Kinetic Studies of *Thermobifida fusca* Cel9A Active Site Mutant Enzymes†

Weilin Zhou,‡ Diana C. Irwin, Jose Escovar-Kousen,§ and David B. Wilson*

Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853

Received March 29, 2004; Revised Manuscript Received April 29, 2004

ABSTRACT: *Thermobifida fusca* Cel9A-90, an unusual family 9 enzyme, is a processive endoglucanase containing a catalytic domain closely linked to a family 3c cellulose binding domain (Cel9A-68) followed by a fibronectin III-like domain and a family 2 cellulose binding domain. To study its catalytic mechanism, 12 mutant genes with changes in five conserved residues of Cel9A-68 were constructed, cloned, and expressed in *Escherichia coli*. The purified mutant enzymes were assayed for their activities on (carboxymethyl)cellulose, phosphoric acid-swollen cellulose, bacterial microcrystalline cellulose, and 2,4-dinitrophenyl β-D-cellobioside. They were also tested for ligand binding, enzyme processivity, and thermostability. The results clearly show that E424 functions as the catalytic acid, D55 and D58 are both required for catalytic base activity, and Y206 plays an important role in binding, catalysis, and processivity, while Y318 plays an important role in binding of crystalline cellulose substrates and is required for processivity. Several amino acids located in a loop at the end of the catalytic cleft (T245−L251) were deleted from Cel9A-68, and this enzyme showed slightly improved filter paper activity and binding to BMCC but otherwise behaved like the wild-type enzyme. The FnIII-like domain was deleted from Cel9A-90, reducing BMCC activity to 43% of the wild type.

Cellulose, an insoluble polymer composed of long chains of β-1,4-linked glucose residues associated in microfibrils, is the major component of plant biomass and the most abundant organic compound in the biosphere (*1*). Biomass cellulose can be degraded to soluble sugars by synergistic mixtures of microbial cellulases and other cell wall degrading enzymes and is an important potential renewable source of energy and chemicals. However, the molecular mechanism of crystalline cellulose degradation by glycoside hydrolases still is not well understood.

The filamentous soil bacterium *Thermobifida fusca* produces many kinds of hydrolytic enzymes that degrade cellulose and other plant cell wall polymers (*2*, *3*). Characterized cellulases from *T. fusca* include three endocellulases, Cel9B, Cel6A, and Cel5A (*4*, *5*), two exocellulases, Cel6B and Cel48A (*6*, *7*), and an unusual cellulase, Cel9A (*5*, *8*). Cel9A (formerly called E4) has both endocellulolytic and processive, exocellulolytic activity on cellulose. It has relatively high activity on bacterial microcrystalline cellulose (BMCC)[1] and shows synergism with both endocellulases and exocellulases (*9*). Cel9A is a 90.4 kDa secreted protein and contains four domains: an N-terminal family 9 catalytic domain, a family 3c cellulose binding domain (CBD), a fibronectin III (FnIII) like domain, and a C-terminal family 2 CBD. Limited proteolysis of Cel9A produced Cel9A-68, and the three-dimensional crystal structure of this enzyme has been determined at 1.9 Å resolution by X-ray crystallography (*10*). This protein consists of an $(\alpha/\alpha)_6$ barrel family 9 catalytic domain rigidly attached to a family 3c CBD. The catalytic domain has a shallow open active site cleft with six glucose binding sites, numbered from −4 to +2, aligned with the flat binding surface of the CBD (Figure 1). All family 9 members, including Cel9A, perform catalysis with inversion of the anomeric configuration (*11*). Hydrolysis of the glycosidic bond is believed to involve a direct displacement of the leaving group by water, via an oxocarbonium-like transition state, and is aided by a catalytic acid and a catalytic base (*12*). On the basis of the structure of Cel9A-68 combined with sequence alignments with other family 9 members, E424, D55, and D58 were identified as the putative catalytic acid and bases (*10*). In this work, these residues and two other catalytic cleft residues, Y206 and Y318, have been changed in Cel9A-68 by site-directed mutagenesis, and the expressed mutant enzymes have been assayed for activity and substrate binding. Cel9A-68 was also mutated by removing a group of amino acids located just beyond the nonreducing end of the active cleft and changing R252 to K, Δ(T245−L251) R252K (Figure 1A). A mutation of Cel9A-90, which eliminated the FnIII-like domain was also investigated. These experiments serve to emphasize that a number of residues are very important for efficient catalysis and their role should be considered along with identification

† This research was supported by Grant NYC-165522 from the USDA.

* To whom correspondence should be addressed. Phone: (607) 255-5706. Fax: (607) 255-2428. E-mail: dbw3@cornell.edu.

‡ Present address: Department of Neurology, University of California at Irvine Medical Center, Orange, CA 92868.

§ Present address: Martek Biosciences, 555 Rolling Hills Lane, Winchester, KY 40391.

[1] Abbreviations: CMC, (carboxymethyl)cellulose; DNPC, 2,4-dinitrophenyl β-D-cellobioside; DNS, dinitrosalicylic acid; G3, cellotriose; G4, cellotetraose; G5, cellopentaose; SW, phosphoric acid-swollen cellulose; BMCC, bacterial microcrystalline cellulose from *Acetobacter xylinum*; FP, filter paper; MUG2, 4-methylumbelliferyl β-cellobioside; MUG3, 4-methylumbelliferyl β-cellotrioside; CBD, cellulose binding domain; CD, circular dichroism; Fn, fibronectin; TLC, thin-layer chromatography.

