Towards a molecular-level theory of carbohydrate processivity in glycoside hydrolases

Gregg T. Beckham[1,2,*], Jerry Ståhlberg[3,*], Brandon C. Knott[1], Michael E. Himmel[4], Michael F. Crowley[4], Mats Sandgren[3], Morten Sørlie[5], Christina M. Payne[6,7,*]

1. National Bioenergy Center, National Renewable Energy Laboratory, Golden CO, 80401, United States
2. Department of Chemical Engineering, Colorado School of Mines, Golden CO, 80401, United States
3. Swedish University of Agricultural Sciences, SE 75007, Uppsala, Sweden
4. Biosciences Center, National Renewable Energy Laboratory, Golden CO, 80401, United States
5. Department of Chemistry, Biotechnology, and Food Science, Norwegian University of Life Sciences, N-1432 Ås, Norway
6. Department of Chemical and Materials Engineering, University of Kentucky, Lexington, KY, 40506, United States
7. Center for Computational Sciences, University of Kentucky, Lexington, KY, 40506, United States

To whom correspondence should be addressed:
Gregg T. Beckham (Gregg.Beckham@nrel.gov), Jerry Ståhlberg (Jerry.Stahlberg@slu.edu), Christina M. Payne (Christy.Payne@uky.edu)

Abstract

Polysaccharide depolymerization in nature is primarily accomplished by processive glycoside hydrolases (GHs), which abstract single carbohydrate chains from polymer crystals and cleave glycosidic linkages without dissociating after each catalytic event. Understanding the molecular-level features and structural aspects of processivity is of importance due to the prevalence of processive GHs in biomass-degrading enzyme cocktails. Here, we describe recent advances towards the development of a molecular-level theory of processivity for cellulolytic and chitinolytic enzymes, including the development of novel methods for measuring rates of key steps in processive action and insights gained from structural and computational studies. Overall, we present a framework for developing structure-function relationships in processive GHs and outline additional progress towards developing a fundamental understanding of these industrially important enzymes.

Introduction

Structural polysaccharides, such as cellulose and chitin, typically arrange in insoluble, polymeric crystals that form significant components of plant, fungal, and algal cell walls. Microorganisms have evolved suites of enzymatic machinery to degrade these polysaccharides to soluble units for food and energy. These enzyme cocktails are primarily composed of various glycoside hydrolases (GHs) with synergistic functions to efficiently cleave the glycosidic linkages [1,2]. More recently, additional enzymatic functions beyond the canonical GH enzyme battery have been discovered including oxidative enzymes that selectively cleave glycosidic bonds [3-7]. GH cocktails contain enzymes typically delineated into two broadly defined classes: cellobiohydrolases (CBHs) and endoglucanases (EGs) for cellulose depolymerization, or chitobiohydrolases and endochitinases for chitin depolymerization. EGs are thought to randomly hydrolyze glycosidic linkages primarily in amorphous regions of polymer fibers. Alternatively, CBHs are able to attach to carbohydrate chains and processively hydrolyze disaccharide units from the end of a chain without dissociation after each catalytic event. Processivity is traditionally thought to be a means of conserving energy during enzymatic function, and is a general strategy used in the synthesis, modification, and depolymerization of many natural biopolymers [8]. It is this ability to act processively that imparts significant hydrolytic potential to CBHs from various GH families such as GH Family 6, 7, 18, and 48 and typically makes them the most abundant enzymes in natural secretomes of many microorganisms. Thus, GHs are the focus of intense protein engineering efforts for the biofuels industry [9*,10].

Here, we aim to briefly summarize developments in understanding GH processivity from the last several years via biophysical, structural, and modeling approaches for several illustrative GH families. In particular, we focus on developments in GH families for which substantial work has been conducted including GH Family 7 and 6, both of which are common fungal cellulolytic enzymes that depolymerize cellulose from the reducing and non-reducing ends, respectively. These systems, along with GH Family 18 chitinases [11], serve as well characterized models from which a concise theory of carbohydrate processivity that accounts for thermodynamics and kinetics can be developed. This, in turn, will enable the development of more comprehensive structure-function relationships in important biomass-degrading enzymes.

Definitions of GH processivity

Many methods used to determine the degree of processivity, the quantitative approximation of processive ability, describe 'apparent processivity'. The formal mathematical definition of apparent processivity is the number of catalytic events an enzyme performs divided by the number of times the enzyme initiates a processive run, i.e. acquires a chain end [12]**. Though a seemingly simple definition, apparent processivity can be difficult to accurately measure, particularly in systems that exhibit biphasic kinetics in their substrate degradation. Additionally, apparent processivity is highly dependent upon the substrate [13**,14**]; thus in practice, apparent processivity can be thought of as the actual processive ability of an enzyme acting on a particular substrate at a given set of conditions. This definition of processivity has utility in comparing degree of processivity across experiments conducted under the same or nearly similar conditions. However, given the variety of

methods developed for measuring this quantity and the numerous possible variations in conditions and substrates, comparison of apparent processivity across studies is often not straightforward.

An alternative definition of degree of processivity has emerged describing the theoretical potential for processive ability of GHs, or 'intrinsic processivity' [12**,14**]. Intrinsic processivity is primarily formulated in probabilistic terms and was first developed to describe the processive mechanism of nucleic acid polymerases [15]. McClure and Chow defined steady-state polymerase processivity as a distribution of probabilities defining the likelihood that the polymerase, upon catalysis, will translocate forward rather than dissociate from the newly formed strand. Later, Lucius *et al*. extended this probability-based definition to a kinetic description of helicase action [16]. Kurašin and Väljamäe further extended applicability of this definition to processive GHs, approximating intrinsic processivity as the catalytic rate constant, k_{cat}, divided by dissociation rate constant, k_{off} [14]**, which assumes that for processive enzymes, the probability of dissociation from the substrate is exceedingly low. Using a mathematical formalism, one can consider intrinsic processivity as the limit of apparent processivity as the polymeric substrate approaches ideality. This definition of processivity is potentially advantageous in the development of structure-function relationships, given its direct correlation to measurable kinetic variables and connection to structural features of GH enzymes and substrates.

Methods to examine GH processivity
As GHs are the primary components of cellulolytic and chitinolytic enzyme cocktails, myriad research approaches have been undertaken to understand how their enzymatic cycles occur at the molecular level, including biophysical measurements, structural biology efforts, and various types of modeling. Currently, several standard approaches for measuring apparent processivity have been described, most of which capitalize on the consistent nature of a processive GH product profile. During a processive cycle, GHs primarily produce disaccharides of cellulose or chitin (cellobiose or chitobiose), with relatively few odd-numbered saccharides [17], and thus, an efficient approach to measure apparent processivity for a given enzyme acting on cellulose or chitin is to measure the ratio of disaccharide units produced to the sum of mono- and trisaccharide units. This measurement technique is readily conducted using standard chromatographic methods [18,19]. However, assumptions regarding the initial binding mode, and thus the initial product profile, as well as the presence of intermediate products longer than dimers, can lead to misinterpretation or overestimation of processivity values [20*,21].

A second method for measuring GH processivity involves simultaneously determining the ratio of soluble to insoluble reducing ends [22-26]. Processive GHs produce significantly higher quantities of soluble reducing ends compared to non-processive GHs because they primarily liberate soluble products. To determine the ratio of soluble to insoluble reducing ends, the supernatant and substrate are separately assayed for reducing ends using somewhat standard analytical methods. As with the product ratio method described above, this measurement technique also requires assumptions regarding enzyme mechanisms that may bias interpretation of the results. Exo-glycosidases, as well as a potential preference of endo-glycosidases for more easily accessible chain ends, yield soluble reducing ends without processive action. Furthermore, this method is particularly sensitive to the type of substrate used [27], where an abundance of available free chain ends may result in unusually high values of soluble reducing ends from non-processive enzymes.

