THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Binding Site Dynamics and Aromatic–Carbohydrate Interactions in Processive and Non-Processive Family 7 Glycoside Hydrolases

Courtney B. Taylor,[†] Christina M. Payne,[‡,§] Michael E. Himmel,[‡] Michael F. Crowley,[‡] Clare McCabe,*[,†,∥] and Gregg T. Beckham*[,⊥,○]

[†]Department of Chemical and Biomolecular Engineering, Vanderbilt University, Nashville, Tennessee 37235, United States

[‡]Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado 80401, United States

[§]Department of Chemical and Materials Engineering, University of Kentucky, Lexington, Kentucky 40506, United States

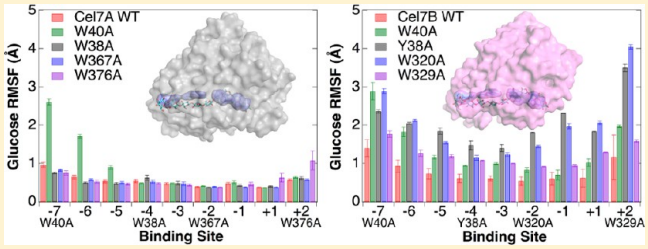
[∥]Department of Chemistry, Vanderbilt University, Nashville, Tennessee 37235, United States

[⊥]National Bioenergy Center, National Renewable Energy Laboratory, Golden, Colorado 80401, United States

[○]Department of Chemical Engineering, Colorado School of Mines, Golden, Colorado 80401, United States

Supporting Information

ABSTRACT: In nature, processive and non-processive cellulase enzymes deconstruct cellulose to soluble sugars. From structural studies, the consensus is that processive cellulases exhibit tunnels lined with aromatic and polar residues, whereas non-processive cellulases exhibit open clefts with fewer ligand contacts. To gain additional insight into the differences between processive and non-processive cellulases, we examine the glycoside hydrolase family 7 (GH7) cellobiohydrolase, Cel7A, and the endoglucanase, Cel7B, from *Trichoderma reesei* with molecular simulation. We compare properties related to processivity and compute the binding affinity changes for mutation of four aromatic residues lining the Cel7A active site tunnel and Cel7B cleft to alanine. For the wild-type enzymes, dissimilar behavior is observed at nearly every glucopyranose-binding site from −7 to +2, except in the −2 site, suggesting that the structural differences directly around the catalytic center and at the active site tunnel entrances and exits may all contribute to processivity in GH7s. Interestingly, the −2 site is similar in both enzymes, likely due to the significant conformational change needed in the cellodextrin ligand near this site for catalysis. Moreover, aromatic residue mutations in the Cel7A and Cel7B active sites display only small differences in binding affinity, but the ligand flexibility and enzyme–ligand interactions are only locally affected in Cel7A, whereas the entire ligand is significantly affected when any aromatic residue is mutated in Cel7B.



INTRODUCTION

Cellulose is the linear homopolymer of β1,4-D-glucose, which is characteristically resistant to degradation in the biosphere as a result of stable networks of hydrogen bonds and stacking interactions.[1−5] In nature, fungi and bacteria secrete cocktails of glycoside hydrolases (GHs) and oxidative enzymes to break down cellulose into soluble sugars for food sources.[6−11] The GHs that break down cellulose, generally referred to as cellulases, are classified on the basis of their mechanism of catalytic action and structure. These classifications are broadly delineated into three categories by function: (1) cellobiohydrolases (CBHs), which act in a processive manner on reducing and non-reducing ends of cellulose chains (a reducing end is able to form an open-chain aldehyde, whereas a non-reducing end is not able to isomerize to an open-ring form), (2) endoglucanases (EGs), which cleave glycosidic bonds at disordered sites on cellulose substrates in a non-processive fashion, and (3) β-glucosidases (cellobiases), which cleave soluble cellobiose to glucose in solution.[6,7] Processive and non-processive GHs are broadly characterized by their respective structural aspects wherein enzymes that exhibit a higher degree of processivity typically exhibit tunnels or deep cleft geometries in the ligand binding sites, and non-processive enzymes exhibit more open, solvent-exposed clefts or grooves for substrate binding.[7,12−17] Understanding the molecular-level basis for GH processivity is central to understanding carbohydrate metabolism in the biosphere.[18] Furthermore, GH processivity has emerged as a key focus in the design of enzymes for the burgeoning biofuels industry.[19,20]

CBHs and EGs must perform work to decrystallize polymer chains from the surfaces of cellulose, which is hypothesized to be offset by the ligand binding free energy of the substrate to the enzyme.[3,4,17,21] The magnitude of the ligand binding free energy for a given enzyme, in turn, will dictate the locations wherein an enzyme can decrystallize and hydrolyze polymer chains from the surfaces of cellulose crystals. As mentioned,

Received: February 7, 2013
Revised: March 24, 2013
Published: March 26, 2013

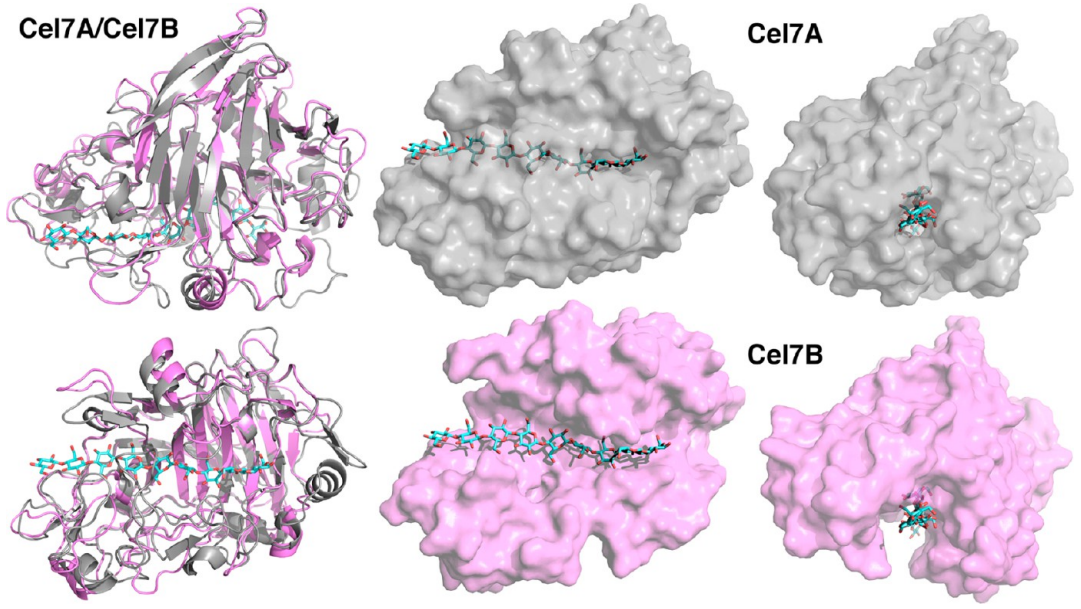


Figure 1. Comparison of *T. reesei* Cel7A and Cel7B catalytic structures. (left) Cel7A (gray) and Cel7B (pink) are aligned in cartoon format. The ligand is shown in red and cyan in stick format. (top middle and right) Cel7A shown along the tunnel and down the tunnel from the −7 site, respectively, with the ligand from the 8CEL crystal structure.[13] (bottom middle and right) Cel7B shown along the cleft and down the cleft from the −7 site, respectively, with the ligand from the Cel7A 8CEL structure docked into the cleft.[13]

structural studies have broadly suggested that the respective functions of CBHs and EGs are imparted by the formation of tunnels or deep clefts in CBHs relative to the more open substrate binding sites in EGs. However, several studies have provided examples that such a clear structural delineation of GH processivity may not be as straightforward for enzymes that act on cellulose or other similar recalcitrant polysaccharides. For example, the GH family 6 cellulase, Cel6B, from *Humicola insolens* is an EG, despite exhibiting a tunnel-like ligand binding site in the crystal structure.[22] The GH family 6 enzyme Cel6A from *Thermobifida fusca* is converted from a processive enzyme to a non-processive enzyme with the removal of a single tunnel loop.[23] Additionally, a GH family 18 chitinase, ChiB, from *Serratia marscesens* is converted from a processive chitinase to an endochitinase with only a single mutation of a tryptophan residue to alanine.[24] Thus, the causal basis for GH processivity may stem from subtler differences in structural and dynamical features that impact the ligand binding free energy including flexibility and solvation of the ligand and the flexibility of the catalytic machinery.[17]

