Active-Site Binding of Glycosides by *Thermomonospora fusca* Endocellulase E2[†]

Brian K. Barr,[‡,§] David E. Wolfgang,[‡] Kathleen Piens,[||] Marc Claeyssens,[||] and David B. Wilson*[,‡]

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853, and Department of Biochemistry, Physiology and Microbiology, State University of Ghent, K. L. Ledeganckstraat 35, B-9000, Ghent, Belgium

Received January 9, 1998; Revised Manuscript Received April 20, 1998

ABSTRACT: The determination of the high-resolution structure of the *Thermomonospora fusca* endocellulase E2 catalytic domain makes it ideal for exploring cellulase structure−function relationships. Here we present binding parameters (K_d, $\Delta H°$, and $\Delta S°$) describing the interaction of E2 with 4-methylumbelliferyl glycosides, determined by titrating the quenching of ligand fluorescence in equilibrium binding experiments. Quenched MU$(Glc)_2$/E2 complexes were used as indicators in displacement titrations to measure the binding of natural glycosides and also of a nonhydrolyzable cellotetraose analogue. Binding of MU$(Glc)_2$ and cellotriose were also determined by titration calorimetry. The results show that E2 binds glycosides exclusively in its active-site cleft, with high affinity and specificity. The observed patterns of ligand hydrolysis and the results with MU$(Glc)_2$ as a substrate indicated that ligands bound to E2 with their nonreducing ends in position −2, consistent with the position of cellobiose in the E2cd structure. Polymerase chain reaction (PCR) mutagenesis of the conserved residue Tyr 73 (in E2 binding subsite −1) to Phe and Ser produced enzymes with lower activity but higher binding affinities, indicating that the volume of the subsite −1 binding pocket is crucial for enzyme function. Similarly, MUXylGlc (with its xylosyl unit located in position −1) bound with 100-fold higher affinity than MU$(Glc)_2$. These results are similar to those for the related *Trichoderma reesei* exocellulase CBH II. The binding data were compared with that previously reported for CBH II and interpreted in terms of the functional differences between endo- and exocellulases.

Cellulases are enzymes that degrade insoluble cellulose into small glucose oligomers (primarily cellobiose) by hydrolysis of $\beta(1\rightarrow4)$ glycosidic bonds. Cellulases are also known as polysaccharide hydrolases (*1*), an enzyme group that includes amylases, xylanases, chitinases, and lysozymes. Recently the three-dimensional structures of many polysaccharide hydrolases have been determined by X-ray crystallography, including those of at least eight cellulases. This expanding collection of high-resolution structural information permits the identification of specific enzyme regions involved in cellulose binding and catalysis, providing critical information for detailed cellulase mechanistic analysis.

Thermomonospora fusca E2 is a 43 kDa[1] bacterial endocellulase with optimal activity at 60 °C (*2*). The gene coding for E2 has been cloned, sequenced, and overexpressed in *Streptomyces lividans* (*3*). E2 contains an N-terminal, 286 amino acid (34 kDa) family 6 catalytic domain (*4*) connected by a ∼28-residue linker region to a C-terminal, family IIa cellulose-binding domain (*5*). Purified E2 has high activity on a wide range of cellulosic substrates (*6*, *7*). In cellulase mixtures (*6*), E2 acts synergistically with both reducing end-directed and nonreducing end-directed exocellulases (*8*) as well as an unusual *T. fusca* endocellulase (E4) but not with other *T. fusca* endocellulases.

The E2 catalytic domain (E2cd) has been crystallized, and its three-dimensional structure has been determined by X-ray crystallography at 1.8 Å (*9*), 1.18 Å (*10*), and finally at 1.0 Å resolution (*11*). E2cd is an α,β-barrel of dimensions 53 × 38 × 36 Å with a large cleft running across the C-terminal end of the barrel. The E2 active site is in the cleft and contains binding subsites for at least four glucosyl units (labeled −2, −1, +1, and +2 from the nonreducing end). Cellobiose binds to the E2cd crystal in subsites −2 and −1, as shown in Figure 1 (*9*).

E2 hydrolyzes cellulose with inversion (*12*) at the glycosidic bond between the −1 and +1 subsites. Mutagenesis studies have shown that catalysis involves Asp 117 acting as a general acid (D. Wolfgang, unpublished results). Another residue near the site of hydrolysis is Tyr 73, whose OH group is within hydrogen-bonding distance of the O-6 hydroxyl of the glucose residue bound in subsite −1 (*10*).

[†] This work was supported by Grant DE-FG02-84ER13233 from the U.S. Department of Energy to D.B.W.

* Address correspondence to this author at Cornell University, 458 Biotechnology Bldg., Ithaca, NY 14853. Fax: (607) 255-2428. E-mail: dbw3@cornell.edu.

[‡] Cornell University.

[§] Present address: Department of Chemistry, Loyola College, 4501 North Charles St., Baltimore, MD 21210. E-mail: bbarr@loyola.edu.

[||] State University of Ghent.

[1] Abbreviations: kDa, kilodalton; MW, molecular weight; E2cd, catalytic domain of cellulase E2; MUF, 4-methylumbelliferone; MU, 4-methylumbelliferyl; K_d, dissociation constant; K_{eq}, equilibrium constant; K_m, Michaelis constant; k_{cat}, turnover number; CD, circular dichroism; Glu, glucose; CB, cellobiose; CTri, cellotriose; S-CTet, methyl 4-*S*-cellobiosyl β-cellobioside; MUGlc, 4-methylumbelliferyl β-glucopyranoside; MU$(Glc)_2$, 4-methylumbelliferyl β-cellobioside; MU$(Glc)_3$, 4-methylumbelliferyl β-cellotrioside; MU$(Xyl)_2$, 4-methylumbelliferyl β-(1→4)-xylobioside; MUXylGlc, 4-methylumbelliferyl 4-*O*-(β-D-glucopyranosyl)-β-D-xylopyranoside; MULac, 4-methylumbelliferyl β-lactoside; MU$(Glc)_2\alpha$Glc, 4-methylumbelliferyl 4-*O*-(α-D-glucopyranosyl)-β-cellobioside; *F*, relative fluorescence; CMC, (carboxymethyl)cellulose; SC, phosphoric acid-swollen cellulose; FP, filter paper; SA, specific activity.

S0006-2960(98)00063-4 CCC: $15.00 © 1998 American Chemical Society
Published on Web 06/12/1998

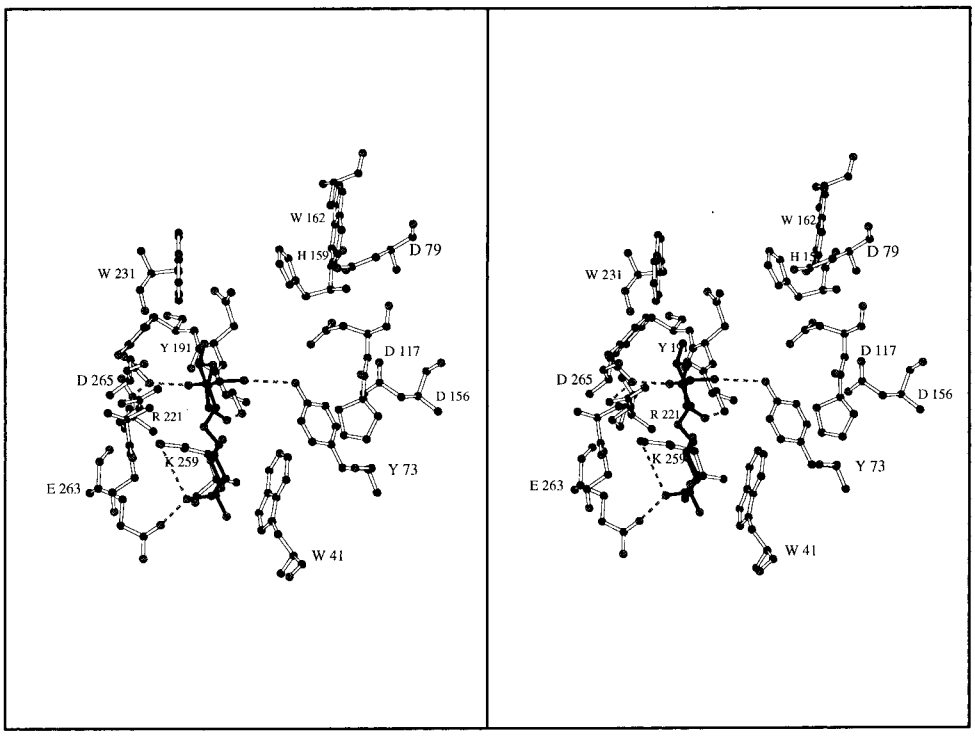

FIGURE 1: Portion of the E2cd active-site cleft showing the interactions with bound cellobiose. This stereoview shows all E2 residues (open lines) within 5 Å of the cellobiose (solid lines) bound in subsites −2 and −1 of the E2 cleft. The long dimension of the cleft is vertical, with −2 at the bottom. Labeled residues are conserved in all members of cellulase family 6. Potential hydrogen bonds (interatom distance ≤ 3.4 Å) are shown (dashed lines). This diagram was drawn with MOLSCRIPT (*41*). Adapted from Spezio (*9*).

