JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Comparison of Cellulose Iβ Simulations with Three Carbohydrate Force Fields

James F. Matthews,*[,†] Gregg T. Beckham,[‡,§] Malin Bergenstråhle-Wohlert,[∥,⊥] John W. Brady,[∥] Michael E. Himmel,[†] and Michael F. Crowley*[,†]

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

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

[∥]Department of Food Science, Cornell University, Ithaca, New York, United States

[⊥]Wallenberg Wood Science Center, Royal Institute of Technology, Stockholm, Sweden

Supporting Information

ABSTRACT: Molecular dynamics simulations of cellulose have recently become more prevalent due to increased interest in renewable energy applications, and many atomistic and coarse-grained force fields exist that can be applied to cellulose. However, to date no systematic comparison between carbohydrate force fields has been conducted for this important system. To that end, we present a molecular dynamics simulation study of hydrated, 36-chain cellulose Iβ microfibrils at room temperature with three carbohydrate force fields (CHARMM35, GLYCAM06, and Gromos 45a4) up to the near-microsecond time scale. Our results indicate that each of these simulated microfibrils diverge from the cellulose Iβ crystal structure to varying degrees under the conditions tested. The CHARMM35 and GLYCAM06 force fields eventually result in structures similar to those observed at 500 K with the same force fields, which are consistent with the experimentally observed high-temperature behavior of cellulose I. The third force field, Gromos 45a4, produces behavior significantly different from experiment, from the other two force fields, and from previously reported simulations with this force field using shorter simulation times and constrained periodic boundary conditions. For the GLYCAM06 force field, initial hydrogen-bond conformations and choice of electrostatic scaling factors significantly affect the rate of structural divergence. Our results suggest dramatically different time scales for convergence of properties of interest, which is important in the design of computational studies and comparisons to experimental data. This study highlights that further experimental and theoretical work is required to understand the structure of small diameter cellulose microfibrils typical of plant cellulose.

INTRODUCTION AND BACKGROUND

Carbohydrates are a primary energy storage mechanism in nature and are also useful motifs in signaling and biological recognition events via glycosylation. Additionally, carbohydrate polymers can serve as the foundation for complex structural composites such as those found in the cell walls of plants, fungi, and bacteria and in the exoskeletons of insects and crustaceans. The most abundant carbohydrate molecule on earth is cellulose, which is the primary structural carbohydrate in plant cell walls.[1,2] Cellulose represents a major component of the global carbon cycle and a viable resource for renewable fuels and chemicals.[3,4] Also, there is growing interest in cellulose-based nanomaterials.[5] The well-ordered fraction of plant cellulose consists of two crystalline forms, cellulose Iβ and Iα, with the Iβ form being the predominant polymorph.[6−8] The crystal structures of these two polymorphs have been solved through high-resolution X-ray and neutron fiber diffraction experiments, with the sample material consisting of large diameter (10−20 nm) algal or tunicate microfibrils cast into films.[9,10] A striking feature of both the cellulose Iβ and the Iα structures is that both form layers of two-dimensional hydrogen-bonded sheets, with many weak but no strong directional interactions between layers.[11,12] This observation is in contrast to the common synthetic polymorphs, namely cellulose II and III_I, which form both intralayer and interlayer hydrogen bonds.[13,14] Both the cellulose Iα and Iβ crystal structures exhibit primary alcohol groups in the trans−gauche (TG) conformation exclusively. The stacking arrangement of two-dimensional hydrogen-bonded layers differs between the Iα and Iβ polymorphs such that Iβ has $P2_1$ symmetry and alternates layer displacement along the *c*-axis between +c/4 and −c/4 (defined as center and origin chains). Iα has *P*1 symmetry, and layers are consistently displaced by +c/4. In the Iβ and Iα crystal structures at 273 K, two mutually exclusive hydrogen-bond networks were resolved, denoted the "**A**" and "**B**" hydrogen-bond patterns.[9,10] Subsequently, the less populated **B** pattern in Iβ was interpreted to likely be an artifact due to the presence of conformational disorder, when an additional Iβ structure obtained at 15 K was unable to resolve a single predominant hydrogen-bond network, but the nature of this disorder is still an unresolved question.[15,16]

It is important to note that these details of the bulk structure of cellulose crystals have been determined using large diameter (10−20 nm) microfibrils from algae and tunicates, but cellulose microfibrils from plants can be significantly smaller in diameter (2−3.5 nm).[17−23] It is uncertain how many cellodextrin chains are produced by a single cellulose synthase terminal complex,