Recently, new techniques based on substrate labeling have been developed to overcome the limitations presented by the more traditional approaches to measure processivity. One method, termed the single-hit approach, again makes use of the fact that processive enzymes produce more soluble than insoluble reducing ends. In this method, the insoluble reducing end fraction is more accurately quantified through fluorescence-based labeling of reduced cellulose [14**,28,29]. Released soluble reducing ends represent the number of catalytic events, and when reduced cellulose is used as the substrate, the insoluble reducing ends encompass the number of initiation events. The fluorescent labeling of insoluble reducing groups in reduced cellulose allows for visualization of the aldehydes generated upon cellulolytic cleavage. While this method is significantly more accurate than those previously described, it also is not without limitations. Currently, this method is only effective for determining insoluble reducing ends from reducing end-specific cellulases, e.g., Family 7 CBHs, and the presence of non-reducing end-specific cellulases will result in overestimation of apparent processivity.

A second label-based method, termed the single-turnover approach, uses ^{14}C-labeled cellulose to quantify apparent processivity of cellulases [12]**. This analysis technique is designed to allow the cellulases to begin a processive run on the ^{14}C-labeled cellulose [29], but an excess of fluorescently labeled, 'trap' substrate is subsequently added after a short period to prevent unbound cellulases from continuing to attack the ^{14}C-labeled substrate. The radioactivity of the supernatant, combined with the concentration of trapped enzymes [30] allows determination of apparent processivity and the rate of a processive run. As with the single-hit approach, the single-turnover approach is currently limited to reducing-end specific cellulases by virtue of the need to use fluorescent-labeled reporter molecules, such as methylumbelliferyl lactoside to quantify the population of trapped cellulases [30].

The development of these more accurate substrate labeling-based techniques has had a significant impact on our understanding of GH processivity. A primary finding from this work was that dissociation rate, k_{off}, of the CBH from the substrate is potentially the key rate-limiting factor in processive turnover of cellulose by Family 7 GHs, whereas catalytic rate constants within the same GH family were nearly constant [14**,30]. Additionally, it was shown that the *Trichoderma reesei* Family 7 CBH (Cel7A) was able to perform an average of ~61 cuts per chain initiation event whereas the more open *Phanerochaete chrysosporium* Family 7 CBH (Cel7D) performs approximately 52 cuts (measured as apparent processivity on crystalline bacterial cellulose) [14]**. However, the k_{off} of *P. chrysosporium* Cel7D is approximately 4 times that of *T. reesei* Cel7A, and with similar k_{cat} values, this results in an intrinsic processivity approximately 4 times as high in *T. reesei* Cel7A relative to *P. chrysosporium* Cel7D. The vast difference in intrinsic and apparent processivity suggest that GH Family 7 CBHs become stalled on the cellulose surface as a result of the substrate. Jalak *et al.* from the same group subsequently demonstrated that at the optimal ratio of CBHs and EGs, the steady state rate of cellulose hydrolysis becomes limited by the processivity of *T. reesei* Cel7A [31]*. The authors interpret this result in that EGs not only provide new places for attachment of CBHs, as is commonly stated, but also that EGs provide points for CBH detachment. Another noteworthy result from these studies is that the rate of endo-initiation is quite high in both CBH enzymes, with *T. reesei* Cel7A conducting endo-initiation in more than 50% of its initial binding events and *P. chrysosporium* Cel7D doing so up to 80% of the time. Taken together, these results are the first glimpse of two CBHs from the same GH family directly compared in terms of both apparent and intrinsic processivity, and offer a reinterpretation of CBH-EG synergy from a mechanistic perspective. Overall, these studies offer a step towards development of structure-processivity relationships in these enzymes.

The finding that k_{off} is likely rate limiting has been corroborated by a separate analytical technique focused on elucidating *T. reesei* Cel7A kinetics in the pre-steady state regime [32]*. Cruys-Bagger *et al.* developed an elegant amperometric technique to examine Family 7 CBHs that employs cellobiose dehydrogenase to oxidize the cellobiose product of GH7 action in solution. This method is able to measure several aspects of processivity at high temporal resolution and at very low cellobiose concentrations [32]*. Coupling these measurements on Family 7 CBH action to mass-action kinetic models, they determined that *T. reesei* Cel7A conducts approximately four hydrolytic reactions per second on an insoluble substrate and that the off-rate was the rate-limiting step, both findings are in agreement with results from Kurašin and Väljamäe described above [14]**. As the amperometric method requires a sophisticated setup, Cruys-Bagger *et al.* went on to develop a new mass-action kinetic model based on a modified Michaelis-Menten approach for CBH action on insoluble substrates from which one can extract processivity parameters directly from short-time activity data [33]. This new model assumes that cellulases are acting at quasi-steady-state and provides a 'kinetic processivity coefficient' that is directly related to the intrinsic processivity of a given cellulase and offers a convenient approach to characterize processivity with a tractable, straightforward approach.

Processivity in Family 18 chitinases, with their unique substrate-assisted catalytic mechanism [34,35], can be more easily quantified with chitosan-based methods of detection. Processive chitinase action on chitosan (partially-deacetylated chitin) produces a product profile dominated by even-numbered oligomers which results from the requirement for an acetamido group in the -1 subsite enabling productive binding [36]. Processive chitinases proceed forward until an acetamido group in the -1 subsite is available to productively bind the substrate, which must occur on alternating *N*-acetyl glucosamine monomers. Non-processive enzymes dissociate from the substrate when confronted with a non-productively bound complex, which results in a random, non-uniform product distribution. Standard chromatographic and direct reducing end assays allow for determination of the number of cleaved glycosidic linkages, which can be compared against the number of endo-initiated events determined through active monitoring of substrate viscosity. This method of approximating apparent processivity is particularly useful for comparisons of variant chitinases [37] and has great potential for studying processivity in GHs generally in this model system [38].

Cellulase processivity has also been examined using a novel high-speed atomic force microscopy (HS-AFM) method that enables nanoscale spatial resolution with sub-second temporal resolution in pioneering work from Igarashi *et al.* [13**,39]. This approach to understanding processivity is unique in its direct visualization of enzyme action on polymeric substrates and has resulted in some revelatory observations regarding cellulase processivity. Namely, Igarashi *et al.* demonstrated that *T. reesei* Cel7A is able to translate along the cellulose surface at a rate of ~7 nm/s during processive action, which is in agreement with the results from Kurašin and Väljamäe [14**] and Cruys-Bagger *et al.* [32]*. Igarashi *et al.* also found that the enzymes eventually became 'jammed' or stalled during cellulose hydrolysis, presumably due to surface heterogeneity. This observation lends direct visual support to the finding that the off-rate is rate-limiting in processive cellulose hydrolysis. Additionally, it was also shown that *T. reesei* Cel6A may not be as conventionally 'processive' as *T. reesei* Cel7A, considering that the former was not observed to translate along the surface [13**], as indirectly observed in a much earlier study on a similar GH6 CBH [40].

In summary, each of the above-described methods quantifies apparent processivity and includes contributions from the substrate. Intrinsic processivity, as a theoretical value, cannot be directly measured. Rather, intrinsic processivity can be

estimated from measurements of k_{cat} and k_{off} [14]**. Horn *et al.* present a complete discussion of the difficulties associated with measuring k_{off} on insoluble, heterogeneous substrates [12]**.