The filamentous fungus *Trichoderma reesei* (or *Hypocrea jecorina*) produces some of the most well characterized GH family 7 cellulases, including a CBH, Cel7A, and an EG, Cel7B.[12−14,16] Both enzymes are multidomain proteins with a family 1 carbohydrate-binding module (CBM) and a catalytic domain (CD) connected by flexible, glycosylated linkers of similar length.[12−14,25−28] Cel7A and Cel7B exhibit approximately 45% sequence identity and similar structural features, as shown in Figure 1, but serve complementary, synergistic functions when degrading cellulose.[14,20,29−31] Cel7A adsorbs to crystalline cellulose and, either through endo-initiation[32] or binding to a free reducing end, hydrolyzes glycosidic linkages in a processive manner from the reducing end down a cellulose chain until it dissociates or the chain terminates,[33] while Cel7B, like other EGs, is hypothesized to attack disordered portions of the cellulose, exposing new chain ends for association and dissociation of CBHs.[7,8,15,29,30,34,35]

While Cel7A and Cel7B share the same β-jelly roll backbone structure characteristic of the family 7 GH fold, notable structural variations around the active site differentiate the two GHs and likely contribute in part to functionality (Figure 1). The Cel7A CD contains four surface loops that create a 50 Å tunnel for ligand binding.[7,12,13,35,36] In contrast, Cel7B has shorter surface loops creating a more open cleft or groove structure with many fewer substrate contacts.[7,12,14,15,37] Divne et al. determined that Cel7A exhibits up to 10 binding sites for a cellodextrin ligand, 3 putative product subsites and 7 for the substrate, with glycosidic cleavage occurring between the +1 and −1 sites, as for all GHs.[13] Kleywegt et al. considered nine of these sites, from −7 to +2, in their structural comparison of Cel7A and Cel7B.[14] A comparison of the two CD structures with a cellononaose chain bound in the active site of each is shown in Figure 1.

An additional structural feature of interest in GHs is the ubiquitous lining of the active sites with aromatic amino acids.[12,14,24,38−42] In Cel7A, four tryptophan (Trp) residues are distributed along the tunnel and participate in ligand binding at the −7, −4, −2, and +1 binding sites, as shown in Figure 2A.[13] In Cel7B, three Trp residues and one tyrosine (Tyr) residue are similarly distributed, as shown in Figure 2B.[14] Biochemical and structural studies to date have investigated the role some of these residues play in ligand binding and processivity. Koivula et al. investigated the role of Trp-272 at the tunnel entrance in the CBH Cel6A from *T. reesei* and found mutagenesis at this site caused a dramatic decrease in activity on crystalline cellulose but did not impact activity on amorphous cellulose or protein stability.[41] In a recent computational study, Payne et al. predicted that for *T. reesei* Cel6A residues associated with ligand acquisition (Trp-272) and product stabilization (Trp-135) had the greatest impact on the ligand binding free energy.[38] Igarashi et al. recently used high-speed atomic force microscopy to study processivity and found a velocity of 3.5 nm/s for *T. reesei* Cel7A degrading a cellulose chain from the crystal surface, but mutation from Trp-40 to alanine (Ala) near the entrance

dx.doi.org/10.1021/jp401410h | J. Phys. Chem. B 2013, 117, 4924−4933

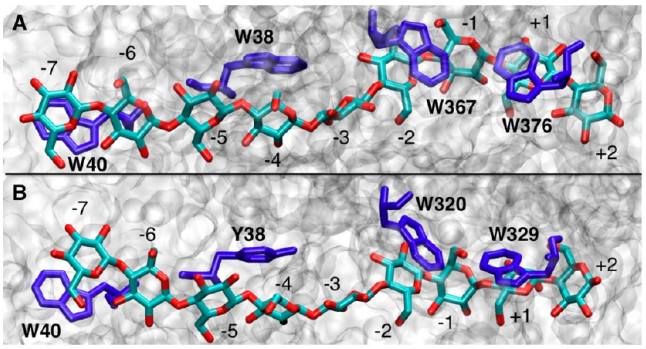


Figure 2. *T. reesei* Cel7A (A) and Cel7B (B) ligand binding sites with aromatic residues of interest shown in blue. Binding sites are labeled from the entrance at −7 to the exit at +2. The glycosidic bond cleavage occurs between the −1 and +1 sites.

binding site resulted in an enzyme that displayed repeated attachment and detachment from the surface without translation across the cellulose surface,[33] suggesting that entrance residues are important for both substrate recognition and processivity. Similarly, a recent theoretical study showed that mutation of Trp-40 to alanine inhibits the initial processivity events in Cel7A.[43] In a study of the processive chitinase B from *Serratia marcescens*, Horn et al. found that mutation of a Trp just upstream of the catalytic site (Trp-97) reduced processivity and activity on insoluble chitin but dramatically increased activity on chitosan, the soluble, deacetylated form of chitin.[24] On crystalline substrates, binding of the ligand by aromatic residues is potentially favorable for maintaining separation of individual decrystallized chains from the polymer surface, but for soluble substrates, tight binding could reduce the dissociation rate, as illustrated for chitinase B.[24] However, Payne et al. and von Ossowski et al. suggest that conclusions related to processivity and the role of aromatic−carbohydrate interactions for one GH family may not be universally applicable to all GH families, and that comparison within GH families might be of significant interest to characterize carbohydrate processivity.[36,38]

The goal of the current study is to examine structural and dynamical differences in the CBH Cel7A and EG Cel7B from *T. reesei*. Molecular simulations can offer insights into the energetics and structure−function relationships related to processivity in carbohydrate-active enzymes.[17,44,45] Additionally, by alchemically mutating aromatic amino acids to alanine individually in a processive and non-processive enzyme from the same GH family, we can also investigate how the relative binding affinity and interactions between the ligand and protein are impacted across these two enzyme classes. The MD simulations are used to quantify energetic and dynamical differences between the two systems, both wild-type and with aromatic-to-alanine mutations, whereas the thermodynamic integration (TI) data quantify how the binding affinity and ligand structure are impacted by mutation in both systems.

METHODS

Cel7A and Cel7B each exhibit at least nine binding sites numbered −7 (at the entrance of the tunnel or cleft near the substrate) to +2 (upstream of the catalytic site), as shown in Figure 2. Molecular dynamics (MD) simulations of the two wild-type systems and of both enzymes with each aromatic residue of interest mutated individually to alanine (Figure 2) have been performed. Using TI, the relative binding affinity changes for the aromatic mutations to alanine have also been calculated. We note that the Cel7A wild-type MD results shown here were reported in an earlier study comparing structural dynamics of GH7 cellobiohydrolases from *Heterobasidion irregulare*, *Phanerochaete chrysosporium*, and *H. jecorina*.[46]

All-Atom MD Simulations of the Cel7A and Cel7B Catalytic Domains. CHARMM[47] was used to build, solvate, and minimize Cel7A and Cel7B from the crystal structures, 8CEL and 1EG1, respectively.[12,14] The Cel7B backbone was glycosylated at Asn-56 and Asn-182 in accordance with experimental structural data,[14,48,49] though MD simulations suggests that the presence of *N*-glycans has little impact on the protein structure (Figure S1, Supporting Information). The starting coordinates of the ligand for both systems were taken from the modeled cellononaose ligand in Cel7A, and the catalytic residues for Cel7A/Cel7B, Glu-212/196, and Glu-217/201 were placed in the catalytically active conformation.[13] NAMD[50] was used for the MD simulations and TI calculations in the *NVT* ensemble. VMD[51] was used for all trajectory visualization. CHARMM[47] and in-house codes were used for analysis. The CHARMM27 all-atom force field with the CMAP correction[47,52,53] was used to describe the proteins. The ligands were modeled using the CHARMM carbohydrate force field,[54,55] and water was modeled with the modified TIP3P force field.[56,57] The solvated systems each contained approximately 55 500 atoms. The system preparation and computational methods are described in more detail in the Supporting Information.