Received: October 31, 2011
Published: January 7, 2012

with estimates from NMR in the range of 18−36 chains[24−28] in general agreement with AFM height measurements of single microfibrils in the range of 2−3.5 Å.[20,29,30] Estimates of microfibril widths calculated with the Scherrer equation are on the same order of magnitude,[21] but these estimates are affected by association with noncellulosic components in the cell wall[31] and can differ significantly depending on details in the analysis.[2,32] This size difference lends uncertainty to the direct extension of observations made on large model systems containing hundreds to thousands of cellodextrin chains to much smaller diameter cellulose microfibrils, where nanoscale structure may deviate from bulk structure due to the small crystallite size.[33,34] Examining small diameter cellulose has relied on AFM,[20,29,30] NMR,[24−28] and vibrational spectroscopy,[35−38] which are experimental methods where it is difficult to unambiguously distinguish conformational features of both interior and surface simultaneously. These experimental measurements have highlighted that for a wide range of cellulose I morphologies, primary alcohol conformations are often not in the TG conformation at surfaces,[26,39,40] and to some extent in the microfibril interior.[28,30,41] Small diameter cellulose I has been called cellulose IV_I,[42,43] but earlier investigations found cellulose IV_I is a distinct polymorph which can be prepared under specific conditions.[44−48] In some studies of cellulose IV_I, solid-state ^{13}C NMR chemical shifts at C6 suggest the GT[49] or GG[50] conformation is present, and Raman spectroscopy suggests that conformations similar to both cellulose I and cellulose II are present[46] (i.e., TG and GT conformations, but it is not clear whether this material was cellulose IV_I or IV_{II}). From the extensive literature on this subject, it is clear that characterization of smaller diameter cellulose microfibrils remains a significant experimental challenge.

Generally, simulation of cellulose at the atomistic scale can offer a complement to these types of experimental findings, as simulation can probe details of cellulose surface and microfibril interior structure and dynamics directly. Computational approaches can include quantum mechanical calculations[51−53] and simulation with atomistic[54−59] and coarse-grained models[60−63] at various resolutions. However, using simulation to probe the structure and dynamics of a system of interest is conditional upon several factors, including using a potential energy function able to accurately reproduce the properties of interest, obtaining sufficient sampling to explore relevant conformational space, and implementing appropriate simulation procedures to avoid the introduction of computational artifacts. While recent advances enable larger simulation sizes with ab initio methods,[51] classical potential energy functions (force fields) are more appropriate to reach the relevant temporal and spatial scales for many questions of interest. Moreover, classical force fields can be delineated into atomistic and coarse-grained resolutions, with the former treating each atom explicitly and the latter treating groups of atoms at some predefined resolution as a single particle.

Several recent carbohydrate force fields at the atomistic level (or united atom where aliphatic hydrogen atoms are not treated explicitly) are applicable to cellulose. These include the CHARMM C35 force field,[64,65] the GLYCAM06 force field,[66] and the Gromos 45a4 force field.[67] Other recent force fields exist which have been applied to cellulose, including OPLS[58,68,69] and PCFF,[59] but are not included in this comparison study. Also, a revision to the Gromos carbohydrate force field used here was recently published, Gromos $56A_{CARBO}$.[70]

Recent examples of molecular dynamics (MD) cellulose simulations using atomistic force fields include calculating the decrystallization free energy as a function of crystal morphology in cellulose polymorphs and in chitin,[12,71,72] examining room- and high-temperature behavior of cellulose crystals,[54,55,73−78] investigating the detailed interactions between cellulose and cellulase enzymes,[79−82] probing cellulose phase transformations,[83] examining cellodextrins or cellulose crystals in the presence of ionic liquids,[84−86] and interpreting experimental observations due to the presence of conformational fluctuations.[87]

In these previous MD simulations, cellulose microfibrils have been constructed and simulated in different fashions, which will affect the observed behavior of the systems. One of the most prevalent considerations in MD simulations of cellulose is the issue of finite and "infinite" microfibrils. Several previous studies have modeled finite length crystals with explicit chain ends, meaning that the cellulose chains are not covalently bonded across the periodic boundary of the simulation box.[54,55,58,77,88] In these simulations the microfibrils can adopt twist along the chain axis. Microfibril twist is the result of chirality amplification.[54] Twist is not present in the cellulose I crystal structures and is observed experimentally in certain materials but is absent in others.[29,89−94] An alternative simulation approach is to model a periodic infinite crystal, where chains are covalently linked across the periodic boundary such that there are no chain ends. These constrained boundary conditions fix the amount of twist along the microfibril axis.

Here, we apply three carbohydrate force fields, C35, GLYCAM06, and Gromos 45a4, to model finite microfibrils of cellulose Iβ. Previous simulation studies compared these and other force fields, but those studies were applied to much smaller systems, such as disaccharides, and to ring puckering of monomeric carbohydrates.[95−98] For cellulose, we use a fully hydrated, 36 chain microfibril as in our previous study,[54,55] with chains of 40 anhydroglucose units long (degree of polymerization 40). The MD simulations are conducted up to the near-microsecond time scale, representing some of the longest MD simulations on cellulose to date. We find behavior of cellulose with the Gromos 45a4 force field significantly diverges from the initial cellulose Iβ crystal structure and from behavior with the other two force fields. For GLYCAM06 and C35, we find that the microfibrils initially twist while retaining the TG conformation in the microfibril interior. After several tens to hundreds of nanoseconds, both microfibrils untwist in events concomitant with changes in primary alcohol conformations and changes in unit cell parameters, yielding structures similar to those found in simulations at high temperature (near 500 K), denoted I-HT.[55,73] The loss of twist is due to the formation of a three-dimensional hydrogen-bond pattern, which is not present in the cellulose Iβ crystal structure. These events are seemingly initiated by surface-mediated conformational changes, which produce both GT and GG conformers and an expanded unit cell. These computationally predicted I-HT structures using the C35 and GLYCAM06 force fields are consistent with experimental observations including anisotropic thermal expansion coefficients,[99] hydrogen/deuterium exchange along preferred lattice planes,[55,100] and the phase transformation of cellulose Iα to Iβ through the high-temperature intermediate I-HT.[101]