Thermodynamics of CBH processivity

Though the processive mechanism of GHs has been extensively studied by both biochemical and structural studies [8,13**,14**,18,30,32,33,37,39,41-50], the direct connection of structure to processive function beyond the broad topological categorizations remains unclear. We recently described the hypothesized elementary steps involved in the catalytic cycle of a processive GH as a free energy profile along the enzymatic reaction coordinate [51]. The sum of the individual steps – surface binding, substrate recognition, initial loading of the polymer chain into the catalytic tunnel, and the processive catalytic cycle – must be energetically favorable overall. Here, we further suggest that the free energy of binding the polymeric substrate to the enzyme active site is proportionally related to processivity through the association and dissociation rate constants, k_{on} and k_{off}, respectively. This hypothesis, illustrated in Figure 1, can be employed to connect the structural features of GHs to the substrate-binding equilibrium constant assuming processivity approximates a steady-state process. As described above, apparent processivity is difficult to express as basic kinetic rates, and it is not immediately clear how to directly relate apparent processivity to binding free energy. Intrinsic processivity, on the other hand, with its simplified formulation in terms of k_{cat} and k_{off}, can be mathematically related to binding free energy. We propose that that this formulation of intrinsic processivity in terms of experimentally measurable kinetic and thermodynamic parameters may serve as the foundation for developing a comprehensive structure-function relationship, which is essential to the successful design of enhanced enzyme cocktails for biomass conversion.

Recent structural studies in which we examined key dynamical properties using complementary molecular dynamics simulations suggest hallmarks exist that may be associated with processive ability [47*,52,53]. In *Serratia marcescens* chitinases, we found that ligand fluctuations and solvation as well as fluctuation of the localized catalytic residues correlated well with previously measured values of apparent processivity between endochitinases and chitobiohydrolases [47]*. In our comparison of the *Heterobasidion irregulare* Cel7A structure to *T. reesei* Cel7A and *P. chrysosporium* Cel7D, we discovered that active site loop regions of the more processive *T. reesei* Cel7A open and close at a significantly lower rate than the less processive *P. chrysosporium* Cel7D and putatively less processive *H. irregulare* Cel7A suggesting the latter two enzymes may perform endo-initiation events more often than *T. reesei* Cel7A, which can affect processive ability [52]. Each of these studies qualitatively highlighted the potential contributions of enzyme dynamics to processive ability. The free energy of ligand binding quantitatively captures these molecular-level details and relating this term to processivity represents a potential protocol for predicting relative processive ability. Robust, enhanced sampling free energy methods capable of treating flexible carbohydrate ligands are required for calculating binding free energy [54,55]. In GHs from the same family, a relative comparison of the calculated binding free energy could provide a good 'first-pass' at screening modified processivity variants. Additionally, isothermal titration calorimetry measurements on catalytically inactive mutants may offer another means to measure the ligand binding free energy directly [56,57].

General mechanisms of CBH processivity

In terms of the mechanism (kinetics) of processive action, CBHs typically initiate substrate attack via two modes, namely in 'exo' or 'endo' mode. Exo-mode initiation is thought to occur from the entrance of the active site tunnel in CBHs where a chain is threaded from the entrance site through the tunnel to form a Michaelis complex and initiate a processive cycle. Aromatic residues at the entrances of tunnels in GH7 CBHs have been shown to be directly involved in exo-mode initiation by molecular dynamics simulations and indirectly in biochemical experiments [58,59]. The importance of aromatic residues in exo-mode initiation was originally inferred by mutating the entrance tryptophan residue in the *T. reesei* Cel6A CBH. Specifically, Koivula *et al.* demonstrated that Trp272 was important for activity against crystalline cellulose by mutation to alanine, which did not adversely affect activity on amorphous cellulose, but dramatically reduced activity on a crystalline substrate [18]. Endo-initiation mode, conversely, occurs when the active-site tunnel loops are able to open sufficiently to directly complex with chains along the active site tunnels to form a Michaelis complex. Endo-initiation mode has been shown for GH7 CBHs to occur more frequently in more open active-site tunnels [14]**. These two modes of CBH complexation are likely to be the extremes of the feasible initiation mechanisms with a distribution of intermediate complexation mechanisms in between.

After formation of a Michaelis complex and initiation of a processive run, a CBH processive cycle will at minimum include catalysis, product expulsion, and translation along the polymer chain by a disaccharide unit to reform the Michaelis complex. However, the molecular-level details of processive motion along the active site tunnel will be a manifestation of structural features in each GH family. Below, we describe evidence for the detailed processive mechanisms in GH7 and GH6 enzymes that have been observed via structural biology approaches.

Mechanisms of processivity in GH Family 7 CBHs

As GH7 CBHs are the cornerstones of many fungal secretomes in nature and the primary enzymes in many industrial biofuels-related enzyme cocktails, multiple structures have been solved to date with various stages of ligand complexation. The substrate-binding sites in GH7 cellulases are characterized from the non-reducing end at the -7 subsite to the reducing end at the +2 subsite. Most solved structures indicate relatively high uniformity in substrate binding in the -7 to -3 sites, and the majority of the significant differences are found in the -2 and -1 and product sites (+1/+2). GH Family 7 enzymes are known to employ two-step retaining mechanisms that consist of the formation of a glycosyl-enzyme intermediate (GEI) [60-62]. From the structural diversity in the -2 to +2 subsites, we propose that the processive cycle of a GH7 CBH proceeds as shown in Figure 2. Starting from the top left frame of Figure 2, a cellulose chain first slides through the binding tunnel (or more precisely, the enzyme slides along a cellulose chain). After sliding, there is sufficient room in the enzyme active site tunnel for the cellulose chain to fill all nine binding subsites in the stable chair conformation (Figure 2, *top right frame*). We propose that the cello-oligomer chain will slide along a straight pathway into the tunnel forming a complex that may resemble the glucosyl binding characteristics in the *T. reesei* CBH Cel7A structure featuring two cellotetraose chains on either side of the vacant -3 binding subsite (PDB code 5CEL, shown in Figure 2, *top right frame*) [44]. All eight glucosyl residues are in the chair conformation, and an intact chain would likely feature a glucosyl residue in subsite -3 also exhibiting a chair conformation. The product binding sites (+1/+2) are occupied in the so-called 'Slide' mode in this structure [63]. This conformation is not active for hydrolysis, as the distance from the -1 anomeric carbon to the nucleophile is 7.1 Å. Before the chemical steps may proceed, the -1 glucosyl residue must be 'activated' by rotating nearly 90° along the axis of the cellulose chain, translating towards the nucleophile, and assuming a distorted non-chair configuration. This ring distortion allows the catalytic nucleophile access to the anomeric carbon reaction center and results in the Michaelis complex. Although no crystal structure for the Michaelis complex of a GH7 CBH has been published, crystal structures from other retaining GHs suggest this conformation may resemble a boat [35,64,65], skew-boat [66-68], or the similar half-chair or envelope configurations. This catalytically-active complex may resemble the theoretical model of the Michaelis complex of *T. reesei* Cel7A (PDB code 8CEL, Figure 2, *bottom right*). Immediately following the first catalytic step (Glycosylation), the cellobiose product resides in what we refer to as the 'Unprimed glycosyl-enzyme intermediate' (GEI) mode (seen in PDB structures 6CEL, 7CEL, 1Q2E, 1Z3W, 2RFY, 4HAP, 4IPM, and exemplified by the 7CEL structure in Figure 2, *lower left frame*) [21,44,63,69,70]. As with the Michaelis complex, no GEI crystal structure has been published to date for a GH Family 7 member. Nevertheless, various GH7 crystal structures reveal the existence of two distinct product binding modes [63]. The 'Primed GEI' mode (which essentially overlays 'Slide' mode in the +1/+2 subsites) results from a slight shift in the cellobiose product from the 'Unprimed GEI' mode towards the tunnel exit (seen in PDB structures 3CEL, 1Z3T, 1Z3V, and exemplified by the 3CEL structure in Figure 2, *bottom left*) [63,71]; this movement creates space for a water molecule (the nucleophile of the Deglycosylation step) to move into the active site in between the two chemical steps [63]. The 'priming' thus refers to the preparation of the cellobiose product and the nucleophilic water for the Deglycosylation reaction. Graphical depictions of the Koshland mechanism of retaining GHs [72] often portray the release of product as preceding the Deglycosylation step, but the processive cycle we propose here based on GH7 crystal structures would allow Deglycosylation to proceed without the prior removal of the cellobiose product. Deglycosylation produces the 'Substrate-Product complex' wherein the cellobiose product likely sits in essentially the same position as in the 'Primed GEI'. The processive cycle is completed by product expulsion [73], resulting in a vacant +1/+2 sites (Figure 2, *top left*) and a cellulose chain extending from the -1 site to beyond the -7 site and out into solution.

