Structure of the *Fusarium oxysporum* Endoglucanase I with a Nonhydrolyzable Substrate Analogue: Substrate Distortion Gives Rise to the Preferred Axial Orientation for the Leaving Group[†,‡]

Gerlind Sulzenbacher,[§] Hugues Driguez,[∥] Bernard Henrissat,[∥] Martin Schülein,[⊥] and Gideon J. Davies*[,§]

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

Received August 5, 1996; Revised Manuscript Received September 18, 1996[⊗]

ABSTRACT: Endoglucanase I (EG I) is a cellulase, from glycosyl hydrolase family 7, which cleaves the β-1,4 linkages of cellulose with overall retention of configuration. The structure of the EG I from *Fusarium oxysporum*, complexed to a nonhydrolyzable thiooligosaccharide substrate analogue, has been determined by X-ray crystallography at a resolution of 2.7 Å utilizing the 4-fold noncrystallographic symmetry present in the asymmetric unit. The electron density map clearly reveals the presence of three glucosyl units of the inhibitor, consistent with the known number of sugar-binding subsites, located at the active site of the enzyme in the −2, −1, and +1 subsites, *i.e.*, actually spanning the point of enzymatic cleavage. The pyranose ring at the point of potential enzymatic cleavage is clearly distorted from the standard $^4C_1$ chair as was originally suggested for β-retaining enzymes by Phillips [Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C., & Tijan, T. (1974) *J. Mol. Biol. 88*, 349−371]. The distortion observed goes beyond the "sofa" conformation observed in previous studies and results in a conformation whose salient feature is the resulting quasi-axial orientation for the glycosidic bond and leaving group, as predicted by stereoelectronic theory. An almost identical conformation has recently been observed in a complex of chitobiase with its unhydrolyzed substrate [Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K. S., & Vorgias, C. E. (1996) *Nat. Struct. Biol. 3*, 638−648]. The striking similarity between these two complexes extends beyond the almost identical pyranose ring distortion. The overlap of the two respective sugars places the enzymatic nucleophile of endoglucanase I in coincidence with the C2 acetamido oxygen of *N*-acetylglucosamine in the catalytic site of the chitobiase, substantiating the involvement of this group in the catalytic mechanism of chitobiase and related chitinolytic enzymes. The endoglucanase I complex with the thiosaccharide substrate analogue clearly illustrates the potential of nonhydrolyzable sulfur-linked oligosaccharides in the elucidation of substrate binding and catalysis by glycosyl hydrolases.

Glycosyl hydrolases, those enzymes hydrolyzing the glycosidic bond, have long been the subject of X-ray structural studies since the pioneering structure determination of hen egg-white lysozyme in the late 1960's (Blake et al., 1965, 1967). Since then, the massive structural and functional diversity of oligosaccharides, and hence of the enzymes hydrolyzing these compounds, has become apparent. Over 57 sequence-based families of glycosyl hydrolases are now recognized (Henrissat, 1991; Henrissat & Bairoch, 1993, 1996), with structural representatives known for at least 21 of these (Davies & Henrissat, 1995). Despite this seeming wealth of knowledge, many aspects of glycosyl hydrolase function and catalytic mechanism are still unclear.

Cellulases (cellobiohydrolases and endoglucanases) are glycosyl hydrolases able to cleave the β-1,4 linkages of cellulose. They attracted much attention during the first oil crisis due to the predominance of cellulose in plant biomass and the quest for alternative energy sources. More recently, environmental concerns and the potential role of cellulases in the recycling of municipal waste and in the foodstuff, textile, and paper industries have refuelled interest in these enzymes. Cellulases are usually multidomain proteins consisting of a catalytic core domain, linked to a cellulose-binding domain *via* a flexible linker region (Gilkes et al., 1991). The endoglucanase I (EG I)[1] from the thermophilic fungus *Fusarium oxysporum* consists of a 411 amino acid catalytic core (Sheppard et al., 1994) but is unusual in that it has neither a linker nor a cellulose-binding domain. The enzyme belongs to glycosyl hydrolase family 7, which also includes the cellobiohydrolase I from *Trichoderma reesei* whose structure has been determined (Divne et al., 1994). The *F. oxysporum* EG I shows the greatest sequence similarity to the *H. insolens* EG I with which it shares 57%

[†] This work was funded, in part, by the Biotechnology and Biological Sciences Research Council, Novo-Nordisk a/s, and the European Union (Contract BIO2-CT94-3018). G.J.D. is a Royal Society University Research Fellow.

[‡] Coordinates for the structure described in this paper have been deposited with the Brookhaven Protein Data Bank under accession numbers 10VW.

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

[§] University of York.

[∥] CERMAV, Grenoble, affiliated with the Joseph Fourier University.

[⊥] Novo Nordisk a/s.

[⊗] Abstract published in *Advance ACS Abstracts,* November 15, 1996.

[1] Abbreviations: APLH, anti-periplanar lone-pair hypothesis; CBH I, cellobiohydrolase I; HEWL, hen egg-white lysozyme; EG I, endoglucanase I; MOPS, 3-(*N*-morpholino)propanesulfonic acid; NAM, *N*-acetylmuramic acid; NAG, *N*-acetylglucosamine; NCS, noncrystallographic symmetry; PEG, polyethylene glycol.

S0006-2960(96)01946-0 CCC: $12.00

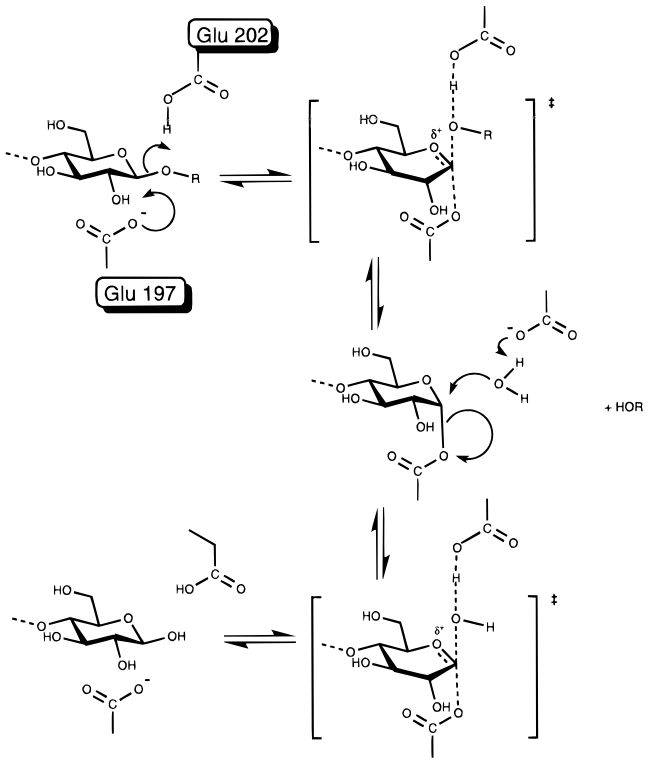

FIGURE 1: Double displacement catalytic mechanism as applied to the EG I from *F. oxysporum*.