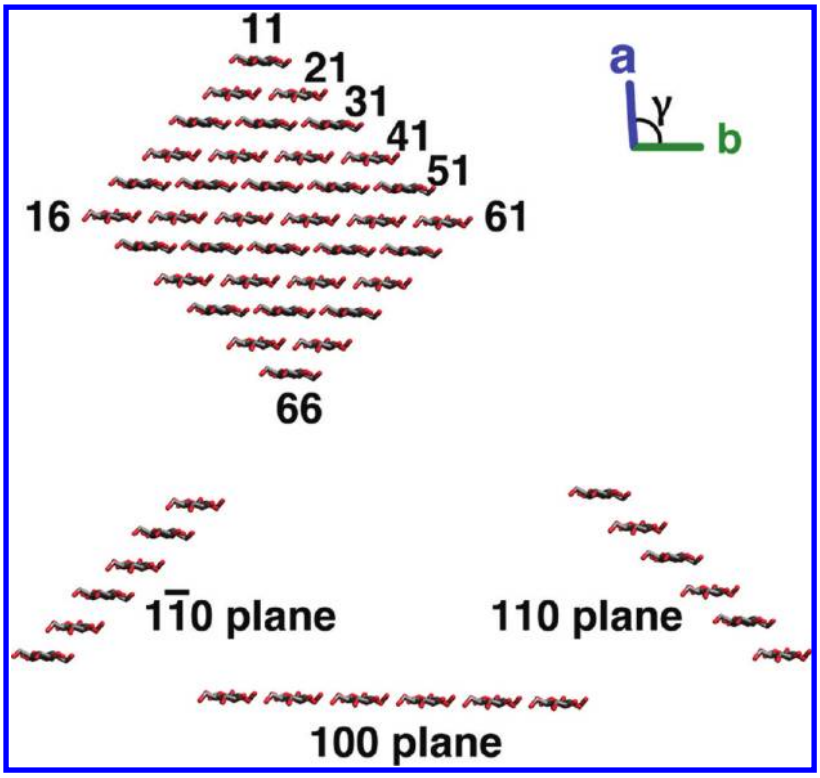

Figure 1. Cellulose Iβ microfibril cross-section and chain numbering scheme. This microfibril shape contains 6 chains on each 1$\bar{1}$0 and 110 face. We number the 36 chains by location (11−66), with the first digit referring to 1$\bar{1}$0 planes and the second digit referring to 110 planes. For each hydrogen-bonded layer in the initial structure (100 planes), the sum of the first and second chain digits is constant, so we number these layers from 2 to 12 (e.g., layer 7 includes chains 16 and 61, and layer 10 contains chains 46, 55, and 64).

For the GLYCAM06 and C35 force fields, we examine the sensitivity of these results to starting hydrogen-bond configurations, and for GLYCAM06, to the 1,4 electrostatic scaling factors. We demonstrate that for GLYCAM06, both the initial hydrogen-bond pattern and 1,4 electrostatic scaling factors increase the rate of structural divergence from the Iβ structure, while for C35 the rate of structural divergence is not significantly affected. We discuss differences between these force fields and how they may affect the structural differences found here. This study highlights the need for both additional experimental work to elucidate the structure and dynamics of small diameter cellulose microfibrils and additional theoretical work to more effectively model carbohydrate behavior at the molecular level.

METHODS

In this section, we describe the simulation protocol for the MD simulations and the measurement schemes for pertinent order parameters from the MD simulations. The system setup and simulation analysis closely follows our previous work examining the high-temperature behavior of cellulose Iβ with C35 and GLYCAM06.[55]

Cellulose Structure. The cellulose structure for all of the simulations was taken from the cellulose Iβ crystal structure.[9] The primary simulation data presented (800 ns simulations) are started in the **A** hydrogen-bond pattern (hbond-**A**). A subset of simulations started in the **B** pattern (hbond-**B**) is described below. The microfibrils in all simulations consist of a 36 chain, diamond-shaped crystal with chains having DP 40, as shown in Figure 1. Chains are numbered on a 6 × 6 grid from 11 to 66, with the first digit referring to 1$\bar{1}$0 planes and the second digit referring to 110 planes. With this indexing scheme, the chains in each 100 layers have constant sum of the first and second digit, so we name these 100 layers from 2 (chain 11) to 12 (chain 66). Although there are other cross-section shapes for a 36 chain microfibril, we chose to use the same shape as in our previous studies of cellulose.[54,55] For all simulations the box sizes are 60 × 60 Å by 233 Å and contain 15 Å of water solvation in each dimension.

C35 and GLYCAM06 Simulation Protocol. Simulations were conducted at 300 K with the C35[64,65] and GLYCAM06[66] (initial release) force fields similar to our previous simulations of cellulose at 500 K,[55] wherein we used the CHAMBER program[102] to convert CHARMM[103] input and parameter files for use in the AMBER PMEMD module.[104] We used the TIP3P water model,[105,106] the SHAKE algorithm to constrain covalent bond lengths to hydrogen,[107] and a 2 fs time step for the dynamics. The particle mesh Ewald (PME) method was used to treat the long-range electrostatics[108] with an approximate FFT grid spacing of 1 Å and Gaussian distribution width, κ, of 0.275. We applied a nonbonded cutoff distance of 10 Å. Temperature was controlled with Langevin dynamics with a collision frequency of 1 ps^{-1}.[109] In both the C35 and GLYCAM06 simulations, the systems were minimized for 100 steps with the steepest descent minimization scheme, heated to 300 K in 50 K increments over 300 ps, and then equilibrated in the NVT ensemble for 5 ns. Production simulations were then conducted in the NVT ensemble for a subsequent 800 ns.

