGV is, therefore, likely to proceed *via* a transition state with some degree of oxocarbonium ion character, suggesting that the S_N2 model suggested by the spatial arrangement of the catalytic residues is not the full story. It is possible that Asp 114 plays a role in maintaining an appropriate environment around the proton donor. In the absence of any mechanistic data, we are not able to speculate further.

We can see, therefore, that the active site geometry of EGV is consistent with a single displacement inverting mechanism. The active site of EGV, however, bares scant relationship to that first observed for an inverting cellulase CBH-II. In this structure, the active site again contains the typical aspartate and glutamate residues, but their separation is much greater than that observed here. The disposition of active site residues of EGV is more similar to that observed for another inverting cellulase, CelD from *C. thermocellum*. This suggests that the mode of catalysis by the family 6 enzymes may be substantially different from that performed by EGV

and CelD. This is backed up by mechanistic data on the family 6 cellulases which suggests that catalysis may proceed in a more complex manner than that described above (Konstantinidis et al., 1993). In order to further our knowledge on catalysis by family 45 enzymes, mechanistic data and the design of appropriate inhibitors will be necessary. This work is currently in progress.

ACKNOWLEDGMENT

The authors thank Igor Tvaroska for stimulating discussions and the staff at Chamrousse for invaluable assistance.

REFERENCES

Åkervall, K., & Strandberg, B. (1971) *J. Mol. Biol. 62*, 625−627.
Béguin, P., & Aubert, J.-P. (1994) *FEMS Microbiol. Rev. 13*, 25−58.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. T., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.
Christensen, T., Wöldike, H., Boel, E., Mortensen, S. B., Hjortshøj, K., Thim, L., & Hansen, M. T. (1988) *Bio/Technology 6*, 1419−1422.
Collaborative Computational Project, Number 4. (1994) *Acta Crystallogr. D50*, 760−763.
Coughlan, M. P. (1985) *Biochem. Soc. Trans. 13*, 405−407.
Damude, H. G., Withers, S. G., Kilburn, D. G., Miller, R. C., Jr., & Warren, R. A. J. (1995) *Biochemistry 34*, 2220−2224.
Davies, G., & Henrissat, B. (1995) *Structure 3*, 853−859.
Davies, G., Dodson, G., Moore, M. H., Tolley, S. P., Dauter, Z., Wilson, K. S., Rasmussen, G., & Schülein, M. (1996) *Acta Crystallogr. D* (in press).
Davies, G. J., Dodson, G. G., Hubbard, R. E., Tolley, S. P., Dauter, Z., Wilson, K. S., Hjort, C., Mikkelsen, J. M., Rasmussen, G., & Schülein, M. (1993) *Nature 364*, 362−364.
Divne, C., Ståhlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K. C., Teeri, T. T., & Jones, A. (1994) *Science 265*, 524−528.
Dominguez, R., Souchon, H., Spinelli, S., Dauter, Z., Wilson, K. S., Chauvaux, S., Béguin, P., & Alzari, P. M. (1995) *Nat. Struct. Biol. 2*, 569−576.
Ducros, V., Czjzek, M., Belaich, A., Gaudin, C., Fierobe, H.-P., Belaich, J.-P., Davies, G. J., & Haser, R. (1995) *Structure 3*, 939−949.
Gessler, K., Krauss, N., Steiner, T., Betzel, C., Sandmann, C., & Saenger, W. (1994) *Science 266*, 1027−1029.
Gilbert, H. J., Hall, J., Hazlewood, G. P., & Ferreira, L. M. A. (1990) *Mol. Microbiol. 4*, 759−767.
Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., Jr., & Warren, R. A. J. (1991) *Microbiol. Rev. 55*, 303−315.
Hendrickson, W. A., & Konnert, J. H. (1980) in *Biomolecular structure, conformation and evolution* (Srinivasan, E., Ed.) Pergamon, Oxford.
Henrissat, B. (1991) *Biochem. J. 280*, 309−316.
Henrissat, B. (1994) *Cellulose 1*, 169−196.
Henrissat, B., & Bairoch, A. (1993) *Biochem. J. 293*, 781−788.
Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
Juy, M., Amit, A. G., Alzari, P. M., Poljak, R. J., Claeyssens, M., Bguin, P., & Aubert, J.-P. (1992) *Nature 357*, 89−91.
Konstantinidis, A. K., Marsden, I., & Sinnott, M. L. (1993) *Biochem. J. 291*, 883−888.
Koshland, D. E. (1953) *Biol. Rev. 28*, 416−436.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J., & Gronenborn, A. M. (1989) *Biochemistry 28*, 7241−7257.
Lamzin, V. S., & Wilson, K. S. (1993) *Acta Crystallogr. D49*, 129−147.
Ludvigsen, S., & Poulsen, F. M. (1992) *Biochemistry 31*, 8783−8789.
McCarter, J. D., & Withers, S. G. (1994) *Curr. Opin. Struct. Biol. 4*, 885−892.
Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
Nelson, P. M., & Long, G. L. (1989) *Anal. Biochem. 180*, 147−151.
Qian, M., Haser, R., Buisson, G., Dueé, E., & Payan, F. (1994) *Biochemistry 33*, 6284−6294.
Raymond, S., Heyraud, A., Tran Qui, D., Kvick, A., & Chanzy, H. (1995) *Macromolecules 28*, 2096−2100.
Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J. K. C., & Jones, T. A. (1990) *Science 249*, 380−386.
Saloheimo, A., Henrissat, B., Hoffrén, A.-M., Teleman, O., & Penttilä, M. (1994) *Mol. Microbiol. 13*, 219−228.
Schou, C., Rasmussen, G., Kaltoft, M.-B., Henrissat, B., & Schülein, M. (1993) *Eur. J. Biochem. 217*, 947−953.
Schülein, M., Tikhomirov, D. F., & Schou, C. (1993) in *Proceedings of the second TRICEL symposium on Trichoderma reesei cellulases and other hydrolases* (Suominen, P., & Reinikainen, T., Eds.) pp 109−116, Foundation for Biotechnical and Industrial Fermentation Research, Espoo, Finland.
Sheppard, P. O., Grant, F. J., Oort, P. J., Cindy, A. S., Foster, D. C., Hagen, F. S., Upshall, A., McKnight, G. L., & O'Hara, P. J. (1994) *Gene 150*, 163−167.
Spezio, M., Wilson, D. B., & Karplus, P. A. (1993) *Biochemistry 32*, 9906−9916.
Strynadka, N. C. J., & James, M. N. G. (1991) *J. Mol. Biol. 220*, 401−424.
Tanaka, Y., Tao, W., Blanchard, J. S., & Hehre, E. J. (1994) *J. Biol. Chem. 269*, 32306−32312.
White, A., Withers, S. G., Gilkes, N. R., & Rose, D. R. (1994) *Biochemistry 33*, 12546−12552.
Wood, T. M. (1992) *Biochem. Soc. Trans. 20*, 46−53.
Yanisch-Perron, C., Viera, J., & Messing, J. (1985) *Gene 33*, 103−119.

BI951510M