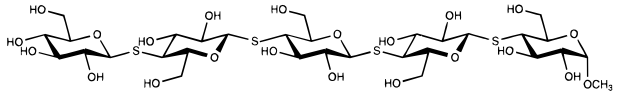

FIGURE 2: Chemical structure of the thio-DP5 inhibitor [methyl *S*-β-D-glucopyranosyl-(1→4)-*S*-4-thio-β-D-glucopyranosyl-(1→4)-*S*-4-thio-β-D-glucopyranosyl-(1→4)-*S*-4-thio-β-D-glucopyranosyl-(1→4)-4-thio-α-D-glucopyranoside].

sequence identity and whose structure has also recently been determined (Davies & Schülein, 1995). EG I is classified as an endocellulase as, unlike the cellobiohydrolases from this family, it is active on carboxymethyl-substituted cellulose but is unable to hydrolyze intact crystalline substrates. Kinetic studies on the highly homologous enzymes from *H. insolens* (Schou et al., 1993a) and *T. reesei* (Biely et al., 1991) have shown that EG I has four kinetically significant subsites for saccharide binding, designated −2, −1, +1, and +2, with enzymatic cleavage taking place between subsites −1 and +1. EG I, in common with all enzymes from family 7, performs catalysis with a net retention of anomeric configuration (Schou et al., 1993a) *via* a double displacement mechanism as first suggested by Koshland (1953) (Figure 1). The proton donor and catalytic nucleophile have been proposed as glutamates 202 and 197, respectively, based on similarity with the homologous CBH I and related *Bacillus* 1,3-1,4 glucanase structures (Divne et al., 1994; Keitel et al., 1993).

In this paper, we present the structure of the *F. oxysporum* EG I complexed to a nonhydrolyzable thiooligosaccharide inhibitor, methyl *S*-β-D-glucopyranosyl-(1→4)-*S*-4-thio-β-D-glucopyranosyl-(1→4)-*S*-4-thio-β-D-glucopyranosyl-(1→4)-*S*-4-thio-β-D-glucopyranosyl-(1→4)-4-thio-α-D-glucopyranoside (hereafter called thio-DP5) (Figure 2). Thio-DP5 is a competitive inhibitor of EG I with a K_i of approximately 30 μM (Schou et al., 1993b). The unbiased averaged electron density map clearly reveals density for the intact oligosaccharide, spanning the point of enzymatic cleavage. The glucopyranose unit in the −1 subsite is clearly distorted from a standard $^4C_1$ chair toward a boat conformation. This results in a quasi-axial orientation for the glycosidic bond and leaving group more consistent with the expected transition-state structure and compatible with a stereoelectronic role in glycoside catalysis.

MATERIALS AND METHODS

The gene encoding the EG I from *F. oxysporum* had previously been cloned (Sheppard et al., 1994). EG I was expressed in, and secreted from, *Aspergillus oryzae* essentially as described previously (Christensen et al., 1988). The protein was initially purified by ion-exchange chromatography, deglycosylated by treatment with endoglycosidase F, and then finally purified by hydrophobic interaction chromatography on phenyl-Sepharose. Crystals were obtained by the hanging-drop vapor diffusion method from 14 mg mL^{-1} enzyme solution buffered with 100 mM MOPS, pH 6.5, and using 22% PEG 8K and 200 mM $MgCl_2$ as precipitants. Crystals were soaked for 1 h in a stabilizing solution containing an elevated concentration of PEG 8K (typically 25−30%) and with the addition of 5 mM thio-DP5. X-ray diffraction data from a single crystal were measured to 2.7 Å using an RAXIS II imaging plate system with a Cu rotating anode operating at 50 kV and 100 mA together with focusing X-ray optics. The crystal was mounted in a rayon fiber loop, and data were collected employing cryogenic techniques, in a stream of boiling N_2 at 120 K. Data were processed and reduced with DENZO and SCALEPACK (Z. Otwinowski, unpublished), and all subsequent computing used the CCP4 (Collaborative Computational Project, Number 4, 1994) suite of programs unless otherwise stated.

The structure was solved by molecular replacement using the program AMoRe (Navaza, 1994) with the native *F. oxysporum* EG I structure (Davies and Schülein, unpublished results) as a search model. The asymmetric unit contains four molecules of EG I (named A−D). The native Patterson, calculated at 6 Å resolution, showed a peak at 0, 0, 1/2 with a peak height 34% of the origin. Molecules A and B are in roughly the same orientation as each other, but separated by a translation of approximately 0, 0, 1/2. A and B are then related by approximate 2-fold rotation axes to their "counterparts" C and D, the consequence of which is that C and D also lie in a similar orientation to each other but again are translated by approximately 0, 0, 1/2. This highly pseudo-symmetric cell cannot be expressed as a higher symmetry space group. The orientations of the four molecules in the asymmetric unit were refined by treating each molecule as a single rigid body. Following this, 5% of the data were set aside for cross-validation analysis (Brünger, 1992), and the model was refined as a single molecule with constrained noncrystallographic symmetry together with a standard XPLOR slow cool protocol (Brünger et al., 1987). After this step, the model had an R_{cryst} of 0.27 and an R_{free} of 0.31. Manual correction to the atomic model was performed with the program O (Jones et al., 1991) using $2F_o - F_c$ electron density maps averaged according to the 4-fold improper NCS with the program RAVE (Kleywegt & Jones, 1994). At this point, the "unbiased" averaged electron density revealed density for three covalently-linked glucose units of the inhibitor in the active site (Figure 3a) and for a great number of solvent water molecules. Further refinement was performed with the maximum likelihood program REFMAC