From analysis of the trajectories of 250 ns MD simulations, the fluctuations of the protein backbones, the interaction energy between the protein and each ligand-binding site, and the number of hydrogen bonds formed between the protein and ligand were calculated. Additionally, differences in ligand solvation in both enzymes were investigated. VMD was used to determine the average number of waters within 3.5 Å of the ligand as a measure of solvation.[51] Cross-correlation maps were generated for the residues of each wild-type using principle component analysis in Amber12.[58,59] By comparing the wild-type and mutants for each enzyme, we attempt to gain insight into the differing roles of the aromatic residues as a function of position in the catalytic tunnel or cleft on ligand binding and stability between a processive and a non-processive cellulase.

Thermodynamic Integration and Relative Ligand Binding Free Energy. TI was used to calculate the changes in ligand binding free energy for mutation of aromatic residues to alanine in Cel7A and Cel7B.[38,60−62] The Cel7A residues Trp-40, Trp-38, Trp-367, and Trp-376 and the Cel7B residues Trp-40, Tyr-38, Trp-320, and Trp-329 were individually mutated to alanine. Both the Cel7A and Cel7B TI calculations were started using a configuration taken from the 100 ns into the MD simulations described above. NAMD dual-topology (see the Supporting Information) TI simulations[50,61] were equilibrated for 1 ns prior to collection of 10 ns of TI data and 2 ns prior to collection of 15 ns of TI data for electrostatic and van der Waals calculations, respectively. The TI methods including error analysis follow the work of Taylor et al. and Payne et al.[38,63] and are described in full in the Supporting Information (Figure S2).

RESULTS AND DISCUSSION

Comparison of the Cel7A and Cel7B Wild-Type Catalytic Domains. Simulations were conducted to investigate molecular-level differences between the two wild-type

dx.doi.org/10.1021/jp401410h | J. Phys. Chem. B 2013, 117, 4924−4933

GH7 enzymes with a bound ligand from the −7 to +2 sites. The root-mean-square fluctuations (RMSFs) from these simulations are shown in Figure 3, with the appropriate tunnel loop deletions included for Cel7B to achieve proper sequence and structural alignment.[14,36] Both systems exhibit similar RMSFs, with higher fluctuations in the loop regions surrounding the tunnel and the cleft (Figure 3). The tunnel loops not present in Cel7B are quite flexible in Cel7A, even though the ligand remains quite rigid in the tunnel during the simulation, as discussed below. The higher RMSF value in the 94−104 loop region for Cel7B correlates with the structural observation that this region is more open, allowing more access to the active site than the corresponding 96−103 loop in Cel7A.[14] The Cel7B peak from residues 260−280 near the exit of the cleft corresponds to a loop not present in Cel7A. The root-mean-square deviations (RMSDs) for each system are both similar and stable on the time scale of the MD simulations (Figure S3, Supporting Information).

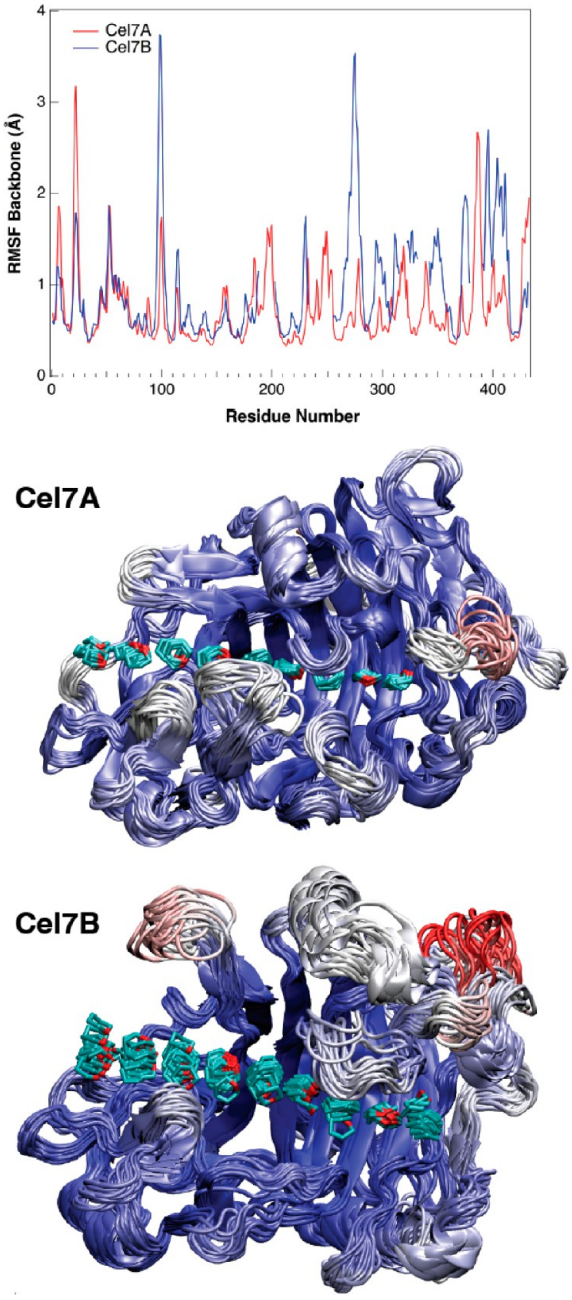


Figure 3. (top) RMSF of the protein backbones with residues 189−202, 234−254, 316−321, 333−341, and 381−390[14] skipped (designated by gaps in the blue line) in Cel7B in accordance with the sequence alignment. (middle) Cluster representations of the Cel7A and (bottom) Cel7B domains, with the ligand oriented from left to right −7 to +2. The ligand ring carbon atoms are cyan, and the oxygen atoms are red. Glycosidic oxygen atoms and hydrogen atoms are omitted for visual clarity in the ligand. The coloring of the CD from blue to white to red represents increasing RMSF scaled to a maximum value of 3.7 Å. *T. reesei* Cel7A results were first presented in ref 46.

The interaction energy and hydrogen bonding (H-bonding) at each binding site are shown in Figure 4 and confirm that the tunnel structure of Cel7A increases both hydrogen bonding to, and interaction with, the ligand over Cel7B, especially at the tunnel entrance and exit. The residue alignment and location of H-bonds do not deviate significantly from crystallographic studies.[13,14,64] With the exception of electrostatic interaction at the −1 site, the interactions and H-bonding directly around the catalytic site (+1 to −2) are equal within error. The significant differences at the −1 site arise from a higher number of longer-lived H-bonds, more rigid catalytic machinery, and longer loops containing polar residues contacting the ligand in Cel7A relative to Cel7B, as discussed further below. The conserved catalytic residues in family 7 GHs[14] comprise two glutamic acid residues (the proton donor for Cel7A/B is Glu-217/201 and the nucleophile Glu-212/196) and one aspartic acid residue (Asp-214/198).[12,15] The initial arrangement of the catalytic residues is consistent with structural studies,[13,64,65] but the Cel7B catalytic center is significantly more flexible overall, as previously found in GH18 chitinases.[17] Interestingly, the −2 sites are quite similar in terms of enzyme−ligand interactions and in the H-bonding pattern. In GH7s, the ligand begins to twist quite significantly at the −2 binding site. This ligand distortion is essential in the formation of the catalytically active conformation, which exposes the anomeric carbon to the nucleophile to form a enzyme-glycosyl intermediate. Previous structural studies have proposed that the −2 to −4 sites are key for stabilizing the ligand and maintaining this twist.[13,66]

