THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

Molecular-Level Origins of Biomass Recalcitrance: Decrystallization Free Energies for Four Common Cellulose Polymorphs

Gregg T. Beckham,*[,†,‡,§] James F. Matthews,[⊥] Baron Peters,[¶,∥] Yannick J. Bomble,[⊥] Michael E. Himmel,[⊥] and Michael F. Crowley*[,⊥]

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

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

[§]Renewable and Sustainable Energy Institute, University of Colorado at Boulder, Boulder, Colorado 80309, United States

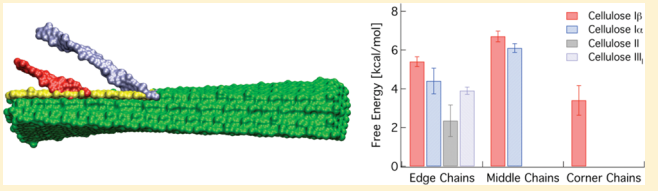
[⊥]Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado 80401, United States

[¶]Department of Chemical Engineering, University of California Santa Barbara, Santa Barbara, California 93106, United States

[∥]Department of Chemistry and Biochemistry, University of California Santa Barbara, Santa Barbara, California 93106, United States

Supporting Information

ABSTRACT: Cellulose is a crystalline polymer of β1,4-D-glucose that is difficult to deconstruct to sugars by enzymes. The recalcitrance of cellulose microfibrils is a function of both the shape of cellulose microfibrils and the intrinsic work required to decrystallize individual chains, the latter of which is calculated here from the surfaces of four crystalline cellulose polymorphs: cellulose Iβ, cellulose Iα, cellulose II, and cellulose III$_I$. For edge chains, the order of decrystallization work is as follows (from highest to lowest): Iβ, Iα, III$_I$, and II. For cellulose Iβ, we compare chains from three different locations on the surface and find that an increasing number of intralayer hydrogen bonds (from 0 to 2) increases the intrinsic decrystallization work. From these results, we propose a microkinetic model for the deconstruction of cellulose (and chitin) by processive enzymes, which when taken with a previous study [Horn et al. Proc. Natl. Acad. Sci. U.S.A. 2006, 103, 18089] identifies the thermodynamic and kinetic attributes of enzyme and substrate engineering for enhanced cellulose (or chitin) conversion. Overall, this study provides new insights into the molecular interactions that form the structural basis of cellulose, which is the primary building block of plant cell walls, and highlights the need for experimentally determining microfibril shape at the nanometer length scale when comparing conversion rates of cellulose polymorphs by enzymes.



INTRODUCTION

The constituent sugars of carbohydrate polymers in plants are vital for both the global carbon cycle and for the production of renewable liquid transportation fuels.[1−6] The main building block in plants is cellulose, which is the linear polymer of β1,4-D-glucose and the most abundant biological material on Earth. Cellulose is a tough, insoluble, crystalline polymer that plants use for defense, structure, and scaffolding. In plants, cellulose is a combination of cellulose Iβ and Iα,[7] which is both held together via a network of hydrogen bonds and hydrophobic interactions,[8−10] making deconstruction of the crystals difficult. This difficulty of deconstructing plant cell polymers to monomers is termed "recalcitrance".[1] To that end, Nature has evolved enzyme cocktails to deconstruct cellulose polymers to glucose, which primarily employ processive and nonprocessive cellulase enzymes. The mechanisms by which these enzyme cocktails break down cellulose is a question of significant scientific importance, which has implications in understanding carbon turnover in the biosphere and for developing an energy industry to utilize plant-based biomass for renewable fuels.

A promising approach for overcoming biomass recalcitrance is a two-step process in which the cell wall is treated thermal-chemically to improve enzymatic accessibility, known as pretreatment. Pretreatment is followed by enzymatic hydrolysis of cellulose to glucose.[1,11,12] Many options for pretreatment exist, including dilute acid hydrolysis, ammonia fiber explosion (AFEX), hot water cooking, and regeneration from novel solvents.[1,11] In most cases, the cellulose polymorph is retained as cellulose I. However, in ionic liquids[13,14] and ammonia-based pretreatments,[15] cellulose II or cellulose III$_I$ are formed, respectively. The mechanism of formation and the ability of enzymes to deconstruct plant cellulose or synthetic cellulose are not well understood at the molecular level.

The second step in biorefining to deconstruct cellulose is enzymatic hydrolysis. Enzyme cocktails can deconstruct intact or pretreated plant cell walls in a synergistic manner. The enzymes responsible for hydrolyzing most of the cellulose are known as "processive" cellulases, which thread single cellulose chains into long hydrophobic tunnels wherein hydrolysis is conducted along a single cellodextrin chain processively.[16] An example of a

Received: November 7, 2010
Revised: January 13, 2011
Published: March 22, 2011

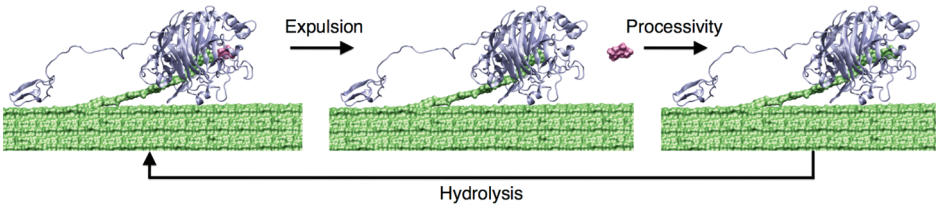


Figure 1. Catalytic cycle of a processive cellulase on cellulose. Processive cellulases (purple) are composed of carbohydrate-binding modules (left) and catalytic domains (right) connected by flexible linkers. The processive cycle of the Family 7 cellobiohydrolase from *Trichoderma reesei* (*Hypocrea jecorina*) acting on cellulose (green) is shown. The product (cellobiose) is shown in pink after expulsion from the catalytic domain. From left to right, the steps are hydrolysis, product expulsion, and threading (or processivity) of another catalytic unit into the enzyme to form the catalytically active complex.

processive cellulase is shown in Figure 1. The primary hypothesis for the mechanism of processive cellulase action to date is as follows: after binding to cellulose via a carbohydrate-binding module (CBM), processive enzymes locate a free chain end of cellulose on the surface of the crystal, and thread the chain into the tunnel of the catalytic domain (CD) to form a catalytically active complex.[12,16−20] The catalytically active complex is illustrated in Figure 1 for the Family 7 cellobiohydrolase from *Trichoderma reesei*, which is referred to here as Cel7A.[16] The catalytic cycle of a processive enzyme involves hydrolyzing the glycosidic linkage in cellulose, expelling cellobiose (the catalytic product, dimer of glucose), decrystallizing another cellobiose unit on the chain from the crystal, and threading the chain into the CD tunnel. For crystalline cellulose, a processive cellulase must perform work to decrystallize the polymer. This step is likely a contributor to the overall activation barrier because it is an inherent, molecular measure of "accessibility".[21] The decrystallization work will contribute to the overall recalcitrance (along with microfibril shape and chain population) and will likely vary depending on the cellulose polymorph and on chain location within the surface. If one can design processes to reduce the decrystallization free energy that enzymes must overcome, then conversion rates can be improved.

To examine the molecular origins of recalcitrance of cellulose and to calculate a key thermodynamic quantity in cellulase action, we compute the decrystallization free energy of cellulose chains from cellulose I ($I\beta$ and $I\alpha$) and two synthetic polymorphs (cellulose II and cellulose III_I). To our knowledge, only one type of study has examined this type of quantity before, but from oligomer chains placed on hydrophilic surfaces of cellulose $I\beta$ only.[22] We examine cellulose chains on the edge of a crystalline microfibril and chains from the middle of crystal faces, and in one case, a corner chain on the surface with no intralayer hydrogen bonds. These results highlight the intrinsic recalcitrance of cellulose as a function of polymorph and location of chains. The results indicate that edge chains in cellulose $I\beta$ and $I\alpha$ are more thermodynamically stable than in cellulose II and III_I. These results are a necessary component (with microfibril shape and populations of chain types) to understand the difference in enzymatic conversion rates observed between these cellulose polymorphs.[23−29] We propose a microkinetic model for cellulose hydrolysis from this work, which will drive development of enhanced enzyme systems tailored to pretreatment strategies, and also inform quantitative, morphology-based models of enzymatic deconstruction of cellulose.[30−32] From a plant biology standpoint, this study is the first measure of the physical integrity of cellulose crystals at the molecular scale.

