JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Carbohydrate–Protein Interactions That Drive Processive Polysaccharide Translocation in Enzymes Revealed from a Computational Study of Cellobiohydrolase Processivity

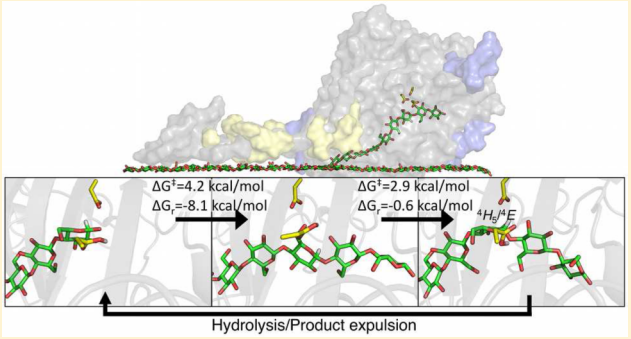
Brandon C. Knott,[†] Michael F. Crowley,[‡] Michael E. Himmel,[‡] Jerry Ståhlberg,*[,§] and Gregg T. Beckham*[,†]

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

[§]Department of Molecular Biology, Swedish University of Agricultural Sciences, SE-750 07 Uppsala, Sweden

Supporting Information

ABSTRACT: Translocation of carbohydrate polymers through protein tunnels and clefts is a ubiquitous biochemical phenomenon in proteins such as polysaccharide synthases, glycoside hydrolases, and carbohydrate-binding modules. Although static snapshots of carbohydrate polymer binding in proteins have long been studied via crystallography and spectroscopy, the molecular details of polysaccharide chain processivity have not been elucidated. Here, we employ simulation to examine how a cellulose chain translocates by a disaccharide unit during the processive cycle of a glycoside hydrolase family 7 cellobiohydrolase. Our results demonstrate that these biologically and industrially important enzymes employ a two-step mechanism for chain threading to form a Michaelis complex and that the free energy barrier to chain threading is significantly lower than the hydrolysis barrier. Taken with previous studies, our findings suggest that the rate-limiting step in enzymatic cellulose degradation is the glycosylation reaction, not chain processivity. Based on the simulations, we find that strong electrostatic interactions with polar residues that are conserved in GH7 cellobiohydrolases, but not in GH7 endoglucanases, at the leading glucosyl ring provide the thermodynamic driving force for polysaccharide chain translocation. Also, we consider the role of aromatic–carbohydrate interactions, which are widespread in carbohydrate-active enzymes and have long been associated with processivity. Our analysis suggests that the primary role for these aromatic residues is to provide tunnel shape and guide the carbohydrate chain to the active site. More broadly, this work elucidates the role of common protein motifs found in carbohydrate-active enzymes that synthesize or depolymerize polysaccharides by chain translocation mechanisms coupled to catalysis.



INTRODUCTION

Polysaccharides serve myriad functions in cell biology, and their formation has been proposed to be a prerequisite for the evolution of life.[1] Concomitant to the complexity and diversity of carbohydrates, a massive battery of enzymes is required for their synthesis, modification, and depolymerization.[2] During both polysaccharide synthesis and depolymerization, carbohydrate polymers are threaded through tunnels and clefts in proteins. These carbohydrate–protein interaction sites are commonly rich in aromatic and polar residues that interact with the carbohydrate ligands.[3,4] For example, the recently solved structure of the bacterial cellulose synthase revealed a large, multidomain protein complex wherein a glycosyltransferase links β-glucose to the growing end of a cellulose chain, which is subsequently threaded through a long transmembrane domain to an extracellular protein complex with multiple carbohydrate recognition motifs.[5] Additionally, many families of carbohydrate-binding modules (CBMs), which serve a wide range of carbohydrate recognition functions, exhibit clefts for binding and threading carbohydrate chains.[6] Similarly, cellulose, chitin, and other structural polysaccharides are depolymerized in nature by processive glycoside hydrolases (GHs) that thread a single polymer chain through an enzyme tunnel for hydrolysis.[7−10] Crystal structures solved in all of these cases demonstrate that a common set of carbohydrate–protein binding motifs are employed in nature, suggesting that recognition of polysaccharide chains is likely mediated by a similar mechanism across a broad spectrum of proteins. However, despite the wealth of structures solved to date, the molecular-level steps involved in polysaccharide translocation have not yet been revealed.

Cellobiohydrolases (CBHs) processively cleave cellobiose from cellulose chains and constitute excellent model systems to

Received: April 23, 2014
Published: May 28, 2014

 © 2014 American Chemical Society dx.doi.org/10.1021/ja504074g | J. Am. Chem. Soc. 2014, 136, 8810–8819

study polysaccharide translocation in proteins. *Trichoderma reesei* (*Hypocrea jecorina*) Cel7A (*Tr*Cel7A) from glycoside hydrolase family 7 (GH7) is an exceptionally well-characterized CBH.[7,11−16] Biomass-degrading fungi[2] commonly use GH7 enzymes for cellulose depolymerization, and given their high activity toward cellulose, they are also commonly employed in industrial biofuels applications.[10,17] The structure of *Tr*Cel7A was first reported in 1994,[11] and multiple ligand-bound structures were published in 1998.[7] The latter study revealed at least 9 binding sites from the tunnel entrance at the −7 subsite to the +2 subsite (Figure 1). Catalysis takes place between the −1 and +1 subsite.

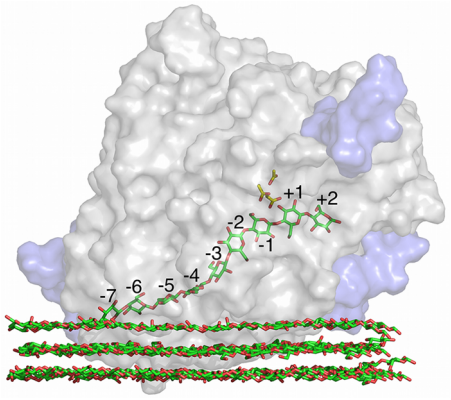


Figure 1. *Tr*Cel7A catalytic domain (CD) acting at the crystalline cellulose surface. *Tr*Cel7A CD is shown in gray with N-glycosylation in blue. The cellulose surface is shown in green and red "sticks". The catalytic residues are shown in yellow and red sticks in between the −1 and +1 subsites.

*Tr*Cel7A depolymerizes cellulose via a processive mechanism wherein a single polymer chain of cellulose binds to the enzyme from the reducing end, which is threaded into the approximately 50 Å enzyme tunnel (Figure 1).[7] Processive motion of this enzyme on the cellulose surface has been observed directly using high-speed atomic force microscopy (HS-AFM),[9,18,19] which revealed a velocity of approximately 3−7 nm/s, but the rate-limiting step in the processive cycle was not directly elucidated from these measurements. More recently, new crystal structures of *Tr*Cel7A were presented wherein an entire cellononaose chain was bound to a catalytically inactive mutant and another structure that revealed the nature of the glycosyl-enzyme intermediate.[16] Transition path sampling simulations were also employed in the same study, which showed that glycosylation is the rate-limiting step of the hydrolytic reaction and that the reaction rate constant is approximately 11 s^{-1}.[16] This study did not consider processivity, however, thus further work is necessary to understand these results in the greater context of the CBH processive cycle.

It has recently been proposed[20] that the processive cycle of GH7 CBHs involves the following elementary steps (illustrated in Figure 2). Subsequent to initial chain loading (or product expulsion), the product sites (+1/+2) are vacant. A cellulose chain is then processed across the active site in a noncatalytically active configuration, wherein the anomeric carbon reaction center is far from the nucleophile (∼7 Å) and all glucose residues are in the stable chair configuration. Catalytic activation follows, which involves a twisting of the cellulose chain toward the active site and a distortion of the −1 sugar ring (into an "envelope" or "half-chair"), which primes the glycosidic bond for catalysis.[21,22] Within the processive cycle, hydrolysis and product expulsion follow, and the processive cycle repeats. Experimentally, *Tr*Cel7A may repeat this processive cycle 10−60 times before disassociating,[15,23−26] though this value is dependent both on the substrate, the presence of endoglucanases (EGs), and the method of calculating processivity. Molecular simulation has provided insight into some of these individual steps, including the initial threading of a cellulose chain (advancing from site −7 to −5),[27,28] the relationship between oligosaccharide binding free energy and processivity,[29] pyranose ring distortion,[21] hydrolysis,[16,30] and product expulsion.[31,32] Docking studies have also rationalized the forward processive motion of CBHs on the basis of a net force imparted by the enzyme in the direction of motion.[33,34]

Here, we employ molecular dynamics (MD) simulations and free energy calculations in *Tr*Cel7A with the aim of elucidating the carbohydrate−protein interactions that drive polysaccharide translocation in enzymes. We compute free energy barriers for both the "processive motion" chain threading and the "catalytic activation" steps that bridge the product expulsion and hydrolysis steps. Comparison with past work on the other elementary steps allows for determination of the rate-limiting step in the processive cycle. More broadly, this work may serve as a roadmap for discovery of the molecular underpinnings of polysaccharide translocation in a broad array of proteins that synthesize, modify, bind, and deconstruct natural polymers.