Differences in interaction and H-bonding arise in the entrance and exit regions as well. The increase in Cel7A interaction energy results from greater electrostatic interaction with the ligand; the van der Waals interactions are roughly equivalent for the two enzymes. This observation can be primarily attributed to more charged and polar residues in contact with ligand due to the Cel7A loops. Recent simulation studies from Lin et al. showed that ligand binding weakens toward the cleft entrance in Cel7B when the enzyme was bound to a microfibril.[67] Our results show similar behavior, in that the aforementioned open active site tunnel from −7 to −3 in Cel7B results in almost no electrostatic interactions or H-bonds with the ligand. Interaction energy and hydrogen bonding are both decreased in Cel7B compared to Cel7A for the −3 site, partially due to the fact that the deleted loop 189−201 in Cel7A can intermittently form H-bonds with the −3 and −4 sites. Conserved residues Arg-107/108, Ser-365/318, and Tyr-145/146 (Cel7A/B) can H-bond with the glucopyranose rings in the −2 and −3 sites.[13,64] While the RMSF values for Arg-107 (Cel7A) and Arg-108 (Cel7B) are nearly equal, the fluctuations of the residues directly before the aforementioned 94−104 loop region are higher in Cel7B (Figure 2), which correlates with the observed changes in the interaction energy

dx.doi.org/10.1021/jp401410h | J. Phys. Chem. B 2013, 117, 4924−4933

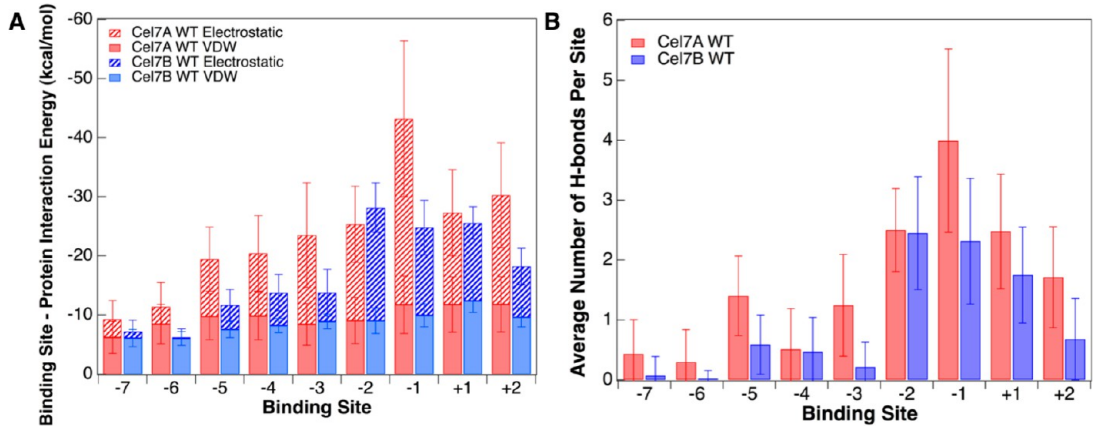


Figure 4. Interaction energy (A) and H-bonds (B) between the protein and each glucose moiety in each binding site. Interaction energy error bars were calculated using block averaging. Hydrogen bonds were calculated using a distance criterion cutoff of 3.0 Å and an angle of <60° from horizontal. H-bond error bars indicate the standard deviation.

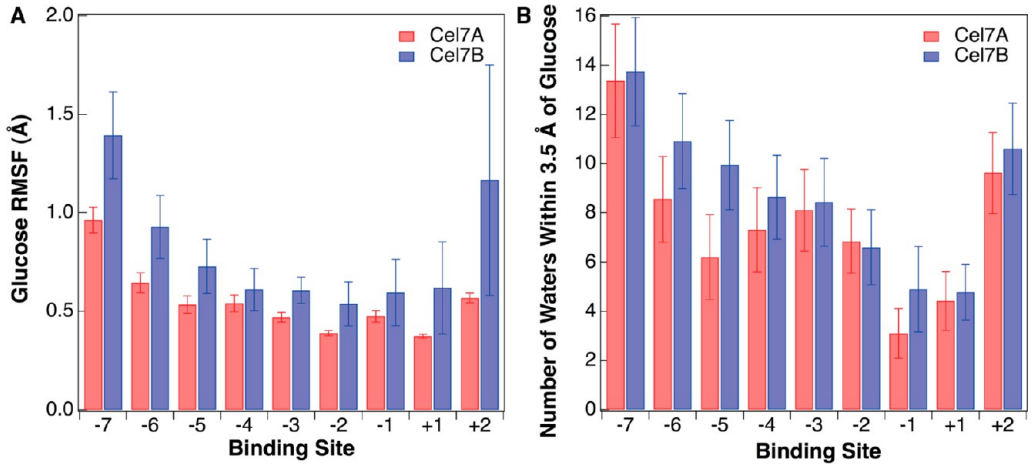


Figure 5. RMSF of the bound ligand (A) and the solvation, or number of water molecules within 3.5 Å of the bound ligand (B), for Cel7A and Cel7B over a 250 ns MD simulation. Errors were calculated using block averaging.

and loss in electrostatic interaction and H-bonds for the −3 site in Cel7B.

Lastly, the product sites exhibit noteworthy differences. The Cel7A loop 234−254 contacts the ligand in the +2 and +1 sites, allowing for increased interaction and formation of H-bonds with the +2 site in particular (Figure 4). The structure of the loops surrounding the product +1 and +2 sites in Cel7A and Cel7B are similar but differ in sequence: Cel7A forms H-bonds with three Arg residues (Arg-251, 267, and 394) as reported experimentally,[13] while Cel7B forms H-bonds with an alanine (Ala-222) and two glycines (Gly-223 and 225). The change in sequence also results in a significant decrease in total interaction energy at the +2 and +1 sites in Cel7B and contributes to the decreased H-bonding at the +2 site for Cel7B.[64,68]

Comparison of the Bound Ligand in the Cel7A and Cel7B Wild-Type. Figure 5 shows the RMSF and solvation differences around the ligand binding site in the two wild-type CDs. From Figure 5A, the fluctuations of each ligand are similar and correlate with the energetic and H-bond profiles (Figure 4). The Cel7B loops are more open than in Cel7A from the −3 to −7 sites, and as such, the ligand RMSF values in Cel7B are higher at the active site cleft entrance (Figure 5A). It is also noted that the +2 product site exhibits a larger degree of error than in Cel7A, corresponding to lower interaction energy and H-bonding in Cel7B (Figure 4B). As expected, the solvation for Cel7B increases over Cel7A from the entrance to the −3 site, where the tunnel loops are removed (Figure 5B). The number of water molecules around the ligand in the tunnel and cleft remain stable after equilibration, and the residence times of individual water molecules around each site are not significantly different for Cel7A and Cel7B (data not shown).