Before discussing the methods and results, we first review the structure of the cellulose polymorphs studied here. The primary structural studies have been conducted in the past decade from pioneering diffraction studies, which form the basis of our study.[8,9,15,33]

Cellulose I. Plant cellulose is composed of cellulose $I\beta$ and $I\alpha$,[7] which form parallel chains arranged in 36-chain microfibrils.[10] Both cellulose $I\beta$ and $I\alpha$ exhibit 2-dimensional intralayer hydrogen-bonding networks.[8,9] Cellulose $I\alpha$ and $I\beta$ differ primarily in the hydrogen-bonded layer offset with cellulose $I\beta$ having "center" and "origin" chains (here we consider only origin chains). Based on data regarding synthesis of cellulose and microscopy observations, the shape of cellulose microfibrils from plants has been proposed to be either diamond-shaped 36-chain microfibrils or hexagonal-shaped 36-chain microfibrils.[10,34] Because neither of these has been definitively elucidated for plant cellulose, here we examined cases from both the diamond- and hexagonal-shaped microfibrils of plant cellulose.

Cellulose II. Ionic liquids and other novel solvents can dissolve cellulose completely.[13,14] Upon precipitation of cellulose, cellulose II is often formed. Cellulose II arranges in antiparallel sheets and exhibits a 3-dimensional network of both intralayer and interlayer hydrogen bonding.[33] The microfibril shape for cellulose II is not known and has not been proposed to our knowledge. Thus, here we assume a rectangular crystal of cellulose II for our study, which is found to be stable.

Cellulose III_I. Treatment of cellulose I with ammonia produces cellulose III_I.[15] Cellulose III_I forms both intralayer and interlayer hydrogen bonds in a manner similar to cellulose II, but the chains are parallel as in cellulose I. It has been observed that cellulose III_I is significantly less recalcitrant to deconstruction even though the binding of cellulase enzymes is reduced compared to cellulose I.[23] Like cellulose II, there are no models of microfibril shapes to our knowledge. Thus, we assumed a rectangular fibril for cellulose III_I like that from Yui et al.[35−37] However, unlike cellulose II, simulations of finite crystals of cellulose III_I in water are somewhat unstable.[35−37] This observation is likely a result of both using short cellulose microfibrils and not knowing the stable shape of cellulose III_I microfibrils, as has been observed in small-molecule organic crystals with simulation.[38] As these questions are outside of the scope of our study (and the latter is an experimental question), we restrained the bottom layer of the cellulose III_I structure during equilibration and decrystallization.

METHODS

We conducted simulations with CHARMM[39] using the C35 force field[40,41] and TIP3P water.[42,43] We decrystallized "origin" layer chains in cellulose $I\beta$ as "center" layer chains in cellulose $I\beta$ are similar to the displacement between layers in cellulose $I\alpha$. The microfibrils for cellulose II and cellulose III_I were constructed

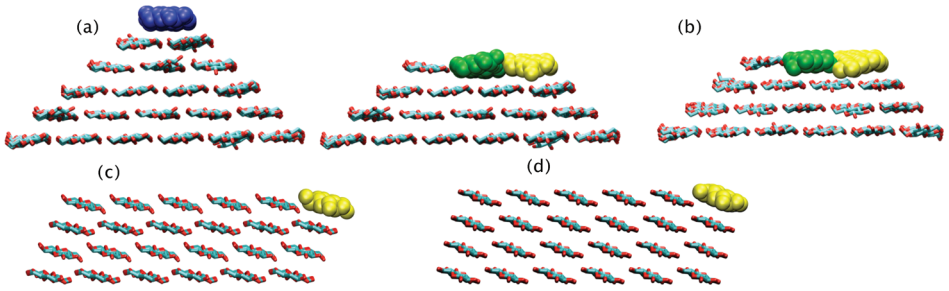


Figure 2. Decrystallization scenarios. (a) Cellulose Iβ: a "corner" chain (blue), and an "edge" chain (yellow) and a "middle" chain (green). Each highlighted chain is an independent simulation set. (b) Decrystallization scenarios for cellulose Iα, (c) cellulose II, and (d) cellulose III_I. An edge chain and middle chain were examined for cellulose Iβ and Iα and an edge chain for cellulose II and III_I. All bottom layers were restrained.

to examine a surface edge chain. As recently noted, it is difficult to assign the experimental microfibril shape for synthetic cellulose polymorphs because of the inherent difficulty in determining their shape and morphology experimentally.[44] However, since enzymatic deconstruction is a surface-mediated process, there will likely always be an edge chain on a surface of cellulose crystals. Thus, the results will be valid regardless of the actual shape of the cellulose microfibrils as long as accessible surface area exists.

For each polymorph, the microfibrils were 24 glucose units long. Each microfibril was equilibrated for 10 ns in the *NVT* ensemble after a 1 ns density equilibration in the *NPT* ensemble. Hydrogen bonds, primary alcohol conformations, and the root-mean-square fluctuations (RMSF) of each chain of interest were monitored. Hydrogen bond cutoffs were defined as within 3.4 Å of the donor and acceptor and 60° from linear. Primary alcohol conformations were defined as in Matthews et al.[10] The RMSF of each glucose unit was measured with only the heavy ring atoms. Following equilibration, the "bottom" half of each of the microfibrils was removed and the bottom layer of cellulose was harmonically restrained at 2.0 kcal/mol/Å^2. All MD simulations used a 2 fs time step. Temperature was maintained via the Nosé–Hoover thermostat.[45] Particle mesh Ewald was used for long-range electrostatics.[46] We used a cutoff of 12 Å for nonbonded interactions, and SHAKE was used to fix the bond distances to hydrogens.[47]

The free energy of decrystallization was calculated with umbrella sampling.[48,49] In each simulation, we decrystallized the chain of interest by 5 cellobioses (~5 nm). The reaction coordinate for all of the simulation sets was the fraction of native contacts for the chain of interest with the crystal.[50] We denote the reaction coordinate by ρ.[50] We define ρ such that $\rho = 0$ when all of the native contacts are present (i.e., the chain of interest is in the crystal) and $\rho = 1$ when all of the native contacts are broken (i.e., 5 nm of the chain of interest is decrystallized),[50] and we note that ρ does not contain any dynamical information.

Native contacts in the chain of interest were calculated from the equilibrations. Native contacts are defined as the center of mass of glucose ring atoms in the crystal within 12 Å of the center of mass of glucose ring atoms in the chain of interest. The 12 Å value for the reaction coordinate cutoff was chosen because it corresponds to the nonbonded cutoff distance so that there is no interaction with the fully decrystallized chain and the cellulose crystal, and it is far enough from the surface of the crystal to not affect the solvation. The PMF could be affected by the choice of the cutoff distance for ρ if it is less than the nonbonded cutoff distance so that the decrystallized chain can interact with the crystal. Moreover, increasing the cutoff distance beyond 12 Å will extend the number of glucose units that are decrystallized and increase the slope of the free energy curves. In the current study, the range of ρ from 0 to 1 corresponds to decrystallizing 5 cellobiose units.