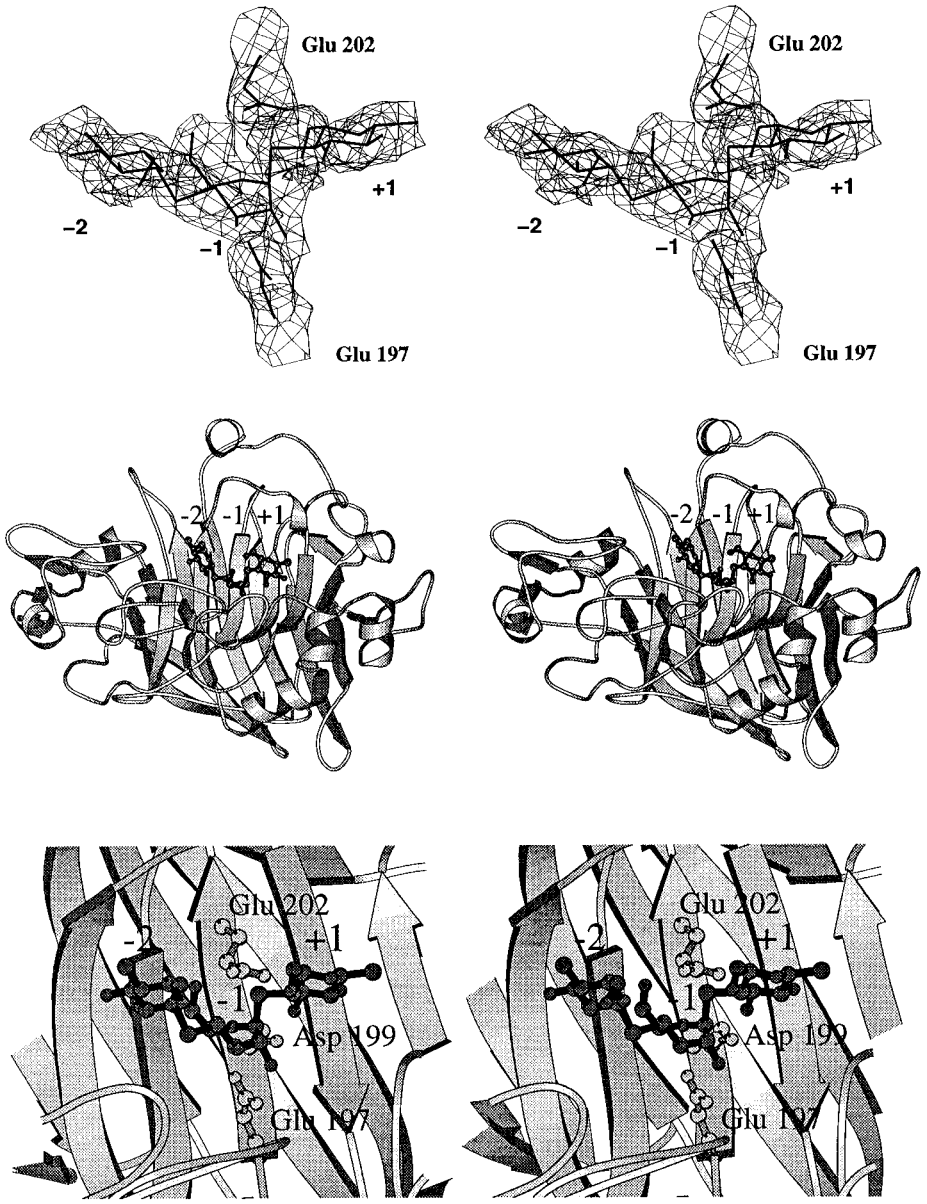

FIGURE 3: (a, top) Stereo electron density for three sugar units of the thio-DP5 inhibitor, plus the two catalytic residues Glu 202 and Glu 197. The refined structure is shown, but the density is the $2F_o - F_c$, α_{calc} map averaged according to the 4-fold NCS calculated *prior* to the incorporation of thio-DP5 or solvent water molecules in the refinement. This map has been chosen so as to give the least biased view of the electron density at this point. **(**b, middle**)** A stereo MOLSCRIPT (Kraulis, 1991) diagram of *F. oxysporum* EG I showing the location of the thio-DP5 inhibitor in the active-site cleft **(**c, bottom**)** stereo MOLSCRIPT diagram showing the distorted conformation for the −1 subsite sugar, together with the active site residues Glu 202 (proton donor), Glu 197 (nucleophile), and Asp 199.

(Murshudov et al., 1996). The behavior of R_{free} was used to direct various aspects of the refinement such as the incorporation of solvent and weighting schemes for the NCS restraints. When the behavior of R_{free} was ambiguous, the maximum likelihood/sigmaA (Read, 1986) figure of merit, calculated on the cross-validation "free" subset of reflections (Free_FOM), was used to monitor various refinement strategies and proved to be a useful indicator of the progress of the refinement. As all observed data (15−2.7 Å) were used for the refinement, a two Gaussian bulk solvent correction was applied according to Babinet's principle (Murshudov et al., 1996). Water molecules were built into $F_o - F_c$ difference maps averaged with the RAVE program, inspected manually, and then expanded to cover all four independent molecules. Tight NCS restraints for both positional and thermal parameters were maintained throughout. The ligand was then built into the averaged map, and additional solvent molecules were included where appropriate. Accurate bond distances for the thiosaccharide were obtained from the crystal structure of cellotetraose (Raymond et al., 1995) with thioglycosidic bond parameters from the structure of 2,3,4,6-tetra-*O*-acetyl-1-*S*-benzhydroximoyl-α-D-glucopyranose (Durier et al., 1992). The thiooligosaccharide structure was refined without torsional angle restraints to avoid imposition of a given pyranose ring conformation.