Relative Binding Free Energies of Aromatic Residues in the CD Tunnels. Table 1 summarizes the TI results

Table 1. Relative Binding Free Energy Changes for Cel7A and Cel7B per Binding Site as a Result of Aromatic Acid Mutation to Alanine

	mutation	binding site	$\Delta\Delta G$ (kcal/mol)
Cel7A	W40A	−7	3.9 ± 0.1
	W38A	−4	3.8 ± 0.1
	W367A	−2	5.0 ± 0.1
	W376A	+1	3.4 ± 0.1
Cel7B	W40A	−7	2.8 ± 0.2
	Y38A	−4	2.7 ± 0.5
	W320A	−2	4.5 ± 0.2
	W329A	+1	6.5 ± 0.2

dx.doi.org/10.1021/jp401410h | J. Phys. Chem. B 2013, 117, 4924−4933

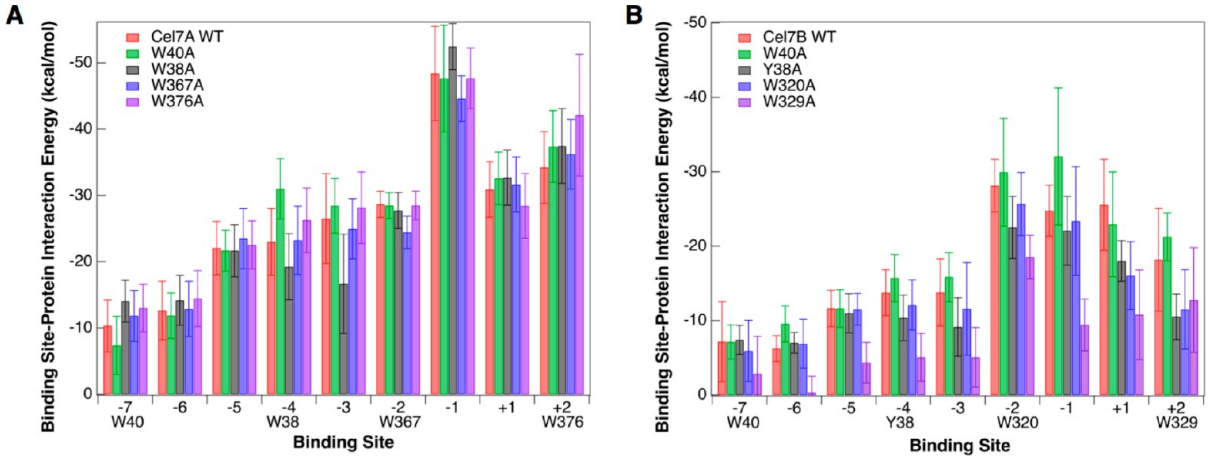


Figure 6. Interaction energy of the protein with each binding site in the ligand for Cel7A (A) and Cel7B (B). Mutations to Ala produce a more marked response in Cel7B.

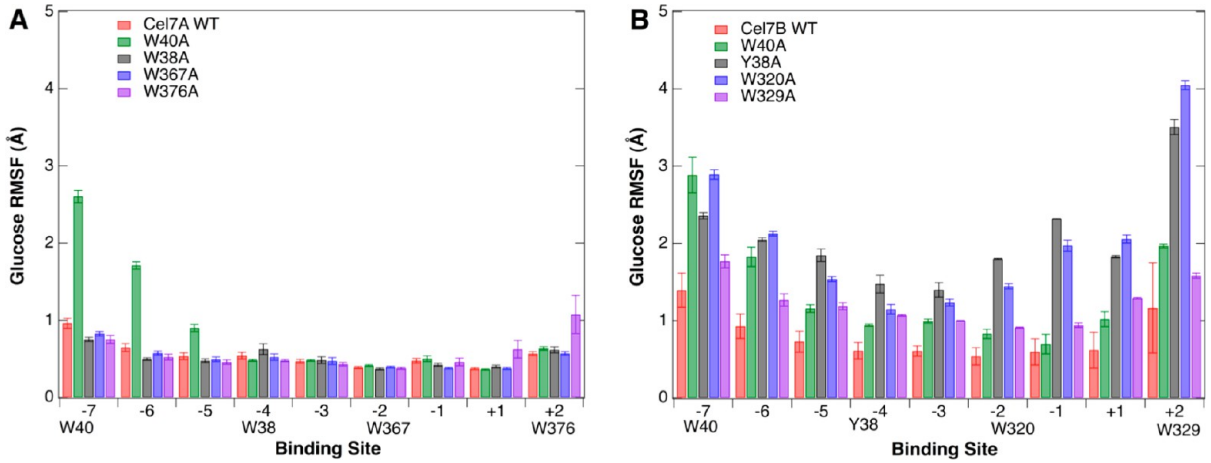


Figure 7. RMSF at each binding site resulting from aromatic mutation for Cel7A (A) and Cel7B (B).

wherein the aromatic residues shown in Figure 2 are mutated to alanine (A). The electrostatic and van der Waals results are reported in Tables S1 and S2 (Supporting Information). Mutation to alanine is found to be detrimental to ligand binding in all cases for the four residues lining the Cel7A and Cel7B ligand binding sites. Additionally, it is noted that the binding affinity change in Cel7A is essentially uniform across the entire enzyme, in stark contrast to the processive GH6 enzyme from the same organism.[38] In Cel7B, the −2 site (W320A) and +1 site (W329A) exhibit higher binding affinity changes than the entrance residues.

To investigate the impact of these mutations, 250 ns MD simulations were conducted to examine the enzyme−ligand interactions (Figure 6) and the ligand fluctuations (Figure 7). For Cel7A, Figures 6A and 7A and Figure S4A (Supporting Information) suggest that all four mutations in Cel7A produce only localized changes in the enzyme−ligand interactions and ligand RMSF, respectively. Mutation of the Trp-40 residue in Cel7A has been predicted recently to impede initial chain processivity,[43] which was hypothesized in previous experimental studies.[33,36,41] The results in Figures 6A and 7A suggest that the W40A mutation, like W272A at the tunnel entrance of *T. reesei* Cel6A,[38] only affects the entrance region. Moreover, it is noted that the $\Delta\Delta G$ for W272A in Cel6A was found to be 3.8 ± 0.4 kcal/mol, which is within error of the results found for W40A in Cel7A (Table 1), suggesting a similar role. Additionally, W38A in Cel7A exhibits a similar $\Delta\Delta G$ as W40A, and as with W40A, the changes in ligand interactions are localized to the −4 and −3 binding sites. The RMSF values are only slightly higher at the −4 site. Similarly, W367A and W376A yield only small changes in the ligand fluctuations near their site of mutation, and only minor changes in the enzyme−ligand interactions. Both the binding affinity changes and the localized changes in interactions and fluctuations suggest that aromatic−carbohydrate interactions in Cel7A are distinctly different than those in Cel6A, wherein aromatic residues in the middle of the tunnel were shown to rigidify the entire ligand from the +4 to −2 subsites, and the range of binding affinities was quite significant (~10 kcal/mol). In Cel7A, conversely, the data in Figures 6A and 7A seem to suggest that the aromatic−carbohydrate interactions play a role in local ligand binding and structuring and do not affect the overall stability of the ligand in terms of impacting the fluctuations along the entire length of the active site tunnel as the tryptophan residues do in Cel6A at the +1 and +2 binding sites. This direct comparison of two processive GHs from the same organism across two different families corroborates previous suggestions that the nature of ligand binding in GHs is significantly more complex than can be observed thorough a simple site-specific model comparison.

dx.doi.org/10.1021/jp401410h | J. Phys. Chem. B 2013, 117, 4924−4933

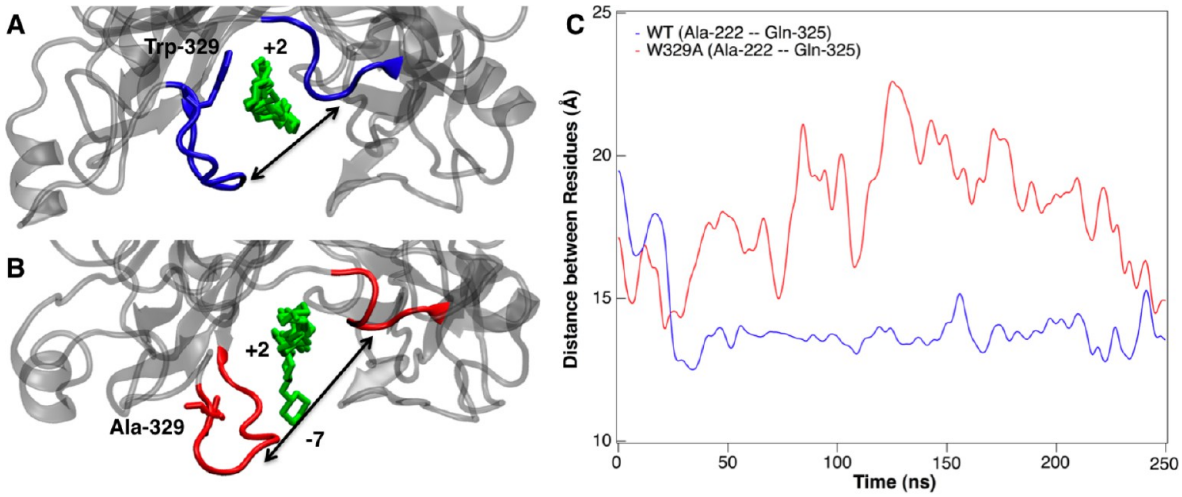


Figure 8. Cel7B wild-type and W329A MD simulations at 250 ns. Entrance view of ligand and position of Trp-329 in wild-type (A) and ligand and position of Ala-329 in the W329A mutation (B) showing the change in position of residue 329. Trp-329 and Ala-329 are shown using the licorice rendering in VMD. The arrows correspond to the distance between the tip of the loops surrounding the +2 and +1 sites, measured by the distance between the α-carbons in Ala-222 on the left and Gln-325 on the right. The time series of this distance over 250 ns is shown in part C and illustrates the significance of W329 in structure stabilization.