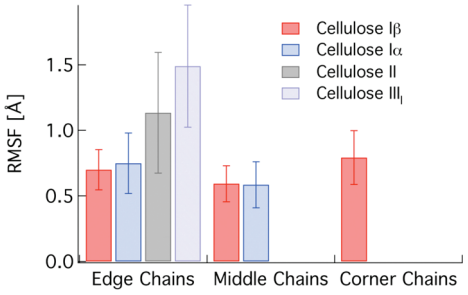


Figure 3. Chain RMSF from equilibration averaged over the entire chain of interest.

Each umbrella sampling window was run for 10 ns in the *NVT* ensemble at 300 K. ρ was sampled from 0 to 1 in intervals of 0.025 for adequate statistics. Each decrystallization scenario was thus 410 ns. The harmonic force constants on ρ were 2000 kcal/mol for edge and corner chains and 4000 kcal/mol for middle chains. The weighted-histogram analysis method was used to construct the free energy curves and the error analysis was conducted with bootstrapping.[51]

The 5 nm length scale was chosen because processive enzymes, like *T. reesei* Cel7A, have tunnels about 5 nm long into which cellodextrins thread.[16] Although it is possible to conduct many types of these free energy calculations from different crystal faces, we chose only to decrystallize chains from the hydrophobic faces of cellulose because it is known that CBMs bind to and diffuse along hydrophobic faces preferentially.[52−54] Here we calculate the work that such an enzyme must perform to decrystallize cellulose to form a catalytically active complex and deconstruct cellulose to sugars.[55,56]

Figure 2 shows the decrystallization scenarios. We show the crystals before they have been equilibrated for visual clarity, and we also show them with the bottom half already removed to show the microfibril sizes used in the umbrella sampling simulations.

To examine the differences in free energy between the different polymorphs and different chains within a given polymorph, we examined multiple quantities. We measured the occupancy of the chain of interest as it was being decrystallized at a mass-weighted isovalue of 1%, which contributes to the entropy term of the decrystallization work. We also examined the

Table 1. Average Number of Hydrogen Bonds (μ) for the Chains of Interest during Equilibration Averaged over the Entire Chain Reported with 1 Standard Deviation (σ)

chain of interest	location	μ	σ
Iβ edge	self	51.0	4.6
	intralayer	19.4	3.0
	interlayer	3.9	1.6
Iβ middle	self	48.5	4.6
	intralayer	40.0	4.4
	interlayer	2.5	1.3
Iβ corner	self	34.3	3.4
	intralayer	–	–
	interlayer	4.1	2.3
Iα edge	self	31.2	3.4
	intralayer	22.1	2.1
	interlayer	1.8	1.2
Iα middle	self	38.7	3.3
	intralayer	45.8	3.3
	interlayer	4.4	2.1
II edge	self	26.0	3.2
	intralayer	19.6	1.9
	interlayer	3.3	1.7
III_I edge	self	25.5	3.9
	intralayer	12.9	5.3
	interlayer	3.2	2.5

solvent-accessible surface area (SASA) of the crystal surfaces to determine if there was a difference in the amount of SASA exposed between the different scenarios. Additionally, we examined the RMSF of adjacent chains in the cellulose microfibrils to determine if the decrystallization affected the surrounding lattice.

RESULTS

Equilibration. We monitored the hydrogen-bonding behavior, the primary alcohol conformations, and the RMSF of the chains of interest during the equilibration. Figure 3 shows the average RMSF for each chain of interest averaged over the entire chain. The RMSF for edge chains in cellulose Iβ and Iα are approximately equal as are middle chains between the two natural cellulose polymorphs. The RMSFs of the middle chains, as expected, are slightly lower than edge chains and the corner chain in Iβ is slightly higher than the edge chain. The edge chains in cellulose II and III_I both exhibit a larger RMSF than edge chains in the natural cellulose polymorphs, which indicates that the cellulose edge chains in cellulose II and III_I may be less thermodynamically stable.

Table 1 shows the hydrogen bonds during the last 8 ns of the equilibrations for each chain of interest. We examined hydrogen bonds within a given chain of interest (self), within the same layer (intralayer), and between layers (interlayer). For cellulose Iβ, the main hydrogen bonds are within the given chain, and for the edge and middle chains there is approximately one intralayer hydrogen bond per residue. In cellulose Iβ and Iα, there are very few hydrogen bonds to the layers below, as expected. For cellulose Iα, there are fewer self-chain hydrogen bonds than in cellulose Iβ, and roughly the same number of hydrogen bonds in the same layer. In the cases of cellulose II and III_I, there is approximately one self-hydrogen bond per glucose and fewer intralayer hydrogen bonds than in the natural cellulose polymorphs. Especially in the edge chain in cellulose III_I, there are few hydrogen bonds (roughly one per every other residue) to the neighboring chains in the crystal, which is likely reflected in the RMSF in Figure 3.

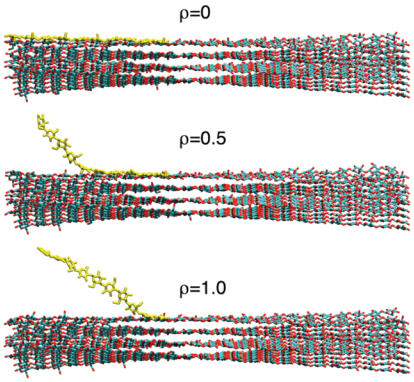


Figure 4. Snapshots of the cellulose Iβ decrystallization trajectory for the edge chain shown at the end points ($\rho = 0, 1$) and midpoint ($\rho = 0.5$) of the free energy calculation. The chain being decrystallized is shown in yellow.

We examined the primary alcohol conformations in the chains of interest in the four cellulose polymorphs as these conformations are related to the hydrogen bonds and potentially to the stability of the lattices. Figure S1 (in the Supporting Information) shows that for cellulose Iβ and Iα there is a considerable shift from the TG conformation. The cellulose II and III_I edge chains are predominantly in the GT conformation.

Decrystallization Simulations. From Figure 2, we examined for cellulose Iβ a corner chain with no intralayer hydrogen bonds (blue), an edge chain containing a single set of intralayer hydrogen bonds (yellow), and a middle chain containing intralayer hydrogen bonds on both sides (green). The corner chain was chosen to quantify the contribution of interlayer interactions. The edge and middle chains combine the interlayer interactions and intralayer hydrogen bonds. Implicit in all of the decrystallization simulations is the contribution of the solvation free energy of the newly exposed faces. These contributions are different in each case, and are difficult to measure directly.

For cellulose Iα, we examined an edge chain and a middle chain. Since both cellulose Iβ and Iα exhibit 2-dimensional hydrogen-bonding patterns, we anticipate that a cellulase may be able to extract middle chains in these polymorphs. In cellulose II and III_I, however, we measured the decrystallization work of an edge chain only. This is because extracting a middle chain from a surface in either of these polymorphs would first require lifting an adjacent chain. Thus, these calculations would require computing the work derived from the steric hindrance from adjacent chains that cover the middle chains in cellulose II or III_I.