Tyr 73 is strictly conserved among cellulases in family 6, and molecular dynamics calculations have suggested that it closely approaches the site of cellulose cleavage (*13*). In the family 6 exocellulase CBH II from *Trichoderma reesei*, the residue corresponding to Tyr 73 (Tyr 169) has been proposed to assist catalysis by distorting the sugar bound in the −1 subsite (*14*).

E2 is a prime candidate for detailed mechanistic analysis. Its high activity, thermostability, and synergistic activity make it an appealing target for modification by protein engineering, while its high-resolution structure makes it ideal for exploring cellulase structure−function relationships through site-directed mutagenesis. This detailed mechanistic analysis requires extensive characterization of the wild-type enzyme, to provide data for comparison with data from the mutants.

A powerful technique for investigating carbohydrate-binding proteins involves measuring the active-site binding affinities for small, soluble ligands by fluorescence spectroscopy. Glycosides containing the highly fluorescent reporter group 4-methylumbelliferone (MU glycosides) (*15*) have been used extensively to measure binding to lectins such as peanut agglutinin (*16*) and to cellulases such as *T. reesei* CBH II (*14, 17, 18*).

In this study we investigated the binding of E2 to a series of MU glycosides by fluorescence titration and to several nonfluorescent glycosides by displacement titration. To explore the importance of Tyr 73 in E2 function, it was mutated to Phe and Ser and the mutants were characterized by cellulase assays and ligand binding measurements. These data provide a basis for structure−function analysis of mutant forms of E2 and are compared with the binding data from the related family 6 exocellulase, *T. reesei* CBH II (*14, 18*).

EXPERIMENTAL PROCEDURES

Bacterial Strains and Growth Conditions. Escherichia coli DH5α (*19*) was used for all routine plasmid isolation and cloning procedures. All *E. coli* strains harboring plasmids were grown in Luria broth (LB) or on LB plates containing 100 μg/mL of ampicillin. The protease-negative strain *S. lividans* TKM31 (*20*) was used as the host cell for protein expression. Media for the preparation and regeneration of *S. lividans* protoplasts were as described (*21*). Cultures of *S. lividans* were grown in tryptone soya broth (TSB) or on TSB plates (*2*), and strains containing plasmids were grown in the presence of thiostrepton (Tsr, 5−10 μg/mL in broth and 50 μg/mL in plates).

Creation of E2 Mutants at Position 73. Tyr 73 was mutated to Phe and Ser by using the PCR overlap extension method as described (*22, 23*). Plasmid pSZ13 (*22*) containing the wild-type E2 gene was used as template. The flanking primers were 5′-GGCGAAAGGGGGATGTGC-3′ (downstream primer) and 5′-GCCGGAAGCATAAAGTGTA-3′ (upstream primer). Degenerate mutagenic primers were designed to mutate Tyr 73 to Phe, Ile, Ser, and Thr. The mutagenic primers were 5′-GGTCGTGWYCAACGCACCGGGCC-3′ (forward primer) and 5′-GGCCCGGTGCGTTGRWCACGACCAGG-3′ (reverse primer), where W = T or A, Y = T or C, and R = A or G. The underlined bases represent silent mutations introduced to destroy a *Sma*I site in pSZ13, to facilitate identification of the mutant products.

Overlap PCR products were digested with *Pst*I and *Sac*I, cloned into the *E. coli*/*S. lividans* shuttle vector pSES1 (*22*), and transformed into *E. coli*. Single colonies were purified and plasmid DNA was isolated and sequenced (on an ABI

Model 373A DNA Sequencer at the Cornell Biotechnology Center Sequencing Facility) to identify each mutation. Mutants E2(Y73F) and E2(Y73S) were identified, shown to have no extra mutations, and cloned into *S. lividans* TKM31 as described (*22*). No E2(Y73I) or E2(Y73T) mutants were found. The E2(Y73S) strain gave a very poor yield of purified enzyme, so it was recloned into *S. lividans* behind the high-expression pLT1 promoter (*24*). To do this, the following three fragments were ligated into *Sac*I/*Hin*dIII-digested pSES1: (1) a 435 bp, *Ban*II/*Hin*dIII fragment from pBW1 (D. Irwin, unpublished result) containing the pLT1 promoter; (2) a *Ban*II/*Sph*I fragment from the E2(Y73S) plasmid containing the first 480 bp of the E2 coding sequence; and (3) a *Sac*I/*Sph*I fragment from pSZ61 (*22*) containing the last 900 bp of the E2 coding sequence. The resulting plasmid (pDW1) was introduced into *E. coli* and *S. lividans* as described.

Production and Purification of E2 and E2 Mutant Proteins. *S. lividans* strains containing the appropriate plasmids were grown in TSB + Tsr at 30 °C for at least 60 h, either in 10 L cultures in a 14 L fermentor (New Brunswick Scientific) or in a 2 L culture aliquoted into four 3 L flasks. E2, E2-(Y73F), and E2(Y73S) were purified from the *S. lividans* overexpression strains as described (*6*) and concentrated using a PTGC 10 000 NMWL membrane (Millipore). A proteolysis product, E2cd, was generated by all *S. lividans* strains in varying amounts. It was separated from E2 during elution from phenyl-Sepharose, and was purified in the same manner as the intact enzyme. The concentrations of the purified enzymes were determined by measuring their absorbance at 280 nm, using the extinction coefficients $\epsilon =$ 81 800 M^{-1} cm^{-1} for E2, 57 580 M^{-1} cm^{-1} for E2cd, and 78 560 M^{-1} cm^{-1} for E2(Y73F) and E2(Y73S) [calculated from the protein sequence as described (*25*)]. Circular dichroism (CD) spectra of the purified enzymes were determined as described (*22*), at 37.0 ± 0.1 °C in 50 mM potassium phosphate, pH 7.0. The CD spectra of E2(Y73S) and E2(Y73F) were indistinguishable from that of E2, indicating that the specified mutations do not cause any global conformational changes in the mutant enzymes.

Substrates and Ligands for Fluorescence Experiments. Glucose, cellobiose, and lactose were from Sigma (St. Louis, MO). Cellotriose was from Dr. H. J. Strobel, University of Kentucky. Concentrations of stock solutions of these ligands were determined by the 2,2′-bicinchoninate reducing sugar method (*26*). A nonhydrolyzable oligosaccharide analogue, methyl 4-*S*-cellobiosyl-β-cellobioside (S-CTet), was a gift from H. Driguez, CERMAV (Centre de Recherches sur les Macromolécules Animales et Vegétales), Grenoble, France. 4-Methylumbelliferone (MUF) was from Sigma (St. Louis, MO). Glycosidic ligands containing MUF (MU glycosides) were prepared as described (*15*). These included 4-methylumbelliferyl β-D-glucopyranoside (MUGlc), 4-methylumbelliferyl β-cellobioside [$MU(Glc)_2$], 4-methylumbelliferyl β-cellotrioside [$MU(Glc)_3$], and 4-methylumbelliferyl β-D-lactoside (MULac). 4-Methylumbelliferyl β-(1→4)-xylobioside [$MU(Xyl)_2$] and 4-methylumbelliferyl 4-*O*-(β-D-glucopyranosyl)-β-D-xylopyranoside (MUXylGlc) were prepared by Dr. Wim Nerinckx at the Department of Biochemistry, Physiology, and Microbiology (University of Ghent). 4-Methylumbelliferyl 4-*O*-(α-D-glucopyranosyl)-β-D-cellobioside [$MU(Glc)_2\alpha Glc$] was prepared enzymatically (K. Piens, unpublished results). Concentrations of MU glycoside stock solutions were calculated from their absorbance at 316 nm using the extinction coefficient $\epsilon =$ 13 600 M^{-1} cm^{-1} (*15*).

