Comparative NMR Analysis of Cellooligosaccharide Hydrolysis by GH9 Bacterial and Plant Endo-1,4-β-glucanases[†]

Ulla J. Rudsander,[‡] Corine Sandstrom,[§] Kathleen Piens,[‡,||] Emma R. Master,[‡,⊥] David B. Wilson,[#] Harry Brumer III,[‡] Lennart Kenne,[§] and Tuula T. Teeri*[,‡]

KTH Biotechnology, Swedish Center for Biomimetic Fiber Engineering, AlbaNova, SE-10691 Stockholm, Sweden, Department of Chemistry, Swedish University of Agricultural Sciences, P.O. Box 7015, SE-750 07 Uppsala, Sweden, and Department of Molecular Biology and Genetics, Cornell University, 458 Biotechnology Building, Ithaca, New York 14853-2703

Received November 2, 2007; Revised Manuscript Received March 2, 2008

ABSTRACT: ^{1}H NMR spectroscopy has been used to analyze the product profiles arising from the hydrolysis of cellooligosaccharides by family GH9 cellulases. The product profiles obtained with the wild type and several active site mutants of a bacterial processive endoglucanase, *Tf* Cel9A, were compared with those obtained by a randomly acting plant endoglucanase, *Ptt*Cel9A. *Ptt*Cel9A is an orthologue of the *Arabidopsis* endocellulase, Korrigan, which is required for efficient cellulose biosynthesis. As expected, poplar *Ptt*Cel9A was shown to catalyze the degradation of cellooligosaccharides by inversion of the configuration of the anomeric carbon. The product analyses showed that the number of interactions between the glucose units of the substrate and the aromatic residues in the enzyme active sites determines the point of cleavage in both enzymes.

Thermobifida fusca cellulase Cel9A (*Tf*Cel9A) (*1*) and a homologous enzyme *Ptt*Cel9A from *Populus tremula x tremuloides* (*2*) represent two extremes of the family GH9 endo-1,4-β-glucanases. *Tf*Cel9A is an efficient, processive endoglucanase (*2*), whereas *Ptt*Cel9A is a randomly acting, membrane-bound endoglucanase (*2*). *Tf*Cel9A is a well-studied member of the family GH9 cellulases, and its X-ray three-dimensional structure has been solved in complex with a substrate (*1*). Site-directed mutagenesis has been used previously to determine the roles of a number of key catalytic residues in this enzyme (*2, 3*). An analysis of the enzymes described in the carbohydrate active enzymes database (www.CAZy.org) shows that most plant endo-1,4-β-glucanases belong in the family GH9 while just a few have been identified in family GH5. *Ptt*Cel9A is homologous to the *Arabidopsis* Korrigan, another family 9 cellulase that has been shown to be required for cellulose biosynthesis (*4*). No three-dimensional structures are available for these enzymes, but the substrate specificity (*2, 5–8*) and kinetic constants (*2*) have been determined for a handful of plant GH9 endo-1,4-β-glucanases. However, many questions remain concerning the substrate specificities and roles of the wide variety of plant enzymes in this family.

Specific interactions between carbohydrates and aromatic amino acid residues in the enzyme hydrolyzing them were originally observed by analyzing the three-dimensional structure of lysozyme (*9*). The nonpolar patches of both faces of the pyranose rings can be involved in aromatic stacking interactions with nonpolar amino acid side chains. Later, this type of interactions has been established as a common feature in carbohydrate recognition (*10*). Since long polymeric and especially insoluble substrates are difficult to use for detailed structure–function studies of enzymes, short oligosaccharides have been commonly used as model substrates for structural and biochemical studies of the active sites of cellulases and other carbohydrate active enzymes (*11–16*). The catalytic domain of *Tf*Cel9A has six glucosyl binding sites (Figure 1). The tryptophan residues W256, W209, and W313 in the active site of *Tf*Cel9A form direct interactions with the glycosyl residues of the substrate at the positions −4, −3, and −2, respectively, while Y420 and H125 perform aromatic stacking interactions with the substrate at subsite +1 (*1*). The positions corresponding to W256, W313, and W209 in *Tf*Cel9A are occupied by glycine, proline, and threonine, respectively, in *Ptt*Cel9A (*2*). We have shown before that *Ptt*Cel9A has a catalytic rate on cellohexaose that is over 2 orders of magnitude lower than that for *Tf*Cel9A and that the removal of these aromatic platforms one by one from the active site of *Tf*Cel9A leads to significantly decreased catalytic rate (*2*). In the present study, the effect of these aromatic platforms on the binding and hydrolysis of short cellooligosaccharides by *Tf*Cel9A, its three active site mutants (W313G, W256A, and W209S), and *Ptt*Cel9A was further investigated using ^{1}H NMR spectroscopy. Since both enzymes operate by the inverting reaction mechanism, the newly formed products carry an anomeric carbon in the α configuration. By using ^{1}H NMR spectroscopy, these newly

[†] This work was supported by the Swedish Research Council (VR).

* To whom correspondence should be addressed. E-mail: tuula@biotech.kth.se. Tel.: +46-8-790 8794. Fax: +46-8-5537 4868.

[‡] Swedish Center for Biomimetic Fiber Engineering.

[§] Swedish University of Agricultural Sciences.

[||] Present address: Laboratory for Protein Biochemistry and Biomolecular Engineering, Department of Biochemistry, Physiology and Microbiology, University of Ghent, K.L.-Ledeganckstraat 35, 9000 Ghent, Belgium.

[⊥] Present address: Department of Chemical Engineering and Applied Chemistry, University of Toronto, 200 College St., WB420C Toronto, Ontario, Canada M5S 3E5.

[#] Cornell University.