10.1021/bi049394n CCC: $27.50
Published on Web 07/03/2004

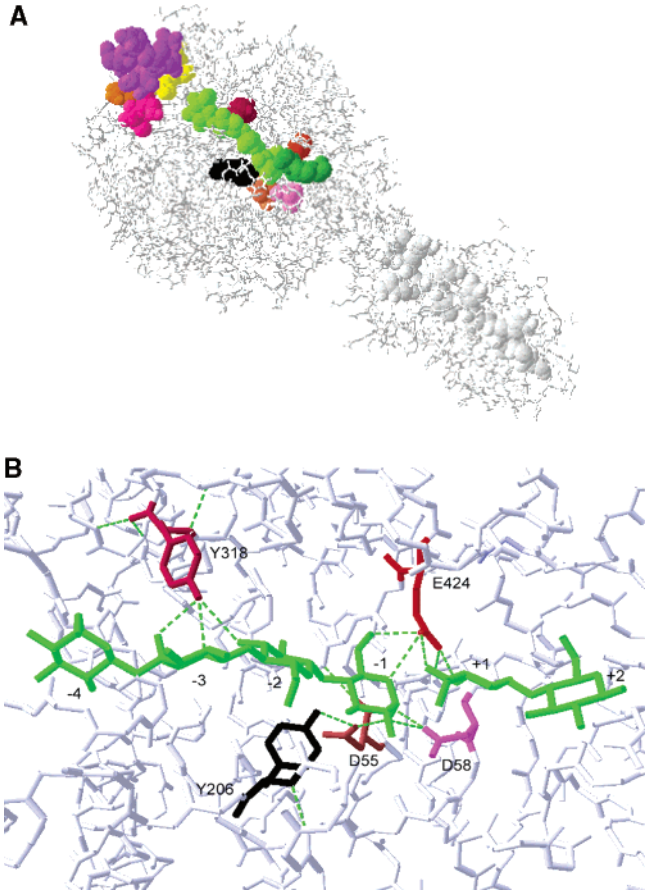

FIGURE 1: The Cel9A-68 product complex structure (*10*). This figure was made with PDB file 4TF4 using a Swiss PDB Viewer. (A) Glc(−1) to Glc(−4) are in light green; Glc(+1) and Glc(+2) are in dark green. The catalytic cleft mutated residues are D55 (orange), D58 (pink), Y206 (black), E424 (red), and Y318 (maroon). The residues removed for the ΔT245−L251 R252K mutant are T245−L251 (purple) and R252 (orange); the adjacent residues L243 and S244 are in dark pink and S253 and Y254 are in yellow. These amino acids and the β strand H556−G566 of the family 3c CBD are shown with van der Waals surfaces displayed. (B) Close-up of the active site cleft showing the interactions of the selected residues with the sugar chain. The glucose residues bound in the active site are labeled as −4 through +2. Hydrogen bonds (as predicted by Swiss PDB Viewer for the residues chosen for mutation) are shown as green dotted lines. They are as follows: D55 Oδ2:Y206 OH, D55 Oδ2:Glc(−1) O1, D58 Oδ1:A56 N, D58 Oδ1:Glc(−1) O1, D58 Oδ2:H125 Nϵ2; E424 Oϵ2:Glc(+1) O3, E424 Oϵ1:Glc(+1) O4, E424 Oϵ1:Glc(−1) O5, E424 Oϵ1:Glc(−1) O6, E424 Oϵ2:G57 N, Y206 OH:D55 Oδ2, Y318 OH:Glc(−3) O6, Y318 OH:Glc(−3) O5, and Y318 OH:Glc(−2) O3.

of the catalytic acid and catalytic base in order to understand the mechanism of cellulose degradation.

EXPERIMENTAL PROCEDURES

Strains and Plasmids. Escherichia coli $DH_{5\alpha}$ was used as the host strain for plasmid isolation (*13*), and *E. coli* BL21-gold(DE_3) from Stratagene was used as the host strain for protein expression. Plasmid pJE2, encoding the *T. fusca* Cel6A signal sequence (MRMSPRPLRALLGAAAAALV-SAAALAFPSQAA) followed by the mature Cel9A-68 gene from pSZ46 (*9*) (EPAFN...WGTAPEEGEEPGG-613 aa) in the pET26b+ vector (Novagen), was used as the template for mutagenesis (*14*).

Mutagenesis and DNA Manipulations. Standard nucleic acid techniques were used as described by Sambrook et al. (*15*). DNA sequence analysis was performed by the Cornell Biotechnology Resource Center using a Perkin-Elmer/Applied Biosystems automated sequencer. Unless otherwise noted all mutant genes were generated using the QuikChange site-directed mutagenesis method (Stratagene). Two oligonucleotide primers for each mutation were designed that are complementary to each other and to opposite strands at the site of the desired mutation. Both primers created or removed a restriction site and were 25−40 bases in length. PCR of the entire plasmid was performed under the following conditions: denaturation at 95 °C, addition of GC-rich polymerase (Roche) followed by 16 cycles of denaturation at 95 °C for 30 s, annealing at 55 °C for 60 s, and extension at 68 °C for 15 min. For the double mutant gene, 18 cycles were used. The sample (50 μL) was then incubated with 20 units of *Dpn*I for 3 h at 37 °C to digest and remove the methylated template. Finally, the ends of the PCR molecules were polished by incubation with Pfu polymerase (Stratagene) at 37 °C for 30 min and at 72 °C for 30 min. The PCR product was transformed into *E. coli* $DH_{5\alpha}$ or BL21-gold(DE3) (Stratagene), and plasmid minipreps of individual transformants were screened by restriction enzyme digestion for the added or removed site. The entire Cel9A-68 gene of the selected transformant was then sequenced to confirm the presence of the desired mutation and to be sure that no other mutations were present. Sequence-confirmed plasmid DNA was transformed into *E. coli* BL21-gold(DE3) (Stratagene) for expression.

The Cel9AΔFnIII-80 plasmid was constructed using pJE1 (Cel9A cloned into pET26b+ with its own signal sequence). PCR fragments were made that created an *Eco*RI site at the C terminus of the family IIIc CBD and at the N terminus of the family II CBD. These two fragments were gel purified and digested with *Kpn*I and *Eco*RI and with *Eco*RI and *Xho*I, respectively. They were then ligated to pJE1, which had been digested with *Kpn*I and *Xho*I. The resulting plasmid, pWL8, was transformed into BL21codon plus RP DE3 (Stratagene). The mature protein (N terminus EPAFNY...) consisted of amino acids 1−619 + Ile + amino acids 728−834.

The Cel9AΔ(T245−L251) R252K-68 deletion removing the block at one end of the catalytic cleft was constructed from pSZ46 (Cel9A-68 in the *Streptomyces*−*E. coli* shuttle plasmid pSES1) (*9*) using the overlap primer extension method as previously described (*16*) to eliminate the desired codons. The PCR product and the wild-type plasmid were digested with *Acs*I and *Fse*I, ligated, and transformed into $DH_{5\alpha}$, and the sequence of the PCR product was confirmed. The plasmid was then transformed into *Streptomyces lividans* TKM31 as previously described (*9*) to produce strain S227.

Enzyme Production and Purification. E. coli BL21-gold(DE3) strains were grown at 37 °C overnight in 100 mL of LB with 60 $\mu g \cdot mL^{-1}$ kanamycin. Thirty milliliters of the overnight culture was transferred into each liter of M9 medium containing 0.5% glucose and 60 $\mu g \cdot mL^{-1}$ kanamycin. After growth at 30 °C for about 5 h to an OD_{600} of about 0.8, isopropyl thio-β-D-galactoside was added to the culture to 0.8 mM and growth continued at 30 °C for 16 h. *Streptomyces* strain S227 was grown at 30 °C for 35 h in tryptic soy broth plus thiostrepton as previously described (*9*).

Wild-type and mutant proteins were purified from the culture supernatants by a published procedure (*9*). Enzyme

purity was assessed on SDS gels, and protein concentrations were determined by absorbance at 280 nm using extinction coefficients calculated from the predicted amino acid compositions.

Circular Dichroism and Mass Spectrometry. Circular dichroism (CD) spectra were run by the Circular Dichroism Spectroscopy Facility, University of Medicine and Dentistry of New Jersey, using an AVIV Model 62D spectropolarimeter and were recorded for both wild-type and mutant enzymes (Y206F, Y206S, and D55A/D58A) over the range 190−260 nm with point spacing of 0.25 nm and 1−2 s average times. The resulting CD spectra, obtained with an enzyme concentration of 0.15 mg·mL^{-1} in potassium phosphate buffer (10 mM, pH 7.0), were analyzed for percent secondary structure using CDNN CD spectra deconvolution developed by Böhm et al. (*17*).

Mass spectrometry was performed with a Mald/I-TOF Bruker Biflex III using a sinapinic acid matrix. The expected error for this system is about 1% of the mass of the protein.