To test the sensitivity of the simulation results to initial coordinates, additional simulations were run with part or all of the microfibril starting in the proposed hbond-**B** pattern for the GLYCAM06 and C35 force fields. We describe these simulations in the Results and Discussion Sections.

dx.doi.org/10.1021/ct2007692 | J. Chem. TheoryComput. 2012, 8, 735−748

Additionally, the GLYCAM06 force field uses a 1,4 electrostatics scaling factor of 1, while the AMBER protein force field scales these interactions by a factor of 5/6. While AMBER11 can properly accommodate the different scaling factors, simulations in previous versions of AMBER using a mixed system of carbohydrate and proteins required using the incorrect scaling factor for either the carbohydrate or the protein component. The recommended choice from the GLYCAM developers, as indicated in the parameter files, was to use the correct protein scaling factors, so an additional simulation (hbond-**A**) was run to determine whether using the inappropriate scaling factor for GLYCAM06 changes the simulated behavior of cellulose Iβ.

Gromos 45a4 Simulation Protocol. Simulations with the Gromos 45a4 carbohydrate force field and the compatible SPC water model[110,111] were performed with the program GROMACS 4.5.1.[112] A time step of 2 fs was used, and a temperature of 300 K was held constant using velocity rescaling.[113] Pressure was kept at 1 atm with a Parrinello–Rahman barostat,[114] with compressibility 4.5×10^{-5} and $t = 5.0$. While this ensemble differs from the ensemble used in the C35 and GLYCAM06 simulations, the structural changes we observe occur so rapidly they are not likely to be due to this difference. Bonds involving hydrogen within the cellulose molecules were constrained with the P-LINCS algorithm.[115] Water molecules were kept rigid during the simulations with SETTLE.[116] PME[117] was applied for the long-range electrostatic interactions outside a cutoff of 8 Å. A twin-range cutoff was applied for van der Waals interactions, with inner cutoff of 8 Å and outer of 14 Å.

Order Parameter Calculation. For the MD simulation analysis, we computed several order parameter time series relevant to the structure and dynamics of the cellulose microfibrils, including the primary alcohol conformation for each anhydroglucose monomer, the overall microfibril twist, and the unit cell parameters for every chain. The primary alcohol conformations are shown with the GT (green), TG (yellow), and GG (blue) notation as described in Figure 2. We use this color scheme in all subsequent figures. Twist was measured as the dihedral angle between four C1 atoms of the end anhydroglucose monomers in the middle layer (layer 7), in chains 25 and 52. Unit cell dimension time series was measured between C1 atoms of monomer 20 in each chain.

RESULTS

Long MD Simulation Results. We first describe the results for all 3 force fields started in the hbond-**A** pattern in 800 ns simulations. Figures 3 and 4 show snapshots along the 800 ns trajectories for each force field with monomers colored as shown in Figure 2. Figure 3 is a view down the cellulose microfibril chains (*c*-axis), whereas Figure 4 shows the individual 100 layers from each microfibril at the same time intervals as Figure 3 (also Figure S1, Supporting Information shows snapshots viewed across $1\bar{1}0$ layers), with the nonreducing chain ends at left and reducing ends at right. Figure 5 shows the populations of TG/GT/GG orientations in chains 42 and 51 of the GLYCAM06 hbond-**A** simulation, and Figure 6 shows this same information for every chain in the three 800 ns simulations as a function of time.

Several relevant observations can be made from these results for each force field. The C35 simulation shows an initial twist in the microfibril in the first several nanoseconds, but the microfibril then untwists around 400 ns (Figures 3 and 7).

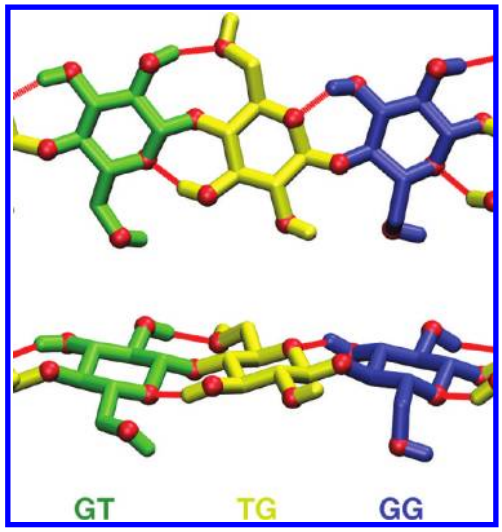

Figure 2. Color scheme to indicate primary alcohol group conformation used in Figures 3–7. Three anhydroglucose monomers in a single chain are shown from two directions, with GT colored green, TG yellow, and GG blue. The first letter corresponds to the relative O6–O5 orientation as trans or gauche, and the second corresponds to the relative O6–C4 orientation. Aliphatic hydrogen atoms are omitted, and both oxygen atoms and hydrogen bonds are indicated in red.