10.1021/bi702193e CCC: $40.75
Published on Web 04/11/2008

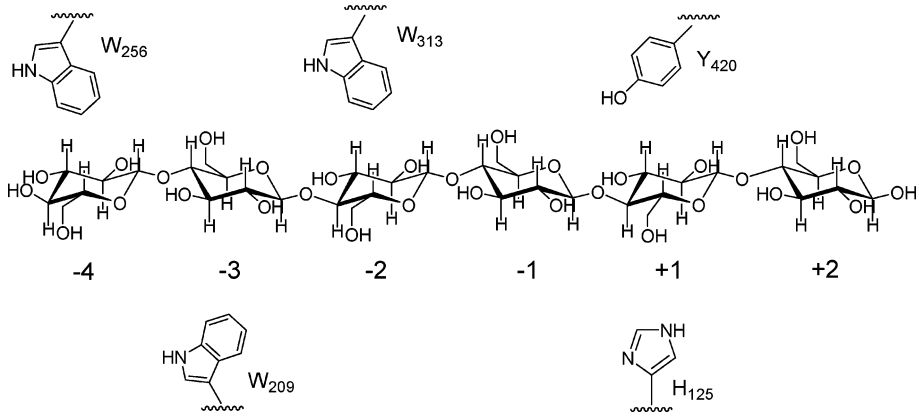

FIGURE 1: Schematic representation of the active site of *Tf*Cel9A. The subsites are numbered using the position of cleavage of the glycosidic bond at −1/+1 as the reference point. The sites are numbered by increasing negative integers toward the nonreducing end of a carbohydrate chain and by increasing positive integers toward the reducing end. Aromatic stacking platforms are shown in subsites −4 to +1 for a cellohexaose substrate.

formed products can be identified and quantified, thus permitting detailed analysis of the active site function of the enzymes.

EXPERIMENTAL PROCEDURES

*Enzyme Preparations. Ptt*Cel9A, *Tf*Cel9A, and the mutant enzymes were expressed and purified as described (*2*). For the proton NMR experiments, buffer exchange was performed by dialysis in deuterated 50 mM sodium phosphate buffer, pH 6.0. All reactions were performed in 30 mM $CaCl_2$. A total sample solution of 0.1 mL was used, and enzymatic hydrolysis was carried out in the NMR tube.

^{1}H NMR Spectroscopy. ^{1}H NMR spectra were recorded on a Bruker DRX 600 spectrometer operating at 600.13 MHz for proton observation and using a 2.5 mm inverse-detection microprobe equipped with a *z*-gradient. The spectra were acquired at 27 °C with the HDO[1] signal as internal reference (δ 4.70 ppm). Prior to the enzymatic reaction, a reference spectrum of the substrate oligosaccharide was acquired. The sample was then removed from the probe, enzyme was added, the tube was placed into the probe, and hydrolysis was monitored as a function of time. The first spectrum was acquired within 10 min after the reaction began, and from then on spectra were acquired every 3 or 6 min depending on the substrate concentration. A sweep width of 6000 Hz was used, and 16 or 32 FID's of 64 K data points were acquired with a recycle delay of 5 s. The enzyme and substrate concentrations were chosen so that the reaction was fast enough to allow determination of the degradation patterns before mutarotation occurred. Proton signals were integrated to determine the concentrations of the substrates and products as a function of time.

Enzyme Activities by High-Performance Anion-Exchange Chromatography (HPAEC) with Pulsed Amperometric Detection (PAD). Enzyme activities were also assayed by HPAEC-PAD using a CarboPac-100 column (Dionex), as described earlier (*2*) except that the formation of G_5 from the hydrolysis of G_6 by *Tf*Cel9A and the W209S mutant was detected after 1:10 dilution of the reaction products, using six data points. In the case of the mutant W313G, no G_2 was formed from G_5 hydrolysis. Instead, formation of G_4 was detected after 1:10 dilution of the reaction products, using four data points. The injection volume was 10 μL in each case, and the reaction products were quantified relative to known standards.

RESULTS

Terminology of Protons and Assignment of NMR Resonances. The proton NMR spectra of the oligosaccharides cellotetraose (G_4), cellopentaose (G_5), and cellohexaose (G_6) were assigned as described (*17*). The resonances of the anomeric protons lie in a well-separated region between δ 4.5 and 5.3 ppm and can be divided into three different categories: reducing end, internal, and terminal protons (Figure S1, Supporting Information). In solution, the α- and β-anomeric forms of the reducing end are in equilibrium with 35% as α-anomers and 65% as β-anomers. The reducing end anomeric α protons in G_1-G_6 on one hand and the reducing end β protons in G_2-G_6 on the other hand have the same chemical shift values (δ 5.224 and 4.662 ppm, respectively; Figure S2a, Supporting Information). Only H-1 of β-G_1 gives a separated signal at δ 4.639 ppm. The internal and terminal H-1 signals have chemical shifts at δ 4.536 and 4.511 ppm, respectively. Signals from H-3 to H-6 lie in a crowded region, whereas the H-2 signals can be useful as reporter signals (Figure S1, Supporting Information). The anomeric configuration of the oligosaccharide also affects the H-2 chemical shifts (Figure S2b, Supporting Information). For example, H-2 of β-G_1 gives a separated signal at δ 3.238 ppm, and H-2 of α-G_1 is observed in the NMR spectra by a signal at δ 3.528 ppm. The α-G_2 terminal H-2 has a signal at δ 3.323 ppm, and this can be distinguished from other reducing end and terminal signals. The oligosaccharides G_3-G_6 give signals at δ 3.368 ppm caused by their H-2 in the internal residue adjacent to the reducing end in the α-form. This signal is shifted slightly more downfield than other internal H-2 signals.