Activity Assays. The activities of Cel9A and the mutant enzymes were determined on CMC (low viscosity, Sigma), 10 mg·mL^{-1}; bacterial microcrystalline cellulose (BMCC) (a gift from Monsanto, now CP Kelco), 2.5 mg·mL^{-1}; phosphoric acid-swollen Sigmacell (SW), 2.5 mg·mL^{-1}; and Whatman filter paper no. 1 (FP), 8 mg·mL^{-1}. All assays were run in a reaction volume of 0.4 mL in triplicate at 50 °C for 16 h in 0.05 M NaOAc, 0.02% sodium azide, and 0.015 M $CaCl_2$ (pH 5.5) using from 0.5 to 600 pmol of protein as needed. Reducing sugars were measured using dinitrosalicylic acid reagent (DNS) (*18*) and a glucose standard (Sigma) curve. The main product of Cel9A hydrolysis is cellobiose with some glucose and cellotriose. In these calculations the OD values were converted to micromoles of cellobiose by comparing a cellobiose standard curve to the routine glucose standard curve. Using these assay conditions, if the micromoles of Glu/OD = 1.6, then the micromoles of cellobiose/OD = 1.06. The SW activity assays were repeated for all of the E424, D55, and D58 mutants and wild type using sodium/potassium phosphate buffer (0.05 M, pH 6.8) compared to sodium acetate (0.05 M, pH 6.2) plus sodium azide buffer to be sure that the acetate and azide anions were not increasing activity. The differences between the two buffers were random and well within the average coefficient of variation for SW assays (13%).

All mutant enzyme activities were determined concurrently with wild type. In the case of wild type and mutants with activity, the nanomoles of protein used (X) was plotted versus the micromoles of reducing sugar produced (Y) and fitted to the equation $Y = m_1X/(m_2 + X)$ using the program KaleidaGraph from Synergy Software. This curve was used to determine the amount of enzyme required to digest the substrate to the target value of 5% for filter paper and CMC and 10% for swollen cellulose and BMCC. If the target digestion could not be reached, the activity was calculated from the micromoles of cellobiose produced by 0.6 nmol of enzyme (1.5 nmol·mL^{-1}). To determine the distribution of reducing ends between filter paper (FP, insoluble) and supernatant (soluble), assays were run on FP using 0.6 nmol of enzyme per reaction; the supernatant was separated from the filter paper circle, and the reducing sugar content of each fraction was determined by the DNS reagent as described previously (*8*).

Activity assays with 2,4-dinitrophenyl β-D-cellobioside (DNPC), a gift from Dr. Stephen Withers, University of British Columbia, Canada, were run in 0.05 M NaP_i buffer, pH 5.8 at 50 °C. The cuvettes were removed from the water bath, and readings at an absorbance of 400 nm were taken approximately every 7 min over a 60 min time period. The DNP produced was calculated using an extinction coefficient of 10910 M^{-1} cm^{-1} (*19*). Blanks were determined by measuring the uncatalyzed rate of DNP production at values from 20 to 1800 μM DNPC. These rates formed a straight line with DNPC concentration, and the slope was used to calculate the blanks, which were subtracted from the slopes of the catalyzed reactions. The rates of DNP production with respect to time were linear for at least 1 h.

Thin-layer chromatography was used to visualize the products of hydrolysis of cellotetraose (G4) and cellopentaose (G5) as described previously (*2*, *20*). Glucose oligomers were obtained from Sigma or Seikagaku America.

Binding Assays. Binding to BMCC was measured in siliconized 1.5 mL Eppendorf tubes by adding 0.5 nmol of enzyme to 2.5 mg·mL^{-1} BMCC in a total of 0.5 mL of 50 mM NaOAc buffer, pH 5.5, plus 10% glycerol. Tubes were rotated end to end at 4 °C for 1 h and centrifuged at 12000*g* in a microfuge for 5 min. The supernatants were filtered through 0.45 μm cellulose acetate centrifuge Costar filters which had been pretreated with 300 μL of bovine serum albumin at 1 mg·mL^{-1} and then rinsed with buffer three times. The amount of unbound Cel9A-68 was determined by the protein concentration calculated from the A_{280}. Control reactions without BMCC and also with BMCC but no protein were performed for every assay. The reducing sugars produced in the binding reactions were determined with the *p*-hydroxybenzoic acid hydrazide (PAHBAH) reagent according to the published procedure (*21*).

Ligand Binding Studies. Dissociation constants (K_d) for the binding of 4-methylumbelliferyl β-cellotrioside (MUG3) to Cel9A-68 and other mutant proteins were determined by direct fluorescence quenching titrations as was done for Cel6A using an Aminco SLM 8000C spectrofluorometer and were calculated as described (*22*). The initial concentration of MUG3 was 1.8 μM in 903 μL of sodium acetate buffer (0.05 M, pH 5.5). Protein (100 μM) was added at 3.5 μL/min. The temperature was maintained at 6.5 °C. Excitation was at 316 nm, and emission was measured at 360 nm. K_d values for cellotriose (G3) and cellotetraose (G4) were measured by displacement titration experiments using the Cel9A-68 titrations with MUG3 as the starting point and adding 16.5 mM G3 or 10 mM G4 (*22*). The G3 and G4 K_d's were calculated using a program from Mathematica Software, Wolfram Research, using equations reported previously (*23*).

Thermostability Study. Cel9A-68 and the mutant enzymes, at a concentration around 100 μM in 50 mM NaOAc, pH 5.5, buffer and 10% glycerol, were preincubated at temperatures ranging from 14 to 75 °C for 15 h and then diluted to 0.2−1.2 μM and assayed on SW at 50 °C (*6*). The relative activity of each enzyme was calculated as a percentage of the activity of the enzyme sample preincubated at 14 °C.

Accession Numbers. The sequence files for Cel9A (formerly called E4) are L20093 and P26221. The structure files for Cel9A-68 are 1JS4, 1TF4, 3TF4, and 4TF4.