Figure 4 shows the decrystallization trajectory for the cellulose Iβ polymorph edge chain at $\rho = 0$, 0.5, and 1. The cellulose chain being decrystallized has significant conformational flexibility in solution with the native contacts reaction coordinate, as it should.[57] Movies of the decrystallization trajectories are provided in the Supporting Information. Additionally, since the Iβ microfibril is not bonded across the periodic image, it adopts a twist as observed previously.[10]

In Figure 5, we show the free energy of decrystallization as a function of ρ for all scenarios. Figure 5a summarizes the results on a cellobiose basis with errors bars from the standard error of

dx.doi.org/10.1021/jp1106394 |J. Phys. Chem. B 2011, 115, 4118–4127

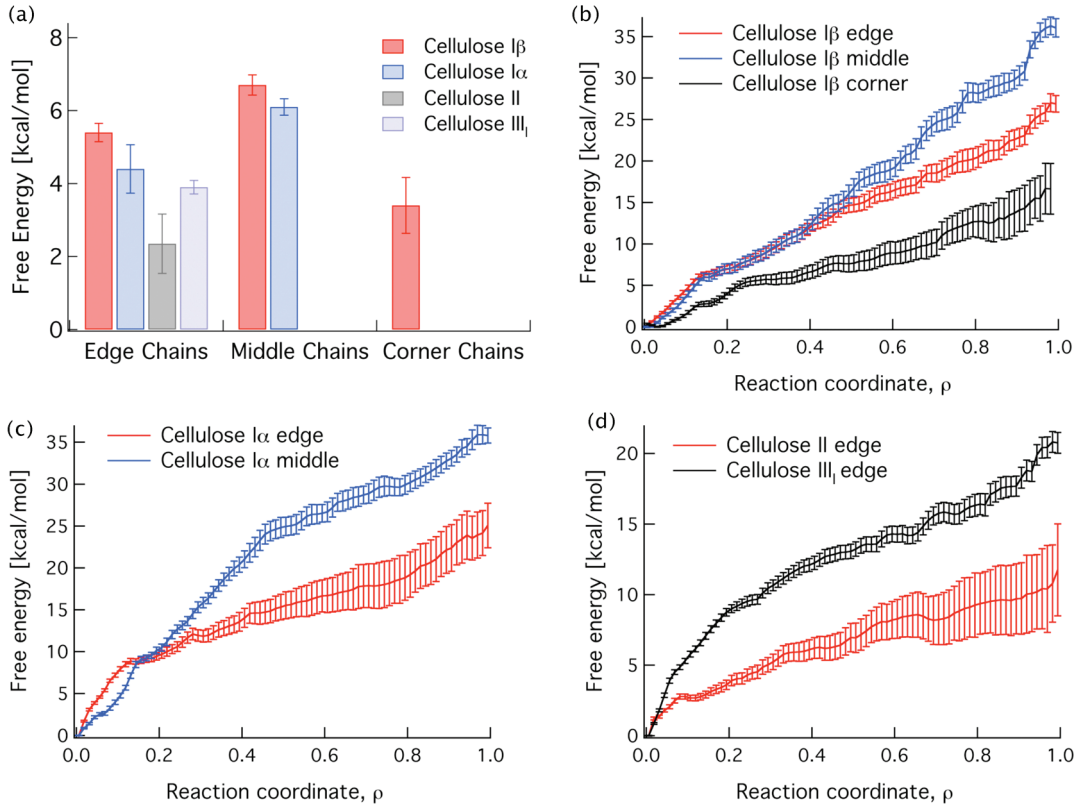


Figure 5. Free energy results for the decrystallization simulations: (a) free energy results per cellobiose unit for four cellulose polymorphs at different locations; (b) free energy as a function of ρ for cellulose Iβ; (c) free energy as a function of ρ for cellulose Iα; and (d) free energy as a function of ρ for cellulose II and III$_I$.

the mean. These quantities are calculated by dividing the free energy at $\rho = 1.0$ by the number of cellobiose units decrystallized (so the units are kcal/mol of cellobiose). For cellulose Iβ, the contribution of interlayer interactions (corner chain) is 3.4 kcal/mol/cellobiose. The addition of breaking intralayer hydrogen bonds increases the work to 5.4 kcal/mol for an edge chain and 6.7 kcal/mol/cellobiose for a middle chain. Slightly lower decrystallization free energies are obtained for cellulose Iα.

The edge chains in cellulose II and cellulose III$_I$ are easier to decrystallize by 1−2 kcal/mol/cellobiose than an edge chain in cellulose I. Cellulose II and III$_I$ have been shown to be less recalcitrant to enzymatic deconstruction than cellulose I.[23−29] The simulation results presented here thus offer insights into the reasons for the observed macroscopic differences between conversion rates. We stress, however, that it is not possible to delineate the contributions of the intrinsic work to decrystallize a single chain calculated here from the potential differences in chain populations caused by differences in the microfibril shapes. The latter is an experimental question which when combined with the present results will enable a more complete thermodynamic and kinetic description of biomass recalcitrance.

Several interesting observations were made from the simulations. For example, in the decrystallization of the edge chains and the corner chain, as illustrated in Figure 4 for cellulose Iβ, the chains were observed to peel up from the end of the crystals. These processes resemble an unzipping process as the chains are peeled from the surface from the end. In the cases of middle chains and the edge chains from the synthetic polymorphs, the free energy curves exhibit an inflection at low values of ρ. This is because in equilibrium MD simulations of cellulose microfibrils, regardless of polymorph, the ends will fray and fluctuate more significantly than the chains in the interior (Figure S2 in the Supporting Information). For cellulose II and III$_I$, up to a cellobiose or more can readily come dislodged from the crystal. Thus, during these decrystallization trajectories, the inflection is a result of the adjacent chain pulling out of the microfibril by a short distance (usually ∼2−3 glucoses) with the chain of interest. Eventually, the adjacent chain will anneal back into the crystal when the chain of interest no longer has significant interaction with the adjacent chain end. Thus, the initial and final states of the free energy profile represent the chain of interest in the crystal as the initial state and the chain of interest 5 nm out of the crystal fully flexible in solution as the final state.

Also, we note that in Figure 5b−d the free energy at small values of ρ ($\leq$0.2) are quite similar. All of the chains that were decrystallized were at the ends of microfibrils. During the simulations, the ends of finite cellulose microfibrils will "fray" in that small segments of cellodextrin chains on the surface will undergo fluctuations up to a cellobiose unit out of the lattice. This is the case for the edge chains and middle chains reported here, and hence explains the similar values of free energy for edge and middle chains at small values of ρ ($\leq$1 cellobiose unit) along the decrystallization trajectories.

We also note that, although the microfibril shapes are not known definitively from experiments, during the surface ablation process by which enzymes deconstruct cellulose, "edge" chains and "middle" chains will exist in any given substrate. Thus, because of the differences in the work to deconstruct cellulose,

dx.doi.org/10.1021/jp1106394 |J. Phys. Chem. B 2011, 115, 4118–4127

cellulose II and III_I will be more amenable to enzymatic attack on edge chains.

Decrystallization Simulation Analysis. We examined potential reasons for the differences seen in Figure 5 by measuring order parameters during equilibration and the decrystallization simulations. Overall, it is likely that the differences in free energy arise from a combination of steric effects determined by the location of the chain on the surface, differences in hydrogen bonding, the new faces exposed by decrystallization, the primary alcohol conformations, and the stability of each crystal lattice. During the decrystallization trajectories, we calculated the change in SASA of the intact crystal, the volume sampled by the decrystallized chain, and the RMSF of the adjacent chains. These data are provided in full in the Supporting Information.

From the SASA data (Figure S3 in the Supporting Information), we observe that cellulose Iβ and Iα edge chains have approximately the same SASA exposed. Middle chains in cellulose Iβ and Iα exhibit lower overall SASAs exposed during the decrystallization relative to the edge chains with the Iβ chain exposing slightly more overall SASA in the middle chain case. Additionally, the corner chain in cellulose Iβ and the two edge chains in cellulose II and III_I all expose approximately the same amount of SASA as edge chains in cellulose Iβ and Iα. We calculated the mass-weighted occupancy of the chains during the decrystallization trajectories (Figure S4 in the Supporting Information) to determine if the conformational space searched affected entropic differences in intrinsic work. As hypothesized, the edge chains have a higher occupancy volume than the middle chains in cellulose Iβ and Iα. Additionally, we also examined the RMSF of the adjacent chains over the decrystallization trajectories (Figure S5 in the Supporting Information) to determine if the adjacent chains exposed during decrystallization may contribute entropically to the decrystallization work. These data are shown in the Supporting Information.