Since the chemical shifts of the α- and β-anomeric protons of the reducing ends are well separated, these signals can be used to monitor the stereochemical course of the reaction. The only requirement is that the enzyme-catalyzed reaction

[1] Abbreviations: FID, free induction decay; G_1, glucose; G_2, cellobiose; G_3, cellotriose; G_4, cellotetraose; G_5, cellopentaose; G_6, cellohexaose; GH, glycosyl hydrolase; HDO, natural abundance of ^{2}H NMR signal arising from deuteriums in water; HPAEC, high-performance anion-exchange chromatography; PAD, pulsed amperometric detection.

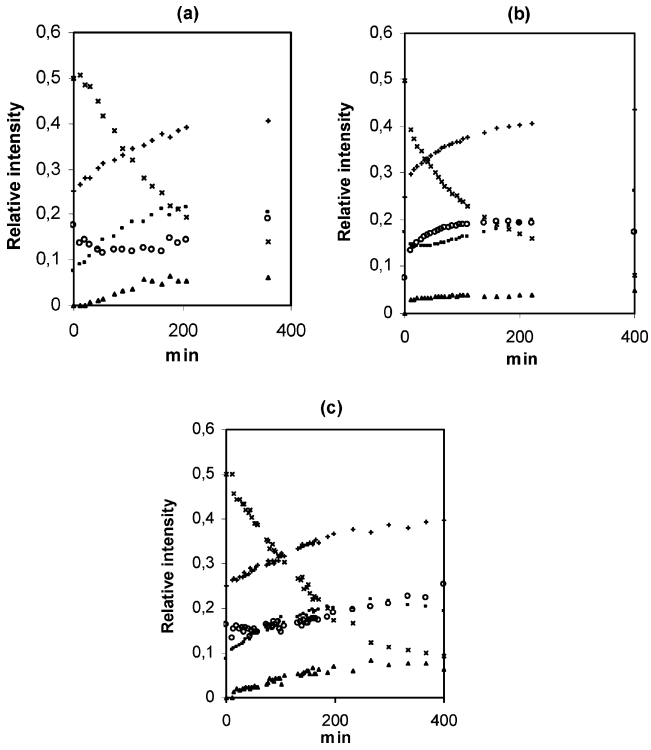

FIGURE 2: Experimental progress curves for the anomeric proton resonances in the hydrolysis of cellotetraose catalyzed by (a) wild-type *Tf*Cel9A, 80 ng/μL, and 0.5 mM G_4, (b) W256A, 40 ng/μL, and 2 mM G_4, and (c) W209S, 400 ng/μL, and 2 mM G_4. The anomeric proton resonances are labeled as follows: (+) terminal, (×) internal, (○) reducing end β (G > 1), (■) reducing end α, and (▲) β-glucose.

is faster than the mutarotation. For *Ptt*Cel9A, which is a slow enzyme, this limitation was overcome by using a high enzyme concentration.

For each cleaved oligosaccharide molecule, a new reducing end is created. If the enzyme has a retaining mechanism, the concentration of reducing end β-anomeric protons increases. If, however, the mechanism is inverting, the concentration of reducing end α-anomeric protons will increase (*18*). The enzymatic degradation of G_4, G_5, and G_6 was followed from a plot of the concentration of H-1 protons as a function of time (*vide infra*) (Figures 2−4). Addition of each enzyme resulted in an initial fast increase in the concentration of the α-anomeric protons whereas the curves for the β-anomeric protons were initially unchanged (Figure S1, Supporting Information). The stereochemistry of the cleaved glucosidic linkages is therefore inverted. After 100 min, the effect of mutarotation is apparent in the progress curves.

Cleavage Modes of TfCel9A. The *Tf*Cel9A enzyme produced about double amounts of G_2 than amounts of G_1 and G_3 from the hydrolysis of G_4, showing that this substrate binds equally well in the −3 to +1 sites as in the −2 to +2 sites (Table 1). G_5 binds with higher affinity to sites −4 to +1 to release α-G_4 (75% in Table 1) than to sites −3 to +2 from which α-G_3 is formed (25% in Table 1). The major binding of G_6 was to the −4 to +2 sites, producing α-G_4 and G_2 (Table 1). The formation of β-G_1 during the first fast phase (Figure 4a) indicates that the initial reaction also includes hydrolysis of G_6 to α-G_5 and G_1. If the G_1 formed initially resulted from hydrolysis of G_4, it should be in the α-form. Another possibility would be that G_1 originates from