Fluorescence Titrations. Fluorescence was measured on either an Aminco SPF-500 spectrofluorometer connected to a recording printer (State University of Ghent), or an Aminco SLM 8000C spectrofluorometer interfaced with an IBM personal computer (Flow Cytometry and Imaging Facility, Cornell University Biotechnology Center). Both instruments were equipped with thermostated copper cuvette holders connected to circulating water baths. MU glycoside fluorescence was excited at 316 nm, with the fluorescence emission measured at 360 nm. The spectral bandwidths varied between experiments, but were generally as follows: excitation = 1 nm and emission = 20 nm for the SPF-500, and excitation = 1 nm and emission = 16 nm for the 8000C.

Preliminary experiments showed that the fluorescence (F) of a solution of $MU(Glc)_2$ was strongly and reversibly quenched upon addition of E2cd. This binding interaction was investigated by direct fluorescence titrations as described for *T. reesei* CBH II (*17*), in 50 mM sodium acetate, pH 5.5. E2 was known to hydrolyze both $MU(Glc)_2$ and $MU(Glc)_3$ at a low rate (*6*). As hydrolysis would interfere with the measurement of binding affinities, fluorescence titrations were routinely performed at 8.5 °C, where cleavage was negligible.

The quenching of ligand fluorescence upon binding to E2 was titrated by continuously adding a concentrated solution of enzyme from a Hamilton model 1750 gastight syringe in a syringe pump to a cuvette containing a continuously stirred solution of MU glycoside. The enzyme concentration used in each experiment was adjusted according to the strength of the binding interaction (weaker binding required higher concentrations) and was in the range of 12−460 μM. With the SPF-500 fluorometer, enzyme was added at 4.08 μL/min to a starting volume of 500 μL in 0.5 × 0.5 × 4.5 cm quartz cuvettes, while with the 8000C fluorometer enzyme was added at 3.5 μL/min to a starting volume of 900 μL in 0.4 × 1.0 × 4.5 cm semimicro quartz cuvettes (Hellma Cells, Forest Hills, NY). Fluorescence was recorded automatically at short intervals (every 6 or 15 s for SPF-500, every 12 s for 8000C) during the course of the titration. For each experiment the data were corrected by a blank titration of enzyme added to a solution of buffer alone. Several titrations were performed on both instruments and gave indistinguishable results.

Displacement titrations (*17*) were used to probe the binding of nonfluorescent sugars by E2 and the E2(Y73) mutants. The indicator ligand was titrated with enzyme and the resulting quenched MU glycoside/enzyme complex was then disrupted by the continuous addition of an excess of competitive ligand from a second syringe. The data were corrected with a blank titration. Fluorescence emission spectra were routinely collected upon completion of the displacement titration to verify that the indicator was not degraded. For S-CTet, displacement titrations were performed by adding 1.0 μL aliquots to the quenched MU glycoside/E2 complex using disposable glass micropipets (Accu-Fill 90, Becton Dickinson-Clay Adams), and manually recording the fluorescence change. The ligand concentra-

tions used were 2.0 mM S-CTet, 16.8 mM cellotriose, 104.3 mM cellobiose, 679 mM lactose and 1.047 M glucose.

Calculation of K_d, $\Delta H°$, and $\Delta S°$. The dissociation constants (K_d) for MU glycosides were determined from direct plots of *F*(corrected) versus volume of enzyme using KaleidaGraph version 3.0.8 (Synergy Software). The binding curves were fitted directly by using the mathematical treatment (available upon request) of de Boeck et al. (*16, 27*), which was adapted and input into the General Curve Fit routine of KaleidaGraph (F. Loontiens, State University of Ghent, Belgium). The data were fitted according to three parameters: K_d, initial fluorescence (F_o), and final fluorescence (F_∞). For displacement experiments, the binding parameters of the indicator ligand were calculated (from repeat titrations) using TITrat, version 1.3 (O. Teleman, VTT Biotechnology, Espoo, Finland). These data were fitted for K_d and ΔF_{max} (the maximal fluorescence intensity change), and the calculated values were input as constants in the displacement titrations. For each displacing ligand, K_d was determined from a plot of *F*(corrected) vs volume of competitor ligand added using TITrat, again by using the mathematical treatment described by de Boeck et al. (*16, 27*).

To determine the thermodynamics of the interaction of $MU(Glc)_2$ with enzyme, the K_d was determined for E2cd at six temperatures in the range of 7−23 °C under conditions where ligand hydrolysis was negligible. For E2(Y73F) and E2(Y73S) [which hydrolyze $MU(Glc)_2$ much more slowly than does E2cd] the K_d was determined at six temperatures in the range from 4 to 30 °C. The thermodynamic constants $\Delta H°$ and $\Delta S°$ were determined by using KaleidaGraph to fit plots of ln K_{eq} versus $1/T$ according to the van't Hoff equation, $\ln K_{eq} = -\Delta H°/RT + \Delta S°/R$ (*28*). The free energy of binding ($\Delta G°$) at 25 °C was then calculated from the expression $\Delta G° = -RT \ln K_{eq}$, using the value for ln K_{eq} at 25 °C derived from the van't Hoff curve fit.

Cellulase Assays. The activities of E2(Y73F) and E2(Y73S) were determined at 50 °C in 50 mM sodium acetate, pH 5.5, on three substrates: soluble (carboxymethyl)cellulose (CMC), amorphous phosphoric acid-swollen cellulose (SC), and filter paper (FP). Production of reducing sugars was measured with the dinitrosalicylic acid reagent (DNS) as described, using cellobiose to generate the DNS standard curve (*6*). The mutant activities were compared with those of wild-type E2 (determined concurrently) and are expressed as percentages. The CMC and SC activity of E2(Y73F) were assayed for 1 h using six amounts of enzyme in the range of 1.0−16.0 μg for E2(Y73F) and 0.1−1.6 μg for E2 and were calculated as described (*22*). For E2(Y73S), the CMC and SC activities of 200 μg of E2(Y73S) were compared to those of 0.2 μg of E2 after 1 h. For both mutants, FP assays were run for 24 h with 40 μg of enzyme.

As previously reported (*6*), E2 slowly degrades $MU(Glc)_2$ to MUF + CB. By using the different fluorescence emission maxima of MUF (λ_{max} = 450 nm) and $MU(Glc)_2$ (λ_{max} = 365 nm), hydrolysis of $MU(Glc)_2$ was quantitated (at 37 °C) in the SLM 8000C spectrofluorometer. Excitation was at 316 nm (2 nm bandwidth), with the MUF emission measured at 455 nm (8 nm bandwidth). Stock solutions of MUF (in 50 mM sodium acetate, pH 5.5) were used to generate a standard curve, which was linear over the range 1−200 nM MUF.

For E2cd, seven different concentrations of $MU(Glc)_2$ (0.99−19 μM) were incubated with enzyme (0.14 μM) in 50 mM sodium acetate, pH 5.5. The release of MUF was followed by recording the fluorescence at 30 s intervals for 30 min. For each concentration the velocity (ΔF/min) was determined from linear curve fits of the data and converted to activity by use of the MUF standard curve. The data were plotted as activity vs $MU(Glc)_2$ concentration, and K_m and k_{cat} were determined from a direct fit of the Michaelis−Menten equation. For the E2(Y73F) mutant, assays were performed with six concentrations of $MU(Glc)_2$ (0.51−5.13 μM) and 0.16 μM enzyme. For each concentration, the amount of MUF product was measured at four different time points (0, 2, 4, and 6 h), by averaging 12 fluorescence measurements (collected at 10 s intervals over 2 min). K_m and k_{cat} were determined in the same manner as with E2cd.