Many surface chains quickly convert to the GT and GG orientations on the microfibril surface within the first several nanoseconds (Figures 3–6). Around 50–100 ns, the C35 microfibril begins to exhibit alternating layers of the GG/GT conformations in center layers (odd-numbered layers). The origin chain layers (even-numbered layers) exhibit mainly the TG orientation in the interior and the GT orientation in solvent-exposed primary alcohol groups. At times beyond 500 ns, origin chain layers gradually convert to a mixture of alternating TG/GT conformations in the interior. Figure 6 shows in detail the primary alcohol conformations along every chain in the microfibril. In C35 origin layers, there is a gradual increase in the population of GT over the whole fibril, which occurs over tens to hundreds of nanoseconds. Primary alcohol conformations dictate the possible hydrogen-bond interactions between 110 and $1\bar{1}0$ planes, as observed in previous high-temperature simulations.[55,73] The gradual untwisting of the fibril is likely related to the shift in the primary alcohol conformations and the development of a regular interlayer hydrogen-bond pattern, which brings the 100 layers into alignment (layers are parallel in Figure 4) and eliminates twist. Figures S2 and S3, Supporting Information, show the unit cell lattice parameters for every chain, illustrating the significant variability in the cellulose Iβ lattice with the C35 force field as chains slide relative to each other by small amounts.

Both in GLYCAM06 hbond-**A** and in the C35 system, the microfibril exhibits an initial twist. In the C35 system, both GT and GG conformations are present in the interior prior to the untwisting event at 400 ns. In contrast, the GLYCAM06 hbond-**A** microfibril slowly untwists over 600 ns as GG conformations appear. In this system, GG conformations first appear in the interior at 200 ns in chains 22, 31, 42, 51, 26, 35, 53, and 62, in what may be a surface-mediated nucleation phenomenon[118–120] over the submicrosecond time scale. Overall, the GLYCAM06 hbond-**A** simulation exhibits relatively few fluctuations in primary alcohol conformations over 800 ns

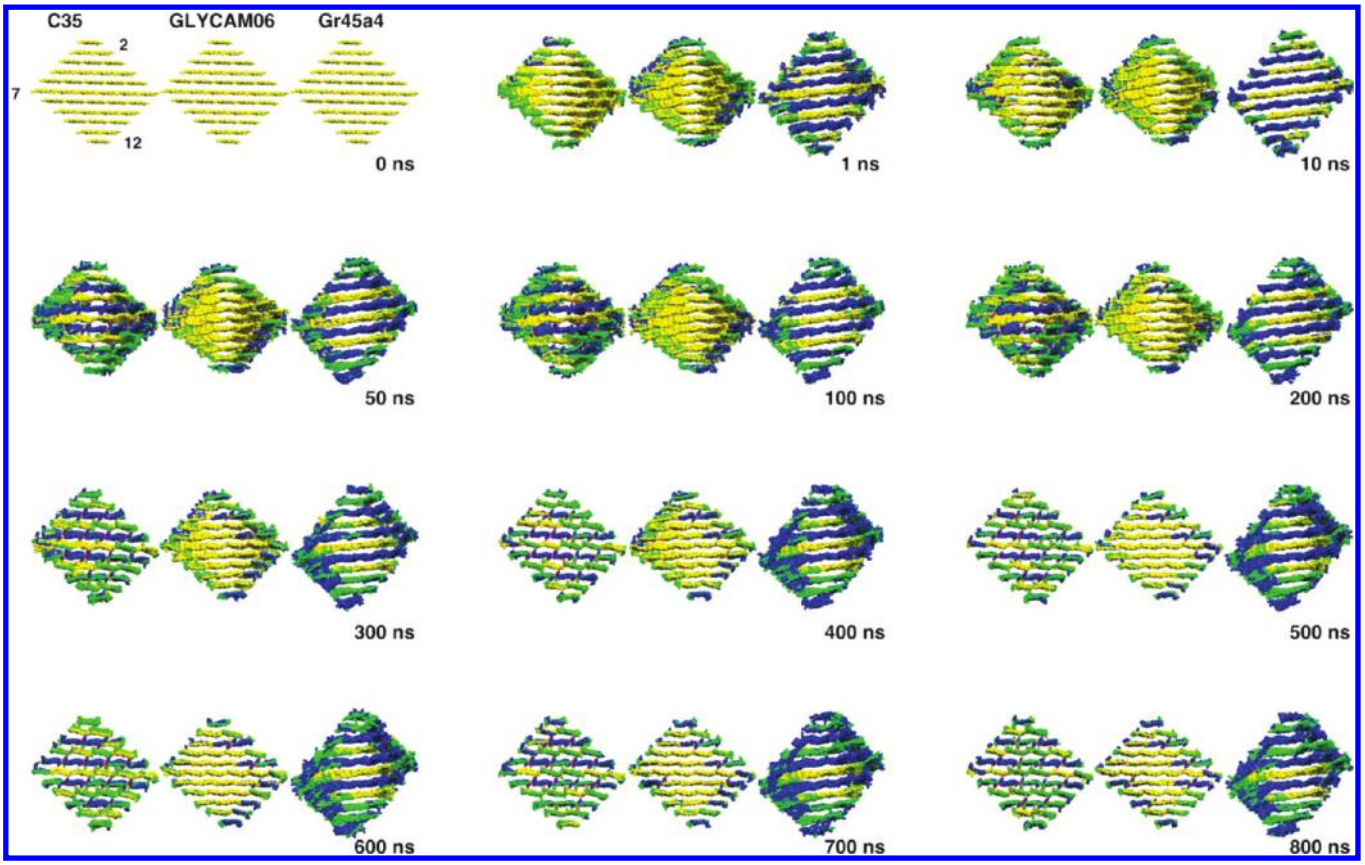

Figure 3. Selected snapshots in cross-section with C35 at left, GLYCAM06 middle, and Gromos 45a4 right. Monomers 11−30 (out of 40) of each chain are shown viewed from the nonreducing end. Monomer color indicates primary alcohol conformation as in Figure 2. Several 100 layers are labeled as described in Figure 1.