DISCUSSION

We calculated the intrinsic work required to decrystallize a chain from the crystal surface of four common cellulose polymorphs and multiple chain morphologies. The equilibration trajectories demonstrate that edge chains in cellulose II and III_I exhibit higher RMSF fluctuations than edge chains or corner chains in plant cellulose, which is likely due to more stable hydrogen-bonding patterns at the crystal surfaces in plant cellulose. As expected from the RMSF data, the cellulose II and III_I edge chains exhibit a lower decrystallization work. We note that a more definitive comparison of these crystalline cellulose polymorphs will require better knowledge of the microfibril shape, and likely MD simulations up to 100s of nanoseconds of the intact crystal. However, here we assume shapes of microfibrils, which in the case of the plant cellulose we have some insights for,[10,34] to determine how much work an enzyme must conduct to pull chains off the hydrophobic surfaces of cellulose crystals.

The results shown in Figure 5 include contributions from hydrophobic interactions and hydrogen bonding, which impart incredible stability to cellulose. Our results suggest that the edge chains in natural plant cellulose, which stack in 2-dimensional hydrogen-bonded layers are more recalcitrant to deconstruction than edge chains in synthetic cellulose polymorphs with a 3-dimensional hydrogen bond network. Along the length of a cellulose microfibril, which is typically on the order of 100s to 1000s of cellobiose units long, the contribution of 1−2 kcal/mol for each cellobiose unit on the surface of plant cellulose will impart significant stability to the polymer crystals. The corner chain in cellulose Iβ is a special case that may not exist in Nature as it is from an inferred model for a plant cellulose microfibril, but it offers insights into interlayer contributions to the decrystallization work. Additionally, even if native cellulose microfibrils exhibit corner chains, it will likely be a very low population chain type.

There are several interesting differences within comparable subsets of data in this study. For cellulose Iβ and Iα edge chains, the Iα chains are slightly easier to decrystallize. This correlates with a slightly higher RMSF in the Iα chain, a higher sampled occupancy of the Iα chain in solution, and a larger amount of flexibility in the adjacent chains during decrystallization. For the cellulose Iβ and Iα middle chains, the same observations hold and offer some insight into the differences in decrystallization work. That the cellulose Iβ corner chain exhibits a lower decrystallization work is not surprising given that it forms few intralayer hydrogen bonds, explores a larger volume due to lower steric hindrance relative to edge and middle chains, and exhibits larger fluctuations when in the crystal. Explaining the difference in the cellulose II and III_I edge chains from the cellulose I is challenging to do definitively because the surfaces exposed are significantly different from one another, and the hydrogen-bonding patterns due to the crystal packing are 3-dimensional rather than the 2-dimensional hydrogen-bonding patterns found in natural cellulose. However, there are some variables that indicate there will be a difference. Most strikingly, the RMSF of the chains in cellulose II and III_I are higher than that in cellulose I, the hydrogen-bonding patterns indicate fewer hydrogen bonds to neighboring chains (for edge chains), and compared to Iβ edge chains, the cellulose II and III_I explore more conformational space during decrystallization. It is likely that the types of surfaces exposed play a significant role in the difference in decrystallization work, which again highlights the need for experimental characterization of microfibril shapes in these different polymorphs.

We again stress here that recalcitrance of cellulose is both a function of the intrinsic work that an enzyme must conduct to decrystallize a single cellulose chain (calculated here) and of the shape of the microfibril. Thus, from these results alone, it is not possible to make claims of differences in recalcitrance that are directly comparable to experiment measured by enzymatic conversion rates. More appropriately, our results indicate the relative difference in work to decrystallize edge chains between the four cellulose polymorphs studied here. If the microfibril shapes are the same between all four polymorphs, our results suggest an order of deconstruction rates. However, it is unlikely given the differences in the unit cells between these four polymorphs that they will exhibit the same microfibril shape or the same chain populations. Thus, with different microfibril shapes between plant celluloses and synthetic celluloses, the relative amounts of edge, corner, and middle chains will be different. This difference in addition to the differences in intrinsic decrystallization work will lead to overall, macroscopic differences in conversion rate, and observations of recalcitrance differences between cellulose polymorphs.

Computational Implications. From a computational standpoint, using the fraction of native contacts as a reaction coordinate yields a decrystallization trajectory that allows significant conformational flexibility of the cellulose chain in solution. Thus, the free energy change we are calculating is the difference of the free energy of the flexible chain in solution (but still bonded to cellulose)

with the new surfaces created on the cellulose crystal from the free energy of the chain in the crystal. In each case, the surfaces that are exposed will be different. For example, in cellulose Iβ, the edge chain and middle chain scenarios will create surfaces with two edge chains or with a water channel between two cellulose chains, respectively. This will in turn change the number of water−cellulose hydrogen bonds between the two scenarios. Thus, the native contacts reaction coordinate used here is more realistic than coordinates designed to pull on a given chain end (e.g., distance-to-a-plane) because pulling cellulose chains by the ends alone limits the conformational flexibility in solution similar to a taut string such that entropic factors are not included appropriately in the decrystallized end state. In preliminary simulations of the cellulose Iβ edge and middle chains, we used a reaction coordinate of the distance of the leading glucose (at the edge of the crystal) in the chain of interest above the plane of the cellulose microfibril, and we observed that the distance-above-a-plane reaction coordinate does not allow the chain to adopt helical conformations in solution, as should be expected from recent replica-exchange MD simulations of cellulose oligomers.[57] This type of reaction coordinate does not include the appropriate entropic factors in the calculated free energies, and was thus not used. In addition, using reaction coordinates that transform the system from the chain in the crystal to a single final conformation will likely overconstrain the product state of the system, and as such these types of approaches should not be used to calculate meaningful free energies for simulations of this type.

The reaction coordinate used here to study the material properties of individual chains in polymer crystals can be readily extended to cellulose in other solvents, such as ionic liquids where cellulose is known to dissolve, or to other biopolymer structures, such as other polymer crystals (e.g., chitin). This approach can be also applied to study the aggregation of cellulose microfibrils at much larger scales, where hydration and dehydration of given cellulose fibrils will likely play a role in self-assembly of cellulose during biosynthesis or during rearrangement of fibrils in enzymatic hydrolysis.[58,59] In addition, this type of thermodynamic approach can be used to study self-assembly and material properties of other types of materials where anisotropic molecular details are relevant.

Implications for Enzymatic Hydrolysis of Cellulose. From the standpoint of biomass deconstruction, our results have implications in both understanding cellulose structure as well as in understanding how processive enzymes deconstruct cellulose. First, we highlight the internal consistency of our results with respect to enzymatic deconstruction. Since we have the largest set of data for cellulose Iβ, it offers a view into the overall free energy of enzymatically deconstructing an entire cellulose layer. Starting with the cellulose Iβ microfibril structure shown in Figure 2 with three chains in the hydrophobic layer, the surface ablation process by which enzymes deconstruct cellulose could occur via two different pathways. One pathway is that an enzyme could acquire the middle chain and convert that first (ΔG_{middle}), which would produce two corner chains. Then, one corner chain can be removed (ΔG_{corner}), resulting in a single corner chain left on the surface. The second pathway would be that an enzyme would acquire an edge chain (ΔG_{edge}) and convert that first, which would leave two edge chains on the surface. The second edge chain could then be acquired and converted (ΔG_{edge}), leaving the same as in the first pathway with a single corner chain on the surface. The work required for these two pathways, which are shown in Figure 6, would be $\Delta G_{middle} + \Delta G_{corner} = 6.7 + 3.4 = 10.1$ kcal/mol or $\Delta G_{edge} + \Delta G_{edge} = 2 \times 5.4 = 10.8$ kcal/mol.