METHODS

Simulation Setup. We aimed to elucidate the discrete processive steps shown in Figure 2 by computing free energy surfaces for processive motion and catalytic activation. To obtain equilibrated

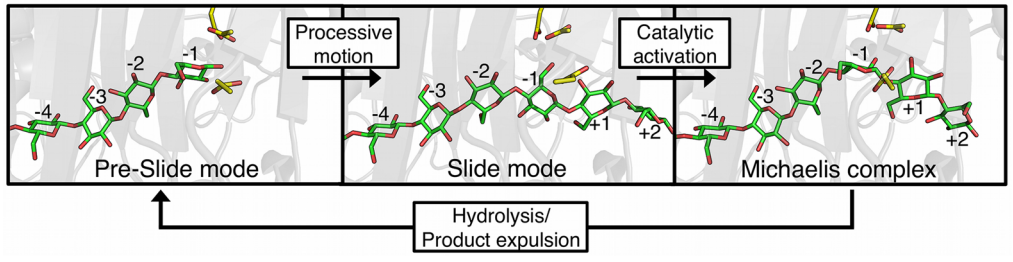


Figure 2. *Tr*Cel7A processive cycle. Structural and computational evidence indicates that the processive cycle proceeds from the Pre-Slide mode (following either the initial chain loading or product expulsion) to Slide mode by advancing by one cellobiose unit and then to the Michaelis complex via a chain twist and −1 sugar distortion.[20] Hydrolysis and product expulsion complete the processive cycle. Each configuration shown is the product of a 500 ns MD simulation. The catalytic residues (Glu212, Asp214, and Glu217) are shown in yellow and red sticks for reference.

dx.doi.org/10.1021/ja504074g | *J. Am. Chem. Soc.* 2014, 136, 8810−8819

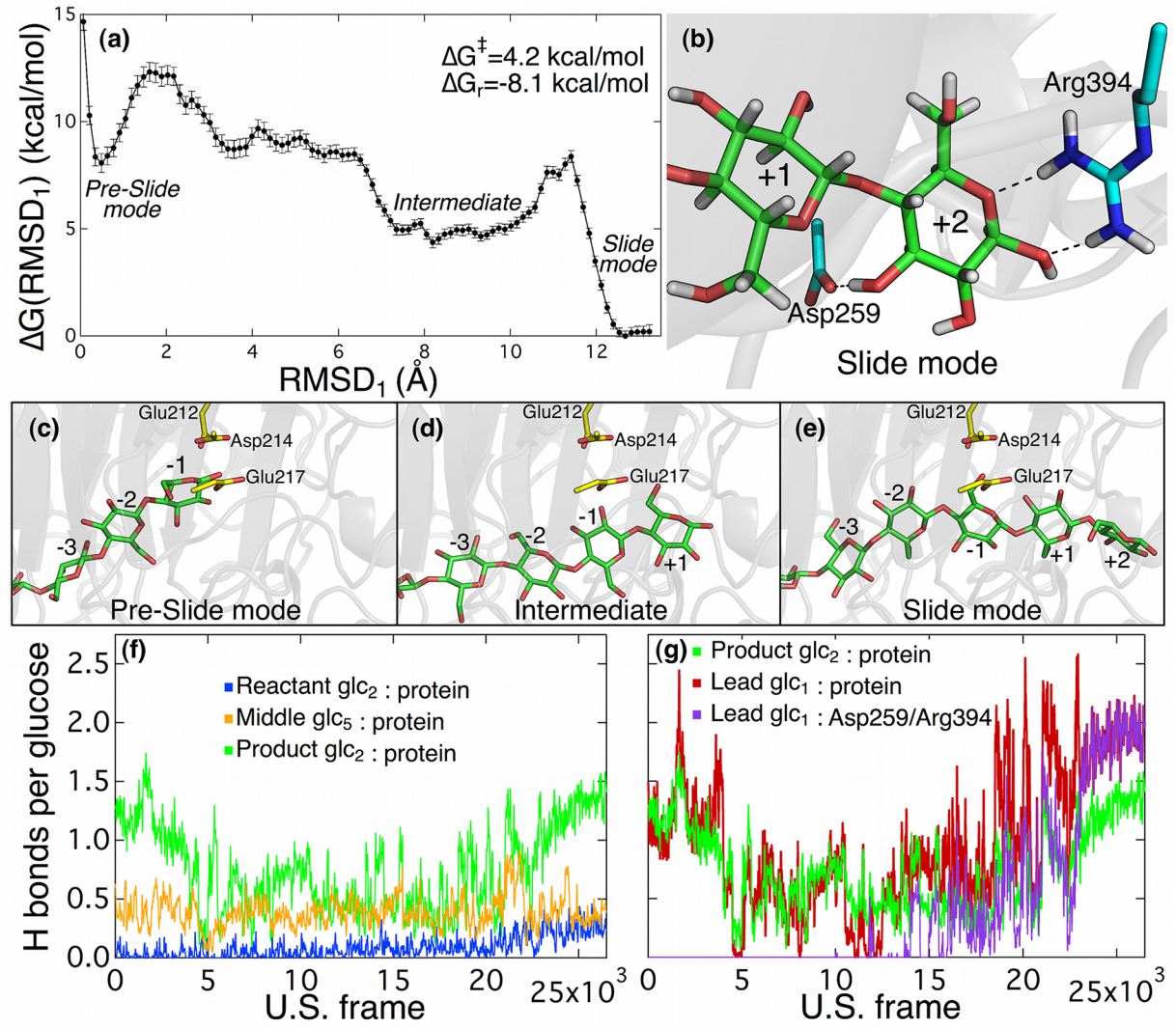


Figure 3. Details of processive motion. (a) The PMF for the first processive step ("processive motion" in Figure 2) reveals two barriers of comparable height, each about 4 kcal/mol, and an overall stabilization of about 8 kcal/mol. (b) The end point of the processive motion (i.e., "Slide mode") is stabilized by strong hydrogen bonds between the leading glucosyl ring and Asp259 and Arg394 (shown in cyan and red sticks). (c) Pre-Slide mode occurs immediately after initial chain threading and also after the cellobiose product is expelled within the processive cycle. (d) A snapshot of a molecular configuration midway between Pre-Slide mode and Slide mode, representative of the intermediate state seen in the PMF. (e) Slide mode is the result of advancing a full cellobiose unit from Pre-Slide mode. (f) The number of hydrogen bonds formed per glucosyl unit is highest for the leading two glucosyl residues, those that fill the product sites in Slide mode. (g) The leading glucosyl residue forms a disproportionately high number of hydrogen bonds with the enzyme, and these can be almost entirely accounted for by Asp259 and Arg394. In (c–e), catalytic residues (Glu212, Asp214, and Glu217) are shown in yellow and red sticks.

input structures in each basin, we first ran unrestrained MD simulations of the three structures depicted in Figure 2 of 500 ns duration. These provide insight into the dynamics at these key stages of the processive cycle and inputs for umbrella sampling (US) simulations to compute potentials of mean force (PMF). Briefly, the starting point for Pre-Slide mode comes from the product of a path sampling simulation for the second catalytic step, deglycosylation,[16] from which the cellobiose product has been removed (the starting structure for the path sampling study was the theoretical model of the Michaelis complex, PDB code 8CEL).[7] The Slide mode configuration features the *Tr*Cel7A crystal structure with two cellotetraose molecules in a noncatalytically active state (PDB code 5CEL)[7] with an additional glucose monomer modeled into the vacant −3 site and bonded to the adjacent glucose units (Figure S1a). The Michaelis complex starting point is the recently reported crystal structure (PDB code 4C4C).[16] The protein and substrate are solvated in an equilibrated, cubic box ∼80 × 80 × 80 Å^3 of TIP3P[35,36] water molecules. The CHARM force field with CMAP correction was used to describe the enzymes,[37−39] and the CHARMM C35 force field described the carbohydrates.[40−42] Overall charge neutrality is achieved by adding Na^+ ions to solution. The total system size is on the order of 50 000 atoms. All unrestrained MD simulations are performed in the isothermal−isobaric ensemble (*NpT*) at 300 K and 1.0 bar in the molecular simulation program NAMD.[43] Further simulation details are given in the SI.