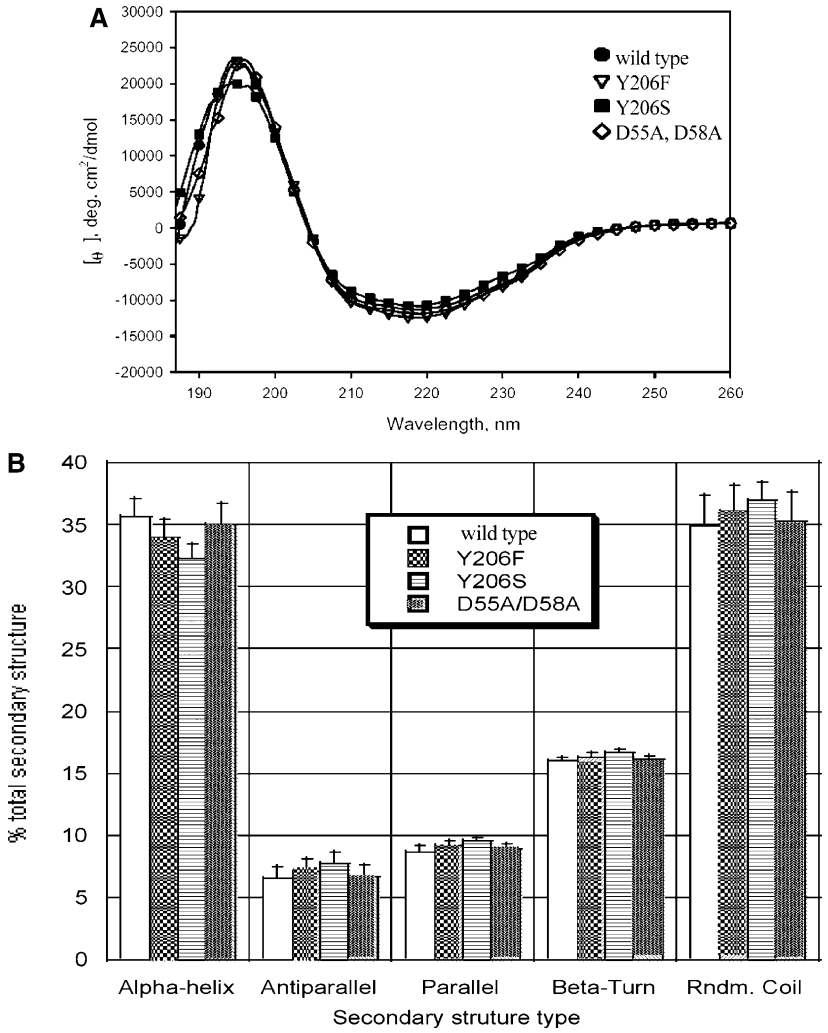

FIGURE 2: Circular dichroism analysis of wild-type and mutant Cel9A. (A) Spectra for wild-type, Y206F, Y206S, and the D55A/D58A double mutant enzymes are shown. The mean residue ellipticity of wild-type and mutant proteins was measured at 0.15 mg/mL in 10 mM potassium phosphate buffer, pH 7.0 at 25 °C. (B) The relative amount of each type of secondary structure is indicated for each enzyme. Error bars represent the standard deviation of the mean of at least three different spectrum ranges.

RESULTS AND DISCUSSION

Selection of Mutation Sites. Twelve mutations of five residues were chosen to begin a study of the Cel9A-68 catalytic mechanism. The structure, 4TF4, with these residues and their respective H-bonds highlighted is shown in Figure 1B. Density for six glycosyl units is present in the active site cleft due to binding of cellopentaose in two ways from Glc(−4) to Glc(+1) and from Glc(−3) to Glc(+2) with cleavage occurring between the −1 and +1 sites (*10*). This structure shows the products still bound in the active site in which D55 and D58, the putative catalytic base(s), are each hydrogen bonded to the Glc(−1) O1 and E424 and the putative catalytic acid is hydrogen bonded to Glc(+1) O4 and to Glc(−1) O5 and O6 as discussed in detail by Sakon et al. (*10*). Y206 is a residue interacting with Glc(−1) and is hydrogen bonded to D55. Y318 was chosen because it has two H-bonds to Glc(−3) and another H-bond to Glc(−2) and the tyrosine ring is not stacked with a substrate sugar ring (*24*). A RPS-BLAST was performed to search the Conserved Domain Database (available at http://www.ncbi.nlm.nih.gov/BLAST) using the Cel9A-68 amino acid sequence. The alignments showed that D55, D58, and E424 are invariant even among the most diverse members of the family 9 glycohydrolase domain. Y206 and Y318 are well conserved, especially among the 25 most similar members of the family, but are sometimes replaced by V, S, R, F, L, D, P or W, H, F, respectively, over the entire family 9.

The loop from T245 to L251 has a high-temperature factor in the structure (*10*), and the residues from 245 to 255 appear to form a block at the far end of the catalytic cleft beyond the −4 sugar binding site (Figure 1A). The mutant Cel9A-68 ΔT245−L251 R252K was made to investigate the role of this block, which is not well conserved and is missing in many family 9 enzymes. The Cel9AΔFnIII-80 mutant was made to investigate the impact of this domain on activity as well as to see if expression in *E. coli* could be improved if the FnIII domain was missing.

Expression and Purification of Mutant Enzymes. All of the Cel9A-68 proteins expressed well in *E. coli* BL21-gold-(DE3) and were secreted into the culture supernatant. All purified proteins (10−25 mg·L^{-1}) were at least 95% pure as shown by SDS−PAGE. The efficiency of signal peptide cleavage was checked by mass spectrometry analysis of Cel9A E424A. Its mass was 67941 ± 679 Da (theoretical mass 67886 Da). The peak was fairly sharp but contained two slight shoulders at higher mass, indicating that some

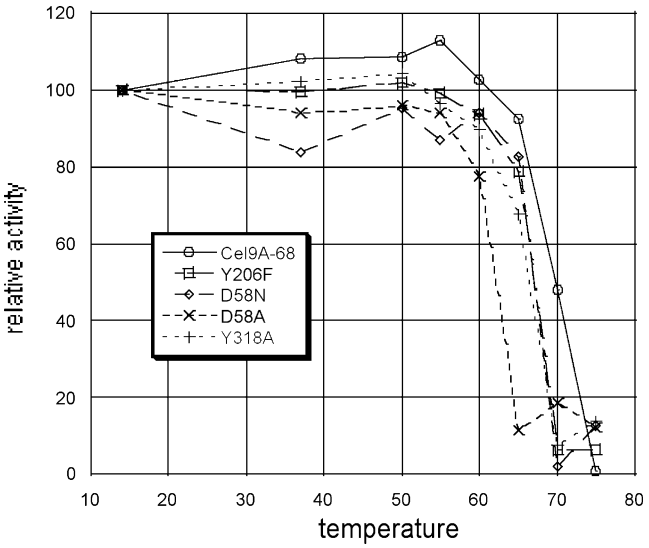

FIGURE 3: Comparison of the thermostability of Cel9A-68 and the mutant enzymes. Samples were pretreated for 15 h at various temperatures (14−75 °C), and then SW activity was assayed at 50 °C. The relative activity of each enzyme was calculated as a percentage of the activity after the 14 °C preincubation.

miscleavage of the signal peptide was occurring to give one or two extra alanines added to the N terminus. Expression of the whole multidomain protein Cel9A-90 in *E. coli* was too low to be useful for protein purification. Exchange of the Cel9A signal sequence for the Cel6A signal sequence did not improve expression of the whole protein. Cel9AΔFnIII-80 was expressed better than Cel 9A-90 but at a much lower level than Cel9A-68 with a yield of only 6 mg from a 6 L preparation.

Integrity of Mutant Protein Folding. The secondary structures of wild-type Cel9A-68 and the Y206F, Y206S, and D55A/D58A mutant enzymes were analyzed using CD spectroscopy. Figure 2 shows that the secondary structure content of all three mutant proteins is similar to that of the wild-type enzyme, although the α-helix content of Y206S is slightly reduced while the content of random coil structure is slightly increased. No significant gross perturbations in secondary structure are evident in the mutant proteins, but it is possible that the mutations cause a localized perturbation of secondary structure or a small conformational change.

The thermostability of all of the mutant enzymes was tested, but only the results for the four mutant enzymes that had measurable activity are shown in Figure 3. At 65 °C the D58A mutant enzyme lost nearly all activity while the Y206F, D58N, and Y318A enzymes retained 70−80% activity and wild type retained >90% activity. At 70 °C only wild-type enzyme had significant activity. Although the mutants are somewhat less thermostable, they were all stable at 50 °C, the temperature at which the assays were performed.