Processivity in GH Family 6 cellobiohydrolases

GH Family 6 CBHs are non-reducing-end specific enzymes often found as major secondary components of fungal and bacterial secretomes after reducing-end-specific CBHs. Similar to GH Family 7 CBHs, many crystallographic snapshots have been captured, particularly of *T. reesei*, *Humicola insolens*, and *Thermobifida fusca* GH6 enzymes, which enable construction of a hypothesized processive cycle. GH Family 6 CBHs employ a single step, inverting mechanism [74] and exhibit at least 6 binding subsites from +4 to -2. The catalytic base is thought to be a conserved aspartate residue that is connected to the nucleophilic water molecule via a water wire observed in multiple crystal structures [2,18,67,75,76]. This water wire is stabilized through hydrogen bonding interactions by a conserved serine residue (Ser181 in *T. reesei* Cel6A) on the so-called 'catalytic center loop' (approximately residues 178 to 183 in *T. reesei* Cel6A) and by the backbone carbonyl atom of Asp401 in *T. reesei* Cel6A. Based on the GH6 CBH structures in various binding poses, we propose a processive cycle, illustrated in Figure 3. The enzyme begins with the product (-1/-2) subsites empty (Figure 3, *top left frame*). To proceed by one cellobiose unit, the catalytic center loop remains open to allow the passage of the substrate through subsite -1. Both the catalytic acid and the catalytic base side chains rotate out of the path of the sliding chain, and form a hydrogen bond with each other, the functional significance of which is described below, and the ligand spans the active site in 'Slide' mode (seen in, e.g., the PDB structure 1QK2 [67]) (Figure 3, *top right*). From Slide mode, the catalytic center loop closes upon isolation of the nucleophilic water and another water molecule from bulk water, to stabilize the water wire to the putative base, and form the Michaelis complex, partially exemplified e.g., by the 1QJW [67] and 4AVO [76] structures (Figure 3, *bottom right*). The 'active' conformation for the catalytic acid (D221 in *T. reesei* Cel6A) is taken from PDB structure 1HGW [77]. From the Michaelis

complex, the hydrolytic reaction occurs to form the Substrate-Product complex with an α-cellobiose as the product, as modeled in a recent study [76]. In various structures with the catalytic center loop in both the 'open' and 'closed' positions, the catalytic acid and putative catalytic base have been observed to form a direct hydrogen bond, perhaps suggesting a mechanism by which proton transfer can occur directly from the catalytic base after the reaction back to the catalytic acid to reset the enzyme during the processive cycle. However, this hypothesis remains untested, and proton transfer can also readily occur through water molecules.

Interestingly, the primary difference between the GH6 and GH7 proposed processivity mechanisms center around which portion of the enzyme-substrate complex must undergo conformational changes. Namely, structural data seem to imply that in GH7 CBHs, processivity occurs with very little conformational changes in the enzyme, but mostly in the substrate. Conversely, in GH6 CBHs, the catalytic center loop seems to open for processivity and to reset the enzyme, while the ligand conformation is relatively stable in terms of its conformation, except for the ring puckering at the -1 subsite. In both cases, these processivity mechanisms can be directly tested with a combination of structural and computational methods [51].

Impact of enzyme modularity on enzyme processivity
Many cellulose and chitin-degrading enzymes are multi-modular, typically combining binding function via carbohydrate-binding modules (CBMs) with catalytic function in the GH domain, connected by linker domains of varying lengths. In fungi, multi-modular enzymes exhibit linkers with significant *O*-glycosylation. Recently, it was shown that linkers between GH6 and GH7 enzymes from fungi are of significantly different length (~50% different on average) [78]. Interestingly, it was also predicted from MD simulation and confirmed experimentally that fungal cellulase linkers add directly to enzyme binding affinity in a dynamic, non-specific way [79]. As mentioned above, Igarashi *et al.* have conducted HS-AFM measurements wherein they showed that GH7 CBHs from *T. reesei* are able to process on the cellulose substrate at a rate of 7 nm/s [13]** and that the catalytic domain alone without a CBM and linker is able to translate at the same speed as a fully intact enzyme. This suggests that the Cel7A linker does not have a significant effect on the processive cycle. However, in the HS-AFM experiments, GH6 enzymes seemingly did not translate significantly on the surface of the substrate when incubated with cellulose alone [13]**. Given that linkers can serve non-specific binding function and that GH6 linkers are on average approximately 50% longer, we propose that GH6 CBHs are potentially more active in a localized region on cellulose perhaps explaining the observed reduction in apparent processivity.

Conclusions and Future Directions
A molecular-level theory describing the thermodynamics and kinetics of processive action in CBHs will aid in the development of more detailed structure-function relationships for these industrially and scientifically important enzymes. New experimental developments such as the development of more effective GH6 small-molecule substrates for measuring GH6 processivity are currently under development [2], which will eventually enable more explicit characterization of the enzyme-generated products remaining in substrate fractions similar to the single-hit/single-turnover approaches for GH7s. Structural biology tools in concert with computational approaches are especially well-suited to study the hypothesized processivity mechanisms shown here, including free energy calculations and path sampling approaches [51]. Taken together, these approaches will help fully uncover the roles of enzyme architecture, composition, and dynamics enabling carbohydrate processivity.

Summaries
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. *Journal of Biological Chemistry* 2012, **287**:18451-18458.
The authors of this study developed a new approach based on amperometric biosensors to understand the biphasic kinetics of processive cellulase depolymerization of insoluble cellulose. Alongside a mathematical model fit to their experimental data, Cruys-Bagger *et al*. delineate a pre-steady state regime in cellulase kinetics and quantify the steady-state apparent processivity of *T. reesei* Cel7A. Kinetic model analysis suggests that dissocation of the cellulase from the insoluble polysaccharide is rate-limiting in steady-state processive cellulase hydrolysis.

*Cruys-Bagger N, Elmerdahl J, Praestgaard E, Borch K, Westh P: **A steady-state theory for processive cellulases**. *Febs Journal* 2013, **280**:3952-3961.