Umbrella Sampling. We perform US along RMSD-based coordinates utilizing the Amber12 utility "targeted MD".[44] For the first step in Figure 2 ("processive motion"), we utilize the RMSD of the ring atoms and glycosidic oxygen molecules of the product side cellobiose unit (those that fill the +1/+2 sites in Slide mode) between Slide and Pre-Slide mode. The input configurations for these US windows come from "pulling" the cellononaose chain (via the coordinate just described) from Slide mode to Pre-Slide mode with all of the protein heavy atoms highly restrained (both the starting and target structures have been equilibrated for 500 ns with unrestrained MD). For the "catalytic activation" PMF, the coordinate is the RMSD of the ring atoms of the −1 glucosyl residue between the Michaelis

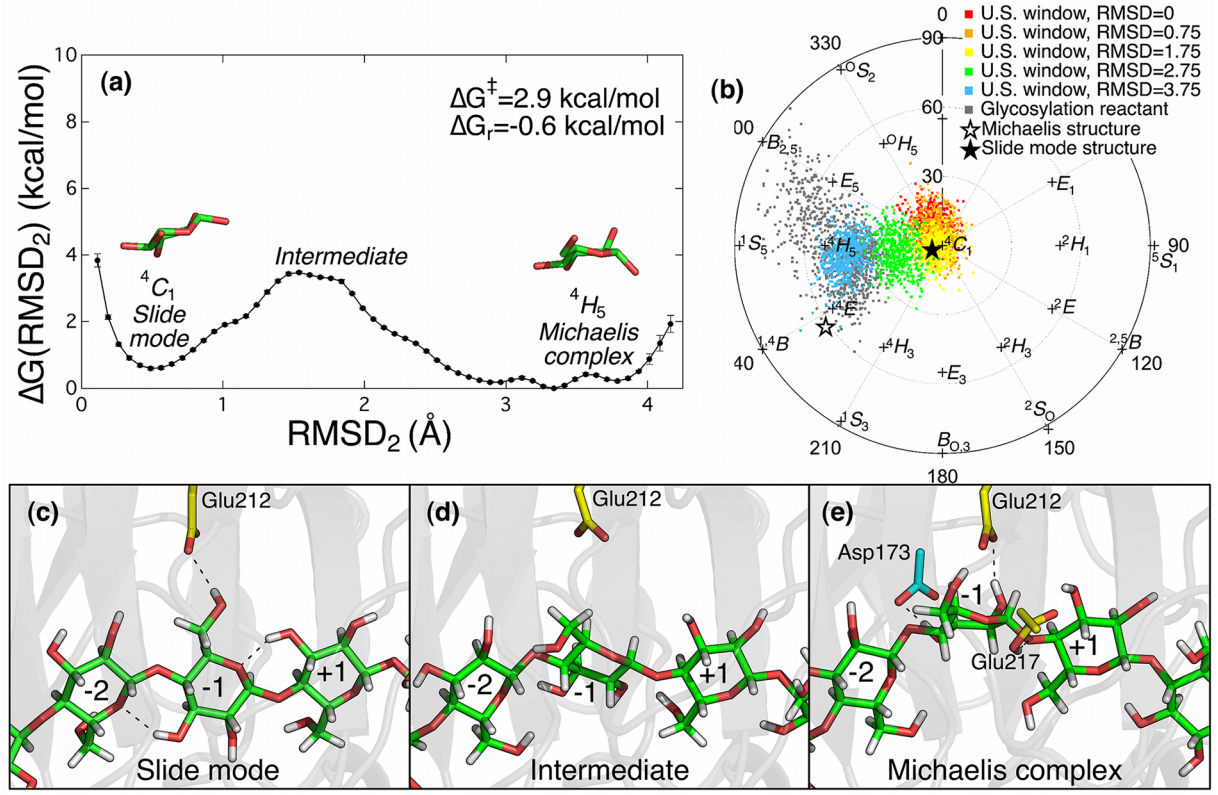


Figure 4. Details of catalytic activation. (a) The PMF for the second processive step ("catalytic activation" in Figure 2) reveals a slightly smaller barrier than that for processive motion and essentially equal stability in Slide mode and the Michaelis complex. (b) The Cremer–Pople parameters detailing the −1 ring puckering progression are shown for a selection of US windows. The "glycosylation reactant" is from a 1 ns simulation of the Michaelis complex from the *Tr*Cel7A path sampling study by Knott et al.[16] The "Michaelis structure" represents the *Tr*Cel7A Michaelis complex crystal structure (PDB code 4C4C),[16] and the "Slide mode structure" represents PDB code 5CEL.[7] (c) The −1 sugar ring is stabilized in Slide mode by a hydrogen bond between the nucleophile Glu212 and the methoxy group from the −1 sugar ring as well as two intracellulose bonds that are completely lost in the Michaelis complex. (d) Midway between Slide mode and the Michaelis complex, the −1 sugar ring is intermediate between $^4C_1$ and $^4H_5$ (see (b)), and many of the stabilizing interactions seen in Slide mode and the Michaelis complex are absent. (e) The Michaelis complex is stabilized by hydrogen bonds between the −1 sugar and catalytic acid/base Glu217, catalytic nucleophile Glu212, and Asp173. Glu217 and Asp173 have been omitted from panels (c) and (d) for clarity. In (c–e), catalytic residues (Glu212 and Glu217) are shown in yellow and red sticks, and Asp173 is shown in (e) in cyan and red sticks.

complex and Slide mode. The input configurations for these US windows come from taking selected configurations from an unrestrained MD simulation of the Michaelis complex (PDB code 4C4C) that spontaneously transitioned to Slide mode.

The PMF for processive motion was constructed from the final 10 ns of 40 ns in each of 53 overlapping US windows. The PMF for catalytic activation was constructed from the final 10 ns of 34 ns in each of 20 US windows.

RESULTS

Processive Motion. The processive motion PMF shown in Figure 3a reveals two barriers of approximately 4 kcal/mol between the Pre-Slide and Slide modes. This movement brings two glucosyl units from solution into the binding tunnel, thus filling the product binding sites (+1/+2), giving a significant stabilization of 8.1 kcal/mol.

The Slide mode configurations from the 500 ns MD simulation, processive motion US ($RMSD_1$ ~12.5 Å in Figure 3a), and catalytic activation US ($RMSD_2$ ~0.5 Å in Figure 4a) are all consistent (Figure S1b). In addition, these configurations agree with the ligand position in the −7 to −4 subsites from the crystal structure (PDB code 5CEL).[7] The ligand in subsites −2 to +2 is shifted about 1.3 Å toward the tunnel exit due to filling of the −3 subsite, which the crystal structure lacks (Figure S1a).

Further analysis of the US simulations reveals molecular details of the carbohydrate–protein interactions that drive processive motion for *Tr*Cel7A and may be relevant for processive enzymes of wide-ranging functions. This analysis points to the particularly strong binding of the carbohydrate chain in the product sites as the "driving force" for cellulose chain processivity (Figures 3b, 3f, and S2). Structural,[7,15,45] mutational,[15] and computational[29,31,34] studies have noted the interactions between polar side chains and the product site glucosyl residues in GH7 CBHs. Consistent with this, we find the number of hydrogen-bond interactions with the protein per glucosyl ring is much higher for the leading two rings than either for the middle five rings or the trailing two rings (Figure 3f; also see Figure S3a–c for hydrogen bonding with water and with adjacent glucosyl rings). The leading glucosyl ring, in particular, has a disproportionate number of hydrogen bonds with the enzyme (Figure 3g), which can be almost entirely accounted for by Arg394 and Asp259 in Slide mode (Figure 3b,g).