Activity of Mutant Enzymes. To elucidate the role of individual amino acids, the mutant enzyme activities were assayed on four cellulosic substrates, and the results are shown in Table 1. CMC is a soluble substrate which endocellulases can easily degrade; the carboxymethyl groups do not allow the enzyme to carry out extensive processsive hydrolysis, and exocellulases have very low activity on this substrate. Phosphoric acid-swollen cellulose (SW) is not soluble, but the individual cellulose chains are much more hydrated than in crystalline cellulose. Thus it is a much easier

Table 1: Activities of Mutant Enzymes on Polysaccharides

enzyme	activity[a] [μmol of cellobiose min^{-1} (μmol of enzyme)$^{-1}$] CMC	SW	BMCC	filter paper	ratio of soluble/insoluble reducing sugar (processivity)
Cel9A-68	216	11.6	2.07	0.280	2.5
D55A	0.58	0.136	0.033		
D55N	0.63	0.189	0.046		
D58A	0.86	0.207	0.073		
D58N	0.99	0.230	0.056		
D55A/D58A	0.02	0.015	0.011		
Y206F	15.2	0.480	0.172	0.132	0.76
Y206S	1.03	0.147	0.037	0.061	0
Y318A	**1193**	**3.29**	0.294	0.171	0.9
Y318F	**1441**	**8.66**	0.344	0.178	1.3
E424A	0.28	0.027	0.009		
E424G	0.71	0.135	0.051		
E424Q	0.10	0.018	0.023		
ΔT245−L251, R252K	**212**	**11.7**	**2.10**	0.359	2.4
Cel9A-90	**221**	**16.0**	**8.3**	0.91	7.4
Cel9A ΔFnIII	**165**	**14.2**	**3.6**	0.63	5.4

[a] Activities in bold were calculated at 10% digestion for SW and BMCC and 5% digestion for CMC and FP. All others were calculated using 1.5 nmol of protein/mL. All assays were done for 16 h at 50 °C. The average coefficients of variation (standard deviation divided by the mean) were 13%, 4.7%, 18.7%, and 2.5% for SW, CMC, BMCC, and FP, respectively.

task for an enzyme to bind a single cellulose molecule into the active cleft. Bacterial microcrystalline cellulose (BMCC) and filter paper (FP) are both crystalline substrates in which the cellulose molecules are highly ordered into microfibrils by hydrogen and hydrophobic, but not covalent, bonding. They differ in that BMCC has been shown to have much more open space between the microfibrils and therefore greater binding capacity than Avicel, which is made from fibrous plants as is filter paper (*25*, *26*).

DNPC is a useful substrate because it has an excellent leaving group with a pK_a of 3.96 and therefore does not require acid catalysis. The Cel9A-68 wild-type DNPC activity is quite low but measurable, and the activities of some of the mutant enzymes are higher as shown in Table 2.

Low activity can be caused by the inability of the enzyme to bind the substrate. The Cel9A-68 mutant enzymes were tested for BMCC binding at 4 °C as a measure of their overall ability to bind to cellulose. The percentages of enzyme bound to 2.5 mg·mL^{-1} BMCC (the concentration in the activity assay) at 4 °C are shown in Table 2. It was previously reported that Cel9A-68 does not bind to BMCC at room temperature although D55C (<1% activity) does bind presumably because once the substrate is bound in an active enzyme, it is quickly hydrolyzed and the product released (*9*). At 4 °C the wild-type enzyme does bind somewhat to BMCC, and measurement of the reducing sugar produced showed that the activity of wild-type Cel9A-68 was approximately 0.88 μmol of cellobiose min^{-1} (μmol of protein)$^{-1}$ (0.9% digestion). The Y318 mutants produced a very small amount of reducing sugar, 0.6% digestion, and none of the other mutants had detectable activity on BMCC under these conditions. Thus, there is some hydrolysis taking place even at 4 °C, and the BMCC binding results represent a combination of the ability of the mutants to bind cellulose and their activity.

Table 2: Activities and Binding Constants for Cel9A-68 Mutant Enzymes on Oligosaccharides

	DNPC activity			BMCC binding	ligand binding, K_d (μM)		
enzyme	k_{cat}[a] (min^{-1})	K_m (μM)	k_{cat}/K_m ($min^{-1}\,\mu M^{-1}$)	% bound[b]	MUG3[c]	G3[d]	G4[d]
Cel9A-68	1.20 ± 0.09	697 ± 108	0.00172	3−15	0.57 ± 0.08	12.2 ± 0.9	4.4 ± 0.6
D55A	(0)			75	2.21 ± 0.22	39.5 ± 6.2	5.6, 3.6
D55N	(0.015)			72	1.89 ± 0.37	24.9 ± 1.1	7.6 ± 2.2
D58A	(0.02)			21	0.50, 0.39	20.0	2.5
D58N	(0.015)			35	0.27, 0.26	20.0	3.1
D55A/D58A	(0.008)			57	2.24 ± 0.21	35.7	10
Y206F				0	1.57, 1.20	25	6.7
Y206S				0	poor binding		
Y318A				11	poor binding		
Y318F				11	∼∼41		
E424A	4.2 ± 0.5	290 ± 100	0.0146	83	1.23 ± 0.26	33.3, 12.5	1.7
E424G	42.1 ± 6.3	197 ± 92	0.213	83	0.69 ± 0.04	10.0	2.1, 1.7
E424Q	(0.066)			85	0.67, 0.74	19.2	8.3
Cel9A-90					0.50 ± 0.06	15.6 ± 1.4	7.2 ± 0.6

[a] The specific acitivity on DNPC was determined with 1 μM enzyme at 350 μM substrate at 50 °C for enzymes with very low activity on this substrate (shown in parentheses). [b] Binding assays were performed with 1 nmol/mL protein and 2.5 mg/mL BMCC at 4 °C for 1 h. [c] Standard errors for binding K_d's are given for repetitions of three or more; both values are given if the assay was done twice. [d] Some G3 and G4 displacement curves were done only once and must be considered approximate values.

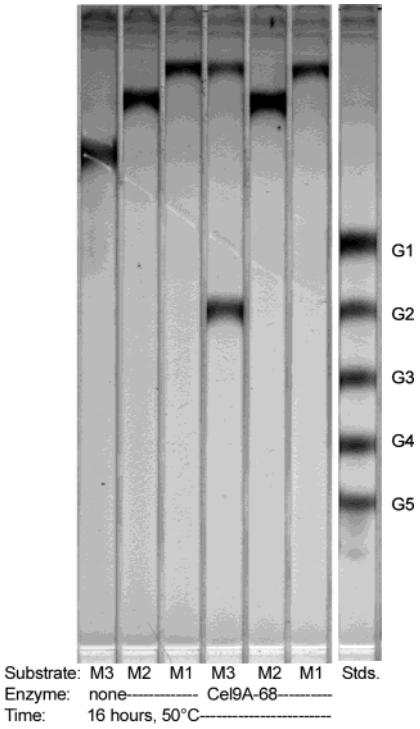

FIGURE 4: TLC showing that Cel9A-68 degrades MUG3 to MUG and cellobiose at 50 °C. Key: M1, MUG; M2, MUG2; M3, MUG3; G1, glucose; G2, cellobiose.