A straightforward, generalized kinetic model for processive enzymes requiring only standard experimentally measurable parameters is reported in this study. Though simplifying assumptions are made for extensibility, the authors demonstrate the model's applicability in both comparing kinetics and critically evaluating the more detailed mechanistic kinetics. In applying the steady-state kinetic model to available cellulase data, Cruys-Bagger *et al*. illustrate their model suggests substrate association and catalysis are not rate-limiting in cellulase processivity. Rather, there appears to be a nearly proportional correlation of the dissociation rate constant with experimentally-measured apparent processivity.

*Dana CM, Saija P, Kal SM, Bryan MB, Blanch HW, Clark DS: **Biased clique shuffling reveals stabilizing mutations in cellulase Cel7A**. *Biotechnology and Bioengineering* 2012, **109**:2710-2719.
Biased clique shuffling was used to develop a library of fungal Family 7 cellulase chimeras to improve thermal stability. The authors report the successful generation of a library containing 51 chimeras demonstrating improved thermal tolerance. Additionally, the improved thermal stability does not appear to come at the cost of abolishing activity with 86% of the chimeras showing activity. The biased clique shuffling technique was significantly more productive in identifying active, thermally stable chimeras than equimolar library generation.

*Fox JM, Levine SE, Clark DS, Blanch HW: **Initial- and Processive-Cut Products Reveal Cellobiohydrolase Rate Limitations and the Role of Companion Enzymes**. *Biochemistry* 2012, **51**:442-452.
The synergistic kinetics of the processive *Trichoderma longibrachiatum* cellobiohydrolase I and *Talaromyces emersonii* endoglucanase II were examined on bacterial microcrystalline cellulose. Kinetic rate parameters, from Michaelis-Menten fits to initial rate data, alongside a mechanistic model of initial- and processive-cut products suggest that cellulase-substrate association is rate-limiting in this system. The authors also suggest the synergistic role of endoglucanase activity may include removing obstacles from the processive cellulase run.

Horn SJ, Sorlie M, Varum KM, Valjamae P, Eijsink VG: **Measuring processivity. *Methods Enzymol* 2012, **510**:69-95.
In this review, the authors provide a detailed overview of several of the currently available methods used to methods to measure processivity. The advantages and limitations of each method are fully outlined making it easy for the reader to select a method appropriate to their needs and available resources. A particularly interesting section of this chapter discusses the advantages of using *S. marcescens* chitinases as a model system for studying processivity with a graphical depiction of how the substrate-assisted mechanism contributes the diagnostic chitosan product profile.

Igarashi K, Uchihashi T, Koivula A, Wada M, Kimura S, Okamoto T, Penttila M, Ando T, Samejima M: **Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. *Science* 2011, **333**:1279-1282.
Igarashi *et al*. report the enhanced temporal (300 ms/frame) and spatial resolution of their original high-speed atomic force microscopy method for visualizing cellulase activity on crystalline cellulose substrates. Processive action of *T. reesei* Cel7A on cellulose Iα and III$_\text{I}$ was visualized with corresponding velocities determined. Alone, *T. reesei* Cel7A appears to be subject to 'traffic jams' where the forward motion of the enzyme is slowed presumably by substrate heterogeneities. Addition of *T. reesei* Cel6A alleviates these obstacles and dramatically improves hydrolytic turnover through synergistic action.

Kurašin M, Väljamäe P: **Processivity of Cellobiohydrolases Is Limited by the Substrate. *Journal of Biological Chemistry* 2011, **286**:169-177.
A new method for determining apparent processivity in reducing-end specific cellobiohydrolases is reported. The authors make use of a clever substrate-labeling technique to accurately represent the number of insoluble reducing groups generated through hydrolytic action. The method is capable of delineating both exo- and endo-mode initiations. Application of this method to study the hydrolytic turnover of *T. reesei* Cel7A and *P. chrysosporium* Cel7D uncovered *T. reesei* Cel7A has much greater processive ability than *P. chrysosporium* Cel7D and that apparent processivity may be at least an order of magnitude lower that intrinsic processivity in each case. Kurašin and Väljamäe also report apparent processivity is substrate-limited and dissociation of the nonproductively bound enzyme is rate-limiting in the processive cycle.

*Momeni MH, Payne CM, Hansson H, Mikkelsen NE, Svedberg J, Engström Å, Sandgren M, Beckham GT, Ståhlberg J: **Structural, Biochemical, and Computational Characterization of the Glycoside Hydrolase Family 7 Cellobiohydrolase of the Tree-killing Fungus *Heterobasidion irregulare***. *Journal of Biological Chemistry* 2013, **288**:5861-5872.
The authors describe the comprehensive characterization of the *Heterobasidion irregulare* Family 7 cellobiohydrolase and use MD simulations to compare dynamics with two other GH7 cellobiohydrolases from *T. reesei* (*H. jecorina*) and *Phanerochaete chrysosporium*. Despite the fact that *H. irregulare* Cel7A exhibits the same structural characteristics as *T. reesei* Cel7A and *P. chrysosporium* Cel7D lending the enzyme its characteristic fold and processive ability, key structural differences exist suggesting *H. irregulare* may be less processive than *T. reesei* Cel7A but more processive than *P. chrysosporium* Cel7D. *H.*

irregulare Cel7A exhibits a unique tyrosine residue at the entrance of the active site that effectively forms an additional binding subsite; however, this contributes to instability of the ligand as the tyrosine freely rotates about it dihedral. MD simulations also suggest *H. irregulare* Cel7A and *P. chrysosporium* Cel7D more readily perform endo-initiation than *T. reesei* Cel7A, as the exo-loops adjacent to the active site open and close with greater frequency. In general, this suggests that the static shape of the active site tunnel in GH Family 7 CBHs is perhaps not the sole determinant of processive ability.

*Payne CM, Baban J, Horn SJ, Backe PH, Arvai AS, Dalhus B, Bjoras M, Eijsink VGH, Sorlie M, Beckham GT, et al.: **Hallmarks of processivity in glycoside hydrolases from crystallographic and computational studies of the *Serratia marcescens* chitinases**. *Journal of Biological Chemistry* 2012, **287**:36322-36330.

This work describes the solution of the nonprocessive chitinase structure, ChiC, from *Serratia marcescens,* which was the final unsolved structure in this host's chitinolytic machinery. This self-consistent suite of processive and nonprocessive chitinases having the same catalytic mechanism represents an ideal model system in which to study the molecular-level details contributing to processivity. Dynamical hallmarks correlated with measured apparent processivity were uncovered using MD simulations to compare active site dynamics of both processive and nonprocessive Family 18 GHs. These hallmarks include magnitude of ligand fluctuation and solvation as well as key catalytic residue fluctuations. Overall, this work suggests a more global definition of carbohydrate processivity based on structural and dynamics characteristics of ligand complexation may exist.