Structural evidence has shown that *Tr*Cel7A Arg394 forms bidentate hydrogen bonds to the +2 O6 and O1 hydroxyl groups,[7,12] allowing for the recognition of the incoming reducing end of the substrate.[7] In these structures, however, the distance from Arg394 to the +2 O6 hydroxyl is only very

dx.doi.org/10.1021/ja504074g | J. Am. Chem. Soc. 2014, 136, 8810–8819

slightly shorter than the distance to the O5 ring oxygen (an observation that was noted for the equivalent residue Arg391 in *Pc*Cel7D structures with various disaccharides).[45] Long-lived hydrogen bonds between Arg394 and the O5 and O1 atoms in the +2 site are formed in our simulations (Figure 3b,g).

The importance of Asp259 in product binding has been previously noted for *Tr*Cel7A structures complexed with the inhibitor IBTC[11] and with the natural product cellobiose[12] and for *Pc*Cel7D in complex with various disaccharides.[45] The backbone oxygen from *Tr*Cel7A Asp259 hydrogen bonds with the O2 hydroxyl of the +1 sugar ring (an interaction that is conserved in *Pc*Cel7D with Asp248[45]). Our processive motion US simulations indicate that this is the most persistent protein−carbohydrate hydrogen bond for the penultimate glucosyl ring that fills the +1 site.

We also find another key interaction involving Asp259 that is less obvious in the crystal structures: the carboxylate group of Asp259 hydrogen bonds to the C3 hydroxyl of the sugar ring in site +2 once the ligand is in Slide mode. Prior to this, Arg251 and Asp259 form a salt bridge (noted for the equivalent residues in *Pc*Cel7D by Ubhayasekera et al.).[45] Although the side chain conformation tends to be different in *Pc*Cel7D than in *Tr*Cel7A crystal structures, the flexibility of this arginine has been noted,[45] and the Arg guanido and Asp carboxylate groups are within 3 Å of each other in various structures of *Tr*Cel7A[7,12,16] and *Pc*Cel7D.[45] A computational study found that mutating either of these residues resulted in weaker binding of the cellobiose product.[31] The R251A mutant gave only slighter weaker binding by 1.3 kcal/mol (−14.4 versus −13.1 kcal/mol), whereas the D259A mutant gave significantly weaker binding by 8.4 kcal/mol (−14.4 versus −6.0 kcal/mol).[31]

The structure of the enzyme was also examined during processive motion to determine if any large-scale enzyme motions correlate with processivity. We were unable to identify behavior of this type that differed significantly from unrestrained MD simulations of *Tr*Cel7A, suggesting that significant structural changes do not occur in the enzyme during processivity.

Catalytic Activation. The PMF for catalytic activation is shown in Figure 4a. The barrier for catalytic activation (formation of the Michaelis complex) is lower than the two barriers for processive motion (Figure 3a) at 2.9 kcal/mol. The Slide mode and Michaelis complex configurations are of essentially equal stability with a ΔG_r of −0.6 kcal/mol. The low barrier and comparable stability are consistent with our observation that an unrestrained MD simulation of the Michaelis complex (PDB code 4C4C) spontaneously transitioned to Slide mode on the time scale of 500 ns (from which we extracted the starting configurations for catalytic activation US).

It has been noted that when the reducing end of a cellulose chain passes the catalytic center and enters the product binding sites (i.e., during processive motion), it still has room for all of the glucosyl rings to maintain a chair conformation.[20,34,45] In order for the nucleophile to access the anomeric carbon of the −1 sugar, the chain must rotate and this sugar must be distorted into a higher-energy nonchair configuration.[20,34,45] The progression for this glucosyl ring distortion in the catalytic activation US simulations is shown in Figure 4b.

The hydrogen bonds formed by the −1 glucosyl ring with the enzyme are central to this rotation and ring distortion. In Slide mode, the O6 hydroxyl at the −1 site hydrogen bonds with nucleophile Glu212 (Figures 4c and S4a). This interaction is present in the *Tr*Cel7A E212Q crystal structure with two cellotetraose molecules (PDB code 5CEL), though the relevant carboxylate oxygen is replaced by nitrogen.[7] Further stabilization is provided by intrachain hydrogen bonds from the +1 C3 hydroxyl to the −1 ring oxygen O5 and from the −1 C3 hydroxyl to the −2 ring oxygen O5 (Figures 4c and S4b).

Most of the interactions that stabilize the −1 sugar conformation and orientation in Slide mode and the Michaelis complex are lost midway along the transition (Figure 4d). Also, the −1 ring pucker adopts a conformation that is midway between $^4C_1$ and $^4H_5$ (Figure 4b).

In the Michaelis complex the −1 glucosyl ring is distorted into a $^4H_5$ half-chair conformation (Figure 4b,e). This distortion is stabilized by hydrogen bonds to Asp173 (with C3 hydroxyl), nucleophile Glu212 (with C2 hydroxyl), and catalytic acid/base Glu217 (with the glycosidic oxygen and the O5 ring oxygen), shown in Figure 4e and quantified in Figure S4a. The nucleophile is further coordinated by Asp214 and Ser174 (Figure S5), consistent with the interactions that are present in the crystal structures.[11,12] The hydrogen bonds made by Glu212 with these two residues persist along the entire catalytic activation process from Slide mode to the Michaelis complex. The intrachain hydrogen bonds present in Slide mode are completely lost in the Michaelis complex (Figure S4b).

Role of Aromatic Residues in Processivity. Lastly, structural studies of GHs have noted the lining of substrate binding tunnels with aromatic residues and hypothesized that they facilitate carbohydrate processivity.[7,11,33,46,47] We have examined the role of these aromatic residues in processivity across the US windows. The four primary aromatic residues that interact with the substrate in *Tr*Cel7A (Trp38, Trp40, Trp367, and Trp376) are shown to be quite stable even as the ligand changes position substantially, throughout both processive motion (Figure 5a) and catalytic activation (Figure 5b). In addition, our analysis indicates a relatively low and constant interaction energy between aromatic residues and the substrate (Figures 5c,d and S6). These dynamical indications of aromatic stability are consistent with what has previously been shown with static crystal structures. For example, *Tr*Cel7A structures bound with a variety of ligands have shown a remarkable consistency in the position of the tryptophan residues (Figure 5e,f), discussed further in the next section.

■ DISCUSSION AND CONCLUSIONS

In the present work, we have considered the steps of the processive cycle of *Tr*Cel7A that occur in between product expulsion and hydrolysis (Figure 2). We have termed the first of these steps "processive motion", wherein the cellulose chain advances by one cellobiose unit, filling the product sites. "Catalytic activation" follows, wherein the −1 sugar ring is distorted into a nonchair conformation and rotates up into the active site. Hydrolysis[16,30] and product expulsion[31,32] in *Tr*Cel7A, as well as cellulose chain decrystallization,[48,49] have been considered previously. In what follows, we discuss our findings for processive motion followed by catalytic activation, relate these to previous work, and consider the potential scope of their applicability in carbohydrate-active enzymes.

The rate-limiting step in the processive cycle of CBHs is still a matter of debate. Decrystallization of a cellulose chain and its processivity into the cellulase active site has been proposed as the rate-limiting step in processive enzymatic hydrolysis and a primary candidate for protein engineering efforts for producing