Mutation of the four aromatic residues in Cel7B produces much larger energetic and structural changes than in Cel7A (Figures 6B and 7B and Figure S4B, Supporting Information). Cel7B binding affinity to the ligand is most negatively impacted by mutation at the product site (W329A) and, like Cel7A, directly upstream of the catalytically active site at the −2 binding site (W320A). The interaction energy drops significantly at all binding sites with the W329A mutation (Figure 6B). The cross-correlation maps from the MD simulation show that Trp-329 is positively correlated with the active site (Glu-196/Asp-198/Glu-201), indicating mutation of this residue could adversely impact catalysis (Figure S5, Supporting Information). Also, in the Cel7B wild-type and W329A simulations, Ala-329 translates much farther away from the product site than Trp-329 (Figure 8). This change in protein structure leads to dramatic instabilities in the cleft width, measured from the tips of the loops surrounding the +2 to −2 sites (Ala-222, a native contact in the wild-type and Gln-325) (Figure 8C). Both the wild-type and W329A mutant exhibit some relaxation of the structure in the first 50 ns, but the wild-type stabilizes, whereas the mutant does not. Similarly, the W320A, Y38A, and W40A mutations also impact the entire cellodextrin chain in terms of inducing significant changes to the enzyme−ligand interaction, ligand fluctuations, and protein backbone fluctuations (Figure S4B, Supporting Information), suggesting that removal of any of these residues would be significantly detrimental to the ligand binding free energy and to the ability of Cel7B to complex cellulose chains.

CONCLUSIONS

We have examined a GH7 CBH and EG, *T. reesei* Cel7A and Cel7B, respectively, to quantify the dynamic and structural differences in ligand−protein interactions, with the overall aim to quantify key properties that are potentially important for GH processivity. First, our results indicate that the Cel7A and Cel7B ligand binding sites behave quite differently at every site from −7 to +2, with the exception of the −2 binding site. The −2 binding site in GH7 enzymes is likely responsible for stabilizing the significant twist in the cellodextrin chain needed for exposure of the anomeric carbon to the catalytic nucleophile for catalysis, and hence, its function is perhaps necessary for both CBH and EG enzymes from this GH family. Moreover, we examined the role of aromatic residues in both Cel7A and Cel7B to understand how these ubiquitous and highly conserved hydrophobic platforms may affect the ability of each enzyme to bind to a cellulose chain and conduct catalysis. Interestingly, we find that the overall $\Delta\Delta G$ values between Cel7A and Cel7B are quite similar, and that in all cases their mutation to alanine reduces the binding affinity to the ligand. However, the structural and dynamical ramifications for aromatic mutations to alanine in the Cel7A and Cel7B are quite different. For Cel7A, the mutations mainly affect the binding and structure in localized regions around the aromatic residue of interest, whereas each mutation is seemingly completely deleterious for ligand binding in Cel7B, such that it is likely that mutation of any aromatic residue will remove the ability to complex cellulose chains effectively from an insoluble substrate where thermodynamic work must be conducted to dislodge chains. The weaker interactions of the Cel7B cleft residues with the ligand support the notion that lack of tight binding may decrease processivity and enhance the detachment of the enzyme on insoluble substrates. Overall, this study highlights the dynamic differences between CBHs and EGs within the same enzyme family and, when compared with earlier results from a cellulase in a different GH family,[38] suggests that the role of aromatic−carbohydrate interactions varies significantly between and within GH families.

ASSOCIATED CONTENT

Supporting Information

Full references for refs 17, 27, 36, 47, 52, and 58. Full details of the MD simulations and TI calculation protocols. Additional data generated from the 250 ns MD production runs for the Cel7A and Cel7B wild-type and mutant systems; this includes the comparison of glycosylated and non-glycosylated wild-type Cel7B (Figure S1), an example TI data set (Figure S2), RMSD and RMSF of Cel7A and Cel7B wild-type and mutated CD (Figures S3 and S4), and principle component analysis for the wild-type structures (Figure S5). The individual electrostatic and van der Waals TI simulation results can be found in Tables

dx.doi.org/10.1021/jp401410h | J. Phys. Chem. B 2013, 117, 4924−4933

S1 and S2. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: (615) 322-6853 (C.M.); (303) 384-7806 (G.T.B.). E-mail: c.mccabe@vanderbilt.edu (C.M.); gregg.beckham@nrel.gov (G.T.B.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the US Department of Energy (DOE) Office of the Biomass Program for funding. Computational time for this research was supported in part by National Science Foundation through XSEDE resources on the Texas Advanced Computer Center Ranger cluster and the National Institute for Computational Science Kraken cluster, under grant TG-MCB090159 and by the National Renewable Energy Laboratory Computational Sciences Center supported by US DOE Energy Efficiency and Renewable Energy under Contract No. DE-AC36-08GO28308. Additional resources were provided through the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the US DOE under Contract No. DE-AC02-05CH11231.

REFERENCES

(1) Nishiyama, Y.; Langan, P.; Chanzy, H. Crystal Structure and Hydrogen-Bonding System in Cellulose 1 Beta from Synchrotron X-Ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.* **2002**, *124*, 9074−9082.

(2) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. Crystal Structure and Hydrogen Bonding System in Cellulose 1 Alpha. *J. Am. Chem. Soc.* **2003**, *125*, 14300−14306.

(3) Beckham, G. T.; Matthews, J. F.; Peters, B.; Bomble, Y. J.; Himmel, M. E.; Crowley, M. F. Molecular-Level Origins of Biomass Recalcitrance: Decrystallization Free Energies for Four Common Cellulose Polymorphs. *J. Phys. Chem. B* **2011**, *115*, 4118−4127.

(4) Payne, C. M.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. Decrystallization of Oligosaccharides from the Cellulose I beta Surface with Molecular Simulation. *J. Phys. Chem. Lett.* **2011**, *2*, 1546−1550.

(5) Gross, A. S.; Chu, J. W. On the Molecular Origins of Biomass Recalcitrance: The Interaction Network and Solvation Structures of Cellulose Microfibrils. *J. Phys. Chem. B* **2010**, *114*, 13333−13341.

(6) Himmel, M. E.; Ding, S. Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* **2007**, *315*, 804−807.

(7) Lynd, L. R.; Weimer, P. J.; van Zyl, W. H.; Pretorius, I. S. Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiol. Mol. Biol. Rev.* **2002**, *66*, 506−577.

(8) Barr, B. K.; Hsieh, Y. L.; Ganem, B.; Wilson, D. B. Identification of Two Functionally Different Classes of Exocellulases. *Biochemistry* **1996**, *35*, 586−592.

(9) Forsberg, Z.; Vaaje-Kolstad, G.; Westereng, B.; Bunaes, A. C.; Stenstrom, Y.; MacKenzie, A.; Sørlie, M.; Horn, S. J.; Eijsink, V. G. H. Cleavage of Cellulose by a CBM33 Protein. *Protein Sci.* **2011**, *20*, 1479−1483.