RESULTS

Crystals of the *F. oxysporum* EG I thiooligosaccharide inhibitor complex are in space group $P2_1$, with cell dimensions $a = 68.2$ Å, $b = 78.3$ Å, $c = 142.5$ Å, and $\beta = 96.9°$. Data merged from 82 217 observations of 38 998 unique reflections, leading to a mean $I/\sigma I$ of 7.2 (2.5 in the outer resolution shell of 2.84−2.7 Å) and are 94.7% complete to 2.7 Å. The asymmetric unit contains four molecules of EG I arranged in a pseudosymmetric manner (see above). The resulting refined model contains 12 088 non-hydrogen protein atoms (398 residues per molecule), 136 ligand atoms, and

1072 water molecules and 8 *N*-acetylglucosamine molecules arising from glycosylation at residues Asn 56 and Asn 247. The *N*-terminal glutamine residue is present as the modified cyclic pyroglutamate ring. All the non-glycine residues have conformational angles (φ,ψ) in permitted regions of the Ramachandran plot (Ramachandran et al., 1963) with 0.3% of these in the generously allowed regions as defined by PROCHECK (Laskowski et al., 1993). The crystallographic *R* factor is 0.19 with an R_{free} of 0.28, and deviations from stereochemical target values of 0.012 Å, 0.038 Å, and 0.038 Å for bonds, angles (1−3 bonding distance), and planes, respectively. The R_{free} and Free_FOM indicated a very tight restraint on the NCS positional and temperature parameters resulting in an rms displacement between all the protein main-chain atoms related by NCS of 0.06 Å. Coordinates have been deposited with the Brookhaven Protein Data Bank (Bernstein et al., 1977).

Despite the intermediate resolution of the data, utilization of the available 4-fold averaging produced a density map of extremely high quality. Solvent water molecules were easily resolved in averaged $F_o - F_c$ difference maps. Three sugar units of the thio-DP5 inhibitor are clearly visible in the electron density map (Figure 3a). They occupy the −2, −1, and +1 subsites, but become disordered beyond +1 and −2, consistent with the presence of approximately four binding subsites on EG I (Schou et al., 1993a) (Figure 3b). The pyranose ring in the −1 subsite is clearly distorted from a standard $^4C_1$ chair conformation. The most important consequence of this distortion is that the glycosidic bond and leaving group adopt a quasi-axial orientation, instead of the equatorial orientation that would result from a $^4C_1$ chair conformation (Figure 3c). The sugars in the −2 and +1 subsites are in the $^4C_1$ chair conformation as expected. The temperature factors are 40, 36, and 48 Å^2 for the sugar rings in the −2, −1, and +1 subsites, respectively. With the current resolution of the crystallographic data, inspection of the electron density in the −1 subsite in isolation would permit interpretation of a number of possible ring conformations at the −1 subsite. We are very fortunate, however, to have covalent links to the adjacent sugars in the −2 and +1 subsites, and thus, the ring conformation at the −1 subsite must be consistent with the maintenance of thioglycosidic linkages to sugars in the −2 and +1 subsites. Very small changes in the ring conformation have a much amplified effect on the position of these adjacent sugars, and consequently very few ring conformations are possible which are compatible with the location of sugars in the adjacent subsites. Indeed, the observed distortion of the sugar group in the −1 subsite causes the direction of the oligosaccharide chain to change markedly and hence the position of the +1 subsite sugar by approximately 8 Å compared to its likely position in a series of $^4C_1$ β-1,4 linked sugars (Figure 4). This anchoring-down of the −1 subsite sugar conformation, by the adjacent sugars, relieves some of the difficulties of modeling the ring distortion that have occurred in other studies where the distorted ring was terminal (and usually associated with disorder and consequently high crystallographic temperature factors).

The salient feature of this complex is the quasi-axial nature of the glycosidic bond, but exact description of the ring distortion is somewhat difficult. Rules for the conformational nomenclature for six-membered ring forms of monosaccharides demand the definition of a four-atom plane (IUPAC−IUB, 1980). The observed conformation can best be described as a distorted $^{1,4}B$ conformation, with C1 much less above the plane than C4. More precise description must await atomic resolution analysis, but the essential nature of the quasi-axial orientation of the glycosidic bond is clear and unambiguous.

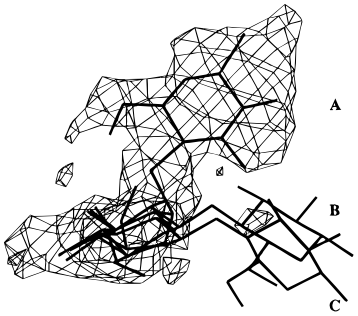

FIGURE 4: $2F_o - F_c$ density, averaged according to the 4-fold NCS and contoured at a level of 0.9 sigma, around the −1 and +1 subsites. Three fits to observed density are shown: (A) the distortion described in this paper; (B) the result of modeling a "sofa" conformation at −1 [based on the coordinates from Strynadka and James (1991)] and (C) with a $^4C_1$ chair conformation at −1. It can clearly be seen that the presence of a linked saccharide in the +1 (and −2) subsite severly limits the number of possibilities for ring conformation, since only a few are consistent with these adjacent subsites.

Family 7 glycosyl hydrolases perform catalysis with a net retention of the anomeric configuration *via* a double displacement reaction mechanism as originally proposed by Koshland (1953). This requires the presence of at least two catalytic groups on the enzyme: a proton donor whose function is to protonate the interglycosidic oxygen and thus allow leaving group departure, and a nucleophile which performs nucleophilic attack on the anomeric C1 atom resulting in a covalent glycosyl-enzyme intermediate of inverted configuration. Following this, the proton donor now acts as a general base, activating an incoming water molecule by deprotonation. The water molecule performs a second inverting nucleophilic attack at C1 resulting in a product of overall retained configuration (Figure 1). The location and interactions of the sugars in the active site of EG I confirm the roles proposed for the active site residues in both the family 7 and structurally similar family 16 enzymes (Divne et al., 1994; Keitel et al., 1993). The proposed proton donor, Glu 202, makes a hydrogen bond of 2.9 Å to the interosidic sulfur atom, while the nucleophile, Glu 197, is poised for nucleophilic attack with the carboxylate group situated some 3.2 Å below the anomeric C1 atom of the sugar in the −1 subsite (Figure 5). The *syn* lone pairs of the OE2 atom of the nucleophile are directed toward the C1 atom of the substrate as suggested on the basis of the native *H. insolens* EG I structure and as expected by theory (Gandour, 1981). Asp 199 makes hydrogen bonds both to the C3 hydroxyl of the +1 subsite sugar and with the catalytic nucleophile Glu 197. The role of Asp 199 is unclear, but it probably functions in the maintenance of the correct pK_a's for the proton donor and nucleophile and in proton shuffling around the active site during catalysis. The active site cleft contains two tryptophan residues, Trp 347 and Trp 356, which make hydrophobic stacking interactions with the sugars in the −2 and +1 subsites, typical of sugar-binding proteins (Vyas, 1991). A greater discussion of the hydrogen bonds involved in the active site of the *F. oxysporum* EG I will be published later when the refined native structure and other complexes become available.