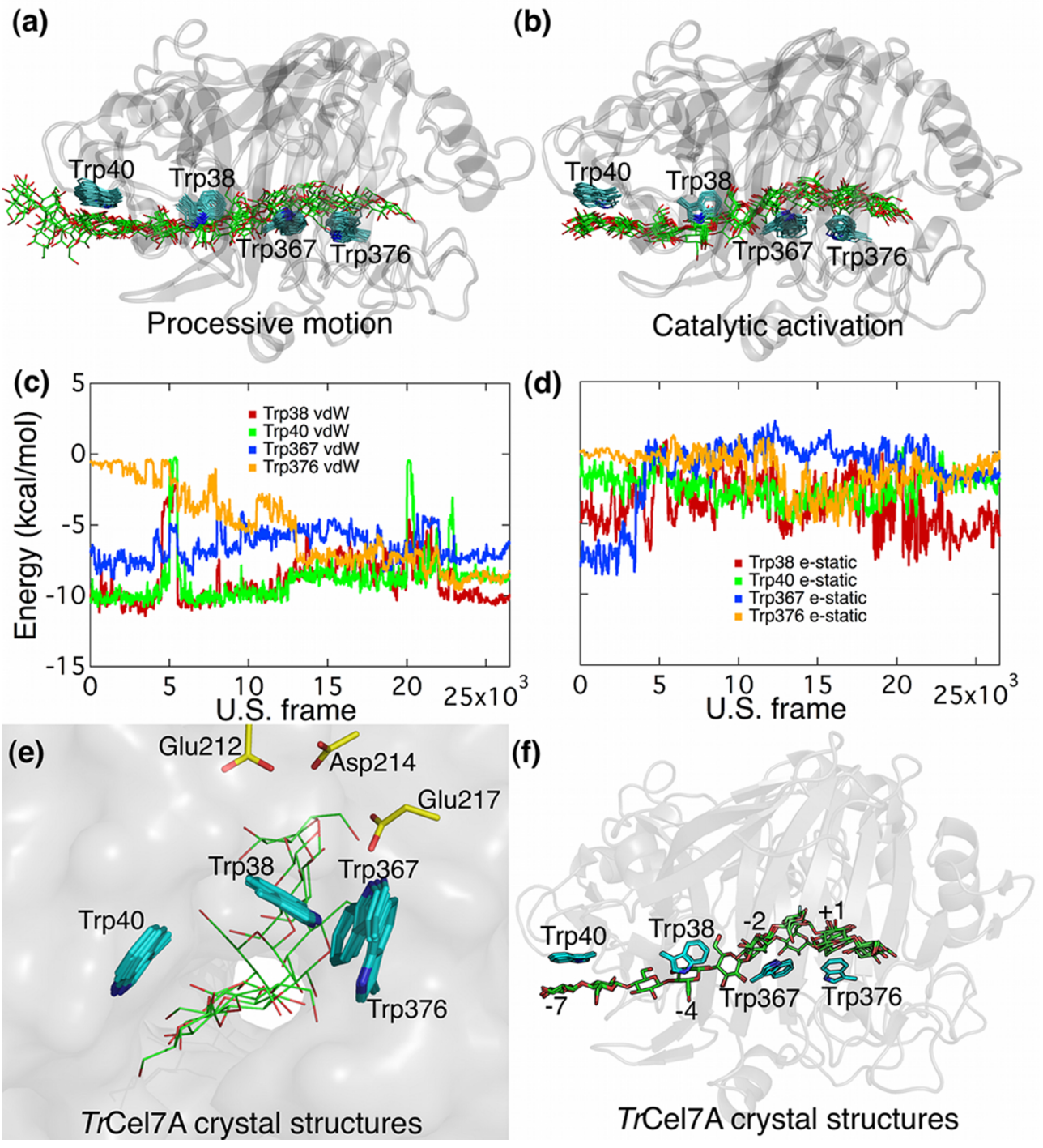


Figure 5. The stability of key aromatic residues in crystal structures and during cellulose processivity. (a) The stability of the aromatic residues during umbrella sampling simulations of processive motion and (b) catalytic activation. For (a) and (b), the cluster frequency for the aromatic residues is once per umbrella sampling window; for the ligand, eight configurations are shown, evenly spaced along the process. (c) The van der Waals and (d) electrostatic contributions of the interaction energy between the key tryptophan residues and the cellononaose chain during processive motion. A similar analysis for the catalytic activation step is shown in Figure S6 (b) and (d). (e) Four tryptophan residues provide the tunnel shape required for the cellulose chain to "twist" and become catalytically activated. The tryptophan residues shown are for *Tr*Cel7A structures with PDB codes 3CEL ("primed" cellobiose[20] in +1/+2),[12] 4CEL (apo),[12] 5CEL (two cellotetraose molecules in −7 to −4 and −2 to +2 analogous to Slide mode),[7] 7CEL (cellohexaose in −7 to −2 and "unprimed" cellobiose[20] in +1/+2),[7] 4C4C (Michaelis complex),[16] and 4C4D (glycosyl−enzyme intermediate).[16] The protein "surface", catalytic triad (yellow and red sticks), and the ligand (green and red sticks) are from the Michaelis complex structure (PDB code 4C4C).[16] (f) Despite considerable variability in substrate binding, the location of the aromatic residues in *Tr*Cel7A crystal structures is remarkably consistent. The ligands and the tryptophans are for the same structures as in (e).

cheaper biofuels.[50,51] The rate constant for processive motion has also been noted as an important and under-characterized parameter that was recently suggested as the source of large differences in enzymatic activity on different cellulose polymorphs.[51] For CBHs acting in isolation, it has been found that the rate-limiting step is dissociation,[23,24,26] but

dx.doi.org/10.1021/ja504074g | J. Am. Chem. Soc. 2014, 136, 8810−8819

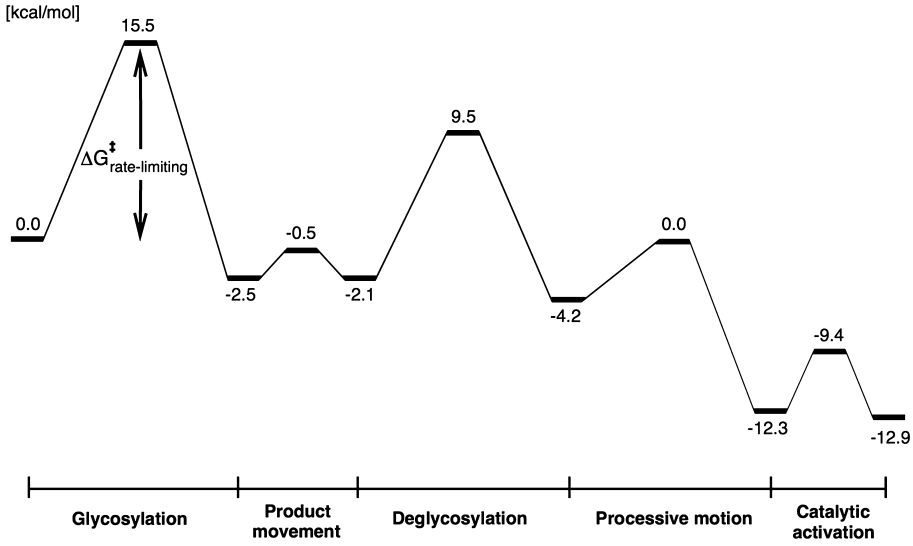


Figure 6. The free energy landscape of the hydrolytic/processive steps in *Tr*Cel7A. The barriers for glycosylation, product movement, and deglycosylation were computed previously by Knott et al.;[16] the barriers for the processive motion and catalytic activation are computed in the present work. These barriers represent all of the key steps in the cellobiohydrolase processive cycle (in between adsorption and desorption) with the exception of product expulsion, which has previously been experimentally ruled out as a rate-limiting factor in these enzymes.[52] Because product expulsion is not explicitly considered here, the overall free energy change for the processive cycle will not be as significant as illustrated.

processive velocity is limiting when EGs are present.[52] A high-speed AFM study comparing CBHs *Tr*Cel7A, *Pc*Cel7D, and *Pc*Cel7C (with no EGs present) concluded that the rate-limiting step in phosphoric acid-swollen cellulose degradation is the processive cycle (including processive motion, hydrolysis, and product expulsion) or dissociation based on the fact that CBH velocity was on the same order of magnitude as the hydrolytic activity and chain dissociation.[19] However, with crystalline cellulose, "picking up" a chain end was considered rate-limiting based on the difference in CBH velocity for cellulose III_I versus cellulose Iα, which differ primarily in the number of free chain ends available.[19]

In the present work, we are not accounting for cellulose decrystallization and thus are assuming that the processivity is not greatly affected by the interaction of the protein with the crystalline cellulose surface. The free energy required to abstract a single cellobiose unit from the cellulose surface is dependent upon the crystalline cellulose polymorph (Iα, Iβ, II, III_I) and the chain location (edge, middle, or corner of the crystal).[48] This quantity ranges from 2.2 kcal/mol (cellulose II edge chain) to 6.7 kcal/mol (cellulose Iβ middle chain); for a corner chain in cellulose Iβ, this number drops to 3.4 kcal/mol.[48] Even if this figure were added to the processive motion free energy barrier, this combined barrier would not approach the barrier for glycosylation.[16]

It should be noted that precise barrier heights, free energy differences, and molecular mechanisms require the elucidation of an accurate reaction coordinate (via an unbiased simulation technique such as transition path sampling[53,54]), which in turn enables kinetically meaningful calculation of free energy barriers and rates. However, neglecting differences in mobility along the reaction coordinate,[55,56] choosing any arbitrary path through a free energy landscape that connects the reactant and product basins (or even along a relevant, but unverified, choice of order parameter) will require traversing a higher barrier than progressing along the "true" reaction path (thus, the value in finding the minimum free energy path[57] for a process). Therefore, the free energy barrier heights that we compute for the processive motion and catalytic activation of *Tr*Cel7A (all less than 4.2 kcal/mol) can be considered upper bounds on the true barrier heights for these processes. In this way, we can conclude that these barriers do not approach the magnitudes of the hydrolytic barriers, which were computed along reaction coordinates that were precisely determined with unbiased path sampling and verified with the p_B histogram test.[16] This study found the free energy barriers for glycosylation and deglycosylation to be 15.5 and 11.6 kcal/mol, respectively (Figure 6).[16] This previous study also computed the kinetic prefactors and reaction rate constants, and the rate constant for glycosylation (the rate-limiting hydrolytic step) was shown[16] to be consistent with HS-AFM processive velocities.[9]