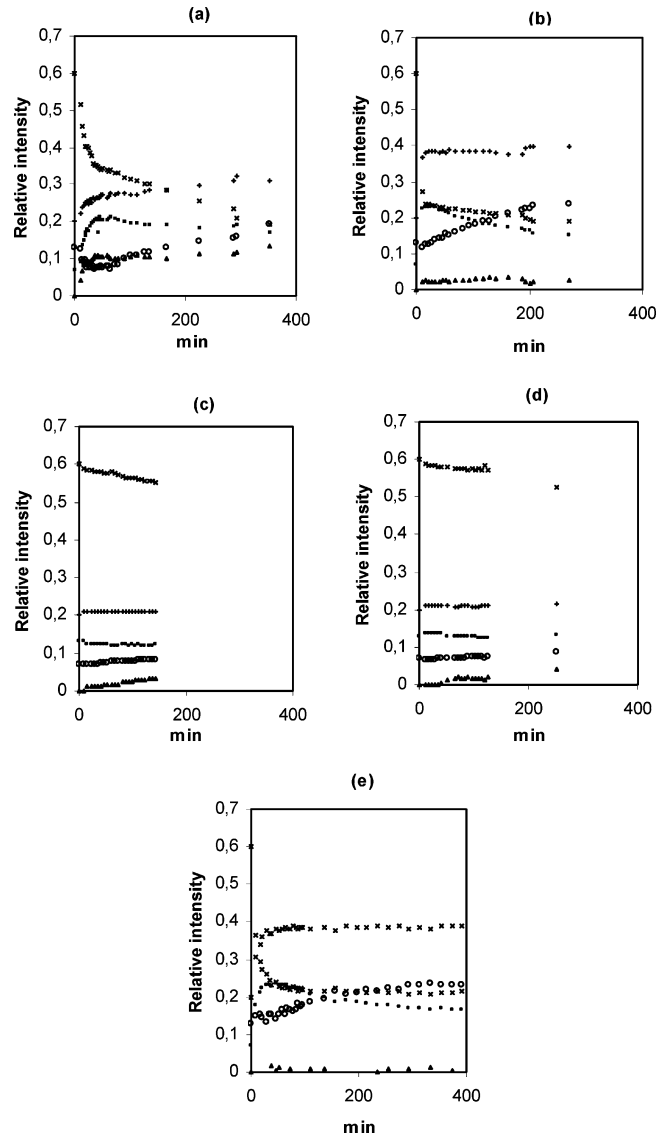

FIGURE 3: Experimental progress curves for the anomeric proton resonances in the hydrolysis of cellopentaose catalyzed by (a) wild-type *Tf*Cel9A, 80 ng/μL, and 0.5 mM G_5, (b) W256A, 40 ng/μL, and 2 mM G_5, (c) W209S, 80 ng/μL, and 2 mM G_5, (d) W313G, 100 ng/μL, and 2 mM G_5, and (e) *Ptt*Cel9A, 4000 ng/μL, and 2 mM G_5. The anomeric proton resonances are labeled as follows: (+) terminal, (×) internal, (○) reducing end β (G > 1), (■) reducing end α, and (▲) β-glucose.

cleavage of G_5, which is present as a 10% contamination in the G_6 substrate. However, an increase in G_5 concentration was observed in HPAEC-PAD experiments with a specific activity of 0.47 ± 0.15 μmol mg^{-1} min^{-1} (data not shown).

Cleavage Modes of the W256A Enzyme. Hydrolysis of G_4 by this enzyme leads to nearly twice as much G_2 than G_1 and G_3, indicating that binding in the −3 to +1 and −2 to +2 sites are equally favored (Table 1). This enzyme hydrolyzed G_5 to give mainly α-G_3 and G_2 as seen from the H-2 resonances in the NMR spectra (Figure S3b,c, Supporting Information): After 10 min, the H-2 internal region shows mainly signals from internal H-2s adjacent to the α-anomeric residue, indicating that mainly α-G_3 is formed. The predominant binding mode of G_6 in the W256A enzyme is at the −4 to +2 sites, producing α-G_4 and G_2 (Table 1). No G_1 formation was observed in the initial phase of hydrolysis. However, after 202 min, G_1 started to appear in the progress

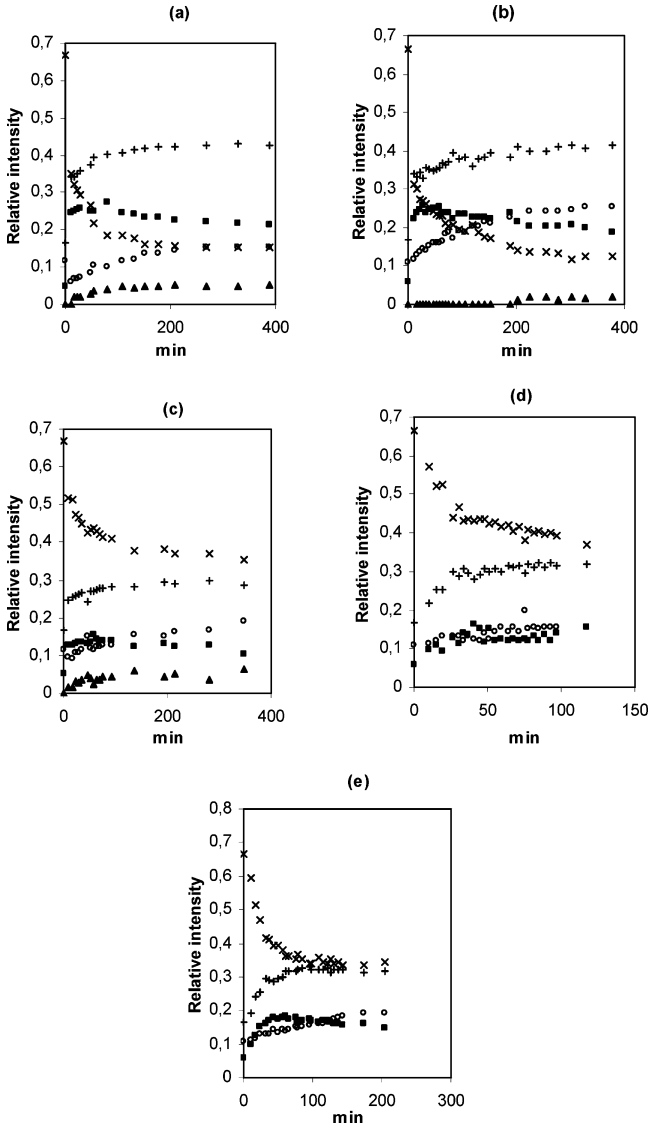

FIGURE 4: Experimental progress curves for the anomeric proton resonances in the hydrolysis of cellohexaose catalyzed by (a) wild-type *Tf*Cel9A, 20 ng/μL, and 0.5 mM G_6, (b) W256A, 40 ng/μL, and 1 mM G_6, (c) W209S, 80 ng/μL, and 0.5 mM G_6, (d) W313G, 80 ng/μL, and 0.5 mM G_6, and (e) *Ptt*Cel9A, 260 ng/μL, and 2 mM G_6. Anomeric proton resonances are labeled as follows: (+) terminal, (×) internal, (○) reducing end β (G > 1), (■) reducing end α, and (▲) β-glucose.

curve, most likely resulting from cleavage of G_4 (Figure 3b). Since G_4 is formed in only small amounts and since G_3 and G_2 are not substrates for the enzyme, no large changes are seen in the intensities of the signals of the internal and terminal protons after no G_5 remained.