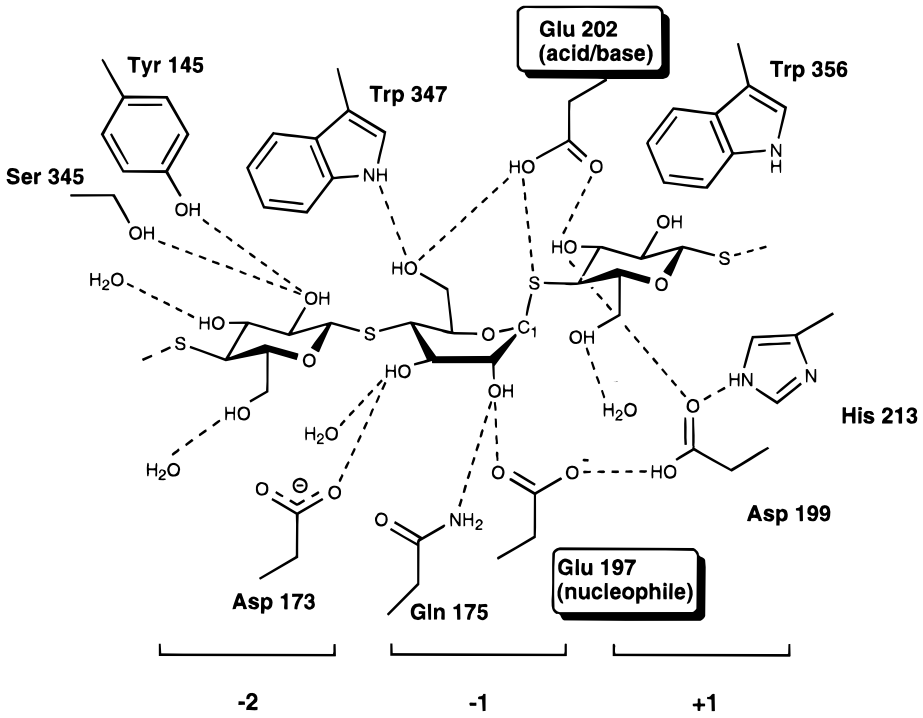

FIGURE 5: Schematic diagram of the interactions between thio-DP5 and EG I.

DISCUSSION

Despite the fact that the first 3-D structure of an enzyme ever solved was that of a glycosyl hydrolase, many aspects of the function and catalytic mechanism of these enzymes still remain controversial. One such aspect is the distortion of the pyranosyl ring at the point of cleavage. The hydrolysis of an equatorial glycosidic bond *via* a double displacement reaction (Koshland, 1953; McCarter & Withers, 1994) proceeds with transition states having substantial oxocarbonium character, with an absolute requirement for planarity at the anomeric carbon. Distortion of the pyranosyl ring from its standard $^4C_1$ conformation into one more resembling the planar transition state is therefore widely accepted as a mechanism utilized by these enzymes to promote catalysis. This distortion was originally proposed for hen egg-white lysozyme, by Phillips and co-workers, based on a complex structure with a lactone derivative of chitotetraose (Ford et al., 1974). The authors found that the best interpretation of the electron density at the −1 subsite arose when the modified sugar ring was modeled as a "sofa" conformation. The structure remained unrefined, but later these proposals were verified by the high-resolution structure determinations of the HEWL NAM−NAG−NAM complex (Strynadka & James, 1991) and a mutant T4 lysozyme product complex (Kuroki et al., 1993), both of which revealed the presence of "sofa" conformations for the pyranosyl ring in the −1 subsite. Both of these structures, however, represent product or pseudoproduct complexes since there is no intact glycosidic linkage accross the point of cleavage. In addition to ring distortion favoring the transition state, glycosidic bond cleavage is only possible because electron donation from the ring oxygen atom helps to relieve the resultant positive charge development at the anomeric carbon. In the ground state, the hydrolyzable C−O bond and hence aglycon in a β-glycosidic linkage are in the equatorial position. Advocates of the antiperiplanar lone pair hypothesis (Deslongchamps, 1983) have suggested that a more stereoelectronically favorable arrangement for aglycon departure would involve distortion of the pyranosyl ring from the standard $^4C_1$ chair form toward a boat conformation resulting in a quasi-axial orientation for the glycosidic bond and leaving group. This would place one of the sp^3 lone pairs on the ring oxygen antiperiplanar to the glycosidic linkage and therefore facilitate electron donation to the transition state (Kirby, 1984). Critics point out, however, that this sort of change in ring conformation when applied to a retaining β-glycosidase could require "quite implausible contortions of the pyranose ring" (Sinnott, 1990). Until very recently, 3-D structural evidence has not been readily forthcoming, and despite numerous attempts, no β-retaining enzyme structures have been published which show sugar units in both the −1 and +1 sites linked by an intact glycosidic bond.