Titration Calorimetry. Calorimetric measurement of the interactions of $MU(Glc)_2$ and CTri with E2cd were performed on a Micro Calorimetry System (MCS) from MicroCal, Inc. (Northampton, MA) (*29*). Titrations were performed according to published procedures (*30*) and the manufacturer's instructions. In the first experiment 50.5 μM E2cd was titrated with 15 5.0−6.5 μL injections of 1.49 mM $MU(Glc)_2$ at 9.6 °C, while in the second experiment 397 μM E2cd was titrated with 30 3.0−4.0 μL injections of 10.06 mM CTri at 10.9 °C. For each experiment, the data were corrected by a blank titration of ligand added to a solution of buffer (50 mM sodium acetate, pH 5.5).

The titration data were plotted as a thermogram (microcalories per second versus time), and analyzed by use of the ORIGIN software package supplied with the calorimeter. The peaks of the thermogram were integrated, and the resulting isotherm (kilocalories per mole versus ligand/protein ratio) was fitted with the ORIGIN data analysis routine. The curve fitting method was comparable to that used to fit the fluorescence binding data, and gave values for K_{eq}, $\Delta H°$, and n (the stoichiometry of binding). $\Delta G°$ and $\Delta S°$ were then calculated from the standard thermodynamic expressions, $\Delta G° = -RT \ln K_{eq}$ and $\Delta G° = \Delta H° - T\Delta S°$.

RESULTS

Suitability of E2 for Fluorescence Binding Studies. The requirements for quantitative ligand binding analysis are that MU glycosides undergo a measurable fluorescence change upon binding to E2, that this change is specific and reversible, and that ligands are not hydrolyzed during the course of the titrations. Preliminary experiments showed that the fluorescence emission of $MU(Glc)_2$ is strongly quenched upon binding to E2cd and that quenching is reversed upon addition of a saturating amount of cellobiose. Also, fluorescence quenching did not cause a shift in the maximum $MU(Glc)_2$ emission wavelength of 370 nm.

Previously, it has been shown (*6*) that E2 slowly degrades $MU(Glc)_2$ and $MU(Glc)_3$ at 50 °C. At 37 °C, HPLC analysis of this activity (as in ref *15*) showed that E2 cleaves $MU(Glc)_3$ to MUGlc + CB but is nearly inactive on $MU(Glc)_2$. Nevertheless the more sensitive fluorescence assay showed that E2cd cleaves a significant amount of $MU(Glc)_2$ to MUF + CB after 15 min at 23 °C, but not after 1 h at 8.5 °C. Since ligand hydrolysis interferes with data interpretation, all binding titrations were routinely performed at 8.5 °C in less than 1 h.

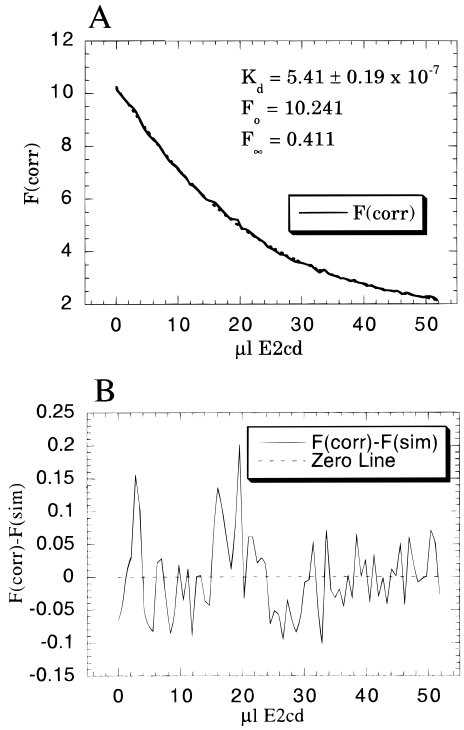

FIGURE 2: Continuous binding titration of $MU(Glc)_2$ with E2cd. In panel A, 1.49 μM $MU(Glc)_2$ (903 μL) was titrated with 63.4 μM E2cd at 3.25 μL/min for 15 min. The resulting binding curve *F*(corrected) is shown (solid line), along with the curve fit (dashed line). A graph of the error in the experimental data is shown in panel B.

Determination of K_m and k_{cat} for E2cd on $MU(Glc)_2$. From the curve fit, the calculated kinetic constants at 37 °C are $K_m = 3.12 \pm 0.27$ μM and $k_{cat} = 1.39 \pm 0.05 \times 10^{-6}$ min^{-1}. By this method the activity of E2cd on MUXylGlc was too low to be accurately measured. It was estimated to be <0.1% of the E2cd activity on $MU(Glc)_2$, based on a comparison at a substrate concentration of 2.0 μM.

MU Glycoside Fluorescence Titrations. The binding of MU glycosides to E2 was determined as described under Experimental Procedures. The rate of mixing was a limiting factor that precluded use of higher rates of enzyme addition. Ligands bound to E2 rapidly, as the quenching for any amount of E2 added was complete as quickly as fluorescence could be measured (ca. 2 s). Figure 2 is a sample binding curve showing the results from a continuous titration of $MU(Glc)_2$ with E2cd. As seen in Figure 2A, ligand fluorescence was strongly quenched (~95%) upon binding to E2cd. This level of quenching was routinely observed in most fluorescence titrations. The experimental data showed that $MU(Glc)_2$ bound to E2cd with a high affinity ($K_d = 0.54 \pm 0.019$ μM). The error value in Figure 2A describes the precision of the parameter fit, which was generally accurate to within a few percent. This can be seen in Figure 2B, where the difference between the experimental data and the best-fit line is small (<2%) and random. Differences in K_d values between duplicate experiments were generally within 10−15%, as was found for CBH II (*14*).

The binding constants for the interaction of MU glycosides with E2 are shown in Table 1. The presence of the E2 cellulose-binding domain had no significant impact on ligand binding, as E2 and E2cd gave similar K_d values for several ligands. This indicates that the ligands are binding only to E2cd, presumably in its active-site cleft. Binding affinities for MU glycosides varied over 5 orders of magnitude, from 10^{-4} M for MUGlc to 10^{-9} M for MUXylGlc. The K_d value for MUGlc in Table 1 is an estimate, as this interaction was too weak for the curve fitting program to converge. Nevertheless the program successfully calculated the K_d for $MU(Xyl)_2$ and MULac, even though their binding was only slightly tighter than that of MUGlc. At the other extreme, MUXylGlc bound too tightly to permit an accurate determination of its K_d. A fluorescence titration was also attempted with $MU(Glc)_2\alpha Glc$ (which has an α-(1→4) glycosidic bond adjacent to its nonreducing end) but it bound very weakly.

The data in Table 1 for E2cd shows that the binding affinity increased with larger MU glycosides, over 500-fold from MUGlc to $MU(Glc)_2$ and another 10-fold from $MU(Glc)_2$ to $MU(Glc)_3$. The data for CBH II showed the same trend, but the increases were smaller. [Note that the CBH II data were collected at a higher temperature (25 °C) than the E2 data; see Discussion.] For E2cd binding to MU disaccharides, changing the ligand from $MU(Glc)_2$ to MULac (with the nonreducing O-4 hydroxyl now in an axial position) caused a 300-fold decrease in binding affinity, while changing it to $MU(Xyl)_2$ (removing both C-6 groups) decreased it 350-fold. In contrast, the structurally intermediate ligand MUXylGlc had the highest affinity of all ligands studied, binding over 250-fold tighter than $MU(Glc)_2$ and 10^5 tighter than $MU(Xyl)_2$. A similar result with MUXylGlc has been reported for CBH II (*17*).

Displacement Experiments. $MU(Glc)_2$ was used as an indicator ligand to measure the binding affinities of several nonfluorescent glycosides. Natural glycosides longer than cellotriose could not be used in these experiments, since they are readily hydrolyzed by E2 (*7*). However the binding affinity of the cellotetraose substrate analogue S-CTet was measured, as this compound is not degraded.