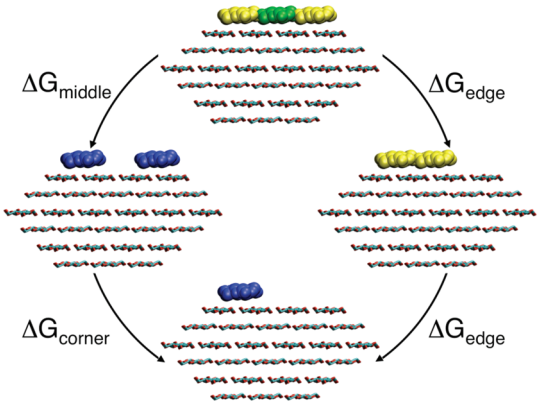


Figure 6. Two distinct pathways for the conversion of a single layer in cellulose Iβ. The work required to reach the endstates is equal based on the simulations conducted here, which serves as an internal consistency check. The chains are colored the same as in Figure 2.

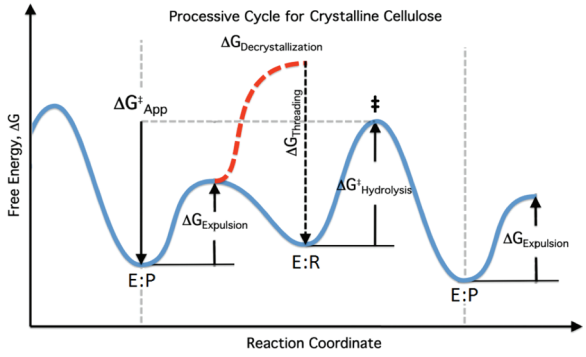


Figure 7. Proposed free energy diagram of processive cellulase action on cellulose as a function of a generalized reaction coordinate. E:P denotes the enzyme + product (after the hydrolysis step has occurred). E:R denotes the catalytically active complex of the enzyme prepared to conduct hydrolysis. There is work associated with expelling the product ($\Delta G_{expulsion}$), decrystallizing cellulose (shown as a dotted red line, $\Delta G_{decrystallization}$) and a favorable ligand binding free energy ($\Delta G_{threading}$) in the enzyme.

That the work for either pathway is equivalent validates the convergence of the results and indicates internal consistency.

Substantial efforts are being conducted on pretreatment and enzyme science for overcoming biomass recalcitrance. Thus, many groups are attempting to describe the molecular action of cellulase (and chitinase) action on cellulose (and chitin) with the aim of rational engineering of improved catalysts.[16−19,52−55,60,61] Here, we have computed a key thermodynamic parameter for deconstruction of cellulose. For processive enzymes (Figure 1), the specific thermodynamic contributions to deconstruction of cellulose are illustrated in eq 1. The overall free energy change from the processive cycle, ΔG_{cycle}, must be free energetically downhill (e.g., because there is no ATP hydrolysis or similar reactions when processive enzymes work on cellulose), which includes the following contributions:

$$\Delta G_{cycle} = \Delta G_{decrystallization} + \Delta G_{threading} + \Delta G_{hydrolysis} + \Delta G_{expulsion} \quad (1)$$

In addition to the overall free energy change, the sequence of events in the processive cycle influences the rate of cellulose

conversion. For a given cellulase, the hydrolysis free energy ($\Delta G_{hydrolysis}$) and the product expulsion free energy ($\Delta G_{expulsion}$) will likely be constant at the molecular level for all forms of cellulose, whether soluble oligomers, amorphous cellulose, or crystalline cellulose. Therefore, observations that hydrolysis rates depend on the polymorph likely imply a sequence of events as shown in Figure 7. The free energy diagram in Figure 7 is tantamount to a microkinetic model for the processive cycle of enzymatic cellulose deconstruction.[62,63] A given substrate will have an inherent decrystallization free energy (more rigorously a characteristic distribution of decrystallization free energies depending on available reactive surface chains for catalysis). The processive cycle of a cellulase or chitinase can then be analyzed in terms of two states: E:P = the enzyme with the hydrolyzed cellobiose (or chitobiose) product still in the tunnel, and E:R = the enzyme with the unhydrolyzed and threaded cellulose (or chitin) prepared for chemical hydrolysis. This picture assumes that decrystallization and threading occur in a concerted fashion. The unfavorable decrystallization step (red dotted line in Figure 7) is thus compensated by the favorable threading process, circumventing the barrier to perform the decrystallization step separately. If chemical hydrolysis is the rate-limiting step, then $\Delta G_{decrystallization}$ can still affect the overall kinetics, providing a molecular level explanation for commonly observed substrate dependence beyond available surface area for chemical reactions.[21,64] Quantitatively, the rate law from Figure 7 is rate = $k_{hydrolysis}K/([\text{cellobiose}] + K)$, where K is the equilibrium constant for E:P ↔ E:R + cellobiose. Therefore, we suggest that the kinetic parameter K is $-RT \ln K = \Delta G_{0,expulsion} + \Delta G_{0,decrystallization} + \Delta G_{0,threading}$. Although expulsion, threading, and decrystallization may occur as concerted processes, simulation can separately identify their contributions to the overall work and probe substrate effects. We have done this analysis here for the decrystallization work as a function of both chain location on the surface of a cellulose crystal and as a function of cellulose polymorph. We also note that even when hydrolysis is the rate-limiting step, substrate decrystallization free energy will influence the overall rate as long as $K < [\text{cellobiose}]$.

The three terms that enter the equilibrium constant K each represent promising targets for designing pretreament processes and for engineering new enzymes. The $\Delta G_{decrystallization}$ term is a target variable in designing pretreatment processes[1,11] and disruptor proteins.[65−67] If one can increase the molecular accessibility of cellulose, by transforming the crystal structure either to a different polymorph or to amorphous cellulose or to a microfibril shape with higher surface area, the available reactive surface area for enzyme attack will increase. The threading free energy ($\Delta G_{threading}$), which is related to the ligand binding free energy, is an important contributor to the $\Delta G_{decrystallization}$ term such that, if $\Delta G_{decrystallization}$ term is lowered, the $\Delta G_{threading}$ term could also likely be reduced. The $\Delta G_{threading}$ term is linked to the rate of threading a new monomer unit into the tunnel and represents significant research in enzyme engineering for improved catalysts. This was demonstrated in a landmark study from Horn et al. for a processive chitinase,[64] wherein aromatic residues lining the catalytic tunnel were mutated to alanine. They found that activity on crystalline chitin decreased, while activity on soluble chitosan increased by a factor of 30 for one amino acid mutation. The authors used additional experimental methods to interpret their findings as increased processivity rates in the alanine mutants, which will be directly linked to the quantity $\Delta G_{threading}$. Because chitin and cellulose and the enzymes that deconstruct them are quite similar, the results from Horn et al. will likely extend to cellulose and processive cellulases. The observation of increased processivity rates was later confirmed for a cellulase by Vuong and Wilson.[68] Taken together, the current study and the study from Horn et al. represent two key pieces of the thermodynamics and kinetics of processive enzyme action that offer a more detailed explanation of substrate dependence at the molecular level. The hydrolysis free energy, $\Delta G_{hydrolysis}$, is a constant for all forms of cellulose, as is the product expulsion free energy, $\Delta G_{expulsion}$, although both quantities can be altered via protein engineering. These quantities will contribute to the overall rate of the processive step, which is a difficult quantity to measure. A recent attempt to measure the rate of the processive step used a novel high-speed atomic force microscope for *T. reesei* Cel7A, and estimated the rate of movement on cellulose to be on the order of 7 nm/s.[18] As the authors note, it is difficult to ascertain if the enzymes in question were "productively" bound to cellulose (i.e., hydrolyzing cellulose), or simply translating down a cellulose microfibril during the period of observation.[18] Considering that the enzymes are observed to translate only in one direction, however, lends merit to observation that the processive enzymes are productively bound and that the rate measurement is reasonable.