The EG1 4−thiooligosaccharide inhibitor complex spans the active site occupying the −2, −1, and +1 subsites and has an intact thioglycosidic linkage. The pyranosyl ring in the −1 subsite is clearly distorted from a $^4C_1$ conformation, resulting in a quasi-axial glycosidic bond, closer to that expected to occur in the transition state and as predicted by ALPH effects. This distortion in the −1 subsite must have an energetic cost in terms of the binding energy, which must be accounted for by the additional favorable noncovalent interactions of the sugar hydroxyl functions permitted at the distorted transition state (Namchuk & Withers, 1995). Similar substrate distortion has recently also been observed with an unhydrolyzed substrate complex of chitobiase (Tews et al., 1996). Although the reason for nonhydrolysis of the chitobiose substrate is unclear, the high resolution complex clearly reveals distortion of the pyranose ring to yield a quasi-axial orientation for the glycosidic bond exactly as observed here. Indeed, the two structures show a striking similarity, (Figure 6) in a distorted boat conformation in which the C1 atom is only slightly above the plane formed by C2, C3, C5, and O5. Tews and co-workers suggest that this ring conformation results in a planar anomeric C1 atom, but this is clearly not the case. The C1 atom has sp^3 hybridization, and so even if the tetrahedral geometry were somewhat

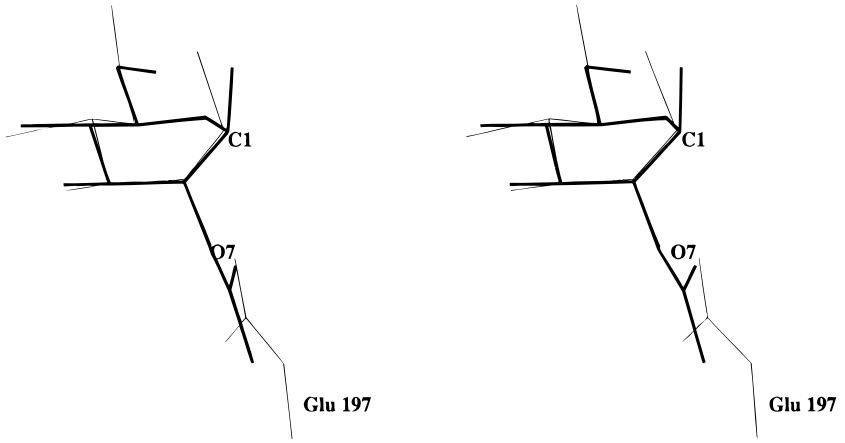

FIGURE 6: Stereoview of the overlap of the −1 subsite rings from the EG I−thio-DP5 complex (this study, faint lines) and the chitobiase−chitobiose complex (Tews et al., 1996). The six ring atoms overlap with an rms of only 0.1 Å. This overlap then places the enzymatic nucleophile of EG I in coincidence with the *N*-acetyl oxygen of the NAG residue in the −1 subsite of chitobiase (see text for details).

distorted, planarity at this atom is chemically implausible. Both these conformations are, however, much closer to the likely transition-state than the standard $^4C_1$ chair conformation.

As described above, precise definition of the substrate distortion is difficult, requiring the definition of an exact four-atom plane. The C2, C3, C5, and O5 atoms deviate from the "best" plane formed by these atoms with an average mean displacement of approximately 0.03 Å, and in this case, the C4 and C1 atoms lie approximately 0.8 and 0.3 Å above this plane, respectively. Although at the resolution of this study we cannot describe more precisely the exact nature of this distorted conformation, the quasi-axial location of the glycosidic bond and resultant marked change in the position of the leaving group are evident. The revealing comparison between the −1 subsite saccharide observed here and the chitobiase−chitobiose complex of Tews and co-workers goes beyond the ring distortion and axial leaving group orientation itself. Overlap of the respective −1 subsite sugars (Figure 6) has the result of placing the enzymatic nucleophile of EG I (OE2 of Glu 197) in almost direct coincidence with the carbonyl oxygen of the C2 acetamido group present on the *N*-acetylglucosamine in the chitobiase structure. Recent work on the hydrolysis of chitosaccharides by various families of β-retaining enzymes, such as soluble lytic transglycosylase (Thunnissen et al., 1994), hevamine (Terwisscha van Scheltinga et al., 1995), and chitobiase (Tews et al., 1996), has strongly implicated neighboring group participation in catalysis by these enzymes. In particular, the lack of an obvious enzymatic nucleophile in these structures has rekindled the proposal that the C2 *N*-acetamido oxygen functions as the nucleophile in the retaining hydrolysis of *N*-acetylglucosamine-based saccharides, as originally suggested in the 1960's (Capon, 1969; Lowe & Sheppard, 1968; Piszkiewicz & Bruice, 1968). Further evidence for this proposal comes from the observation that retaining chitinases are unable to hydrolyze deacetylated chitosaccharides, but inverting chitinases, which must function by a different mechanism, are still able to hydrolyze such substrates (Ohno et al., 1996). Indeed, neighboring group participation which was at one time proposed to be important in the catalytic mechanism of HEWL (but which was later opposed on the basis of X-ray structural data) may yet be shown to be relevant to that enzyme. The spatial equivalence of the *N*-acetamido oxygen of the −1 sugar in the chitobiase complex with the enzymatic nucleophile of EG I strongly supports the proposal that this atom functions as the nucleophile in the mechanism of chitobiase and related enzymes. Since EG I is known to perform catalysis *via* a covalent enzyme intermediate, this implies that enzymatic hydrolysis of NAG polymers with anchimeric assistance may indeed go *via* a cyclic oxazolinium intermediate, in which an analogous covalent bond is formed between the nucleophilic oxygen and the C1 atom (Figure 7).

Summary. In the HEWL NAM−NAG−NAM complex, it appears that a ground-state distorted sugar in the −1 subsite is stabilized by additional favorable interactions that this distortion permits (Strynadka & James, 1991). We do not believe that this is the case for the EG I−thio-DP5 complex described here. In the EG I structure, the substrate distortion is, to a large extent, dictated by the shape and direction of the active site cleft and the requirement for aglycon binding in the +1 subsite. A continuation of the substrate beyond the −1 subsite, with a series of $^4C_1$ β-1,4 linked glucopyranosyl units, would simply result in a major steric clash of the +1 sugar into the protein. We believe that, in addition to any extra interactions possible in the transition state, the interactions of the aglycon in the +1 subsite are likely to be crucial for the distortion observed at −1. Thus, in marked contrast to the HEWL case, we envisage no obstacle to a ground-state $^4C_1$ conformation in the −1 subsite in the absence of a link to a sugar at +1. This is, indeed, not surprising since the covalent glycosyl-enzyme intermediate, in the −1 subsite, is most likely to have a $^4C_1$ conformation, as has been observed crystallographically in the *Cellulomonas fimi* xylanase/cellulase Cex structure (White et al., 1996).