The binding affinities of E2 and E2cd for nonfluorescent glycosides were calculated from direct curve fits of the corrected displacement titration data with TITrat and are listed in Table 1. The K_d values listed for cellobiose, lactose and glucose must be regarded as approximate. The curve fit for the cellobiose data contained a large systematic deviation. This was most likely due to a complex displacement process, since a simple 1:1 displacement is assumed for modeling by the fitting routine while the E2 active-site cleft contains multiple glucosyl binding sites. This deviation was also observed with lactose but was not as severe. The CTri and S-CTet data did not show this systematic deviation, consistent with these larger ligands having fewer binding modes.

The uncertainty in the parameter fit for CB, lactose, and glucose was also large due to difficulties with the fitting program. At competitor ligand concentrations greater than 30 mM, the fitting routine was unstable and unable to converge to a reliable value for K_d. To compensate, ligand concentrations were divided by a constant (10 for CB, 25 for lactose, and 40 for glucose). The curve fit was then applied, and the resulting binding constants were multiplied by the appropriate factor to determine K_d. To check the validity of this approach the CTri data were analyzed in the same manner. The analysis showed that this approach introduced a systematic error into the calculated K_d of CTri. With a constant of 10 the introduced systematic error was

Table 1: Binding Constants of E2, E2cd, and CBH II

ligand	K_d (M) E2[a]	E2cd[a]	CBH II[b]
MUGlc		$\sim 3 \times 10^{-4}$	$(5.0 \pm 0.2) \times 10^{-5}$
$MU(Glc)_2$	$(4.00 \pm 0.07) \times 10^{-7}$ [c]	$(5.41 \pm 0.19) \times 10^{-7}$ [(1.64 ± 0.09 × 10⁻⁶][d]	$(5.00 \pm 0.07) \times 10^{-6}$
$MU(Glc)_3$	$(5.33 \pm 0.43) \times 10^{-8}$	$(4.18 \pm 0.26) \times 10^{-8}$	$(1.11 \pm 0.01) \times 10^{-6}$ [e]
$MU(Xyl)_2$		$\sim 1.9 \times 10^{-4}$	
MUXylGlc	$(3.8 \pm 0.5) \times 10^{-9}$	$\sim 2 \times 10^{-9}$	$(2.04 \pm 0.02) \times 10^{-8}$
MULac		$\sim 1.7 \times 10^{-4}$	
S-CTet		$(4.6 \pm 0.2) \times 10^{-6}$	
cellotriose[f]	$(1.30 \pm 0.03) \times 10^{-4}$	$(1.00 \pm 0.04) \times 10^{-4}$	$(1.61 \pm 0.02) \times 10^{-5}$
cellobiose[f]		$\sim 4 \times 10^{-4}$ [g]	$(1.85 \pm 0.02) \times 10^{-3}$
lactose[f]		$\sim 5 \times 10^{-3}$ [g]	
glucose[f]		$\sim 8 \times 10^{-2}$ [g]	$(6.54 \pm 0.08) \times 10^{-3}$ [h]

[a] At 8.5 °C in 50 mM sodium acetate, pH 5.5. [b] At 25 °C in 50 mM sodium acetate, pH 5.0, unless noted otherwise. CBH II data are from ref *18*. [c] Error values describe the uncertainty in the parameter fit. The total uncertainty for MU glycoside experiments was 10−15%, and for CTri and S-CTet ∼20%. [d] Determined by titration calorimetry at 9.6 °C. [e] At 4.8 °C. [f] From a displacement titration with $MU(Glc)_2$. [g] Value is approximate. For explanation, see Results. [h] At 4.3 °C.

6%, and with a constant of 40 it was 27%. As a result the data for CB, lactose, and glucose in Table 1 should only be considered accurate to within a factor of 2.

The data in Table 1 show that both fluorescent and nonfluorescent glycosides display increased binding affinity with increasing chain length, although the nonfluorescent ligands bind more weakly than their fluorescent counterparts. These effects were also observed with CBH II. There was a 200-fold increase in binding affinity going from Glu to CB, a ∼4-fold increase from CB to CTri, and a 20-fold increase from CTri to S-CTet. The K_d for lactose was midway between that for Glu and CB. The error values shown for CTri and S-CTet in Table 1 are for the parameter fit, as was the case for the MU glycosides. The variation between duplicate experiments was larger than that for the MU glycosides, and was estimated to be 20%.

Influence of Glucose on Binding. The affinity of CBH II for $MU(Glc)_2$, CB, and CTri was enhanced in the presence of near-saturating amounts (0.33 M) of glucose (*18*). To investigate this effect with E2, 5.0 μM E2cd was added to a cuvette containing 2.0 μM $MU(Glc)_2$ in the presence and absence of 0.33 M D-glucose at 24 °C. Emission spectra were recorded prior to and at several time points (0.5, 15, 30, and 60 min) after enzyme addition, and compared. In contrast to CBH II, glucose decreased the extent of fluorescence quenching by E2cd (from 62% to 42% of F_o), indicating a decrease in the $MU(Glc)_2$ binding affinity. Glucose appeared to exert its effect by acting as a competitive inhibitor. This was observed from the 60 min emission spectra, where the height of the MUF emission peak at 450 nm (from hydrolysis of $MU(Glc)_2$) was nearly 50% smaller in the presence of glucose.

Ligand Binding at Altered pH Values. E2 has a very broad pH optimum, retaining nearly 100% activity from pH 5.0 to 10.0. To determine how the ligand binding affinities vary with pH, the E2cd binding constants for $MU(Glc)_2$ and CTri were determined at pH 3.0, 5.5, and 10.0. The data were of high quality and are summarized in Table 2. The K_d values varied significantly over this pH range, with a 16-fold change for $MU(Glc)_2$ and a 43-fold change for CTri. For both ligands, the K_d values at pH 5.5 were intermediate between those at pH 3.0 and pH 10.0. A major difference between $MU(Glc)_2$ and CTri was that their binding affinities varied in the opposite direction: $MU(Glc)_2$ had its highest affinity at pH 3.0, while CTri had its highest affinity at pH 10.0.

Table 2: E2cd Binding Constants at Different pH Values[a]

pH	K_d (M) $MU(Glc)_2$	cellotriose[b]
3.0[c]	$(2.45 \pm 0.04) \times 10^{-7}$	$(4.84 \pm 0.19) \times 10^{-4}$
5.5[d]	$(5.41 \pm 0.43) \times 10^{-7}$	$(1.00 \pm 0.04) \times 10^{-4}$
10.0[e]	$(3.86 \pm 0.11) \times 10^{-6}$	$(1.13 \pm 0.12) \times 10^{-5}$

[a] At 8.5 °C. [b] From a displacement titration with $MU(Glc)_2$. [c] In 100 mM citrate/phosphate, pH 3.0. [d] In 50 mM sodium acetate, pH 5.5. [e] In 100 mM glycine/phosphate, pH 10.0.

Determination of $\Delta G°$, $\Delta H°$, and $\Delta S°$ for $MU(Glc)_2$. The thermodynamic parameters for the interaction of $MU(Glc)_2$ with E2cd are shown in Table 3, along with the comparable values for CBH II. The $MU(Glc)_2$ binding enthalpy was significantly more favorable (by −35 kJ/mol) for E2cd than for CBH II. While the entropy of $MU(Glc)_2$ binding for CBH II was positive (favorable), for E2cd it was negative (unfavorable). The E2cd $\Delta H°$ and $\Delta S°$ values were used to calculate the K_d for $MU(Glc)_2$ at 25 °C (1.42 μM), which is 3.5-fold tighter than that of CBH II at the same temperature (5.00 μM, see Table 1). These binding constants were then used to calculate $\Delta G°$ of binding at 25 °C and are listed in Table 3.