References

1. Levasseur A, Drula E, Lombard V, Coutinho P, Henrissat B: **Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes**. *Biotechnol for Biofuels* 2013, **6**:41.
2. Wu M, Nerinckx W, Piens K, Ishida T, Hansson H, Sandgren M, Stahlberg J: **Rational design, synthesis, evaluation and enzyme-substrate structures of improved fluorogenic substrates for family 6 glycoside hydrolases**. *Febs Journal* 2013, **280**:184-198.
3. Beeson WT, Phillips CM, Cate JHD, Marletta MA: **Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases**. *J. Am. Chem. Soc.* 2012, **134**:890-892.
4. Phillips CM, Beeson WT, Cate JH, Marletta MA: **Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa***. *ACS Chem Biol* 2011, **6**:1399-1406.
5. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JCN, Johansen KS, Krogh K, Jorgensen CI, Tovborg M, Anthonsen A, et al.: **Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**:15079-15084.
6. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu ZL, Zhai H, Sorlie M, Eijsink VGH: **An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides**. *Science* 2010, **330**:219-222.
7. Horn S, Vaaje-Kolstad G, Westereng B, Eijsink V: **Novel enzymes for the degradation of cellulose**. *Biotechnology for Biofuels* 2012, **5**:45.
8. Stern R, Jedrzejas MJ: **Carbohydrate Polymers at the Center of Life's Origins: The Importance of Molecular Processivity**. *Chemical Reviews* 2008, **108**:5061-5085.
9. Dana CM, Saija P, Kal SM, Bryan MB, Blanch HW, Clark DS: **Biased clique shuffling reveals stabilizing mutations in cellulase Cel7A**. *Biotechnology and Bioengineering* 2012, **109**:2710-2719.
10. Lantz SE, Goedegebuur F, Hommes R, Kaper T, Kelemen BR, Mitchinson C, Wallace L, Stahlberg J, Larenas EA: **Hypocrea jecorina CEL6A protein engineering**. *Biotechnology for Biofuels* 2010, **3**.
11. Eijsink VGH, Vaaje-Kolstad G, Varum KM, Horn SJ: **Towards new enzymes for biofuels: lessons from chitinase research**. *Trends in Biotech.* 2008, **26**:228-235.
12. Horn SJ, Sorlie M, Varum KM, Valjamae P, Eijsink VG: **Measuring processivity**. *Methods Enzymol* 2012, **510**:69-95.
13. Igarashi K, Uchihashi T, Koivula A, Wada M, Kimura S, Okamoto T, Penttila M, Ando T, Samejima M: **Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface**. *Science* 2011, **333**:1279-1282.
14. Kurasin M, Valjamae P: **Processivity of Cellobiohydrolases Is Limited by the Substrate**. *Journal of Biological Chemistry* 2011, **286**:169-177.
15. McClure WR, Chow Y: **The kinetics and processivity of nucleic acid polymerases**. In *Methods in Enzymology*. Edited by Daniel LP. Academic Press; 1980:277-297.
16. Lucius AL, Maluf NK, Fischer CJ, Lohman TM: **General methods for analysis of sequential "n-step" kinetic mechanisms: Application to single turnover kinetics of helicase-catalyzed DNA unwinding**. *Biophysical Journal* 2003, **85**:2224-2239.
17. Teeri TT: **Crystalline cellulose degradation: new insight into the function of cellobiohydrolases**. *Trends in Biotechnology* 1997, **15**:160-167.

18. Koivula A, Kinnari T, Harjunpaa V, Ruohonen L, Teleman A, Drakenberg T, Rouvinen J, Jones TA, Teeri TT: **Tryptophan 272: An essential determinant of crystalline cellulose degradation by *Trichoderma reesei* cellobiohydrolase Cel6A**. *Febs Letters* 1998, **429**:341-346.

19. Vuong TV, Wilson DB: **Processivity, Synergism, and Substrate Specificity of *Thermobifida fusca* Cel6B**. *Applied and Environmental Microbiology* 2009, **75**:6655-6661.

20. Fox JM, Levine SE, Clark DS, Blanch HW: **Initial- and Processive-Cut Products Reveal Cellobiohydrolase Rate Limitations and the Role of Companion Enzymes**. *Biochemistry* 2012, **51**:442-452.

21. 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, Ce17A. A comparison with *Phanerochaete chrysosporium* Cel7D**. *Journal of Molecular Biology* 2003, **333**:817-829.

22. Doner LW, Irwin PL: **Assay of reducing end-groups in oligosaccharide homologs with 2,2'-bicinchoninate**. *Analytical Biochemistry* 1992, **202**:50-53.

23. Irwin DC, Spezio M, Walker LP, Wilson DB: **Activity studies of 8 purified cellulases - Specificity, synergism, and binding domain effects**. *Biotechnology and Bioengineering* 1993, **42**:1002-1013.

24. Irwin D, Shin DH, Zhang S, Barr BK, Sakon J, Karplus PA, Wilson DB: **Roles of the catalytic domain and two cellulose binding domains of Thermomonospora fusca E4 in cellulose hydrolysis**. *Journal of Bacteriology* 1998, **180**:1709-1714.

25. Zhang S, Irwin DC, Wilson DB: **Site-directed mutation of noncatalytic residues of Thermobifida fusca exocellulase Cel6B**. *European Journal of Biochemistry* 2000, **267**:3101-3115.

26. Watson BJ, Zhang HT, Longmire AG, Moon YH, Hutcheson SW: **Processive Endoglucanases Mediate Degradation of Cellulose by Saccharophagus degradans**. *Journal of Bacteriology* 2009, **191**:5697-5705.

27. Zhang YHP, Lynd LR: **Determination of the number-average degree of polymerization of cellodextrins and cellulose with application to enzymatic hydrolysis**. *Biomacromolecules* 2005, **6**:1510-1515.

28. Kipper K, Valjamae P, Johansson G: **Processive action of cellobiohydrolase Cel7A from Trichoderma reesei is revealed as 'burst' kinetics on fluorescent polymeric model substrates**. *Biochemical Journal* 2005, **385**:527-535.

29. Velleste R, Teugjas H, Valjamae P: **Reducing end-specific fluorescence labeled celluloses for cellulase mode of action**. *Cellulose* 2010, **17**:125-138.

30. Jalak J, Valjamae P: **Mechanism of Initial Rapid Rate Retardation in Cellobiohydrolase Catalyzed Cellulose Hydrolysis**. *Biotechnology and Bioengineering* 2010, **106**:871-883.

31. Jalak J, Kurasin M, Teugjas H, Valjamae P: **Endo-exo Synergism in Cellulose Hydrolysis Revisited**. *Journal of Biological Chemistry* 2012, **287**:28802-28815.

32. 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**. *Journal of Biological Chemistry* 2012, **287**:18451-18458.

33. Cruys-Bagger N, Elmerdahl J, Praestgaard E, Borch K, Westh P: **A steady-state theory for processive cellulases**. *Febs Journal* 2013, **280**:3952-3961.

34. Tews I, van Scheltinga ACT, Perrakis A, Wilson KS, Dijkstra BW: **Substrate-assisted catalysis unifies two families of chitinolytic enzymes**. *Journal of the American Chemical Society* 1997, **119**:7954-7959.

35. van Aalten DMF, Komander D, Synstad B, Gaseidnes S, Peter MG, Eijsink VGH: **Structural insights into the catalytic mechanism of a family 18 exo-chitinase**. *Proceedings of the National Academy of Sciences of the United States of America* 2001, **98**:8979-8984.

36. Sorbotten A, Horn SJ, Eijsink VGH, Varum KM: **Degradation of chitosans with chitinase B from Serratia marcescens - Production of chito-oligosaccharides and insight into enzyme processivity**. *Febs Journal* 2005, **272**:538-549.

37. Horn SJ, Sikorski P, Cederkvist JB, Vaaje-Kolstad G, Sorlie M, Synstad B, Vriend G, Varum KM, Eijsink VGH: **Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides**. *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103**:18089-18094.

38. Vaaje-Kolstad G, Horn SJ, Sorlie M, Eijsink VGH: **The chitinolytic machinery of Serratiamarcescens - a model system for enzymatic degradation of recalcitrant polysaccharides**. *Febs Journal* 2013, **280**:3028-3049.

39. 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**. *Journal of Biological Chemistry* 2009, **284**:36186-36190.