Cleavage Modes of the W209S Enzyme. In the W209S enzyme loss of W209 in the −3 subsite did not affect the preferred binding modes of G_4 compared to those of *Tf*Cel9A, but it did affect the binding of G_5. Hydrolysis of G_4 gave nearly equal amounts of α-G_2 and α-G_3 (Table 1). This enzyme also hydrolyzed G_3, since the internal H-1 signal totally disappeared in the NMR spectrum of G_4 hydrolysis after 24 h (data not shown). This result was confirmed by HPAEC-PAD analysis. No signal from the H-2 proton of terminal α-G_2 was observed in the NMR spectra from hydrolysis of G_5 (Figure S4c, Supporting Information), indicating that G_3 and G_2 were not released as only G_4 with inversion of configuration at the reducing end residue was

Table 1: Favored Binding Modes of Wild-Type *Ptt*Cel9A, Wild-Type *Tf*Cel9A, and Mutant Enzymes W256A, W209S, and W313G for Cleavage of Substrates G_4, G_5, and G_6 after <30 min of Hydrolysis[a]

		subsite							
		−256	−209	−313					
enzyme	−5	−4	−3	−2	−1	1	2	3	pattern (%)
G_4									
*Tf*Cel9A			**G**	**G**	G-α	G			40
				G	G-α	G	G		60
*Tf*Cel9A$_{W256A}$			**G**	**G**	G-α	G			45
				G	G-α	G	G		55
*Tf*Cel9A$_{W209S}$			G	**G**	G-α	G			45
				G	G-α	G	G		55
G_5									
*Tf*Cel9A		**G**	**G**	**G**	G-α	G			75
			G	**G**	G-α	G	G		25
*Tf*Cel9A$_{W256A}$		G	**G**	**G**	G-α	G			18
			G	**G**	G-α	G	G		82
*Tf*Cel9A$_{W209S}$		**G**	G	**G**	G-α	G			100
*Tf*Cel9A$_{W313G}$		**G**	**G**	G	G-α	G			100
*Ptt*Cel9A			G	G	G-α	G	G		95
		G	G	G	G-α	G			5
G_6									
*Tf*Cel9A	G	**G**	**G**	**G**	G-α	G			14
		G	**G**	**G**	G-α	G	G		58
			G	**G**	G-α	G	G	G	28
*Tf*Cel9A$_{W256A}$		G	**G**	**G**	G-α	G	G		80
			G	**G**	G-α	G	G	G	20
*Tf*Cel9A$_{W209S}$		**G**	G	**G**	G-α	G	G		80
	G	**G**	G	**G**	G-α	G			20
*Tf*Cel9A$_{W313G}$		**G**	**G**	G	G-α	G	G		100
*Ptt*Cel9A		G	G	G	G-α	G	G		60
			G	G	G-α	G	G	G	40

[a] The experimental error is 10%, and the number of data points for pattern percentage numbers = 3. The glycosyl residues interacting with tryptophans are represented as **G**.

formed. Examination of the H-2 region in the ^{1}H NMR spectra showed that G_3 was not formed upon hydrolysis of G_6, confirming the result from a previous experiment analyzed by HPAEC-PAD (*2*). The major binding of G_6 to this enzyme was in the −4 to +2 sites (Table 1), producing α-G_4 and G_2. Some G_1 was detected, similar to the amounts detected in the hydrolysis of G_6 by the wild-type *Tf*Cel9A. The specific activity for $G_1 + G_5$ formation by the W209S enzyme, as determined by HPAEC-PAD analysis, was 1.17 ± 0.17 μmol mg^{-1} min^{-1} (data not shown).

Cleavage Modes of the W313G Enzyme. Hydrolysis of G_4 by this enzyme was not detected after overnight hydrolysis at 600 ng/μL enzyme concentration. During hydrolysis of G_5, no signal from the H-2 proton of terminal α-G_2 was observed (Figure S4b, Supporting Information), indicating that G_4 and G_1 were the only products formed. A larger amount of nonhydrolyzed G_5 was found with the W313G enzyme than with the W209S enzyme after 400 min (37% versus 15%). The W313G enzyme produced only G_4 and G_2 from hydrolysis of G_6 (Table 1). The relative intensity of the signal of internal H-2 adjacent to the α-reducing end showed that the formation of G_3 could be excluded or was produced in very small amount. No hydrolysis of G_4 was detected after overnight incubation at the highest enzyme concentration in the progress curve analysis. However, previous biochemical analysis indicates degradation of G_4 with specific activity that is 5-fold lower than that of the wild-type enzyme (Table 2).

Cleavage Modes of PttCel9A. In a previous study, G_4 was found not to be a substrate for *Ptt*Cel9A (*2*). In the current

Table 2: Specific Activities of G_2 Product Formation for *Tf*Cel9A Wild Type and Mutants for Different Oligosaccharides[a]

enzyme	activity on 600 μM G4 (μmol mg^{-1} min^{-1})	activity on 600 μM G_5 (μmol mg^{-1} min^{-1}) ($n = 2$)[b]	activity on 600 μM G_6 (μmol mg^{-1} min^{-1})
*Tf*Cel9A	0.28 ± 0.01	18 ± 4	603 ± 57
*Tf*Cel9A$_{W256A}$	1.1 ± 0.1	37 ± 2	28 ± 3
*Tf*Cel9A$_{W209S}$	0.29 ± 0.06	nd	131 ± 11
*Tf*Cel9A$_{W313G}$	0.053 ± 0.007	nd	7.4 ± 1

[a] Values for G_4 and G_6 are taken from ref *2*. [b] n = number of repeated experiments.