Jalak et al. have recently demonstrated that at optimal substrate/enzyme loadings (and with EGs present) the rate-limiting step in the overall processive cycle (including association, initial chain threading, and dissociation[20]) is the velocity of processive movement[52] (which involves the steps shown in Figure 2). Combined with past findings for cellulose decrystallization,[48,49] hydrolysis,[16] and cellobiose product expulsion[31,32] (which has been previously ruled out as the rate-limiting step in the processive cycle[52]), our findings here suggest that with synergistic endolytic enzymes present, the glycosylation reaction is the rate-limiting step in the enzymatic degradation of cellulose by cellobiohydrolases (Figure 6).

Within the CBH processive cycle, the only step not shown in Figure 6 is product expulsion. The most accurate computational estimate to date for the work to remove the cellobiose product from the +1/+2 sites is 11.2 kcal/mol.[31] However, this calculation may be an overestimate, given that the cellobiose product was expelled directly from the "unprimed" position (rather than the "primed").[20] This may reduce the work required to expel the product. Regardless, with product expulsion considered, the overall free energy change for the processive cycle is expected to still be downhill, albeit smaller in magnitude than −12.9 kcal/mol, as illustrated in Figure 6.

Our results indicate that the ligand-coordinating residues at the product binding sites play a central role in processive

dx.doi.org/10.1021/ja504074g | *J. Am. Chem. Soc.* 2014, 136, 8810−8819

motion, and this is supported by past free energy studies. The free energy change of −8.1 kcal/mol downhill we find for the processive motion agrees well with the results of Payne et al., who found that filling all nine binding sites of *Tr*Cel7A was favored by −11.1 ± 3.13 kcal/mol versus filling only the seven on the "substrate" side of the binding tunnel.[29] This figure is also in agreement with a previous determination of cellobiose binding in the product sites of *Tr*Cel7A by an independent method, which calculated this quantity as −11.2 ± 0.6 kcal/mol.[31] Another significant consequence of this particularly strong binding in the product sites is product inhibition, in which the binding of reaction products glucose or cellobiose retards the overall conversion rate of lignocellulose to glucose.[58] Dramatic differences in product inhibition between *Tr*Cel7A WT, *Tr*Cel7A with "exo" loop deletion, and *Pc*Cel7D were rationalized on the basis of different protein−carbohydrate interactions in the product sites, including Arg251.[15] The deletion mutant lacked this interaction, but *Pc*Cel7D maintains it (Arg240 in *Pc*Cel7D). This results in product inhibition that is significantly reduced for the deletion mutant, but only moderately reduced for *Pc*Cel7D (due to the loss of other interactions).[15]

Our results for the catalytic activation step warrant comparison with past studies of the free energy landscape for various ring puckering conformations of pyranose rings.[21,59−61] We find a −1 puckering progression from $^4C_1$ chair to $^4H_5$ half-chair (Figure 4b). In vacuum, the $^4C_1 \rightarrow {}^4H_5$ transition has been calculated as 8.4 kcal/mol uphill with high-level quantum mechanics (QM).[61] In the active site of *Tr*Cel7A, however, the similar $^4C_1 \rightarrow {}^4E$ transition has been calculated as −5.97 kcal/mol downhill; the $^4H_3$ conformation is 3.4 kcal/mol above the ^4E conformation, thus the $^4H_3$ is 2.57 kcal/mol more stable than the $^4C_1$ conformer.[21] The latter figure was computed with SCC-DFTB, which has been shown to be the most accurate of the semiempirical QM methods for ring puckering.[60]

Carbohydrate−aromatic interactions are a ubiquitous motif in all carbohydrate active enzymes,[4] including GHs, carbohydrate-binding modules, and polysaccharide-synthesizing enzymes. In glycoside hydrolases in particular, these have been considered important to carbohydrate processivity.[7,11,46,47] Mutating aromatic residues in processive cellulases[18,28,62,63] and chitinases[64,65] essentially destroys their processive ability and greatly diminishes catalytic activity on crystalline substrates, but actually increases it on amorphous or soluble substrates. Computational studies[27,66,67] have revealed molecular details of carbohydrate−aromatic interactions, indicating importance for chain acquisition,[27,28,66] product stabilization,[66] priming the substrate for catalysis via ring distortion,[66] and stabilizing the chain twist at the −2 subsite.[7,67] Previous theoretical treatment has indicated that carbohydrate−aromatic interactions in a maltoporin balance the strong hydrogen bonds, giving a smooth energy profile during processivity.[68] *Tr*Cel7A crystal structures suggest these interactions are nonspecific and primarily provide a platform for the glucosyl rings to Slide over.[7] Consistent with this, we propose that the primary function of these aromatic residues lining the binding tunnel is for tunnel shape as the carbohydrate-π stacking[4] "guides" the polysaccharide through the binding tunnel and contorts the chain to its catalytically primed state (Figure 5e).[7] This proposal is supported by the stability of these aromatic residues in *Tr*Cel7A crystal structures (Figure 5e,f). This is in spite of the fact that a wide variety of substrate poses have been captured, indicating different stages of the processive cycle. As noted above, aromatic stability is corroborated by our free energy simulations along the processive cycle (Figure 5a,b). In addition, the low interaction energy between aromatic residues and the substrate (Figures 5c,d and S6) indicate that these particular residues are not the primary "drivers" for forward processive motion.

Sequence alignment of all GH7 CBHs and EGs with solved crystal structures (7 CBHs[11,69−74] and 3 EGs[75−77]) indicates that the residues we have identified as important for processive motion are extremely well-conserved within the CBHs, but almost entirely absent in EGs (Figure S7). Arg 394 is found in all GH7 CBHs with solved structures, but is replaced by alanine in the three EGs. Asp259 is also found in all of these CBHs, with the exception of *Melanocarpus albomyces* Cel7B[70] (where it is replaced by asparagine). The three EGs have Asp259 replaced either by alanine (*Fusarium oxysporum* Cel7B[77] and *Tr*Cel7B[75]) or asparagine (*Humicola insolens* Cel7B[76]). The three key residues for catalytic activation (Glu212, Glu217, and Asp173) are completely conserved in all 10 enzymes. All GH7 CBHs with solved structures possess the four aromatic residues discussed above (Trp38, Trp40, Trp367, Trp376). The latter two of these (those on either side of the active site) are also present in the three EGs. The former two, however, are modified in the EGs. Trp38 and Trp40 are missing in *Fo*Cel7B and *Hin*Cel7B, though *Tr*Cel7B maintains aromatic interactions at both sites (Trp40 and Tyr38). Functionally, the intrinsic processivity of CBHs is much higher than that of EGs,[23] but GH7 CBHs and EGs employ a common catalytic mechanism.[78] Overall, the fact that residues important to processive motion are well-conserved in only CBHs, but those important for catalytic activation are also conserved in EGs is consistent with their well-established functional differences.