40. Boisset C, Fraschini C, Schulein M, Henrissat B, Chanzy H: **Imaging the enzymatic digestion of bacterial cellulose ribbons reveals the endo character of the cellobiohydrolase Cel6A from Humicola insolens and its mode of synergy with cellobiohydrolase Cel7A**. *Applied and Environmental Microbiology* 2000, **66**:1444-1452.

41. Barr BK, Hsieh YL, Ganem B, Wilson DB: **Identification of two functionally different classes of exocellulases**. *Biochemistry* 1996, **35**:586-592.

42. Breyer WA, Matthews BW: **A structural basis for processivity**. *Protein Science* 2001, **10**:1699-1711.

43. Divne C, Ståhlberg J, Reinikainen T, Ruohonen L, Pettersson G, Knowles JKC, Teeri TT, Jones TA: **The 3-dimensional crystal-structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei***. *Science* 1994, **265**:524-528.

44. Divne C, Ståhlberg J, Teeri TT, Jones TA: **High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei***. *Journal of Molecular Biology* 1998, **275**:309-325.

45. Horn SJ, Sorbotten A, Synstad B, Sikorski P, Sorlie M, Varum KM, Eijsink VGH: **Endo/exo mechanism and processivity of family 18 chitinases produced by *Serratia marcescens***. *Febs Journal* 2006, **273**:491-503.

46. Muñoz IG, Ubhayasekera W, Henriksson H, Szabó I, Pettersson G, Johansson G, Mowbray SL, Ståhlberg J: **Family 7 cellobiohydrolases from *Phanerochaete chrysosporium*: Crystal structure of the catalytic module of Cel7D (CBH58) at 1.32 Å resolution and homology models of the isozymes**. *Journal of Molecular Biology* 2001, **314**:1097-1111.

47. Payne CM, Baban J, Horn SJ, Backe PH, Arvai AS, Dalhus B, Bjoras M, Eijsink VGH, Sorlie M, Beckham GT, et al.: **Hallmarks of processivity in glycoside hydrolases from crystallographic and computational studies of the *Serratia marcescens* chitinases**. *Journal of Biological Chemistry* 2012, **287**:36322-36330.

48. Sorlie M, Zakariassen H, Norberg AL, Eijsink VGH: **Processivity and substrate-binding in family 18 chitinases**. *Biocatalysis and Biotransformation* 2012, **30**:353-365.

49. Varrot A, Frandsen TP, von Ossowski I, Boyer V, Cottaz S, Driguez H, Schulein M, Davies GJ: **Structural basis for ligand binding and processivity in cellobiohydrolase Cel6A from *Humicola insolens***. *Structure* 2003, **11**:855-864.

50. Zakariassen H, Aam BB, Horn SJ, Varum KM, Sorlie M, Eijsink VGH: **Aromatic residues in the catalytic center of chitinase A from *Serratia marcescens* affect processivity, enzyme activity, and biomass converting efficiency**. *J. Biol. Chem.* 2009, **284**:10610-10617.

51. Beckham GT, Bomble YJ, Bayer EA, Himmel ME, Crowley MF: **Applications of computational science for understanding enzymatic deconstruction of cellulose**. *Current Opinion in Biotechnology* 2011, **22**:231-238.

52. Momeni MH, Payne CM, Hansson H, Mikkelsen NE, Svedberg J, Engström Å, Sandgren M, Beckham GT, Ståhlberg J: **Structural, Biochemical, and Computational Characterization of the Glycoside Hydrolase Family 7 Cellobiohydrolase of the Tree-killing Fungus Heterobasidion irregulare**. *Journal of Biological Chemistry* 2013, **288**:5861-5872.

53. Taylor CB, Payne CM, Himmel ME, Crowley MF, McCabe C, Beckham GT: **Binding Site Dynamics and Aromatic-Carbohydrate Interactions in Processive and Non-Processive Family 7 Glycoside Hydrolases**. *Journal of Physical Chemistry B* 2013, **117**:4924-4933.

54. Jiang W, Hodoscek M, Roux B: **Computation of Absolute Hydration and Binding Free Energy with Free Energy Perturbation Distributed Replica-Exchange Molecular Dynamics**. *Journal of Chemical Theory and Computation* 2009, **5**:2583-2588.

55. Jiang W, Roux B: **Free Energy Perturbation Hamiltonian Replica-Exchange Molecular Dynamics (FEP/H-REMD) for Absolute Ligand Binding Free Energy Calculations**. *Journal of Chemical Theory and Computation* 2010, **6**:2559-2565.

56. Norberg AL, Karlsen V, Hoell IA, Bakke I, Eijsink VGH, Sorlie M: **Determination of substrate binding energies in individual subsites of a family 18 chitinase**. *FEBS Letters* 2010, **584**:4581-4585.

57. Zolotnitsky G, Cogan U, Adir N, Solomon V, Shoham G, Shoham Y: **Mapping glycoside hydrolase substrate subsites by isothermal titration calorimetry**. *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**:11275-11280.

58. GhattyVenkataKrishna PK, Alekozai EM, Beckham GT, Schulz R, Crowley MF, Uberbacher EC, Cheng XL: **Initial Recognition of a Cellodextrin Chain in the Cellulose-Binding Tunnel May Affect Cellobiohydrolase Directional Specificity**. *Biophysical Journal* 2013, **104**:904-912.

59. Nakamura A, Tsukada T, Auer S, Furuta T, Wada M, Koivula A, Igarashi K, Samejima M: **Tryptophan residue at active-site tunnel entrance of Trichoderma reesei cellobiohydrolase Cel7A is important to initiate degradation of crystalline cellulose**. *Journal of Biological Chemistry* 2013, **288**:13503-13510.

60. Mackenzie LF, Davies GJ, Schulein M, Withers SG: **Identification of the catalytic nucleophile of endoglucanase I from Fusarium oxysporum by mass spectrometry**. *Biochemistry* 1997, **36**:5893-5901.

61. Mackenzie LF, Sulzenbacher G, Divne C, Jones TA, Woldike HF, Schulein M, Withers SG, Davies GJ: **Crystal structure of the family 7 endoglucanase I (Cel7B) from Humicola insolens at 2.2 angstrom resolution and identification of the catalytic nucleophile by trapping of the covalent glycosyl-enzyme intermediate**. *Biochemical Journal* 1998, **335**:409-416.

62. Sulzenbacher G, Schulein M, Davies GJ: **Structure of the endoglucanase I from Fusarium oxysporum: Native, cellobiose, and 3,4-epoxybutyl beta-D-cellobioside-inhibited forms, at 2.3 angstrom resolution**. *Biochemistry* 1997, **36**:5902-5911.

63. Ubhayasekera W, Muñoz IG, Vasella A, Ståhlberg J, Mowbray SL: **Structures of *Phanerochaete chrysosporium* Cel7D in complex with product and inhibitors**. *Febs Journal* 2005, **272**:1952-1964.

64. Tews I, Perrakis A, Oppenheim A, Dauter Z, Wilson KS, Vorgias CE: **Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease**. *Nature Structural Biology* 1996, **3**:638-648.

65. He Y, Macauley MS, Stubbs KA, Vocadlo DJ, Davies GJ: **Visualizing the Reaction Coordinate of an O-GlcNAc Hydrolase**. *Journal of the American Chemical Society* 2010, **132**:1807-+.

66. Sulzenbacher G, Driguez H, Henrissat B, Schulein M, Davies GJ: **Structure of the Fusarium oxysporum endoglucanase I with a nonhydrolyzable substrate analogue: Substrate distortion gives rise to the preferred axial orientation for the leaving group**. *Biochemistry* 1996, **35**:15280-15287.