NMR experiment, a large amount of enzyme was necessary to observe the hydrolysis products of G_5 before mutarotation occurred. α-G_3 and G_2 were produced as major products (Table 1). A small amount of G_1 (~5%) was detected in the NMR spectra, indicating that α-G_4 was also formed (Table 1). Hydrolysis of G_6 proceeded until it was depleted, but no G_1 or α-G_2 was observed in the NMR spectra. The ratio of internal H-2 adjacent to the reducing end in the α-form relative to other internal H-2 showed that α-G_3 plus α/β-G_3 and α-G_4 plus G_2 were formed. This indicates that the preferred binding mode of G_6 was to binding sites −4 to +2 (Table 1). In a previous study where HPAEC-PAD was used for analysis of the products (*2*), more G_4 (45%) than G_2 (29%) was formed. Broadened resonance lines in the ^{1}H NMR spectra prevented determination of the relative amounts of G_2 and G_4, but the HPAEC-PAD experiments showed that G_2 and G_4 resulting from cleavage of G_6 were produced in equal amounts (data not shown).

Kinetic Analyses. The enzyme and substrate concentrations for the ^{1}H NMR analyses were chosen individually for each enzyme to be able to detect the product profiles before mutarotation occurred. Because of the different reaction conditions, the rate of the product formation observed in the progress curves did not allow direct comparison of catalytic rates. However, specific activities of *Tf*Cel9A and the mutant enzymes W256A, W209S, and W313G were determined on 600 μM G_5 and compared with previously obtained data for the same enzymes on G_4 and G_6 (*2*) (Table 2). The data obtained show that the W256A mutant has a higher activity than the wild-type *Tf*Cel9A on both G_4 and G_5 but not on G_6.

DISCUSSION

^{1}H NMR spectroscopy is a valuable technique for studying enzymatic cleavage patterns, since it does not destroy the enzyme and unmodified, natural substrates can be used. The NMR data hereby obtained confirm that, similar to other enzymes in family GH9, both *Tf*Cel9A and *Ptt*Cel9A hydrolyze glycosidic linkages by inversion of the configuration of the anomeric carbon. In attempts to understand why a cellulase might be needed for cellulose synthesis, it has been suggested that they might synthesize glycosidic bonds by way of transferase activity (*19, 20*). However, since enzymes employing the inverting reaction mechanism do not catalyze transglycosylation reactions (*21*), this possibility can now be ruled out at least for the family 9 plant endoglucanases.

The correlation between the presence of tryptophan residues at subsites −4, −3, and −2 and the ability to degrade short oligosaccharides was investigated by comparing the cleavage patterns from *Tf*Cel9A, the W313G, W256A, W209S enzymes, and *Ptt*Cel9A. Mutation of *Tf*Cel9A at subsites −4, −3, or −2 was shown to alter the position for cleavage of oligosaccharides. The preferred binding and degradation pattern for the mutant enzymes was correlated with the maximum number of interactions between the glycosyl residues in the substrate and the tryptophan residues in the enzyme active site (Table 1).

Wild-type *Tf*Cel9A was shown to convert a small amount of G_6 to α-G_5 and G_1, whereas the main product formed was α-G_4. Formation of glucose requires binding of a glucosyl residue at the −5 subsite. The crystal structure of *Tf*Cel9A suggests that there may be "insufficient room for a hypothetical G(−5) binding" (*1*). Therefore, the cellooligosaccharide may have to bend in order to fit into the active site cleft in this binding mode. Conversion of G_6 to α-G_5 and G_1 was also seen with the W209S mutant of *Tf*Cel9A, both by NMR and by HPAEC-PAD analysis.

The mutation W256A changed the cleavage patterns of G_5 and G_6 compared to those seen for wild-type *Tf*Cel9A. Hydrophobic interactions via at least two tryptophan residues thus seem to be necessary for productive binding. With wild-type *Tf*Cel9A, G_5 was mainly hydrolyzed to α-G_4 and G_1, while the W256A mutant produced exclusively α-G_3 and G_2. When *Tf*Cel9A binds the substrate over subsites −4 to +1, interactions with three tryptophan residues are possible. However, substrates binding over subsites −3 to +2 can only interact with two tryptophan residues. When the number of glucosyl−tryptophan interactions is the same for two binding modes, as for the W256A mutant binding G_5 (Table 1), it appears that release of G_2 from the reducing end is preferred over release of G_1. The mutation W256A increased the specific activity of *Tf*Cel9A for both G_5 and G_4 but not for G_6 (Table 2). An intact W256 might thus trap G_4 in a nonproductive binding mode. Hydrolysis of G_5 by the W256A mutant is more efficient than by the wild-type enzyme because substrate binding over subsites −3 to +2 is more favorable for hydrolysis than binding over subsites −4 to +1.

The mutations W313G and W209S had the greatest effect on the hydrolysis of G_5 and G_6. For example, the W209S and W313G mutants produced only α-G_4 and G_1 from G_5, indicating binding in the −4 to +1 subsites that allows interactions between two glucosyl and two tryptophan residues. No α-G_3 and G_2 was produced, showing that binding did not occur in the −3 to +2 subsites, where only one tryptophan interaction is possible with the substrate. In a similar manner, no α-G_3 was produced by these two enzymes upon degradation of G_6 since the required binding gave only one glucosyl−tryptophan interaction. The binding affinity of the cellotetraose analogue, MUG-G3, was shown to be drastically reduced by these mutations, which is consistent with the present results (*3*). In a previous study, *Tf*Cel9A Y318A and Y318F −3 subsite mutations caused less G_4 and more G_2 and G_3 to be produced from hydrolysis of G_5, suggesting that in this case binding in the −3 to +2 subsites was favored (*2*). However, residue Y318 does not form an aromatic stacking interaction but instead has two hydrogen bonds to the −3 subsite and one to the −2 subsite.