Fluorescence titration binding assays were used to measure the ability of the mutant enzymes to bind the oligosaccharides: MUG3, cellotriose (G3), and cellotetraose (G4). Unlike *T. fusca* Cel6A (*27*) no fluorescence quenching was observed after Cel9A-68 was titrated into a MUG2 solution. However, MUG3 did bind to Cel9A-68 as shown by rapid quenching. Cel9A-68 was incubated with MUG3 for up to 2 h at 6.5 °C and also for 2 h at 55 °C, and TLC analysis of the hydrolysis products showed that that there was no hydrolysis at 6.5 °C. However, at 50 °C, the products were cellobiose and MUG (Figure 4), showing that MUG3 must bind into the Glc(−2) to Glc(+2) binding sites with MU in the +2 site at the reducing end. A typical fluorescence quenching binding curve for the titration of MUG3 with

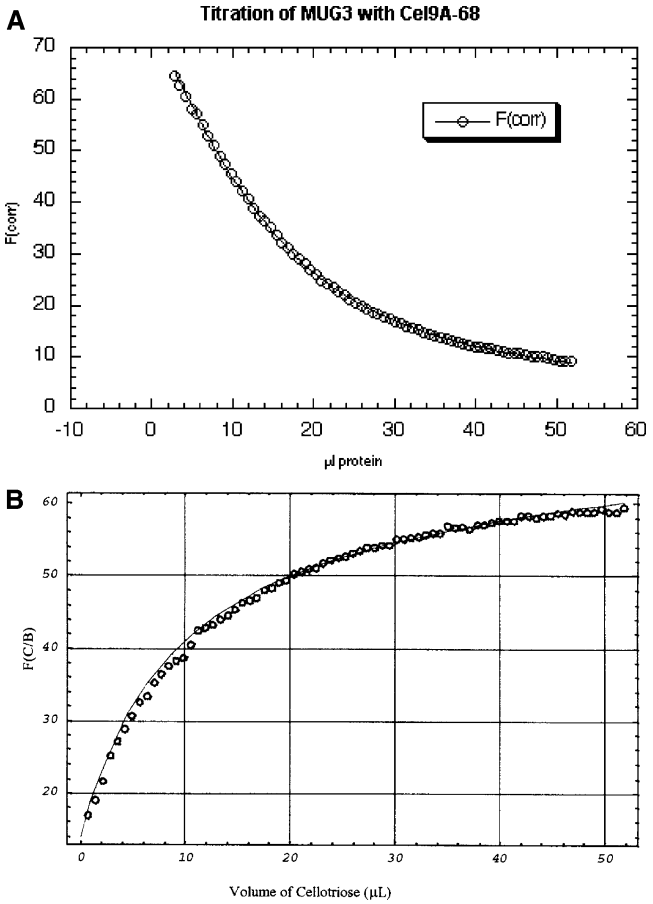

FIGURE 5: Binding and displacement curve of MUG3 and G3 with Cel9A-68. (A) 1.0 μM MUG3 was titrated with 100 μM Cel9A-68 at 3.5 $\mu L \cdot min^{-1}$ for 15 min. The resulting binding curve is shown (circles) along with the curve fit (solid line). (B) Titration displacement of MUG3 by 16.5 mM G3 immediately after the reaction in (A) at 3.5 $\mu L \cdot min^{-1}$ for 15 min.

Cel9A-68 is shown in Figure 5A, and a displacement titration of MUG3 bound to Cel9A-68 with G3 is shown in Figure 5B. Although there is some deviation of the curve fit from the data in both cases, these curve fits were consistent for all of the Cel9A proteins, and they are useful for comparing

the K_d of the various mutants. The binding constants for MUG3, G3, and G4 were determined for all mutants, except those where the binding of MUG3 was too weak to measure (Table 2).

The ratio of soluble to insoluble reducing sugar ends was also determined for mutant enzymes that retained some filter paper activity (Table 1). Exocellulases typically produce more than 90% soluble reducing ends, which reflects their ability to "process" along the cellulose chain (*8*), that is, to make a hydrolytic cleavage, release the product, and then move the same molecule further along the active site tunnel and into position for another hydrolytic cleavage without releasing the substrate. Endocellulases with family 2 CBD's produce about 60% soluble reducing sugar. Cel9A-90 and -68 produce 88% and 71% soluble reducing ends, respectively (Table 1) (*9*).

E424 Mutant Enzymes. The three mutations of E424 dramatically lowered activity on all cellulosic substrates and all exhibited very high binding to BMCC, showing that their low activity was not due to poor binding of the substrate. The K_d's for MUG3, G3, and G4 showed no drastic changes although the E424A mutant enzyme had a K_d for MUG3 that was twice that of wild type. The E424 mutant enzymes showed surprisingly wide variations in activity on DNPC (Table 2) with the k_{cat}/K_m of E424A being almost 10 times that of the wild type and that of E424G being 100 times that of the wild type. This could be due to the size and orientation of the residue 424 side chain, which may interfere with DNP binding productively into the Glc(+1) site; E424A has a much smaller side chain, and E424G has no side chain. As noted above, MUG2 does not bind to Cel9A-68, which also supports the idea that binding of groups such as MU or DNP into the Glc(+1) site is very weak. It was not possible to measure k_{cat} and K_m for E424Q because the activity was so low, and this could be due to a large side chain which does not have the proper orientation and hydrogen bonding to support the necessary transition state. The fact that the E424 mutant enzymes do have activity on DNP, which requires no acid catalysis, but not on cellulose confirms that E424 is the catalytic acid.

D55 and D58 Mutant Enzymes. The D55 and D58 mutant enzymes had very low activity on all the cellulosic substrates, especially on CMC with 0.6% or less activity. The double mutation D55A/D58A further lowered the CMC and SW activities more than an order of magnitude, which is as low as the activity of the E424 mutants. The BMCC binding of the D55 mutants was high, indicating that these mutations had no effect on binding while the D58 mutants and the double mutant had lower BMCC binding, indicating that D58 is involved in the binding of cellulose as well as in the catalytic activity. In contrast, the ligand binding experiments showed that the K_d's for MUG3 binding to the D58 mutant enzymes were very close to those of wild type while the K_d's of the D55 and double mutant enzymes were about three times higher than the wild type. This result might be caused by a small change in the orientation of Y206 due to the loss of the hydrogen bond to D55. None of these mutant enzymes had activity on DNPC as expected. Clearly, both D55 and D58 are very important for catalysis.

Y206 Mutant Enzymes. The Y206S mutant enzyme had very low activity on cellulosic substrates comparable to that of the D55 and D58 mutants while the Y206F mutant had from 4% to 8% of wild-type activity on CMC, SW, and BMCC. Replacement of tyrosine with serine puts the hydroxyl much further away from D55 and from Glc(−1), so it cannot make the appropriate hydrogen bond with D55 Oδ2. Neither of these proteins demonstrated binding to BMCC, but because Y206F has 8% residual activity on BMCC, it must have the ability to bind this substrate into the catalytic cleft long enough for some cleavage to take place. The MUG3 binding of Y206S was below detection levels, but the K_d's for Y206F were approximately double that of wild type, showing that it could bind small oligosaccharides. It is striking that there are no soluble reducing ends produced by the Y206S mutant enzyme even though it retains 21% of the wild-type filter paper activity. Perhaps a long cellulose chain allows some binding to the active cleft, but this is followed by cleavage and immediate dissociation of the product chains. In contrast, wild-type Cel9A produces cellotetrose, which is rapidly hydrolyzed to cellobiose or cellotriose and glucose. These results show that the hydrophobic nature of this residue is playing a role in substrate binding and that the tyrosine hydroxyl group is also very important.

Y318 Mutant Enzymes. The Y318A and Y318F mutant enzymes had very high activity on CMC, showing that Y318 is not a catalytic residue and that removing the hydroxyl group which hydrogen bonds to the Glc(−3) and Glc(−2) residues probably leaves more room in the active cleft for CMC to bind. The SW activity was 28% and 74% of wild-type activity for Y318A and Y318F, respectively. The BMCC and FP activites were nearly the same for both mutants; however, the BMCC activities were only about 15% of wild type while the FP activities were 61% of wild type. BMCC binding was about the same as wild type despite the lower activity, suggesting that binding was impaired. Ligand binding was very poor for both mutants, showing that the hydrogen bonding of the tyrosine side chain to Glc(−2) is very important for oligosaccharide binding. The number of soluble reducing ends produced by both mutants was about equal to the number of insoluble reducing ends, indicating that this residue is involved in processivity. The products of hydrolysis of G4 and G5 by these mutant enzymes are shown by TLC analysis in Figure 6. The wild-type enzyme produces about equal amounts of glucose and G3 or cellobiose from G4 hydrolysis, showing that the G4 binds equally well in the Glc(−3) to Glc(+1) or the Glc(−2) to Glc(+2) positions. The main product of G4 hydrolysis by both mutants is cellobiose, showing that the binding in the Glc(−3) subsite is not as energetically favorable in the mutant enzymes. The G5 hydrolysis pattern also shows that the products are different for the mutant enzymes. It is interesting that the Y318A enzyme retained high activity on G4 even though it had very poor binding to MUG3.