67. Zou J-y, Kleywegt GJ, Ståhlberg J, Driguez H, Nerinckx W, Claeyssens M, Koivula A, Teeri TT, Jones TA: **Crystallographic evidence for substrate ring distortion and protein conformational changes during catalysis in cellobiohydrolase Cel6A from *Trichoderma reesei***. *Structure* 1999, **7**:1035-1045.

68. Davies GJ, Mackenzie L, Varrot A, Dauter M, Brzozowski AM, Schulein M, Withers SG: **Snapshots along an enzymatic reaction coordinate: Analysis of a retaining beta-glycoside hydrolase**. *Biochemistry* 1998, **37**:11707-11713.

69. Parkkinen T, Koivula A, Vehmaanpera J, Rouvinen J: **Crystal structures of *Melanocarpus albomyces* cellobiohydrolase Cel7B in complex with cello-oligomers show high flexibility in the substrate binding**. *Protein Science* 2008, **17**:1383-1394.

70. Kern M, McGeehan JE, Streeter SD, Martin RNA, Besser K, Elias L, Eborall W, Malyon GP, Payne CM, Himmel ME, et al.: **Structural characterization of a unique marine animal Family 7 cellobiohydrolase suggests a mechanism of cellulase salt tolerance**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**:10189-10194.

71. Stahlberg J, Divne C, Koivula A, Piens K, Claeyssens M, Teeri TT, Jones TA: **Activity studies and crystal structures of catalytically deficient mutants of cellobiohydrolase I from *Trichoderma reesei***. *Journal of Molecular Biology* 1996, **264**:337-349.

72. Koshland DE: **Stereochemistry and the mechanism of enzymatic reactions**. *Biological Reviews of the Cambridge Philosophical Society* 1953, **28**:416-436.

73. Bu LT, Nimlos MR, Shirts MR, Ståhlberg J, Himmel ME, Crowley MF, Beckham GT: **Product Binding Varies Dramatically between Processive and Nonprocessive Cellulase Enzymes**. *Journal of Biological Chemistry* 2012, **287**:24807-24813.

74. Rouvinen J, Bergfors T, Teeri T, Knowles JK, Jones TA: **Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei***. *Science* 1990, **249**:380-386.

75. Sandgren M, Wu M, Karkehabadi S, Mitchinson C, Kelemen BR, Larenas EA, Stahlberg J, Hansson H: **The Structure of a Bacterial Cellobiohydrolase: The Catalytic Core of the Thermobifida fusca Family GH6 Cellobiohydrolase Cel6B**. *Journal of Molecular Biology* 2013, **425**:622-635.

76. Wu M, Bu L, Vuong TV, Wilson DB, Crowley MF, Sandgren M, Ståhlberg J, Beckham GT, Hansson H: **Loop motions important to product expulsion in the Thermobifida fusca glycoside hydrolase Family 6 cellobiohydrolase from structural and computational studies**. *Journal of Biological Chemistry* 2013, **in press**.

77. Koivula A, Ruohonen L, Wohlfahrt G, Reinikainen T, Teeri TT, Piens K, Claeyssens M, Weber M, Vasella A, Becker D, et al.: **The Active Site of Cellobiohydrolase Cel6A from Trichoderma reesei:‚Äâ The Roles of Aspartic Acids D221 and D175**. *Journal of the American Chemical Society* 2002, **124**:10015-10024.

78. Sammond DW, Payne CM, Brunecky R, Himmel ME, Crowley MF, Beckham GT: **Cellulase linkers are optimized based on domain type and function: Insights from sequence analysis, biophysical measurements, and molecular simulation**. *Plos One* 2012, **7**:e48615.

79. Payne CM, Resch MG, Chen L, Crowley MF, Himmel M, Taylor LEI, Sandgren M, Ståhlberg J, Stals I, Tan Z, et al.: **Glycosylated linkers in multi-modular lignocellulose degrading enzymes dynamically bind to cellulose**. *Proceedings of the National Academy of Sciences* 2013, **in press**.

Figure Captions:

Figure 1. The proposed processive catalytic cycle of a processive GH, illustrated here using the *T. reesei* Cel7A enzyme as an example. *T. reesei* Cel7A is shown in gray with *N*- and *O*-glycosylation in blue and yellow, respectively. The polysaccharide surface is shown in green. The enzyme first associates with the substrate and then repeatedly cleaves the glycosidic linkage releasing a dimeric product until it eventually dissociates from the polysaccharide surface. We hypothesize that binding free energy, ΔG_b^o, is proportional to the natural log of the association rate constant, k_{on}, divided by the dissociation rate constant, k_{off}. This can be further related to intrinsic processivity through the catalytic rate constant, k_{cat}.

Figure 2. Structural evidence suggests that the processive catalytic cycle of a GH Family 7 cellobiohydrolase consists of at least 6 steps: 1) the starting configuration with product sites initially vacant (*top left frame*), 2) processing of a cellulose chain across the active site (*top right*, showing the 5CEL substrate in green and the 7CEL -3 sugar in blue for reference), 3) rotation/translation of the cellulose chain and distortion of the -1 sugar ring, producing the Michaelis complex (*lower right*, PDB code 8CEL), 4) Step 1 of the hydrolytic cycle (Glycosylation) forming the glycosyl-enzyme intermediate (GEI) with the cellobiose product in 'Unprimed' mode (*lower left*, showing the 7CEL structure in green and a modeled glucosyl residue in the chair conformation in blue), 5) translation of the product towards the tunnel exit producing the 'Primed' GEI (*lower left*, PDB code 3CEL product in orange), and 6) Step 2 of the catalytic cycle (Deglycosylation) breaks the glycosyl-enzyme covalent bond, and product expulsion completes the processive cycle. Each frame also shows the key catalytic residues in yellow: Glu212 (nucleophile, top left residue in each frame), Asp214 (top right), and Glu217 (acid/base, bottom).

Figure 3. (A) Structural evidence suggests that the processive catalytic cycle of a GH Family 6 cellobiohydrolase consists of at least four steps, which involve conformational changes in the catalytic center loop highlighted in blue: 1) the starting configuration with product (-1/-2) subsites initially vacant (*top left frame*), 2) processing of a cellulose chain across the active site to 'Slide' mode (*top right*, showing the 4AVO substrate in green and the 1QK2 enzyme in gray), 3) closing of the catalytic center loop wherein the Michaelis complex is formed (*lower right*, 4AVO substrate in green, the catalytic acid is from the 1HGW structure, and the 1QJW enzyme in gray), 4) hydrolysis occurs to form the Substrate-Product complex with an α-cellobiose unit in the product site (*lower left*, modified 4AVN substrate with a modeled product [76] and the catalytic center loop still closed) and 4) product expulsion completes the processive cycle combined with the catalytic center loop opening. Each panel also shows the key catalytic residues in yellow with the *T. reesei* Cel6A residues: Asp221 (catalytic acid, top left residue), Asp175 (middle residue near the catalytic acid), Ser181 (residue on the catalytic center loop that stabilizes the catalytic water when the loop is closed), and Asp401 backbone atoms (stabilization residue for catalytic water when the catalytic center loop is closed, far right). (B) The 'open' and 'closed' forms of the enzyme based on the catalytic center loop. When the catalytic center loop is closed in the Michaelis complex, Ser181 and Asp401 stabilize the nucleophilic water and the putative water wire for proton transfer during the inverting hydrolysis mechanism.