(10) Vaaje-Kolstad, G.; Westereng, B.; Horn, S. J.; Liu, Z. L.; Zhai, H.; Sørlie, M.; Eijsink, V. G. H. An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. *Science* **2010**, *330*, 219−222.

(11) Horn, S.; Vaaje-Kolstad, G.; Westereng, B.; Eijsink, V. G. Novel Enzymes for the Degradation of Cellulose. *Biotechnol. Biofuels* **2012**, *5*, 45.

(12) Divne, C.; Ståhlberg, J.; Reinikainen, T.; Ruohonen, L.; Pettersson, G.; Knowles, J. K. C.; Teeri, T. T.; Jones, T. A. The 3-Dimensional Crystal-Structure of the Catalytic Core of Cellobiohydrolase-I from *Trichoderma reesei*. *Science* **1994**, *265*, 524−528.

(13) Divne, C.; Ståhlberg, J.; Teeri, T. T.; Jones, T. A. High-Resolution Crystal Structures Reveal How a Cellulose Chain is Bound in the 50 Å Long Tunnel of Cellobiohydrolase I from *Trichoderma reesei*. *J. Mol. Biol.* **1998**, *275*, 309−325.

(14) Kleywegt, G. J.; Zou, J. Y.; Divne, C.; Davies, G. J.; Sinning, I.; Ståhlberg, J.; Reinikainen, T.; Srisodsuk, M.; Teeri, T. T.; Jones, T. A. The Crystal Structure of the Catalytic Core Domain of Endoglucanase I from *Trichoderma reesei* at 3.6 Å Resolution, and a Comparison with Related Enzymes. *J. Mol. Biol.* **1997**, *272*, 383−397.

(15) Mackenzie, L. F.; Sulzenbacher, G.; Divne, C.; Jones, T. A.; Woldike, H. F.; Schulein, M.; Withers, S. G.; Davies, G. J. Crystal Structure of the Family 7 Dndoglucanase I (Cel7B) from *Humicola insolens* at 2.2 Å Resolution and Identification of the Catalytic Nucleophile by Trapping of the Covalent Glycosyl-Enzyme Intermediate. *Biochem. J.* **1998**, *335*, 409−416.

(16) Ståhlberg, J.; Johansson, G.; Pettersson, G.; New, A. Model for Enzymatic-Hydrolysis of Cellulose Based on the 2-Domain Structure of Cellobiohydrolase-I. *Biotechnol. Biofuels* **1991**, *9*, 286−290.

(17) Payne, C. M.; Baban, J.; Horn, S. J.; Backe, P. H.; Arvai, A. S.; Dalhus, B.; Bjørås, M.; Eijsink, V. G. H.; Sørlie, M.; Beckham, G. T.; et al. Hallmarks of Processivity in Glycoside Hydrolases from Crystallographic and Computational Studies of the *Serratia marcescens* Chitinases. *J. Biol. Chem.* **2012**, *287*, 36322−36330.

(18) Stern, R.; Jedrzejas, M. J. Carbohydrate Polymers at the Center of Life's Origins: The Importance of Molecular Processivity. *Chem. Rev.* **2008**, *108*, 5061−5085.

(19) Cruys-Bagger, N.; Elmerdahl, J.; Praestgaard, E.; Tatsumi, H.; Spodsberg, N.; Borch, K.; Westh, P. Pre-Steady-State Kinetics for Hydrolysis of Insoluble Cellulose by Cellobiohydrolase Cel7A. *J. Biol. Chem.* **2012**, *287*, 18451−18458.

(20) Kurasin, M.; Väljamäe, P. Processivity of Cellobiohydrolases Is Limited by the Substrate. *J. Biol. Chem.* **2011**, *286*, 169−177.

(21) Beckham, G. T.; Crowley, M. F. Examination of the Alpha-Chitin Structure and Decrystallization Thermodynamics at the Nanoscale. *J. Phys. Chem. B* **2011**, *115*, 4516−4522.

(22) Davies, G. J.; Brzozowski, A. M.; Dauter, M.; Varrot, A.; Schulein, M. Structure and Function of *Humicola insolens* Family 6 Cellulases: Structure of the Endoglucanase, Cel6B, at 1.6 Å Resolution. *Biochem. J.* **2000**, *348*, 201−207.

(23) Meinke, A.; Damude, H. G.; Tomme, P.; Kwan, E.; Kilburn, D. G.; Miller, R. C.; Warren, R. A. J.; Gilkes, N. R. Enhancement of the Endo-Beta-1,4-Glucanase Activity of an Exocellobiohydrolase by Deletion of a Surface Loop. *J. Biol. Chem.* **1995**, *270*, 4383−4386.

(24) Horn, S. J.; Sikorski, P.; Cederkvist, J. B.; Vaaje-Kolstad, G.; Sørlie, M.; Synstad, B.; Vriend, G.; Varum, K. M.; Eijsink, V. G. H. Costs and Benefits of Processivity in Enzymatic Degradation of Recalcitrant Polysaccharides. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18089−18094.

(25) Kraulis, P. J.; Clore, G. M.; Nilges, M.; Jones, T. A.; Pettersson, G.; Knowles, J.; Gronenborn, A. M. Determination of the 3-Dimensional Solution Structure of the C-Terminal Doman of Cellobiohydrolase-I from *Trichoderma reesei* - A Study Using Nuclear Magnetic Resonance and Hybrid Distance Geometry Dynamical Simulated Annealing. *Biochemistry* **1989**, *28*, 7241−7257.

(26) Harrison, M. J.; Nouwens, A. S.; Jardine, D. R.; Zachara, N. E.; Gooley, A. A.; Nevalainen, H.; Packer, N. H. Modified Glycosylation of Cellobiohydrolase I from a High Cellulase-Producing Mutant Strain of *Trichoderma reesei*. *Eur. J. Biochem.* **1998**, *256*, 119−127.

(27) Beckham, G. T.; Bomble, Y. J.; Matthews, J. F.; Taylor, C. B.; Resch, M. G.; Yarbrough, J. M.; Decker, S. R.; Bu, L. T.; Zhao, X. C.; McCabe, C.; et al. The O-Glycosylated Linker from the *Trichoderma reesei* Family 7 Cellulase Is a Flexible, Disordered Protein. *Biophys. J.* **2010**, *99*, 3773−3781.

(28) Sammond, D. W.; Payne, C. M.; Brunecky, R.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. Cellulase Linkers Are Optimized