While both the EG I−thio-DP5 and chitobiase−chitobiose complexes are consistent with anti-periplanar and stereoelectronic effects in glycoside catalysis, they do not constitute proof. A β-retaining enzyme going through a covalent enzyme intermediate with $^4C_1$ conformation (White et al., 1996) but inverted configuration, *via* a ring conformation in which the glycosidic bond is quasi-axial, would indeed be undergoing a variety of ring contortions as explained by Sinnott (1990). It should be noted that both the chitobiase−chitobiose complex and the EG I−thio-DP5 complex contain unhydrolyzed saccharides, which may perhaps have some bearing on their conformations. Additionally, in this case, the geometry of sulfur-containing linkages is not identical to their oxygen-containing counterparts. The C−S−C angle of 97° and C−S bond length of 1.83 Å differ from the values of approximately 117° and 1.42 Å for the oxygen-containing

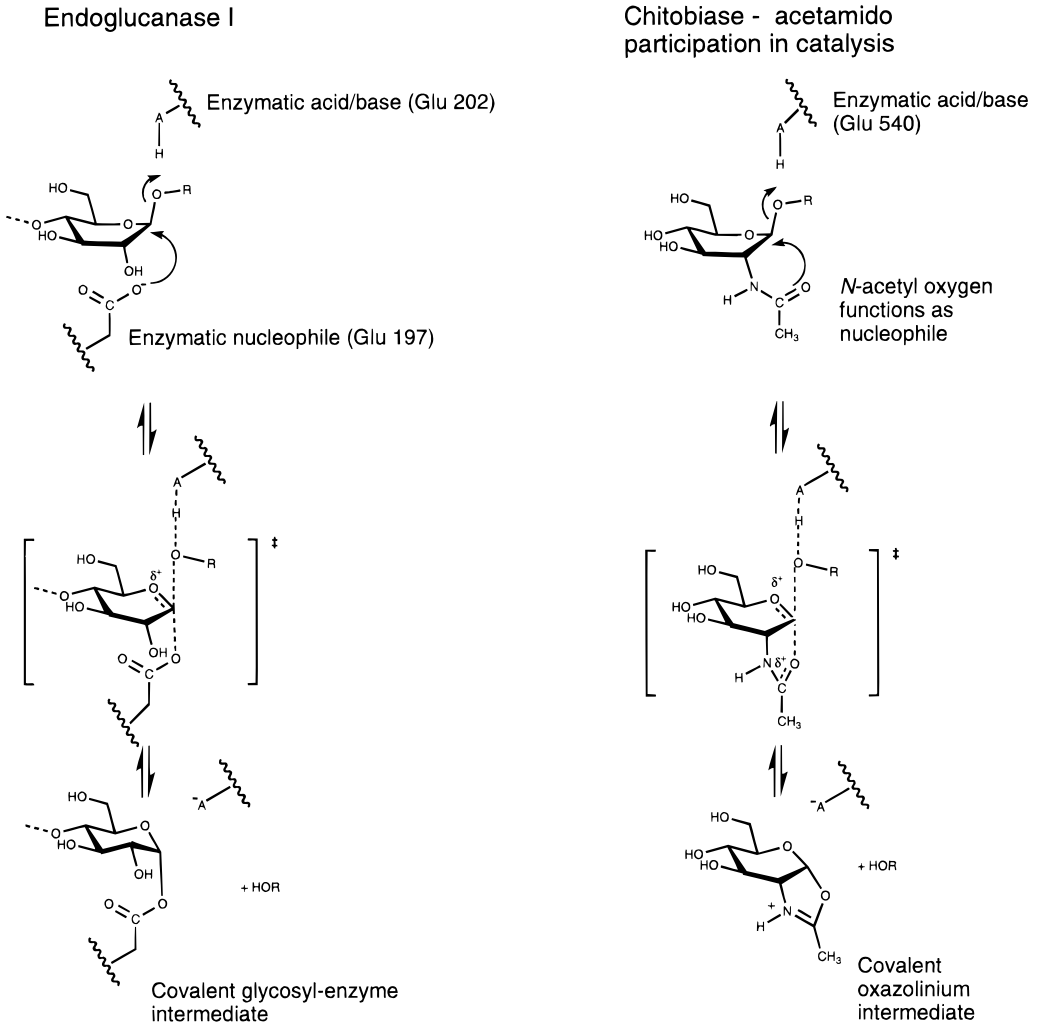

FIGURE 7: Reaction mechanisms for the enzymatic hydrolysis catalyzed by EG I and chitobiase. Only the first half of the reaction pathway, up to the covalent intermediate, is shown. The spatial equivalence of the *N*-acetyl oxygen in the chitobiase complex with the enzymatic nucleophile of EG I (Figure 6) strongly supports the proposal that this atom acts as the nucleophile in the enzymatic hydrolysis of NAG saccharides by chitobiase and related enzymes.

equivalents, although together these differences only result in a difference of 0.35 Å in the positions of adjacent sugars (Driguez, 1995). It remains a possibility that these small geometric differences in some way contribute to a ring distortion in the −1 subsite, although the fact that identical ring distortion is observed in the structure of chitobiase with the natural *O*-linked substrate strongly suggests that this is not so.

Thiooligosaccharides represent a well-established class of inhibitors for glycosyl hydrolases. *O*-Glycosyl hydrolases have evolved to hydrolyze oxygen-containing bonds and are not adapted to the chemistry of sulfur. The interosidic sulfur atom is only weakly basic, and this coupled to the poor leaving group ability of *S*-glycosides results in their resistance to enzymatic hydrolysis by *O*-glycosidases (Capon, 1969). Furthermore, the geometric differences between the sulfur- and oxygen-containing linkages, which are deleterious to binding compared to a natural *O*-glycoside, may be partially advantageous, since geometric imperfections will further reduce the likelihood of hydrolysis. This absence of perfect geometric equivalence presumably accounts for the weaker binding of thio-DP5, in this particular study, and the observation that the electron density becomes disordered beyond the +1 subsite. Thiooligosaccharides have proved extremely useful for the study of sugar−enzyme interactions [for review, see Driguez (1995)]. In this paper, we present the first structure of a complex between a thiooligosaccharide and an endoglucanase. The pyranosyl ring at the point of cleavage is distorted such that the leaving group is in a quasi-axial conformation, as is predicted to exist early in the reaction pathway for enzymes hydrolyzing equatorial glycosidic bonds with retention of configuration [e→e enzymes in the nomenclature of Sinnott (1990)] and which is favored by stereoelectronic effects. In a previous paper, we proposed one mechanism whereby a glycosyl hydrolase could assist in hydrolysis, namely, by favoring the binding of oligosaccharides with an elongated glycosidic bond as found in the transition state (Davies et al., 1995). The structure presented here shows yet another facet of the mechanism used by glycosyl hydrolases to achieve catalysis.