throughout most of the microfibril interior. However, the structure at 800 ns is not converged, and we speculate that the several nucleation events indicate that I-HT-like structure may develop if the simulation was run for significantly longer time. For this reason, we also ran simulations starting in the hbond-**B** pattern with all three force fields to help determine whether the apparent stability of Iβ-like structure with GLYCAM06 is due to kinetic trapping which might be overcome by perturbing the initial coordinates.

In the Gromos 45a4 simulation, Figure 3 shows the γ angle deviates significantly from the cellulose Iβ crystal structure value of 96.5° to near 70° almost immediately (see Figure S3, Supporting Information for additional details). Additionally, the microfibril adopts a twist, which does not equilibrate during the 800 ns MD simulation as shown in Figures 3 and 7. The primary alcohol conformations in center layers adopt the GG conformation within the first 10 ns, and two center chains (55 and 64) adopt the GT conformation at longer times. However, origin layers in the microfibril interior retain the TG conformation throughout the majority of the simulation. For solvent-exposed primary alcohol groups, the GT and GG primary alcohol conformations are adopted, similar to the C35 and GLYCAM06 simulations. However, most monomers in surface chains where the primary alcohol groups are facing the interior of the microfibril also adopt non-TG conformations.

Sensitivity to Initial Structure. To test whether apparent kinetic trapping of Iβ-like structure with GLYCAM06 might be overcome by perturbing the initial coordinates, we probed the microfibril behavior as a function of the initial hydrogen-bond pattern of cellulose Iβ for all three force fields examined here. We used slightly modified initial hbond-**B** pattern coordinates to avoid steric clashes between HO2 and HO6 in origin chains (Figure S4, Supporting Information). After 100 ns, simulations with the C35 and Gromos 45a4 force fields are not sensitive to the initial hydrogen-bond pattern, as the 800 ns hond-**A** simulations with these force fields already significantly diverged from Iβ-like structure (Figure S5, Supporting Information). However, the GLYCAM06 force field simulations are very sensitive to the initial hydrogen-bond pattern, as shown in Figure 8. Specifically, when starting in hbond-**A**, the initial hydroxymethyl conformations in the microfibril interior are stable for several hundred nanoseconds, and only a few nucleation events occur. However, when starting in hbond-**B**, nucleation events occur almost immediately throughout the microfibril, causing rapid convergence to the structures observed with GLYCAM06 at 500−550 K,[55,73] where center layers have the GG conformation, and origin layers either exhibit a mixture of TG and GT (500 K) or entirely GT (550 K). The 800 ns simulation started in hbond-**A** does not appear to be converged after 800 ns. We cannot definitively state what the final structure of the GLYCAM06 hbond-**A** microfibril will converge to, if given sufficient simulation time. It is possible that the simulations using the GLYCAM06 force field at 300 K will converge to the same I-HT-like final structure regardless of the initial hydrogen-bond pattern and that the apparently stable Iβ hbond-**A** structure is a kinetically trapped metastable state with this force field.