The ^{1}H NMR results showed that the W209S mutant can degrade G_3, and this was confirmed by HPAEC-PAD (data not shown). Activity on G_3 is not observed for wild-type

*Tf*Cel9A or for any other subsite mutant. This activity can be explained if G_3 binds nonproductively into the −3 to −1 subsites when W209 is present but productively when the residue in mutated. *T. fusca* Cel9B (*Tf*Cel9B) is a true, randomly acting GH9 endoglucanase (*2*). It is also able to cleave G_3 and lacks aromatic platforms in both the −4 and the −3 subsites (*22*).

The mutation W313G in the −2 subsite had the highest influence on the hydrolysis rates of G_4−G_6. This enzyme does not produce G_2 from G_5, but the specific activity measured for the formation of G_4 was 0.35 ± 0.03 μmol mg^{-1} min^{-1}, which is 2 orders of magnitudes lower than that measured in the same conditions for the wild-type *Tf*Cel9A (30 ± 3 μmol mg^{-1} min^{-1}). For G_6, the decrease in specific activity as measured for the production of G_2 was 90-fold compared to wild-type *Tf*Cel9a (Table 2). The −2 subsite seems to contribute high affinity to the substrate even in other glycosyl hydrolases such as barley α-amylases 1 and 2 (*23*) and a fungal cellobiohydrolase, TrCel6A (*24*). In TrCel6A, residue W135 performs a stacking interaction with the glycosyl ring similar to residue W313 in *Tf*Cel9A. Similar to our present data, the mutations W135F and W135L at the −2 subsite of TrCel6A had catalytic turnover constants for G_4 that were reduced 3-fold and 200-fold, respectively, compared to the wild-type enzyme.

Aromatic stacking interactions are well conserved in many GH9 enzymes. Here we show that the presence of a tryptophan residue in the −2 subsite is a key determinant governing multiple binding modes and fast hydrolysis. The corresponding residue is strictly conserved in all bacterial and animal GH9 sequences so far analyzed (*25, 26*). Inspection of the fully sequenced genome of *Populus trichocarpa* (*27*) contains 26 highly homologous full-length GH9 gene models (J. Takahashi and E. Mellerowicz, manuscript in preparation), but only a fraction of these gene models encode tryptophans in the −2 subsites of the corresponding enzyme. The *Populus* EST database (http://poppel.fysbot.umu.se/) holds over 100000 ESTs and covers all major tissues of the plant. Analysis of the EST sequences corresponding to the family GH9 cellulases containing a −2 tryptophan reveals that these enzymes are primarily expressed in tissues in which the metabolism is shifted toward cellulose degradation, as in flower or fruit tissues (*28, 29*). All of these genes also encode a putative carbohydrate binding module, further suggesting that these enzymes indeed degrade cellulose. However, the majority of the plant family GH9 endoglucanases have no apparent tryptophan residues in the −2 subsite (Figure S5, Supporting Information). Interestingly, these enzymes are primarily expressed during cellulose biosynthesis or rapid cell elongation. It thus seems that enzymes participating in cellulose synthesis have deliberately evolved low rates of cellulose hydrolysis by dropping a key tryptophan in their active site.

ACKNOWLEDGMENT

We thank Weilin Zhou for kindly providing the wild-type enzyme *Tf*Cel9A and the mutant *Tf*Cel9A enzymes W256A, W209S, and W313G.