Titration Calorimetry of E2cd. To confirm the results from the fluorescence titrations, $MU(Glc)_2$ binding to E2cd was independently probed by titration calorimetry. From the raw data, the calculated binding isotherm fit well to a sigmoidal curve based on a 1:1 binding interaction and generated values (at 9.6 °C) for K_d and $\Delta H°$ of 1.64 ± 0.09 μM and −35.14 ± 0.28 kJ/mol, respectively. This dissociation constant is 3-fold weaker than that measured by fluorescence titration, while $\Delta H°$ differs by 25%. The thermodynamic parameters are summarized in Table 3. Titration calorimetry was also tried to measure the interaction of E2cd with CTri. Yet, while CTri binding was confirmed, the calorimetric data (not shown) were uninterpretable, as the calculated isotherm did not correspond to any standard binding model.

Characterization of E2(Y73F) and E2(Y73S). Mutants E2-(Y73F) and E2(Y73S) were generated by site-directed mutagenesis, overexpressed, and characterized by ligand

Table 3: Thermodynamic and Kinetic Constants for the Interaction of $MU(Glc)_2$ with E2Cd, CBH II, E2(Y73F), and E2(Y73S)

	E2cd	CBH II[a]	E2(Y73F)	E2(Y73S)
		Binding		
$\Delta G°$ [b] (kJ/mol)	−33.4 (−31.1)[c]	−30.4	−35.0	−37.8
$\Delta H°$ (kJ/mol)	−47.3 ± 2.9 (−35.14 ± 0.28)[c]	−11.6 ± 1.7	−58.0 ± 0.8	−64.6 ± 0.8
$\Delta S°$ [J/(mol·K)]	−46.8 ± 10.0 (−13.5)[c]	63 ± 6	−77.1 ± 2.8	−89.9 ± 2.6
		Hydrolysis at 37 °C		
K_d[d] (M)	2.96×10^{-6}	5.66×10^{-6}	1.76×10^{-6}	6.30×10^{-7}
K_m (M^{-1})	$(3.12 \pm 0.27) \times 10^{-6}$	*e*	$(1.44 \pm 0.46) \times 10^{-6}$	*f*
k_{cat} (min^{-1})	$(1.39 \pm 0.05) \times 10^{-2}$	*e*	$(1.63 \pm 0.18) \times 10^{-4}$	*f*
k_{cat}/K_m (M^{-1} min^{-1})	4450	*e*	114	*f*

[a] Data are from ref *18*. [b] At 25 °C. [c] Determined by titration calorimetry. [d] Calculated by using binding energies from the top of the table. [e] No detectable activity; see ref *40*. [f] Activity too low to permit accurate measurement; see Results.

Table 4: Binding Constants of E2(Y73F), E2(Y73S), and E2cd[a]

	K_d (M)		
ligand	E2cd	E2(Y73F)	E2(Y73S)
$MU(Glc)_2$	$(5.41 \pm 0.19) \times 10^{-7}$	$(2.07 \pm 0.16) \times 10^{-7}$	$(5.67 \pm 0.18) \times 10^{-8}$
$MU(Glc)_3$	$(4.18 \pm 0.26) \times 10^{-8}$	$(2.10 \pm 0.32) \times 10^{-8}$	$(1.64 \pm 0.18) \times 10^{-8}$
MUXylGlc	$\sim 2 \times 10^{-9}$	$\sim 5 \times 10^{-9}$	$(1.38 \pm 0.10) \times 10^{-8}$
cellotriose[b]	$(1.00 \pm 0.04) \times 10^{-4}$	$(1.38 \pm 0.06) \times 10^{-4}$	$(8.3 \pm 0.6) \times 10^{-5}$

[a] From fluorescence titrations at 8.5 °C in 50 mM sodium acetate, pH 5.5. [b] From a displacement titration with $MU(Glc)_2$.

Table 5: Relative CMC, Swollen Cellulose, and Filter Paper Activities of E2(Y73F) and E2(Y73S)[a]

substrate	E2 SA[b] [μmol of CB/(min·μmol)]	relative SA (% wild-type)	
		E2(Y73F)	E2(Y73S)
CMC	335	8.4	0.022
SC	631	5.7	0.088
FP	1.84	74	14

[a] Calculated as described under Experimental Procedures. [b] Data from ref *22*.

binding and activity assays. The ligand binding data are summarized in Table 4. For $MU(Glc)_2$ and $MU(Glc)_3$, decreasing the size of the residue at position 73 (from Tyr to Phe to Ser) progressively increased the binding affinity, with the binding of $MU(Glc)_2$ to E2(Y73S) being nearly 10-fold tighter than that of the wild type. The thermodynamic parameters for $MU(Glc)_2$ binding (Table 3) follow this trend, as both $\Delta H°$ and $\Delta S°$ became progressively more negative when going from E2 to E2(Y73F) to E2(Y73S). In contrast, a smaller residue size at position 73 gave weaker binding of MUXylGlc (Table 4), although all three forms of E2 bound this ligand more tightly than $MU(Glc)_2$. No trend was apparent in the binding data for CTri.

Activity data for the E2 mutants are shown in Table 5. E2(Y73F) retained 8.4% of wild-type activity on CMC, 5.7% on SC, and 74% on filter paper. In contrast, E2(Y73S) retained only 0.022% activity on CMC, 0.088% on SC, and 14% on FP. Hydrolysis of $MU(Glc)_2$ was investigated in the fluorometer as described for E2cd. For E2(Y73F) the calculated K_m was ∼50% smaller than that for E2cd, matching the change in K_d listed in Table 4. The E2(Y73F) catalytic constant (k_{cat}) was reduced nearly 100-fold, while the specificity constant (k_{cat}/K_m) decreased 40-fold (Table 3). Hydrolysis of $MU(Glc)_2$ by E2(Y73S) was too slow to permit measurement of the kinetic constants. It was estimated that E2(Y73S) hydrolyzed 8.12 μM $MU(Glc)_2$ at ∼1% of the rate catalyzed by E2(Y73F) (or ∼0.01% of the rate catalyzed by E2).

DISCUSSION

These results show that E2 binds glycosides with high affinity and specificity. The MU glycoside binding constants shown are among the tightest known for glycosyl hydrolases and are significantly tighter than those for CBH II (*18*). The only reported enzyme/glycoside binding interaction stronger than that of E2cd and MUXylGlc is the binding of the pseudotetrasaccharide inhibitor acarbose to *Aspergillus niger* glucoamylase, with a K_d of 1×10^{-12} M (*31*, *32*). For E2 much of this high affinity is due to the MUF group, as ligands without it have significantly lower affinities than their labeled counterparts. Yet the sugar portions of the ligands are essential for strong binding, as the K_d for MUGlc is 600-fold weaker than the K_d for $MU(Glc)_2$. Also, MU glycosides display the same trends in binding affinities as do the glycosides. The large increase in binding affinity seen with the MU compounds probably results from tighter stacking between the MU group and active-site tryptophan residues than between glucosyl groups and tryptophan residues. Furthermore, glucosyl groups interact much more with water than does MU, which will also cause the MU group to bind more tightly to E2 than does a glucose residue.

The strict specificity for the β(1→4) linkage can be seen by comparing the dissociation constant of $MU(Glc)_3$ ($K_d = 4.18 \times 10^{-8}$ M, Table 1) with that of $MU(Glc)_2$αGlc ($K_d \geq 10^{-4}$). The K_d for MULac (1.7×10^{-4} M) is 300-fold weaker than that for $MU(Glc)_2$ (5.41×10^{-7} M), corresponding to a $\Delta\Delta G$ of 13.4 kJ/mol. This free energy change is of the same magnitude as the loss of one enzyme/ligand hydrogen bond. This would presumably occur as a result of shifting the nonreducing O-4 hydroxyl from an equatorial to an axial orientation, although unfavorable steric interactions could also play a role (such as with Trp 41; see Figure 1). A similar (but smaller) difference in K_d is observed for cellobiose and lactose. Finally, E2 has only weak affinity for $MU(Xyl)_2$, consistent with its inability to degrade xylan (D. Irwin, unpublished results). Thus the affinity of E2 for

MU glycosides accurately reflects its relative affinity for its natural substrates and products.