Overall, this model highlights the need for codesign of pretreatment technologies and enzymes for overcoming biomass recalcitrance. Our results indicate that cellulose Iβ and Iα, for which natural enzymes evolved, require that enzymes conduct more work to hydrolyze edge chains than cellulose II or III_I in reaching the transition state in the processive cycle. If $\Delta G_{decrystallization}$ is reduced by changing cellulose morphology (or increasing the number of corner and edge chains), the apparent activation energy ($\Delta G^{\ddagger}_{app}$) will be lowered resulting in faster kinetics. Taken with a previous enzyme engineering study,[64] the microkinetic model also suggests that one can achieve an even higher conversion rate by reducing the ligand binding free energy ($\Delta G_{threading}$) when $\Delta G_{decrystallization}$ is reduced. Namely, production of cellulose II and III_I with the same or smaller fibril sizes will likely lead to faster conversion when using engineered enzymes rather than natural cellulases that have evolved to work effectively on cellulose I from plants.

Additionally, here we focus on processive cellulases as they form the majority of the enzymes in natural cocktails. However, the results presented in this study also have implications for constructing models of the action of nonprocessive endoglucanases. We hypothesize that given the cleft nature of most endoglucanases that the ligand binding free energy will be lower (i.e., less favorable), thus endoglucanases will likely only be able to bind to chains with much lower decrystallization free energies. The absolute binding free energies of processive cellulases and nonprocessive endoglucanases will need to be quantified either with absolute binding free energy calculations or experimentally to determine the difference between the two types of enzymes.

Implications for Cellulose Structure. From the standpoint of understanding cellulose structure, the results presented here indicate that surface chains in cellulose Iβ have slightly higher decrystallization free energies than cellulose Iα. In enzymatic hydrolysis of cellulose I, the Iα phase has been observed to be degraded preferentially before the Iβ phase, producing a Iβ-rich fraction as the digestion progresses.[69] It should be noted that it is quite difficult to ascertain the location of cellulose Iβ relative to Iα in plant cellulose,[70] so that the observed differences in hydrolysis rates may not solely be due to differences in decrystallization free energy at the molecular level as shown here. Again,

other factors in the differences could be likely due to differences in polymorph distribution or differences in fibril shape.

In terms of cellulose II and III_I, here we have shown that edge chains in these two synthetic polymorphs are likely to be more amenable to deconstruction by processive enzymes than edge chains in cellulose I. There have been several experimental observations wherein cellulose II and/or III_I are more digestible.[23,26−29,44] As previously mentioned, connecting these experimental observations directly to our simulation results are confounded by differences in microfibril shape, chain population, and polymorph distribution. For example, the transformation of cellulose I to cellulose III_I is reversible, but the crystal diameter is reduced during the reversion of cellulose III_I to cellulose Iβ.[28] This size reduction increases the surface area of the fibrils, likely resulting in more digestible cellulose through ammonia pretreatment as a result of producing a change in crystal shape with more surface chains. Further characterization of the shapes of microfibrils and distributions of surface chains relative to interior chains is needed to fully explain experimental observations between digestion rates of cellulose polymorphs. Here, we have provided molecular insights inaccessible to conventional experimental techniques to aid in the explanation of observations from the past several decades.

CONCLUSIONS

We have measured the intrinsic recalcitrance of cellulose from plant cell walls and two synthetic cellulose polymorphs, and have examined the separate effects of hydrophobic interactions and hydrogen bonding to deconstruction of cellulose. Our results demonstrate that edge chains in plant cellulose are more difficult to decrystallize than edge chains from synthetic cellulose polymorphs. These results offer a molecular explanation of the differences in recalcitrance observed between cellulose I and synthetic celluloses, which should be corrected for differences in microfibril shape and the populations of chain types. We anticipate that this study will lead to experimental measures of cellulose decrystallization free energies, more experimental focus on determining microfibril shapes, and a shift in enzyme engineering strategies for pretreated substrates. The simulation framework developed here can be extended to examine the molecular characteristics of cellulose in other solvents and to other biopolymers, such as chitin. This study if combined with experimental characterization of microfibril shape will elucidate the morphology dependence of enzymatic cellulose deconstruction, which is an essential parameter in models for enzyme action on cellulose.[32] This study with a previous enzyme engineering study[64] also highlights the individual variables for enzyme engineering and substrate modifications that can lead to enhanced conversion rates of recalcitrant polysaccharides.

ASSOCIATED CONTENT

Supporting Information. Movies of the decrystallization trajectories. Additional figures describing order parameters analyzed in the equilibration and decrystallization simulations. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: Gregg.Beckham@nrel.gov (G.T.B.); Michael.Crowley@nrel.gov (M.F.C.).

ACKNOWLEDGMENT

We thank the DOE Office of the Biomass Program and DOE Office of Science ASCR SciDAC program for funding. Computer time was provided by the TACC Ranger cluster under the National Science Foundation Teragrid grant no. MCB090159 and by the NREL Computational Sciences Center supported by the DOE Office of EERE under Contract no. DE-AC36-08GO28308. B.P. is supported by a National Science Foundation CAREER award no. 0955502. We acknowledge Giovanni Bellesia for helpful discussions, use of Alan Grossfield's WHAM code for the construction of the free energy curves and error analysis, and the anonymous reviewers for insightful and helpful comments. All figures and movies of MD trajectories were made with VMD 1.8.6.[71]

REFERENCES

(1) Himmel, M. E.; Ding, S. Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. *Science* **2007**, *315*, 804–807.

(2) Huber, G. W.; Chheda, J. N.; Barrett, C. J.; Dumesic, J. A. *Science* **2005**, *308*, 1446–1450.

(3) Steen, E. J.; Kang, Y. S.; Bokinsky, G.; Hu, Z. H.; Schirmer, A.; McClure, A.; del Cardayre, S. B.; Keasling, J. D. *Nature* **2010**, *463*, 559–U182.

(4) Atsumi, S.; Hanai, T.; Liao, J. C. *Nature* **2008**, *451*, 86–U13.

(5) Lynd, L. R.; Weimer, P. J.; van Zyl, W.; Pretorius, I. S. *Microbiol. Mol. Biol. Rev.* **2002**, *66*, 506–577.

(6) Ragauskas, A. J.; Williams, C. K.; Davison, B. H.; Britovsek, G.; Cairney, J.; Eckert, C. A.; Frederick, W. J.; Hallett, J. P.; Leak, D. J.; Liotta, C. L.; Mielenz, J. R.; Murphy, R.; Templer, R.; Tschaplinski, T. *Science* **2006**, *311*, 484–489.

(7) Atalla, R. H.; Vanderhart, D. L. *Science* **1984**, *223*, 283–285.

(8) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074–9082.

(9) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300–14306.

(10) Matthews, J. F.; Skopec, C. E.; Mason, P. E.; Zuccato, P.; Torget, R. W.; Sugiyama, J.; Himmel, M. E.; Brady, J. W. *Carbohydr. Res.* **2006**, *341*, 138–152.

(11) Mosier, N.; Wyman, C.; Dale, B.; Elander, R.; Lee, Y. Y.; Holtzapple, M.; Ladisch, M. *Bioresour. Technol.* **2005**, *96*, 673–686.

(12) Chundawat, S. P. S.; Beckham, G. T.; Himmel, M. E.; Dale, B. E. *Annu. Rev. Chem. Biomol. Eng*, in press. DOI: 10.1146/annurev-chembioeng-061010-114205.

(13) Swatloski, R. P.; Spear, S. K.; Holbrey, J. D.; Rogers, R. D. *J. Am. Chem. Soc.* **2002**, *124*, 4974–4975.

(14) Ramakrishnan, S.; Collier, J.; Oyetunji, R.; Stutts, B.; Burnett, R. *Bioresour. Technol.* **2010**, *101*, 4965–4970.

(15) Wada, M.; Chanzy, H.; Nishiyama, Y.; Langan, P. *Macromolecules* **2004**, *37*, 8548–8555.

(16) Divne, C.; Stahlberg, J.; Reinikainen, T.; Ruohonen, L.; Pettersson, G.; Knowles, J. K. C.; Teeri, T. T.; Jones, T. A. *Science* **1994**, *265*, 524–528.