dx.doi.org/10.1021/ct2007692 | J. Chem. TheoryComput. 2012, 8, 735−748

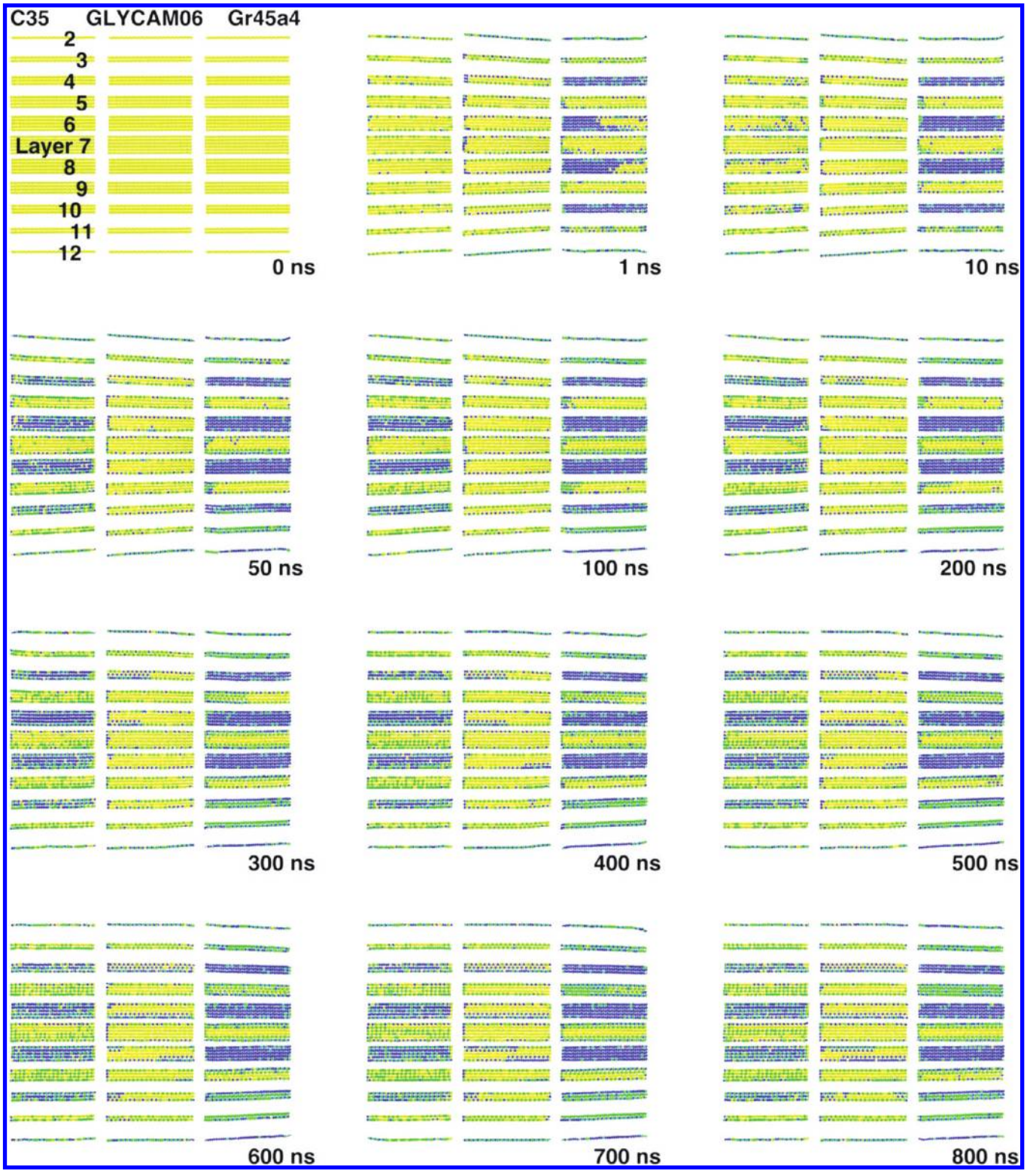

Figure 4. Selected snapshots of every 100 layer of each microfibril model with C35 at left, GLYCAM06 middle, and Gromos 45a4 right. Layer numbers correspond to those in Figure 3. All 40 monomers in each chain are shown, with the nonreducing ends at left and reducing ends at right. Monomer color indicates primary alcohol conformation. Deviation from parallel alignment (most obvious in layers 2 and 12) indicates the presence of twist.

We further investigated the sensitivity to the initial hydrogen-bond pattern in simulations with the GLYCAM06 force field by starting the hbond-**B** pattern either at the nonreducing end half (left) or at the reducing end half (right), as shown in Figure 9. Nucleation events occurred in the portion of the microfibrils started in the **B** pattern, and the I-HT structure forms at that end. This transformation also occurs when the entire microfibril is started in the **B** pattern, as shown in Figure 8 (GLYCAM06 hbond-**B**). Figure 9 suggests that the structural changes initiated by the **B** pattern migrate toward the end started in

dx.doi.org/10.1021/ct2007692 | *J. Chem. Theory Comput.* 2012, 8, 735−748

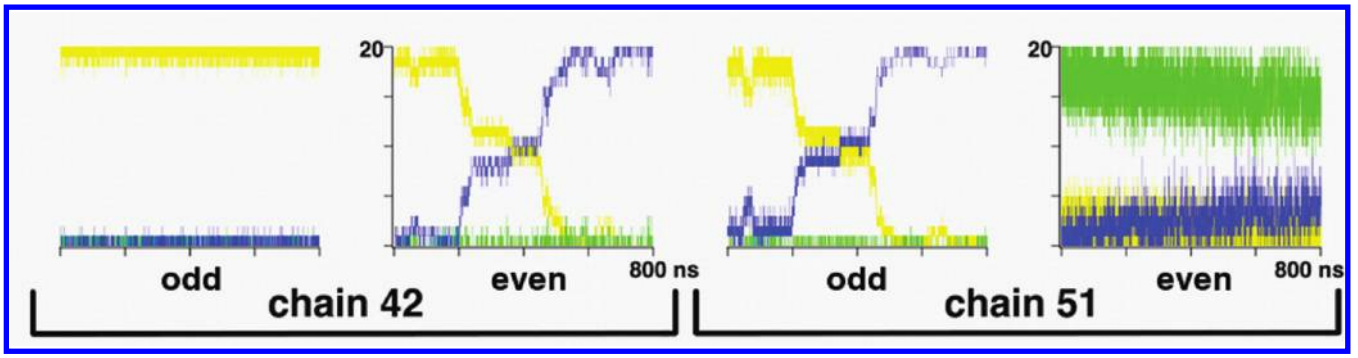

Figure 5. Primary alcohol populations over the 800 ns simulation in the DP 40 chains are shown with TG yellow, GT green, and GG blue. Shown as a guide to Figure 6, where primary alcohol populations in each of the 36 chains are represented with two time series (72 total), this figure shows enlarged time series with labels for two neighboring chains. Each chain has odd (the 20 with left-facing O6) monomers and even (the 20 with right-facing O6) monomers displayed as a separate time series. In the time series shown at left, chain 42_{odd} primary alcohol groups remain almost entirely in the initial TG conformation, with the chain end adopting mostly GG conformations. At right, chain 51_{even} primary alcohol groups are exposed to solvent so fluctuations are larger. Chains 42_{even} and 51_{odd} are part of the same hydrogen-bond network within this layer, and a nucleation event causes a cascade of primary alcohol conformation changes that alternate between these chain pairs as can be seen in Figure 4. These figures are from the GLYCAM06 hbond-**A** 800 ns simulation.