The kinetics of processivity that we have studied here in the context of a model cellobiohydrolase have implications for processivity in polysaccharide-active enzymes of wide-ranging functions and substrates, including DNA glycosylases[79] and endonucleases[80] that "scan" DNA looking for damage, helicases that separate the hydrogen-bonded double helix of DNA or RNA by coupling a chemical reaction (ATP hydrolysis) with directional movement along their polymeric substrate,[81] enzymes that degrade hyaluronan,[82] chitin,[64] cellulose,[9] DNA,[83] RNA,[84] proteins,[85] and peptidoglycan[86] as well as the enzymes that synthesize them.[5,87−90] While these enzymes certainly employ a diversity of mechanisms, common motifs emerge, including the role of tunnel shape and substrate enclosure for processivity, and the role of specific chemical (particularly electrostatic) interactions with the protein residues lining the tunnel for processivity.[3] One notable example of this is the recently solved structure of the *Rhodobacter sphaeroides* cellulose synthase that couples a chemical reaction (glycosyl transfer) with translocation of cellulose across the cell wall.[5] The binding tunnel accommodates 10 glucosyl rings and is lined with large aromatic residues and polar residues that interact directly with the hydroxyl groups of the glucan chain.[5,91] This supports the idea that the findings we have presented here regarding the processive cycle of a cellobiohydrolase, and the protein motifs that drive it, may well have broad relevance across many classes of enzymes.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Simulation details, including preparation of initial configurations for umbrella sampling, umbrella sampling procedures,

PMF calculation, and data analysis. Figures comparing Slide mode from crystal structure and simulations. Graphs showing the key specific hydrogen-bonding, electrostatic, and dispersive interactions along the processive motion and catalytic activation umbrella sampling profiles. Sequence alignment of the 10 GH7 enzymes with solved crystal structures (7 CBHs and 3 EGs). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

Jerry.Stahlberg@slu.se
Gregg.Beckham@nrel.gov

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

B.C.K. and G.T.B. thank the NREL Director's Fellowship Program for funding. M.F.C. and G.T.B. acknowledge funding from the Department of Energy BioEnergy Technologies Office and NREL Laboratory Directed Research and Development funds. The authors thank Prof. Christina M. Payne for creating the GH7 sequence alignment presented in the Supporting Information and Dr. Seonah Kim for creating the free energy profile figure (Figure 6).

REFERENCES

(1) Stern, R.; Jedrzejas, M. J. *Chem. Rev.* **2008**, *108*, 5061.
(2) Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B. *Nucleic Acids Res.* **2014**, *42*, D490.
(3) Breyer, W. A.; Matthews, B. W. *Protein Sci.* **2001**, *10*, 1699.
(4) Asensio, J. L.; Arda, A.; Canada, F. J.; Jimenez-Barbero, J. *Acc. Chem. Res.* **2013**, *46*, 946.
(5) Morgan, J. L. W.; Strumillo, J.; Zimmer, J. *Nature* **2013**, *493*, 181.
(6) Boraston, A. B.; Bolam, D. N.; Gilbert, H. J.; Davies, G. J. *Biochem. J.* **2004**, *382*, 769.
(7) Divne, C.; Stahlberg, J.; Teeri, T. T.; Jones, T. A. *J. Mol. Biol.* **1998**, *275*, 309.
(8) van Aalten, D. M. F.; Komander, D.; Synstad, B.; Gaseidnes, S.; Peter, M. G.; Eijsink, V. G. H. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8979.
(9) Igarashi, K.; Uchihashi, T.; Koivula, A.; Wada, M.; Kimura, S.; Okamoto, T.; Penttila, M.; Ando, T.; Samejima, M. *Science* **2011**, *333*, 1279.
(10) Chundawat, S. P. S.; Beckham, G. T.; Himmel, M. E.; Dale, B. E. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 121.
(11) Divne, C.; Stahlberg, J.; Reinikainen, T.; Ruohonen, L.; Pettersson, G.; Knowles, J. K. C.; Teeri, T. T.; Jones, T. A. *Science* **1994**, *265*, 524.
(12) Stahlberg, J.; Divne, C.; Koivula, A.; Piens, K.; Claeyssens, M.; Teeri, T. T.; Jones, T. A. *J. Mol. Biol.* **1996**, *264*, 337.
(13) Becker, D.; Braet, C.; Brumer, H.; Claeyssens, M.; Divne, C.; Fagerstrom, B. R.; Harris, M.; Jones, T. A.; Kleywegt, G. J.; Koivula, A.; Mahdi, S.; Piens, K.; Sinnott, M. L.; Stahlberg, J.; Teeri, T. T.; Underwood, M.; Wohlfahrt, G. *Biochem. J.* **2001**, *356*, 19.
(14) Stahlberg, J.; Henriksson, H.; Divne, C.; Isaksson, R.; Pettersson, G.; Johansson, G.; Jones, T. A. *J. Mol. Biol.* **2001**, *305*, 79.
(15) von Ossowski, I.; Stahlberg, J.; Koivula, A.; Piens, K.; Becker, D.; Boer, H.; Harle, R.; Harris, M.; Divne, C.; Mahdi, S.; Zhao, Y. X.; Driguez, H.; Claeyssens, M.; Sinnott, M. L.; Teeri, T. T. *J. Mol. Biol.* **2003**, *333*, 817.
(16) Knott, B. C.; Haddad Momeni, M.; Crowley, M. F.; Mackenzie, L. F.; Götz, A. W.; Sandgren, M.; Withers, S. G.; Ståhlberg, J.; Beckham, G. T. *J. Am. Chem. Soc.* **2014**, *136*, 321.
(17) Himmel, M. E.; Ding, S. Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. *Science* **2007**, *315*, 804.
(18) Igarashi, K.; Koivula, A.; Wada, M.; Kimura, S.; Penttila, M.; Samejima, M. *J. Biol. Chem.* **2009**, *284*, 36186.
(19) Nakamura, A.; Watanabe, H.; Ishida, T.; Uchihashi, T.; Wada, M.; Ando, T.; Igarashi, K.; Samejima, M. *J. Am. Chem. Soc.* **2014**, *136*, 4584.
(20) Beckham, G. T.; Ståhlberg, J.; Knott, B. C.; Himmel, M. E.; Crowley, M. F.; Sandgren, M.; Sørlie, M.; Payne, C. M. *Curr. Opin. Biotechnol.* **2014**, *27*, 96.
(21) Barnett, C. B.; Wilkinson, K. A.; Naidoo, K. J. *J. Am. Chem. Soc.* **2010**, *132*, 12800.
(22) Vocadlo, D. J.; Davies, G. J. *Curr. Opin. Chem. Biol.* **2008**, *12*, 539.
(23) Kurasin, M.; Valjamae, P. *J. Biol. Chem.* **2011**, *286*, 169.
(24) Cruys-Bagger, N.; Elmerdahl, J.; Praestgaard, E.; Tatsumi, H.; Spodsberg, N.; Borch, K.; Westh, P. *J. Biol. Chem.* **2012**, *287*, 18451.
(25) Cruys-Bagger, N.; Elmerdahl, J.; Praestgaard, E.; Borch, K.; Westh, P. *FEBS J.* **2013**, *280*, 3952.
(26) Cruys-Bagger, N.; Tatsumi, H.; Ren, G. R.; Borch, K.; Westh, P. *Biochemistry-US* **2013**, *52*, 8938.
(27) GhattyVenkataKrishna, P. K.; Alekozai, E. M.; Beckham, G. T.; Schulz, R.; Crowley, M. F.; Uberbacher, E. C.; Cheng, X. L. *Biophys. J.* **2013**, *104*, 904.
(28) Nakamura, A.; Tsukada, T.; Auer, S.; Furuta, T.; Wada, M.; Koivula, A.; Igarashi, K.; Samejima, M. *J. Biol. Chem.* **2013**, *288*, 13503.
(29) Payne, C. M.; Jiang, W.; Shirts, M. R.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. *J. Am. Chem. Soc.* **2013**, *135*, 18831.
(30) Barnett, C. B.; Wilkinson, K. A.; Naidoo, K. J. *J. Am. Chem. Soc.* **2011**, *133*, 19474.
(31) Bu, L. T.; Beckham, G. T.; Shirts, M. R.; Nimlos, M. R.; Adney, W. S.; Himmel, M. E.; Crowley, M. F. *J. Biol. Chem.* **2011**, *286*, 18161.
(32) Bu, L. T.; Nimlos, M. R.; Shirts, M. R.; Ståhlberg, J.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. *J. Biol. Chem.* **2012**, *287*, 24807.
(33) Mertz, B.; Hill, A. D.; Mulakala, C.; Reilly, P. J. *Biopolymers* **2007**, *87*, 249.
(34) Mulakala, C.; Reilly, P. J. *Proteins* **2005**, *60*, 598.
(35) Durell, S. R.; Brooks, B. R.; Bennaim, A. *J. Phys. Chem.* **1994**, *98*, 2198.
(36) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(37) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545.
(38) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
(39) Mackerell, A. D.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400.
(40) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D. *J. Comput. Chem.* **2008**, *29*, 2543.
(41) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. *J. Chem. Theory Comput.* **2009**, *5*, 2353.
(42) Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W.; MacKerell, A. D. *J. Chem. Theory Comput.* **2011**, *7*, 3162.
(43) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
(44) Case, D. A.; Darden, T.A.; T.E. Cheatham, I.; Simmerling, C.L.; Wang, J.; Duke, R.E.; Luo, R.; Walker, R.C.; Zhang, W.; Merz, K.M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A.W.;