SUPPORTING INFORMATION AVAILABLE

Figure S1, proton NMR spectra of unhydrolyzed and hydrolyzed cellohexaose; Figure S2, chemical shifts for H-1 and H-2; Figure S3, proton NMR spectra for the H-2 region of cellohexaose; Figure S4, proton NMR spectra for the H-2 region of cellopentaose; Figure S5, sequence alignment of *Populus trichocarpa* gene models. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. 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.
2. Master, E. R., Rudsander, U. J., Zhou, W. L., Henriksson, H., Divne, C., Denman, S., Wilson, D. B., and Teeri, T. T. (2004) Recombinant expression and enzymatic characterization of *Ptt*Cel9A, a KOR homologue from *Populus tremula x tremuloides*. *Biochemistry 43*, 10080–10089.
3. Li, Y. C., Irwin, D. C., and Wilson, D. B. (2007) Processivity, substrate binding, and mechanism of cellulose hydrolysis by *Thermobifida fusca* Ce19A. *Appl. Environ. Microbiol. 73*, 3165–3172.
4. Nicol, F., His, I., Jauneau, A., Vernhettes, S., Canut, H., and Hofte, H. (1998) A plasma membrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in *Arabidopsis*. *EMBO J. 17*, 5563–5576.
5. Ferrarese, L., Trainotti, L., Gattolin, S., and Casadoro, G. (1998) Secretion, purification and activity of two recombinant pepper endo-beta-1,4-glucanases expressed in the yeast *Pichia pastoris*. *FEBS Lett. 422*, 23–26.
6. Molhoj, M., Ulvskov, P., and Dal Degan, F. (2001) Characterization of a functional soluble form of a Brassica napus membrane-anchored endo-1,4-beta-glucanase heterologously expressed in *Pichia pastoris*. *Plant Physiol. 127*, 674–684.
7. Woolley, L. C., James, D. J., and Manning, K. (2001) Purification and properties of an endo-beta-1,4-glucanase from strawberry and down-regulation of the corresponding gene, cel1. *Planta 214*, 11–21.
8. Nakamura, S., and Hayashi, T. (1993) Purification and properties of an extracellular endo-1,4-β-glucanase from suspension-cultured poplar cells. *Plant Cell Physiol. 34*, 1009–1013.
9. Blake, C. C. F., Koening, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. (1965) Structure of hen egg white lysozyme. A three-dimensional Fourier synthesis at 2.0 Å resolution. *Nature 206*, 757–761.
10. Vyas, N. K. (1991) Atomic features of protein-carbohydrate interactions. *Curr. Opin. Struct. Biol. 1*, 732–740.
11. Saura-Valls, M., Faure, R., Ragas, S., Piens, K., Brumer, H., Teeri, T. T., Cottaz, S., Driguez, H., and Planas, A. (2006) Kinetic analysis using low-molecular mass xyloglucan oligosaccharides defines the catalytic mechanism of a *Populus* xyloglucan endotransglycosylase. *Biochem. J. 395*, 99–106.
12. Sandgren, M., Berglund, G. I., Shaw, A., Stahlberg, J., Kenne, L., Desmet, T., and Mitchinson, C. (2004) Crystal complex structures reveal how substrate is bound in the −4 to the +2 binding sites of *Humicola grisea* cel12A. *J. Mol. Biol. 342*, 1505–1517.
13. Zolotnitsky, G., Cogan, U., Adir, N., Solomon, V., Shoham, G., and Shoham, Y. (2004) Mapping glycoside hydrolase substrate subsites by isothermal titration calorimetry. *Proc. Natl. Acad. Sci. U.S.A. 101*, 11275–11280.
14. Zou, J.-y., Kleywegt, G. J., Ståhlberg, J., Driguez, H., Nerinckx, W., Claeyssens, M., Koivula, A., Teeri, T. T., and Jones, T. A. (1999) Crystallographic evidence for substrate ring distortion and protein conformational changes during catalysis in cellobiohydrolase Cel6A from *Trichoderma reesei*. *Structure 7*, 1035–1045.
15. 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.
16. Biely, P., Vrsanska, M., and Claeyssens, M. (1991) The endo-1,4-beta-glucanase-I from *Trichoderma reesei*—Action on beta-1,4-oligomers and polymers derived from D-glucose and D-xylose. *Eur. J. Biochem. 200*, 157–163.
17. Harjunpaa, V., Teleman, A., Koivula, A., Ruohonen, L., Teeri, T. T., Teleman, O., and Drakenberg, T. (1996) Cello-oligosaccharide hydrolysis by cellobiohydrolase II from *Trichoderma reesei*. Association and rate constants derived from an analysis of progress curves. *Eur. J. Biochem. 240*, 584–591.
18. Sinnott, M. L. (1990) Catalytic mechanisms of enzymatic glycosyl transfer. *Chem. Rev. 90*, 1171–1202.

19. Brummell, D. A., Catala, C., Lashbrook, C. C., and Bennett, A. B. (1997) A membrane-anchored E-type endo-1,4-beta-glucanase is localized on Golgi and plasma membranes of higher plants. *Proc. Natl. Acad. Sci. U.S.A. 94*, 4794–4799.
20. Matthysse, A. G., White, S., and Lightfoot, R. (1995) Genes required for cellulose synthesis in *Agrobacterium tumefaciens*. *J. Bacteriol. 177*, 1069–1075.
21. Koshland, D. E. (1953) *Biol. Rev. 28*, 416–436.
22. Lin, E., and Wilson, D. B. (1988) Identification of a cele-binding protein and its potential role in induction of the cele gene in *Thermomonospora fusca*. *J. Bacteriol. 170*, 3843–3846.
23. Kandra, L., Abou Hachem, M., Gyemant, G., Kramhoft, B., and Svensson, B. (2006) Mapping of barley alpha-amylases and outer subsite mutants reveals dynamic high-affinity subsites and barriers in the long substrate binding cleft. *FEBS Lett. 580*, 5049–5053.
24. Ruohonen, L., Koivula, A., Reinikainen, T., Valkeajärvi, A., Teleman, A., Claeyssens, N., Szardenings, M., Jones, T. A., and Teeri, T. T. (1993) Active site of *T. reesei* cellobiohydrolase II. *Trichoderma reesei* cellulases and other hydrolases. *Found. Biotech. Ind. Ferment. Res., [Publ.] 8*, 87–96.
25. Watanabe, H., and Tokuda, G. (2001) Animal cellulases. *Cell. Mol. Life Sci. 58*, 1167–1178.
26. Mandelman, D., Belaich, A., Belaich, J. P., Aghajari, N., Driguez, H., and Haser, R. (2003) X-ray crystal structure of the multidomain endoglucanase Cel9G from *Clostridium cellulolyticum* complexed with natural and synthetic cello-oligosaccharides. *J. Bacteriol. 185*, 4127–4135.
27. Tuskan, G. A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R. R., Bhalerao, R. P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G. L., Cooper, D., Coutinho, P. M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Dejardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J. C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D. R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C. J., Uberbacher, E., Unneberg, P., et al. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science 313*, 1596–1604.
28. Bernier, G., Havelange, A., Houssa, C., Petitjean, A., and Lejeune, P. (1993) Physiological signals that induce flowering. *Plant Cell 5*, 1147–1155.
29. Trainotti, L., Spolaore, S., Pavanello, A., Baldan, B., and Casadoro, G. (1999) A novel E-type endo-beta-1,4-glucanase with a putative cellulose-binding domain is highly expressed in ripening strawberry fruits. *Plant Mol. Biol. 40*, 323–332.

BI702193E