ACKNOWLEDGMENT

We acknowledge the work of Charlotte Schou in the preparation of the thiocellooligosaccharide inhibitors and thank Professors A. J. Kirby, F.R.S., and S. Withers for useful discussions.

REFERENCES

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.

Biely, P., Vrsanska, M., & Claeyssens, M. (1991) *Eur. J. Biochem. 200*, 157−163.
Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., & Sarma, V. R. (1965) *Nature 206*, 757−763.
Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C., & Sarma, V. R. (1967) *Proc. R. Soc. London, Ser. B 167*, 378−388.
Brünger, A. T. (1992) *Nature 355*, 472−475.
Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458−460.
Capon, B. (1969) *Chem. Rev. 69*, 407−498.
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., Sect. D50*, 760−763.
Davies, G., & Henrissat, B. (1995) *Structure 3*, 853−859.
Davies, G. J., & Schülein, M. (1995) in *Carbohydrate Bioengineering* (Petersen, S. B., Svensson, B., & Pedersen, S., Eds.) pp 225−237, Elsevier, Amsterdam.
Davies, G. J., Tolley, S. P., Henrissat, B., Hjort, C., & Schülein, M. (1995) *Biochemistry 34*, 16210−16220.
Deslongchamps, P. (1983) *Stereoelectronic Effects in Organic Chemistry*, Pergamon Press, Oxford, U.K.
Divne, C., Ståhlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K. C., Teeri, T. T., & Jones, A. (1994) *Science 265*, 524−528.
Driguez, H. (1995) in *Carbohydrate Bioengineering* (Petersen, S. B., Svensson, B., & Pedersen, S., Eds.) pp 113−124, Elsevier, Amsterdam.
Durier, V., Driguez, H., Rollin, P., Duee, E., & Buisson, G. (1992) *Acta Crystallogr., Sect. C48*, 1791−1794.
Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C., & Tijan, T. (1974) *J. Mol. Biol. 88*, 349−371.
Gandour, R. D. (1981) *Bioinorg. Chem. 10*, 169−176.
Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., Jr., & Warren, R. A. J. (1991) *Microbiol. Rev. 55*, 303−315.
Henrissat, B. (1991) *Biochem. J. 280*, 309−316.
Henrissat, B., & Bairoch, A. (1993) *Biochem. J. 293*, 781−788.
Henrissat, B., & Bairoch, A. (1996) *Biochem. J. 316*, 695−696.
IUPAC−IUB (1980) *Eur. J. Biochem. 111*, 295−298.
Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr., Sect. A47*, 110−119.
Keitel, T., Simon, O., Borriss, R., & Heinemann, U. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5287−5291.
Kirby, A. J. (1984) *Acc. Chem. Res. 17*, 305−311.
Kleywegt, G. J., & Jones, T. A. (1994) in *From first map to final model* (Bailey, S., Hubbard, R., & Waller, D., Eds.) EPSRC, Daresbury, U.K.
Koshland, D. E. (1953) *Biol. Rev. 28*, 416−436.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kuroki, R., Weaver, L. H., & Matthews, B. W. (1993) *Science 262*, 2030−2033.
Laskowski, R. A., McArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 282−291.
Lowe, G., & Sheppard, G. (1968) *J. Chem. Soc., Chem. Commun.*, 529−530.
McCarter, J. D., & Withers, S. G. (1994) *Curr. Opin. Struct. Biol. 4*, 885−892.
Murshudov, G. N., Vagin, A. A., & Dodson, E. J. (1996) *Acta Crystallogr., Sect. D* (in press).
Namchuk, M. N., & Withers, S. G. (1995) *Biochemistry 34*, 16194−16202.
Navaza, J. (1994) *Acta Crystallogr., Sect. A50*, 157−163.
Ohno, T., Armand, S., Hata, T., Nikaidou, N., Henrissat, B., Mitsutomi, M., & Watanabe, T. (1996) *J. Bacteriol. 178*, 5065−5070.
Piszkiewicz, D., & Bruice, T. C. (1968) *J. Am. Chem. Soc. 90*, 2156−2163.
Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963) *J. Mol. Biol 7*, 95−99.
Raymond, S., Heyraud, A., Qui, D. T., Kvick, A., & Chanzy, H. (1995) *Macromolecules 28*, 2096−2100.
Read, R. J. (1986) *Acta Crystallogr., Sect. A 42*, 140−149.
Schou, C., Rasmussen, G., Kaltoft, M.-B., Henrissat, B., & Schülein, M. (1993a) *Eur. J. Biochem 217*, 947−953.
Schou, C., Rasmussen, G., Schülein, M., Henrissat, B., & Driguez, H. (1993b) *J. Carbohydr. Chem. 12*, 743−752.
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.
Sinnott, M. L. (1990) *Chem. Rev. 90*, 1171−1202.
Strynadka, N. C. J., & James, M. N. G. (1991) *J. Mol. Biol. 220*, 401−424.
Terwisscha van Scheltinga, A. C., Armand, S., Kalk, K. H., Isogai, A., Henrissat, B., & Dijkstra, B. W. (1995) *Biochemistry 34*, 15619−15623.
Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K. S., & Vorgias, C. E. (1996) *Nat. Struct. Biol. 3*, 638−648.
Thunnissen, A.-M., Dijkstra, A. J., Kalk, K. H., Rozeboom, H. J., Engel, H., Keck, W., & Dijkstra, B. W. (1994) *Nature 367*, 750−753.
Vyas, N. K. (1991) *Curr. Opin. Struct. Biol. 1*, 732−740.
White, A., Tull, D., Johns, K., Withers, S. G., & Rose, D. R. (1996) *Nat. Struct. Biol. 3*, 149−154.

BI961946H