Cel9A-68 ΔT245−L251, R252K. No significant differences in the CMC, SW, or BMCC activities were found between wild type and Cel9A-68 ΔT245−L251, R252K, which has the loop at the nonreducing end of the active cleft deleted. The products of G5 and G6 hydrolysis were investigated by TLC analysis (data not shown), and they were the same as for the wild-type enzyme. It seems likely that the affinity of the Glc(−4) to Glc(−1) subsites determines where the chain binds during processive cleavage rather than this peptide block. The ΔT245−L251, R252K mutant enzyme bound

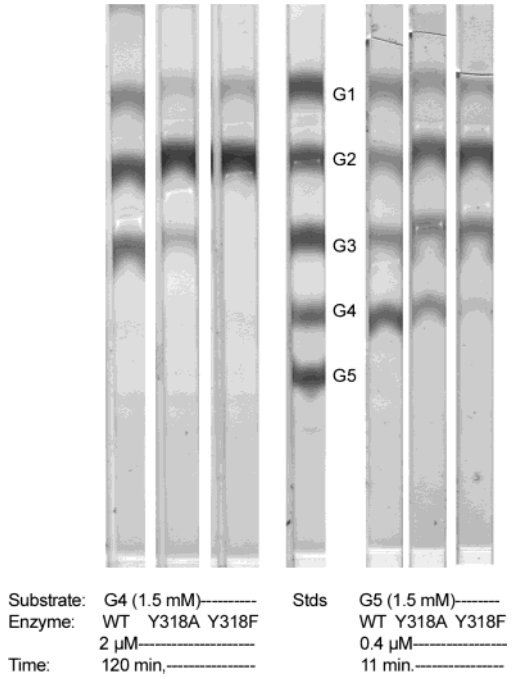

FIGURE 6: TLC of the products of G4 and G5 degradation by the Y318A and Y318F mutant enzymes.

better to BMCC than wild type and had higher filter paper activity, indicating that removing this block may provide better access to the cellulose chain. It will be interesting to see if the complete enzyme incorporating this mutation has higher activity on crystalline cellulose.

Cel9A-80 Δ*FNIII*. Removal of the FnIII domain from Cel9A lowered activity on BMCC more than on CMC or SW and also reduced processivity but not dramatically. A Blast search using Cel9A amino acids 626−730 showed that similar FNIII domains are found in a wide variety of cellulases, chitinases, amylases, and other glycosyl hydrolase enzymes. Deletion of the two FnIII domains from *Clostridium thermocellum* CBHA reduced activity on insoluble cellulose by 50%, and the *C. thermocellum* CBHA $FnIII_{1,2}$ region alone was shown by scanning electron micrographs to significantly erode the surface of cellulose fibers (*28*).

CONCLUSIONS

Family 9 is the most taxonomically diverse family, containing cellulases from plants, bacteria, termites, and slime molds, as well as a few from fungi (*29*). The family includes both processive and many nonprocessive endocellulases and even a rare exocellulase (*30*). Since Cel9A has unusual properties including especially high activity on BMCC, it is of interest to determine its catalytic mechanism.

The high activity of the E424A and E424G mutant enzymes on DNPC, which has such a good leaving group that no acid catalysis is required, proves that E424 is the catalytic acid as predicted from the structure (*10*). The corresponding residues in normal endoglucanases, Cel9A (CelD) from *C. thermocellum* (*31*) and *Clostridium stercorarium* CelZ (*11*, *32*), also have been shown to function as the catalytic acid.

According to the structure and to sequence alignments with other family 9 enzymes, either D55 or D58 could be the putative catalytic base. The distance between each of these residues and E424 is around 10 Å, which is the average distance between the catalytic base and catalytic acid of most inverting glycosidases (*12*, *33*). Comparison of the enzyme−product structure with the unliganded structure revealed a water molecule 0.3 Å from the anomeric hydroxyl group at Glc(−1), which interacts with both D55 Oδ1 and D58 Oδ1 and appears to be the nucleophilic water (*10*). These two residues are conserved in all family 9 cellulases, and when the corresponding residues were mutated in *C. thermocellum* CelD, both mutant enzymes had low activity (*31*). In *C. stercorarium* CelZ only the residue equivalent to D58 was mutated, which greatly reduced activity (*32*). Information about the CelZ mutant equivalent to D55 has not yet been published. In a review article only a single catalytic base was proposed for family 9 enzymes, the equivalent of D58 (*11*). We find no compelling evidence to single out either D55 or D58 as the more important catalytic residue. The double mutant D55A/D58A has more than an order of magnitude lower activity on CMC and SW than either of the single mutants. Although the D58 mutant enzymes have impaired binding to BMCC, indicating that D58 may be more involved in binding than in catalysis, they have wild-type binding to MUG3. The results presented in this paper show that the catalytic mechanism is more complicated than a single residue acting as the catalytic base.

Y206 is clearly also an important residue for catalytic activity; one of its functions is substrate binding, but hydrogen bonding to D55 is also significant. There is a H-bond from D58 Oδ2 to H125 Nϵ2 that may also be part of a catalytic charge network that stabilizes the interactions of D55 and D58 with the nucleophillic water. Although the role of H125 remains to be investigated, one can speculate that a hydrogen-bonding network between Y206, D55, D58, and H125 is important for catalytic function.

Processivity is not yet well understood. It appears that the ability of an enzyme to process is most important for crystalline cellulose substrates. The initial cleavage and release of the product are followed by a sliding of the substrate into position for the next cleavage. It must be energetically favorable for the cellulose molecule to slide from the Cel9A Glc(+1) and Glc(+2) sites into the full catalytic cleft, but at the same time the binding in sites Glc(−1) to Glc(−4) must not be so tight that the product is not released. It is likely that a number of residues are involved in this process including Y206 and Y318. The loss of processivity caused by the Y206 and Y318 mutations probably results from weaker binding of the cellulose chain to the Glc(−1) or Glc(−2) subsites, causing the weakly bound fragment produced by the initial endocellulolytic cleavage to dissociate rather than to slide into the Glc(−4) to Glc(−1) sites.

In a recent paper, Nerinckx et al. identified a hydrophobic residue that stabilizes the transition state in all glycoside hydrolases, and they propose that this function may be served by Y429 in conjunction with A426 in Cel9A (*34*). This tyrosine is oriented perpendicular to the G(−1) sugar ring, and it does not H-bond to the substrate. In *Trichoderma reesei* Cel6A (CBHII), this residue is Y169, and it is homologous to *T. fusca* Cel6A Y73. Barr et al. (*22*) also suggested that Y73 of Cel6A could assist catalysis by stabilizing the oxocarbonium ion intermediate and/ or by sugar ring distortion.