Based on Domain Type and Function: Insights from Sequence Analysis, Biophysical Measurements, and Molecular Simulation. *PLoS One* **2012**, *7*, e48615.
(29) Claeyssens, M.; Vantilbeurgh, H.; Kamerling, J. P.; Berg, J.; Vrsanska, M.; Biely, P. Studies of the Cellulolytic System of the Filamentous Fungus *Trichoderma reesei* QM-9414 - Substrate-Specificity and Transfer Activity of Endoglucanase-I. *Biochem. J.* **1990**, *270*, 251−256.
(30) Jalak, J.; Kurasin, M.; Teugjas, H.; Väljamäe, P. Endo-Exo Synergism in Cellulose Hydrolysis Revisited. *J. Biol. Chem.* **2012**, *287*, 28802−28815.
(31) Murphy, L.; Cruys-Bagger, N.; Damgaard, H. D.; Baumann, M. J.; Olsen, S. N.; Borch, K.; Lassen, S. F.; Sweeney, M.; Tatsumi, H.; Westh, P. Origin of Initial Burst in Activity for *Trichoderma reesei* Endo-Glucanases Hydrolyzing Insoluble Cellulose. *J. Biol. Chem.* **2012**, *287*, 1252−1260.
(32) Ståhlberg, J.; Johansson, G.; Pettersson, G. *Trichoderma reesei* Has No True Exo-Cellulase - All Intact and Truncated Cellulases Produce New Reducing End Groups on Cellulose. *Biochim. Biophys. Acta* **1993**, *1157*, 107−113.
(33) Igarashi, K.; Koivula, A.; Wada, M.; Kimura, S.; Penttila, M.; Samejima, M. High Speed Atomic Force Microscopy Visualizes Processive Movement of *Trichoderma reesei* Cellobiohydrolase I on Crystalline Cellulose. *J. Biol. Chem.* **2009**, *284*, 36186−36190.
(34) Nidetzky, B.; Steiner, W.; Hayn, M.; Claeyssens, M. Cellulose Hydrolysis by the Cellulases from *Trichoderma-reesei* - A New Model for Synergistic Interaction. *Biochem. J.* **1994**, *298*, 705−710.
(35) Väljamäe, P.; Sild, V.; Nutt, A.; Pettersson, G.; Johansson, G. Acid Hydrolosis of Bacterial Cellulose Reveals Different Modes of Synergistic Action between Cellobiohydrolase I and Endoglucanase I. *Eur. J. Biochem.* **1999**, *266*, 327−334.
(36) von Ossowski, I.; Ståhlberg, J.; Koivula, A.; Piens, K.; Becker, D.; Boer, H.; Harle, R.; Harris, M.; Divne, C.; Mahdi, S.; et al. Engineering the Exo-Loop of *Trichoderma reesei* Cellobiohydrolase, Cel7A. A Comparison with *Phanerochaete chrysosporium* Cel7D. *J. Mol. Biol.* **2003**, *333*, 817−829.
(37) Penttila, M.; Lehtovaara, P.; Nevalainen, H.; Bhikhabhai, R.; Knowles, J. Homology Between Cellulase Genes of *Trichoderma reesei* - Complete Nucleotide-Sequence of the Endoglucanase-I Gene. *Gene* **1986**, *45*, 253−263.
(38) Payne, C. M.; Bomble, Y.; Taylor, C. B.; McCabe, C.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. Multiple Functions of Aromatic-Carbohydrate Interactions in a Processive Cellulase Examined with Molecular Simulation. *J. Biol. Chem.* **2011**, *286*, 41028−41035.
(39) Rouvinen, J.; Bergfors, T.; Teeri, T.; Knowles, J. K. C.; Jones, T. A. 3-Dimensional Structure of Cellobiohydrolase II from *Trichoderma reesei*. *Science* **1990**, *249*, 380−386.
(40) van Aalten, D. M. F.; Synstad, B.; Brurberg, M. B.; Hough, E.; Riise, B. W.; Eijsink, V. G. H.; Wierenga, R. K. Structure of a Two-Domain Chitotriosidase from *Serratia marcescens* at 1.9 Å Resolution. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5842−5847.
(41) Koivula, A.; Kinnari, T.; Harjunpaa, V.; Ruohonen, L.; Teleman, A.; Drakenberg, T.; Rouvinen, J.; Jones, T. A.; Teeri, T. T. Tryptophan 272: An Essential Determinant of Crystalline Cellulose Degradation by *Trichoderma reesei* Cellobiohydrolase Cel6A. *FEBS Lett.* **1998**, *429*, 341−346.
(42) Williams, S. J.; Davies, G. J. Protein-Carbohydrate Interactions: Learning Lessons from Nature. *Trends Biotechnol.* **2001**, *19*, 356−362.
(43) GhattyVenkataKrishna, P. K.; Alekozai, E. M.; Beckham, G. T.; Schulz, R.; Crowley, M. F.; Uberbacher, E. C.; Cheng, X. Initial Recognition of a Cellodextrin Chain in the Cellulose-Binding Tunnel May Affect Cellobiohydrolase Directional Specificity. *Biophys. J.* **2013**, *104*, 1−9.
(44) Beckham, G. T.; Bomble, Y. J.; Bayer, E. A.; Himmel, M. E.; Crowley, M. F. Applications of Computational Science for Understanding Enzymatic Deconstruction of Cellulose. *Curr. Opin. Biotechnol.* **2011**, *22*, 231−238.
(45) Chundawat, S. P. S.; Beckham, G. T.; Himmel, M. E.; Dale, B. E. Deconstruction of Lignocellulosic Biomass to Fuels and Chemicals. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 6.1−6.25.
(46) Momeni, M. H.; Payne, C. M.; Hansson, H.; Mikkelsen, N. E.; Svedberg, J.; Engström, Å.; Sandgren, M.; Beckham, G. T.; Ståhlberg, J. Structural, Biochemical, and Computational Characterization of the Glycoside Hydrolase Family 7 Cellobiohydrolase of the Tree-Killing Fungus *Heterobasidion irregulare*. *J. Biol. Chem.* **2013**, *288*, 5861−5872.
(47) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(48) Deshpande, N.; Wilkins, M. R.; Packer, N.; Nevalainen, H. Protein Glycosylation Pathways in Filamentous Fungi. *Glycobiology* **2008**, *18*, 626−637.
(49) Eriksson, T.; Stals, I.; Collen, A.; Tjerneld, F.; Claeyssens, M.; Stalbrand, H.; Brumer, H. Heterogeneity of Homologously Expressed *Hypocrea jecorina* (*Trichoderma reesei*) Cel7B Catalytic Module. *Eur. J. Biochem.* **2004**, *271*, 1266−1276.
(50) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(51) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(52) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(53) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(54) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J. Chem. Theory Comput.* **2009**, *5*, 2353−2370.
(55) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D. Additive Empirical Force Field for Hexopyranose Monosaccharides. *J. Comput. Chem.* **2008**, *29*, 2543−2564.
(56) Durell, S. R.; Brooks, B. R.; Bennaim, A. Solvent-Induced Forces between 2 Hydrophilic Groups. *J. Phys. Chem.* **1994**, *98*, 2198−2202.
(57) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(58) Case, D. A.; Darden, T.; Cheatham, T. E., III; Simmerling, C.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B. P.; et al. *AMBER 12*; University of California: San Fransisco, CA, 2012.
(59) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(60) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Academic Press: San Diego, CA, 2002.
(61) Pohorille, A.; Jarzynski, C.; Chipot, C. Good Practices in Free-Energy Calculations. *J. Phys. Chem. B* **2010**, 10235−10253.
(62) Kollman, P. Free Energy Calculations: Applications to Chemical and Biological Phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.
(63) Taylor, C. B.; Talib, M. F.; McCabe, C.; Bu, L.; Adney, W. S.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. Computational Investigation of Glycosylation Effects on a Family 1 Carbohydrate-Binding Module. *J. Biol. Chem.* **2012**, 3147−3155.
(64) Parkkinen, T.; Koivula, A.; Vehmaanpera, J.; Rouvinen, J. Crystal Structures of *Melanocarpus albomyces* Cellobiohydrolase

Ce17B in Complex with Cello-Oligomers Show High Flexibility in the Substrate Binding. *Protein Sci.* **2008**, *17*, 1383−1394.
(65) Sulzenbacher, G.; Schülein, M.; Davies, G. J. Structure of the Endoglucanase I from *Fusarium oxysporum*, Native, Cellobiose, and 3,4-Epoxybutyl-d-Cellobioside-Inhibited Forms, at 2.3 Å Resolution. *Biochemistry* **1997**, *36*, 5902−5911.
(66) Teeri, T. T.; Koivula, A.; Linder, M.; Wohlfahrt, G.; Divne, C.; Jones, T. A. *Trichoderma reesei* Cellobiohydrolases: Why So Efficient on Crystalline Cellulose? *Biochem. Soc. Trans.* **1998**, *26*, 173−178.
(67) Lin, Y.; Silvestre-Ryan, J.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T.; Chu, J.-W. Protein Allostery at the Solid-Liquid Interface: Endoglucanase Attachment to Cellulose Affects Glucan Clenching in the Binding Cleft. *J. Am. Chem. Soc.* **2011**, *133*, 16617−16624.
(68) Bu, L.; Nimlos, M. R.; Shirts, M. R.; Ståhlberg, J.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. Product Binding Varies Dramatically Between Processive and Non-Processive Cellulase Enzymes. *J. Biol. Chem.* **2012**, *287*, 24807−24813.