(17) Zhong, L. H.; Matthews, J. F.; Hansen, P. I.; Crowley, M. F.; Cleary, J. M.; Walker, R. C.; Nimlos, M. R.; Brooks, C. L.; Adney, W. S.; Himmel, M. E.; Brady, J. W. *Carbohydr. Res.* **2009**, *344*, 1984–1992.

(18) Igarashi, K.; Koivula, A.; Wada, M.; Kimura, S.; Penttila, M.; Samejima, M. *J. Biol. Chem.* **2009**, *284*, 36186–36190.

(19) Divne, C.; Stahlberg, J.; Teeri, T. T.; Jones, T. A. *J. Mol. Biol.* **1998**, *275*, 309–325.

(20) Beckham, G. T.; Bomble, Y. J.; Bayer, E. A.; Himmel, M. E.; Crowley, M. F. *Curr. Opin. Biotechnol*, in press. DOI: 10.1016/j.copbio.2010.11.005.

(21) Jeoh, T.; Ishizawa, C. I.; Davis, M. F.; Himmel, M. E.; Adney, W. S.; Johnson, D. K. *Biotechnol. Bioeng.* **2007**, *98*, 112–122.

(22) Bergenstrahle, M.; Thormann, E.; Nordgren, N.; Berglund, L. A. *Langmuir* **2009**, *25*, 4635–4642.

(23) Igarashi, K.; Wada, M.; Samejima, M. *FEBS J.* **2007**, *274*, 1785–1792.
(24) Li, C. L.; Knierim, B.; Manisseri, C.; Arora, R.; Scheller, H. V.; Auer, M.; Vogel, K. P.; Simmons, B. A.; Singh, S. *Bioresour. Technol.* **2010**, *101*, 4900–4906.
(25) Singh, S.; Simmons, B. A.; Vogel, K. P. *Biotechnol. Bioeng.* **2009**, *104*, 68–75.
(26) Chen, Y.; Stipanovic, A. J.; Winter, W. T.; Wilson, D. B.; Kim, Y. J. *Cellulose* **2007**, *14*, 283–293.
(27) Jung, H.; Yoon, H. G.; Park, W. J.; Choi, C.; Wilson, D. B.; Shin, D. H.; Kim, Y. J. *Cellulose* **2008**, *15*, 465–471.
(28) Chanzy, H.; Henrissat, B.; Vuong, R.; Revol, J. F. *Holzforschung* **1986**, *40*, 25–30.
(29) Weimer, P. J.; French, A. D.; Calamari, T. A. *Appl. Environ. Microbiol.* **1991**, *57*, 3101–3106.
(30) Levine, S. E.; Fox, J. M.; Blanch, H. W.; Clark, D. S. *Biotechnol. Bioeng.* **2010**, *107*, 37–51.
(31) Zhou, W.; Schuttler, H. B.; Hao, Z. Q.; Xu, Y. *Biotechnol. Bioeng.* **2009**, *104*, 261–274.
(32) Bansal, P.; Hall, M.; Realff, M. J.; Lee, J. H.; Bommarius, A. S. *Biotechnol. Adv.* **2009**, *27*, 833–848.
(33) Langan, P.; Nishiyama, Y.; Chanzy, H. *Biomacromolecules* **2001**, *2*, 410–416.
(34) Ding, S. Y.; Himmel, M. E. *J. Agric. Food Chem.* **2006**, *54*, 597–606.
(35) Yui, T.; Hayashi, S. *Biomacromolecules* **2007**, *8*, 817–824.
(36) Yui, T.; Hayashi, S. *Cellulose* **2009**, *16*, 151–165.
(37) Yui, T.; Okayama, N.; Hayashi, S. *Cellulose* **2010**, *17*, 679–691.
(38) Beckham, G. T.; Peters, B.; Variankaval, N.; Starbuck, C.; Trout, B. L. *J. Am. Chem. Soc.* **2007**, *129*, 4714–4724.
(39) 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*, 1545–1614.
(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*, 2543–2564.
(41) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. *J. Chem. Theory Comput.* **2009**, *5*, 2353–2370.
(42) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926–935.
(43) Durell, S. R.; Brooks, B. R.; Bennaim, A. *J. Phys. Chem.* **1994**, *98*, 2198–2202.
(44) Bellesia, G.; Asztalos, A.; Shen, T.; Langan, P.; Redondo, A.; Gnanakaran, S. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2010**, *D66*, 1184–1188.
(45) Frenkel, D.; Smit, B. *Understanding Molecular Simulations: From Algorithms to Applications*, 2nd ed.; Academic Press: New York, 2002.
(46) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8857.
(47) Ryckaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, 23.
(48) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.
(49) Kottalam, J.; Case, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 7690–7697.
(50) Sheinerman, F. B.; Brooks, C. L. *J. Mol. Biol.* **1998**, *278*, 439–456.
(51) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
(52) Lehtiö, J.; Sugiyama, J.; Gustavsson, M.; Fransson, L.; Linder, M.; Teeri, T. T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 484–489.
(53) Beckham, G. T.; Matthews, J. F.; Bomble, Y. J.; Bu, L.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R.; Crowley, M. F. *J. Phys. Chem. B* **2010**, *114*, 1447–1453.
(54) Bu, L.; Beckham, G. T.; Crowley, M. F.; Chang, C. H.; Matthews, J. F.; Bomble, Y. J.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R. *J. Phys. Chem. B* **2009**, *113*, 10994–11002.
(55) Wilson, D. B. *Curr. Opin. Biotechnol.* **2009**, *20*, 295–299.
(56) Heinzelman, P.; Snow, C. D.; Wu, I.; Nguyen, C.; Villalobos, A.; Govindarajan, S.; Minshull, J.; Arnold, F. H. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 5610–5615.
(57) Shen, T. Y.; Langan, P.; French, A. D.; Johnson, G. P.; Gnanakaran, S. *J. Am. Chem. Soc.* **2009**, *131*, 14786–14794.
(58) Cosgrove, D. J. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 850–861.
(59) Santa-Maria, M.; Jeoh, T. *Biomacromolecules* **2010**, *11*, 2000–2007.
(60) 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*, 8979–8984.
(61) Beckham, G. T.; Bomble, Y. J.; Matthews, J. F.; Taylor, C. B.; Resch, M. G.; Yarbrough, J. M.; Decker, S. R.; Bu, L.; Zhao, X.; McCabe, C. M.; Wolhert, J.; Bergenstråhle, M.; Brady, J. W.; Adney, W. S.; Himmel, M. E.; Crowley, M. F. *Biophys. J.* **2010**, *99*, 3773–3781.
(62) Murdoch, J. R. *J. Chem. Educ.* **1981**, *58*, 32.
(63) Dumesic, J. A. *J. Catal.* **1999**, *185*, 496–505.
(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*, 18089–18094.
(65) Eijsink, V. G. H.; Vaaje-Kolstad, G.; Varum, K. M.; Horn, S. J. *Trends Biotechnol.* **2008**, *26*, 228–235.
(66) Vaaje-Kolstad, G.; Horn, S. J.; van Aalten, D. M. F.; Synstad, B.; Eijsink, V. G. H. *J. Biol. Chem.* **2005**, *280*, 28492–28497.
(67) Harris, P. V.; Welner, D.; McFarland, K. C.; Re, E.; Poulsen, J. C. N.; Brown, K.; Salbo, R.; Ding, H. S.; Vlasenko, E.; Merino, S.; Xu, F.; Cherry, J.; Larsen, S.; Lo Leggio, L. *Biochemistry* **2010**, *49*, 3305–3316.
(68) Vuong, T. V.; Wilson, D. B. *Appl. Environ. Microbiol.* **2009**, *75*, 6655–6661.
(69) Hayashi, N.; Sugiyama, J.; Okano, T.; Ishihara, M. *Carbohydr. Res.* **1997**, *305*, 261–269.
(70) Horikawa, Y.; Sugiyama, J. *Biomacromolecules* **2009**, *10*, 2235–2239.
(71) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–&.