The steady increase in binding affinity for ligands of greater chain length [MU glycosides up to $MU(Glc)_3$ and oligosaccharides up to S-CTet] is consistent with the structure of the E2 active-site cleft, which contains binding subsites for at least four glucosyl units (*10*). Mutation of the active-site residue Tyr 73 alters ligand K_d values (Table 4), confirming (since CD spectroscopy shows no evidence of a global conformational change in the mutant) that the ligands are binding directly to the E2 cleft.

Since the E2 cleft contains at least four adjacent glucosyl binding subsites, small ligands could bind to it in several different (overlapping) positions. Small glycosides appear to bind to the E2 cleft with their non-reducing-end sugars bound preferentially in glucosyl subsite −2, as indicated by the observed pattern of MU glycoside hydrolysis. E2 (which cleaves between −1 and +1) hydrolyzed $MU(Glc)_2$ to CB + MUF and $MU(Glc)_3$ to CB + MUGlc. This location is supported by the E2 enzyme structure, where cellobiose is bound to the E2cd cleft in subsites −2 and −1 (Figure 1). Finally, a series of mutations of Glu 263 (a key sugar-binding residue in subsite −2) display greatly (up to 1000-fold) reduced binding of MU glycosides (B. Barr, unpublished results). These mutant results indicate that subsite −2 has a high specificity for glucose and explains why the binding of $MU(Xyl)_2$ is so much weaker than that of MUXylGlc.

While it is possible that ligands bind in one (productive) mode during hydrolysis and a second (nonproductive) mode during the binding titrations, the kinetic characteristics of $MU(Glc)_2$ hydrolysis make this unlikely. As shown in Table 3, the K_m for hydrolysis of $MU(Glc)_2$ at 37 °C (3.12 μM) is slightly larger than the calculated K_d (2.96 μM). This result is consistent with K_m and K_d both describing the same binding interaction, as the mathematical relationship between the two constants does not permit K_m to be less than K_d. This is also true for E2(Y73F).

A more detailed characterization of the E2cd/$MU(Glc)_2$ interaction revealed some interesting features. Thermodynamically, $MU(Glc)_2$ binding is driven by a large favorable binding enthalpy that is partially offset by an unfavorable decrease in binding entropy (Table 3). Also, the binding of both $MU(Glc)_2$ and CTri displays a definite pH dependence (Table 2). The magnitude of this pH effect is larger than the one observed for lysozyme (*33*), although that study did not go above pH 8.0. This pH effect is most likely caused by the titration of ionizable groups in the E2 cleft that form close interactions with the ligands. Since binding constants for $MU(Glc)_2$ and CTri (both of which are neutral molecules) vary in opposite directions, it is likely that the group or groups responsible for the change are in subsite +1, where the MU group appears to bind. Subsite +1 includes His 159 and Asp 117. His 159 is most likely responsible for the observed effects because its p*K* is likely to give the observed changes.

The results from titration calorimetry for $MU(Glc)_2$ agree fairly well with those from fluorescence titrations, differing only by a factor of 3. The measured values of $\Delta H°$ and $\Delta S°$ have an even greater difference in magnitude, although both show an unfavorable change in binding entropy. The reasons for these differences are not clear. While calorimetry confirms that CTri binds to E2cd, the binding constant cannot be determined, most likely because CTri is binding in two or more different modes with similar binding affinities. This disagreement with the fluorescence data is probably due to the different ways that these two techniques measure CTri binding. Calorimetry measures CTri binding directly while fluorescence titrations measure the displacement of $MU(Glc)_2$, whose presence may limit the binding modes available to CTri. Alternatively, since $MU(Glc)_2$ displacement precedes CTri binding, the fluorescence experiments may be unable to distinguish between different CTri binding modes.

T. reesei exocellulase CBH II is in the same gene family as E2, and its ligand-binding properties have been extensively characterized (*14*, *18*). A comparison of the results for these enzymes reveals that E2 binds most MU glycosides with higher affinity than does CBH II. Yet since the measured K_d values are all strongly temperature-dependent, the size of the differences in Table 1 is due in part to the lower temperatures required for the E2 experiments (8.5 °C vs 25 °C for CBH II). Both enzymes bind MUXylGlc with much higher affinity than they bind $MU(Glc)_2$.

E2 has a wider range of affinities for ligands than CBH II. The relative affinities of E2 for MUGlc and MUXylGlc differ by 10^5, compared to 10^3 for CBH II. It is unclear whether these differences between E2 and CBH II are due to differences in active-site accessibility (endocellulase cleft vs exocellulase tunnel) or to differential binding of the MU reporter group. The enhancement of CBH II binding affinities in the presence of glucose (*18*) was not observed with E2.

The negative binding entropy of E2cd for $MU(Glc)_2$ is in contrast with the positive entropy of CBH II and is consistent with a conformational change occurring in E2cd upon ligand binding. Such a conformational change in the loops surrounding the active site would make the E2 cleft resemble the CBH II tunnel, where these loops meet to form a roof over the active site. This type of loop movement has been observed in the structure of a catalytically inactive mutant of EGV from *Humicola insolens* (*34*), where the presence of cellohexaose causes the active-site loops to shift and bury the catalytic site. Significantly, E2 mutants with altered loop flexibility (either more rigid or more flexible) display reduced enzymatic activity and wild-type ligand binding affinities (*22*).

The data reported in this paper show that disruption of the interactions between Tyr 73 and substrate has strong effects on binding and hydrolysis. As shown in Figure 1, this highly conserved residue is tilted at an angle to the bound cellobiose, which prevents it from engaging in hydrophobic stacking interactions with the sugar rings. In the E2cd/cellobiose cocrystal, the OH of Tyr 73 interacts directly with cellobiose by a hydrogen bond to the C-6 hydroxyl of the glucose in position −1. While Tyr 73 undoubtedly engages in direct contact with substrate, the catalytic significance of this specific hydrogen bond is unclear as the E2cd crystals are not enzymatically active (9) and a conformational change in the E2 cleft could cause structural rearrangements around Tyr 73.

Mutation of Tyr 73 to Phe and Ser caused progressively tighter binding of $MU(Glc)_2$ and $MU(Glc)_3$ and progressively lower activity on all substrates. Thus increasing the volume of the subsite −1 sugar binding pocket promotes substrate binding, to the detriment of substrate hydrolysis. The

complementary experiment (decreasing the size of the substrate) gave the same result. The predicted binding mode for MUXylGlc places its xylose residue in subsite −1, where it lacks the C-6 CH_2OH group to interact with Tyr 73. E2 bound this ligand more tightly than $MU(Glc)_2$ and cleaved it slowly (<0.1 wt %). The interaction of E2(Y73F) with $MU(Glc)_2$ is structurally intermediate (in subsite volume) between that of E2 with $MU(Glc)_2$ and E2 with MUXylGlc, and likewise its K_d and MU glycoside activity (k_{cat}) were intermediate. However, there is apparently a point beyond which increased volume at subsite −1 interferes with binding, as the K_ds for MUXylGlc with E2(Y73F) and E2(Y73S) were progressively weaker than with wild type. Thus E2(Y73S) bound MUXylGlc only 4-fold tighter than it bound $MU(Glc)_2$, while for E2 the binding difference between these ligands was ~270-fold. This is most likely due to creation of a hole that destabilizes the local structure of the binding cleft.

These results for MUXylGlc and the E2(Y73) mutants are consistent with those previously described for CBH II (*14*) and can be explained in several interconnected ways. First is the issue of sugar ring distortion. Subsite −1 is directly adjacent to the substrate cleavage site, and traditionally there has been speculation that sugar distortion at this position is a vital part of catalysis by polysaccharide hydrolases (especially lysozyme; see ref *35*). This theory has recently received support from ab initio molecular orbital theory, which demonstrated a clear role for ring distortion in enzymes that hydrolyze equatorial glycosides (*36*). The results presented here support the sugar distortion hypothesis, as disruption of the interaction between substrate and Tyr 73 decreased activity and increased binding affinity while creating room for the sugar at position −1 to bind in an unstrained conformation as well as removing a hydrogen bond that may distort the sugar.