I. Kolossváry; Wong, K.F.; Paesani, F.; Vanicek, J.; Wolf, R.M.; Liu, J.; Wu, X.; Brozell, S.R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D.R.; Mathews, D.H.; Seetin, M.G.; R. Salomon-Ferrer; Sagui, C.; Babin, V.; ; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *Amber 12*; University of California: San Francisco, 2012.
(45) Ubhayasekera, W.; Munoz, I. G.; Vasella, A.; Stahlberg, J.; Mowbray, S. L. *FEBS J.* **2005**, *272*, 1952.
(46) Rouvinen, J.; Bergfors, T.; Teeri, T.; Knowles, J. K. C.; Jones, T. A. *Science* **1990**, *249*, 380.
(47) van Aalten, D. M. F.; Synstad, B.; Brurberg, M. B.; Hough, E.; Riise, B. W.; Eijsink, V. G. H.; Wierenga, R. K. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5842.
(48) Beckham, G. T.; Matthews, J. F.; Peters, B.; Bomble, Y. J.; Himmel, M. E.; Crowley, M. F. *J. Phys. Chem. B* **2011**, *115*, 4118.
(49) Payne, C. M.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. *J. Phys. Chem. Lett.* **2011**, *2*, 1546.
(50) Wilson, D. B. *Curr. Opin. Biotechnol.* **2009**, *20*, 295.
(51) Gao, D. H.; Chundawat, S. P. S.; Sethi, A.; Balan, V.; Gnanakaran, S.; Dale, B. E. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 10922.
(52) Jalak, J.; Kurasin, M.; Teugjas, H.; Valjamae, P. *J. Biol. Chem.* **2012**, *287*, 28802.
(53) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291.
(54) Peters, B. *Mol. Simul.* **2010**, *36*, 1265.
(55) Knott, B. C.; Duff, N.; Doherty, M. F.; Peters, B. *J. Chem. Phys.* **2009**, *131*, 224112.
(56) Peters, B. *J. Chem. Phys.* **2009**, *131*, 244103.
(57) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. *J. Chem. Phys.* **2006**, *125*, 024106.
(58) Andric, P.; Meyer, A. S.; Jensen, P. A.; Dam-Johansen, K. *Biotechnol. Adv.* **2010**, *28*, 308.
(59) Biarnes, X.; Ardevol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2007**, *129*, 10686.
(60) Barnett, C. B.; Naidoo, K. J. *J. Phys. Chem. B* **2010**, *114*, 17142.
(61) Mayes, H. B.; Broadbelt, L. J.; Beckham, G. T. *J. Am. Chem. Soc.* **2013**, *136*, 1008.
(62) Koivula, A.; Kinnari, T.; Harjunpaa, V.; Ruohonen, L.; Teleman, A.; Drakenberg, T.; Rouvinen, J.; Jones, T. A.; Teeri, T. T. *FEBS Lett.* **1998**, *429*, 341.
(63) Vuong, T. V.; Wilson, D. B. *Appl. Environ. Microbiol.* **2009**, *75*, 6655.
(64) Horn, S. J.; Sikorski, P.; Cederkvist, J. B.; Vaaje-Kolstad, G.; Sorlie, M.; Synstad, B.; Vriend, G.; Varum, K. M.; Eijsink, V. G. H. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18089.
(65) Zakariassen, H.; Aam, B. B.; Horn, S. J.; Varum, K. M.; Sorlie, M.; Eijsink, V. G. H. *J. Biol. Chem.* **2009**, *284*, 10610.
(66) Payne, C. M.; Bomble, Y.; Taylor, C. B.; McCabe, C.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. *J. Biol. Chem.* **2011**, *286*, 41028.
(67) Taylor, C. B.; Payne, C. M.; Himmel, M. E.; Crowley, M. F.; McCabe, C.; Beckham, G. T. *J. Phys. Chem. B* **2013**, *117*, 4924.
(68) Meyer, J. E. W.; Schulz, G. E. *Protein Sci.* **1997**, *6*, 1084.
(69) Munoz, I. G.; Ubhayasekera, W.; Henriksson, H.; Szabo, I.; Pettersson, G.; Johansson, G.; Mowbray, S. L.; Stahlberg, J. *J. Mol. Biol.* **2001**, *314*, 1097.
(70) Parkkinen, T.; Koivula, A.; Vehmaanpera, J.; Rouvinen, J. *Protein Sci.* **2008**, *17*, 1383.
(71) Kern, M.; McGeehan, J. E.; Streeter, S. D.; Martin, R. N. A.; Besser, K.; Elias, L.; Eborall, W.; Malyon, G. P.; Payne, C. M.; Himmel, M. E.; Schnorr, K.; Beckham, G. T.; Cragg, S. M.; Bruce, N. C.; McQueen-Mason, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 10189.
(72) Momeni, M. H.; Payne, C. M.; Hansson, H.; Mikkelsen, N. E.; Svedberg, J.; Engstrom, A.; Sandgren, M.; Beckham, G. T.; Stahlberg, J. *J. Biol. Chem.* **2013**, *288*, 5861.
(73) Grassick, A.; Murray, P. G.; Thompson, R.; Collins, C. M.; Byrnes, L.; Birrane, G.; Higgins, T. M.; Tuohy, M. G. *Eur. J. Biochem.* **2004**, *271*, 4495.
(74) Textor, L. C.; Colussi, F.; Silveira, R. L.; Serpa, V.; de Mello, B. L.; Muniz, J. R. C.; Squina, F. M.; Pereira, N.; Skaf, M. S.; Polikarpov, I. *FEBS J.* **2013**, *280*, 56.
(75) Kleywegt, G. J.; Zou, J. Y.; Divne, C.; Davies, G. J.; Sinning, I.; Stahlberg, J.; Reinikainen, T.; Srisodsuk, M.; Teeri, T. T.; Jones, T. A. *J. Mol. Biol.* **1997**, *272*, 383.
(76) Mackenzie, L. F.; Sulzenbacher, G.; Divne, C.; Jones, T. A.; Woldike, H. F.; Schulein, M.; Withers, S. G.; Davies, G. J. *Biochem. J.* **1998**, *335*, 409.
(77) Sulzenbacher, G.; Driguez, H.; Henrissat, B.; Schulein, M.; Davies, G. J. *Biochemistry* **1996**, *35*, 15280.
(78) Schou, C.; Rasmussen, G.; Kaltoft, M. B.; Henrissat, B.; Schulein, M. *Eur. J. Biochem.* **1993**, *217*, 947.
(79) Higley, M.; Lloyd, R. S. *Mutat. Res.* **1993**, *294*, 109.
(80) Carey, D. C.; Strauss, P. R. *Biochemistry* **1999**, *38*, 16553.
(81) Velankar, S. S.; Soultanas, P.; Dillingham, M. S.; Subramanya, H. S.; Wigley, D. B. *Cell* **1999**, *97*, 75.
(82) Pritchard, D. G.; Lin, B.; Willingham, T. R.; Baker, J. R. *Arch. Biochem. Biophys.* **1994**, *315*, 431.
(83) Breyer, W. A.; Matthews, B. W. *Nat. Struct. Biol.* **2000**, *7*, 1125.
(84) Symmons, M. F.; Jones, G. H.; Luisi, B. F. *Structure* **2000**, *8*, 1215.
(85) Akopian, T. N.; Kisselev, A. F.; Goldberg, A. L. *J. Biol. Chem.* **1997**, *272*, 1791.
(86) van Asselt, E. J.; Dijkstra, A. J.; Kalk, K. H.; Takacs, B.; Keck, W.; Dijkstra, B. W. *Struct. Fold Des.* **1999**, *7*, 1167.
(87) Muller, D. K.; Martin, C. T.; Coleman, J. E. *Biochemistry* **1988**, *27*, 5763.
(88) Doublie, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. *Nature* **1998**, *391*, 251.
(89) Merzendorfer, H. *J. Comp. Physiol. B* **2006**, *176*, 1.
(90) Hubbard, C.; McNamara, J. T.; Azumaya, C.; Patel, M. S.; Zimmer, J. *J. Mol. Biol.* **2012**, *418*, 21.
(91) Slabaugh, E.; Davis, J. K.; Haigler, C. H.; Yingling, Y. G.; Zimmer, J. *Trends Plant Sci.* **2014**, *19*, 99.