REFERENCES

1. Walker, L. P., and Wilson, D. B. (1991) Engineering cellulases, *Bioresour. Technol. 36*, 3−14.
2. Irwin, D., Cheng, M., Xiang, B., Rose, J. K. C., and Wilson, D. B. (2003) Cloning, expression and characterization of a family-74 xyloglucanase from *Thermobifida fusca*, *Eur. J. Biochem. 270*, 3083−3091.
3. Irwin, D., Leathers, T. D., Greene, R. V., and Wilson, D. B. (2003) Corn fiber hydrolysis by *Thermobifida fusca* extracellular enzymes, *Appl. Microbiol. Biotechnol. 61*, 352−358.
4. Lao, G., Ghangas, G. S., Jung, E. D., and Wilson, D. B. (1991) DNA sequences of three beta-1,4-endoglucanase genes from *Thermomonospora fusca*, *J. Bacteriol. 173*, 3397−3407.
5. Jung, E. D., Lao, G., Irwin, D., Barr, B. K., Benjamin, A., and Wilson, D. B. (1993) DNA sequences and expression in *Streptomyces lividans* of an exoglucanase gene and an endoglucanase gene from *Thermomonospora fusca*, *Appl. Environ. Microbiol. 59*, 3032−3043.
6. Zhang, S., Lao, G., and Wilson, D. B. (1995) Characterization of a *Thermomonospora fusca* exocellulase, *Biochemistry 34*, 3386−3395.
7. Irwin, D. C., Zhang, S., and Wilson, D. B. (2000) Cloning, expression and characterization of a family 48 exocellulase, Cel48A, from *Thermobifida fusca*, *Eur. J. Biochem. 267*, 4988−4997.
8. Irwin, A., Spezio, M., Walker, L. P., and Wilson, D. B. (1993) Activity studies of eight purified cellulases: specificity, synergism and binding domain effects, *Biotechnol. Bioeng. 42*, 1002−1013.
9. Irwin, D., Shin, D. H., Zhang, S., Barr, B. K., Sakon, J., Karplus, P. A., and Wilson, D. B. (1998) Roles of the catalytic domain and two cellulose binding domains of *Thermomonospora fusca* E4 in cellulose hydrolysis, *J. Bacteriol. 180*, 1709−1714.
10. Sakon, J., Irwin, D., Wilson, D. B., and Karplus, P. A. (1997) Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*, *Nat. Struct. Biol. 4*, 810−818.
11. Schulein, M. (2000) Protein engineering of cellulases, *Biochim. Biophys. Acta 1543*, 239−252.
12. Rye, C. S., and Withers, S. G. (2000) Glycosidase mechanisms, *Curr. Opin. Chem. Biol. 4*, 573−580.
13. Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids, *J. Mol. Biol. 166*, 557−580.
14. Escovar-Kousen, J. M., Wilson, D. B., and Irwin, D. (2004) Integration of computer modeling and site-directed mutagenesis, initial studies to improve cellulase activity on Cel9A from *Thermobifida fusca*, *Appl. Biochem. Biotechnol*. (submitted for publication).
15. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
16. Zhang, S., and Wilson, D. B. (1997) Surface residue mutations which change the substrate specificity of *Thermomonospora fusca* endoglucanase E2, *J. Biotechnol. 57*, 101−113.
17. Bohm, G., Muhr, R., and Jaenicke, R. (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks, *Protein Eng. 5*, 191−195.
18. Ghose, T. K. (1987) Measurement of cellulase activities, *Pure Appl. Chem. 59*, 257−268.
19. Kempton, J. B., and Withers, S. G. (1992) Mechanism of *Agrobacterium* beta-glucosidase: kinetic studies, *Biochemistry 31*, 9961−9969.
20. Chirico, W. J., and Brown, R. D. (1985) Separation of [1-^{3}H] cellooligosaccharides by thin-layer chromatography: assay for cellulolytic enzymes, *Anal. Biochem. 150*, 264−272.
21. Lever, M. (1972) A new reaction for colorimetric determination of carbohydrates, *Anal. Biochem. 47*, 273−279.
22. Barr, B. K., Wolfgang, D. E., Piens, K., Claeyssens, M., and Wilson, D. B. (1998) Active-site binding of glycosides by *Thermomonospora fusca* endocellulase E2, *Biochemistry 37*, 9220−9229.
23. De Boeck, H., Matta, K. L., Claeyssens, M., Sharon, N., and Loontiens, F. G. (1983) Binding of 4-methylumbelliferyl beta-D-galactosyl-1-3-*N*-acetyl-beta-D-galactosaminide to peanut agglutinin. Characterization and application in substitution titrations, *Eur. J. Biochem. 131*, 453−460.
24. André, G., Kanchanawong, P., Palma, R., Cho, H., Deng, X., Irwin, D., Himmel, M. E., Wilson, D. B., and Brady, J. W. (2003) Computational and experimental studies of the catalytic mechanism of *Thermobifida fusca* cellulase Cel6A (E2), *Protein Eng. 16*, 125−134.
25. Jung, H., Wilson, D. B., and Walker, L. P. (2002) Binding mechanisms for *Thermobifida fusca* Cel5A, Cel6B, and Cel48A cellulose-binding modules on bacterial microcrystalline cellulose, *Biotechnol. Bioeng. 80*, 380−392.
26. Bothwell, M., Daughhetee, S., Chaua, G., Wilson, D. B., and Walker, L. P. (1997) Binding capacities for *Thermomonospora fusca* E3, E4 and E5, the E3 binding domain, and *Trichoderma reesei* CBHI on avicel and bacterial microcrystalline cellulose, *Bioresour. Technol. 60*, 169−178.
27. Zhang, S., Barr, B. K., and Wilson, D. B. (2000) Effects of noncatalytic residue mutations on substrate specificity and ligand binding of *Thermobifida fusca* endocellulase cel6A, *Eur. J. Biochem. 267*, 244−252.
28. Kataeva, I. A., Seidel, R. D., III, Shah, A., West, L. T., Li, X. L., and Ljungdahl, L. G. (2002) The fibronectin type 3-like repeat from the *Clostridium thermocellum* cellobiohydrolase CbhA promotes hydrolysis of cellulose by modifying its surface, *Appl. Environ. Microbiol. 68*, 4292−4300.
29. Khademi, S., Guarino, L. A., Watanabe, H., Tokuda, G., and Meyer, E. F. (2002) Structure of an endoglucanase from termite, *Nasutitermes takasagoensis*, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 58*, 653−659.
30. Schubot, F. D., Kataeva, I. A., Chang, J., Shah, A. K., Ljungdahl, L. G., Rose, J. P., and Wang, B. C. (2004) Structural basis for the exocellulase activity of the cellobiohydrolase CbhA from *Clostridium thermocellum*, *Biochemistry 43*, 1163−1170.
31. Chauvaux, S., Beguin, P., and Aubert, J. P. (1992) Site-directed mutagenesis of essential carboxylic residues in *Clostridium thermocellum* endoglucanase CelD, *J. Biol. Chem. 267*, 4472−4478.
32. Riedel, K., and Bronnenmeier, K. (1999) Active-site mutations which change the substrate specificity of the *Clostridium stercorarium* cellulase CelZ: implications for synergism, *Eur. J. Biochem. 262*, 218−223.
33. Zechel, D. L., and Withers, S. G. (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst, *Acc. Chem. Res. 33*, 11−18.
34. Nerinckx, W., Desmet, T., and Claeyssens, M. (2003) A hydrophobic platform as a mechanistically relevant transition state stabilising factor appears to be present in the active centre of *all* glycoside hydrolases, *FEBS Lett. 538*, 1−7.

BI049394N