Tyr 73 may also assist catalysis by pushing the bound glucose into an optimal position relative to the catalytic residues. In the crystal structure of the inverting *T. fusca* endo/exocellulase E4 (*37*), the addition of substrate (cellopentaose) induces a conformational change, closing the active-site cleft. Through this closing, both the glucose at position −1 and the nucleophilic water are pushed toward the E4 catalytic base, to the point that the base deprotonates the water and induces catalysis. In E2, we have not identified a catalytic base, but mutations of Tyr 73 to Phe and Ser would provide room for the sugar in subsite −1 to bind further from catalytic residues, interfering with substrate hydrolysis. Alternatively, Tyr 73 may assist catalysis by stabilizing the oxocarbonium ion intermediate that has been postulated for inverting glycosidases (*38*). The stabilization would occur through a cation−π interaction (*39*) between the oxocarbonium intermediate and the aromatic Tyr 73 side chain, helping to lower the activation energy of the transition state. This model provides an explanation as to why E2(Y73S) (without an aromatic group) has much lower cellulase activity than E2(Y73F).

In CBH II the residue (Tyr 169) corresponding to Tyr 73 was proposed to help keep the conserved residue Asp 175 ionized and thus assist the catalytic acid during hydrolysis (*14*). In the E2cd structure the residue corresponding to Asp 175 (Asp 79) is too distant from Tyr 73 to have its pK_a affected, although the proposed conformational change may move Asp 79 into close proximity with both Tyr 73 and the catalytic acid.

ACKNOWLEDGMENT

We are indebted to Frank Loontiens for adapting the MU glycoside curve fitting routines for use in KaleidaGraph and to Olle Teleman for providing the TITrat fitting program. We thank Joshua Sakon for helpful discussions and for preparing Figure 1, and Andrew Karplus and the MARKEY Program in Structure-Based Drug Design for use of the titration calorimeter. For collection of the CD spectra, we thank Harold Scheraga for use of the instrument and Gary Davenport and Sheng Zhang for technical assistance.

REFERENCES

1. Warren, R. A. J. (1996) *Annu. Rev. Microbiol. 50*, 183-212.
2. Ghangas, G. S., and Wilson, D. B. (1988) *Appl. Environ. Microbiol. 54*, 2521−2526.
3. Lao, G., Ghangas, G. S., Jung, E. D., and Wilson, D. B. (1991) *J. Bacteriol. 173*, 3397−3407.
4. Bairoch, A., Bucher, P., and Hoffman, K. (1997) *Nucleic Acids Res. 25*, 217−221.
5. Tomme, P., Warren, R. A. J., Miller, R. C., Jr., Kilburn, D. G., and Gilkes, N. R. (1995) in *Enzymatic Degradation of Insoluble Carbohydrates* (Saddler, J. N., and Penner, M. H., Eds.) pp 142−163, American Chemical Society, Washington, DC.
6. Irwin, D. C., Spezio, M., Walker, L. P., and Wilson, D. B. (1993) *Biotechnol. Bioeng. 42*, 1002−1013.
7. Barr, B. K. (1997) Ph.D. Thesis, Cornell University, Ithaca, NY.
8. Barr, B. K., Hsieh, Y.-L., Ganem, B., and Wilson, D. B. (1996) *Biochemistry 35*, 586−592.
9. Spezio, M., Wilson, D. B., and Karplus, P. A. (1993) *Biochemistry 32*, 9906−9916.
10. Spezio, M. L. R. (1994) Ph.D. Thesis, Cornell University, Ithaca, NY.
11. Wilson, D. B., Spezio, M. L., Irwin, D. C., Karplus, P. A., and Taylor, J. S. (1995) in *Enzymatic Degradation of Insoluble Carbohydrates* (Saddler, J. N., and Penner, M. H., Eds.) pp 1−12, American Chemical Society, Washington, DC.
12. Gebler, J., Gilkes, N. R., Claeyssens, M., Wilson, D. B., Béguin, P., Wakarchuk, W. W., Kilburn, D. G., Miller, R. C., Jr., Warren, R. A. J., and Withers, S. G. (1992) *J. Biol. Chem. 267*, 12559−12561.
13. Taylor, J. S., Teo, B., Wilson, D. B., and Brady, J. W. (1995) *Protein Eng. 8*, 1145−1152.
14. Koivula, A., Reinikainen, T., Ruohonen, L., Valkeajärvi, A., Claeyssens, M., Teleman, O., Kleywegt, G. J., Szardenngs, M., Rouvinen, J., Jones, T. A., and Teeri, T. T. (1996) *Protein Eng. 9*, 691−699.
15. van Tilbeurgh, H., Loontiens, F. G., de Bruyne, C. K., and Claeyssens, M. (1988) *Methods Enzymol. 160*, 45−59.
16. de Boeck, H., Matta, K. L., Claeyssens, M., Sharon, N., and Loontiens, F. G. (1983) *Eur. J. Biochem. 131*, 453−460.
17. van Tilbeurgh, H., Pettersson, G., Bhikabhai, R., de Boeck, H., and Claeyssens, M. (1985) *Eur. J. Biochem. 148*, 329−344.
18. van Tilbeurgh, H., Loontiens, F. G., Engelborgs, Y., and Claeyssens, M. (1989) *Eur. J. Biochem. 184*, 553−559.
19. Hanahan, D. (1983) *J. Mol. Biol. 166*, 557−580.
20. Zhang, S., Lao, G., and Wilson, D. B. (1995) *Biochemistry 34*, 3386−3395.
21. Hopwood, D. A., Bibb, M. J., Chater, K. F., Kiersen, T., Bruton, C. J., Kieser, H. M., Lydiate, D. J., Smith, C. P., Ward, J. M., and Schrempf, H. (1985) *Genetic Manipulation of Streptomyces−A Laboratory Manual*, The John Innes Foundation, Norwich, England.

22. Zhang, S., and Wilson, D. B. (1997) *J. Biotechnol.* 57, 101−113.
23. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene 77*, 51−59.
24. Fornwald, J. A., Donovan, M. J., Gerber, R., Keller, J., Taylor, D. P., Arcuri, E. J., and Brawner, M. E. (1993) *Bio/Technology 11*, 1031−1036.
25. Beaven, G. H., and Holiday, E. R. (1952) in *Advances in Protein Chemistry* (Anson, M. L., Bailey, K., and Edsall, J. T., Eds.) pp 324, Academic Press, New York.
26. Doner, L. W., and Irwin, P. L. (1992) *Anal. Biochem. 202*, 50−53.
27. de Boeck, H. (1983) Ph.D. Thesis, Rijksuniversiteit Gent, Belgium.
28. Segel, I. H. (1976) *Biochemical Calculations*, John Wiley & Sons, Inc., New York.
29. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L.-N. (1989) *Anal. Biochem. 179*, 131−137.
30. Bundle, D. R., and Sigurskjold, B. W. (1994) *Methods Enzymol. 247*, 288−305.
31. Sigurskjold, B. W., Berland, C. R., and Svensson, B. (1994) *Biochemistry 33*, 10191−10199.
32. Berland, C. R., Sigurskjold, B. W., Stoffer, B., Frandsen, T. P., and Svensson, B. (1995) *Biochemistry 34*, 10153−10161.
33. Schindler, M., Assaf, Y., Sharon, N., and Chipman, D. M. (1977) *Biochemistry 16*, 423−431.
34. Davies, G. J., and Schülein, M. (1995) in *Carbohydrate Engineering: Proceedings of an International Conference, Elsinore, Denmark, April 23−26, 1995* (Petersen, S., Svensson, B., and Pedersen, S., Eds.) pp 225−237, Elsevier Science B.V., New York.
35. Strynadka, N. C. J., and James, M. N. G. (1991) *J. Mol. Biol. 220*, 401−424.
36. Smith, B. J. (1997) *J. Am. Chem. Soc. 119*, 2699−2706.
37. Sakon, J., Irwin, D., Wilson, D. B., and Karplus, P. A. (1997) *Nat. Struct. Biol. 4*, 810−818.
38. Sinnott, M. L. (1990) *Chem. Rev. 90*, 1171−1202.
39. Dougherty, D. A. (1996) *Science 271*, 163−68.
40. Claeyssens, M., van Tilbeurgh, H., Tomme, P., Wood, T. M., and McRae, S. I. (1989) *Biochem. J. 261*, 819−825.
41. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI980063I