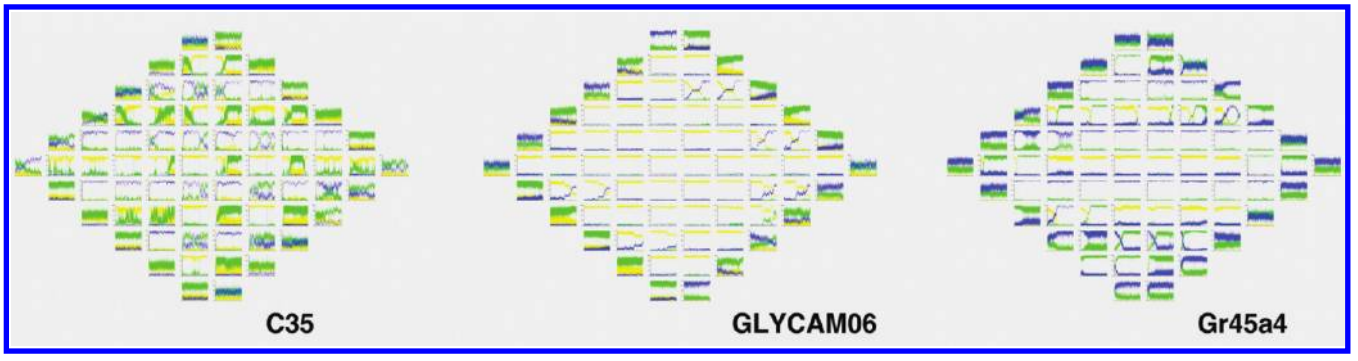

Figure 6. Distribution of primary alcohol conformers over 800 ns in simulations starting in the proposed hbond-**A** pattern of cellulose Iβ. Each chain is represented by two time series (odd and even numbered monomers) as in Figure 5. Changes in primary alcohol conformation are concurrent with changes in microfibril twist. Each time series has the same scales as those in Figure 5.

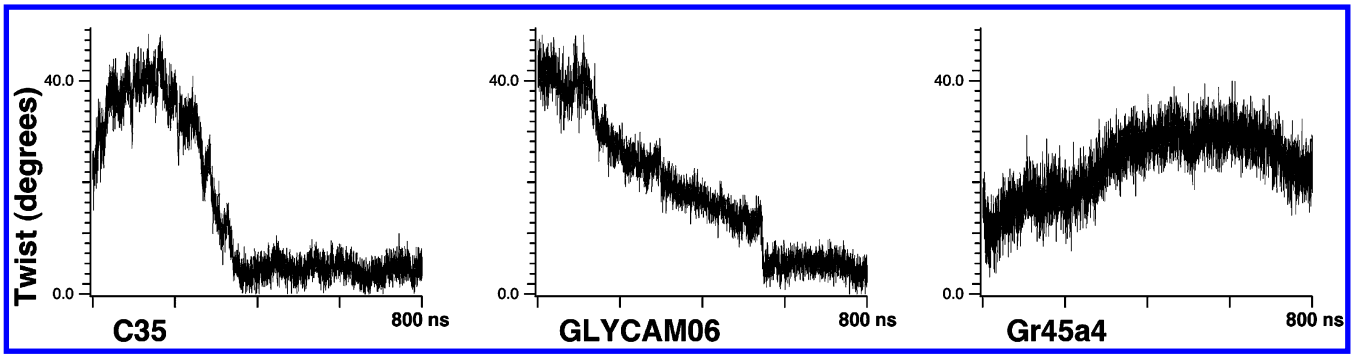

Figure 7. Microfibril twist over 800 ns. Changes in twist occur concurrently with changes in primary alcohol orientation.

the **A** pattern. That is, the I-HT structure propagates at the expense of Iβ, and we do not observe the reverse nucleation of Iβ at the expense of I-HT in these 100 ns simulations.

We also started with the **B** pattern only in origin layers or only in center layers, as shown in Figure 10. We observe that when center layers start in the **B** pattern, the I-HT structure nucleates rapidly in both origin and center layers. However, when origin layers start in the **B** pattern, the origin layers transition toward I-HT conformations, while center layers remain close to the initial Iβ conformations over the 100 ns MD simulation. This indicates that the initial events required to nucleate I-HT-like structure occur in center layers, which is consistent with a proposed mechanism of this transformation pathway.[101]

Sensitivity to Force Field Parameters. We also studied the sensitivity of the cellulose Iβ structure to the 1,4 electrostatic scaling factor for the GLYCAM06 force field. GLYCAM06 is parametrized for no scaling of 1,4 electrostatic interactions.[121] The AMBER protein force fields use a scaling factor of 5/6 for 1,4 electrostatic interactions. Prior to AMBER11, mixed carbohydrate−protein systems required that one of the force fields (either GLYCAM06 or the AMBER protein force field) be run with the incorrect 1,4 electrostatics scaling factor. Thus, we ran MD simulations to

dx.doi.org/10.1021/ct2007692 | J. Chem. TheoryComput. 2012, 8, 735−748

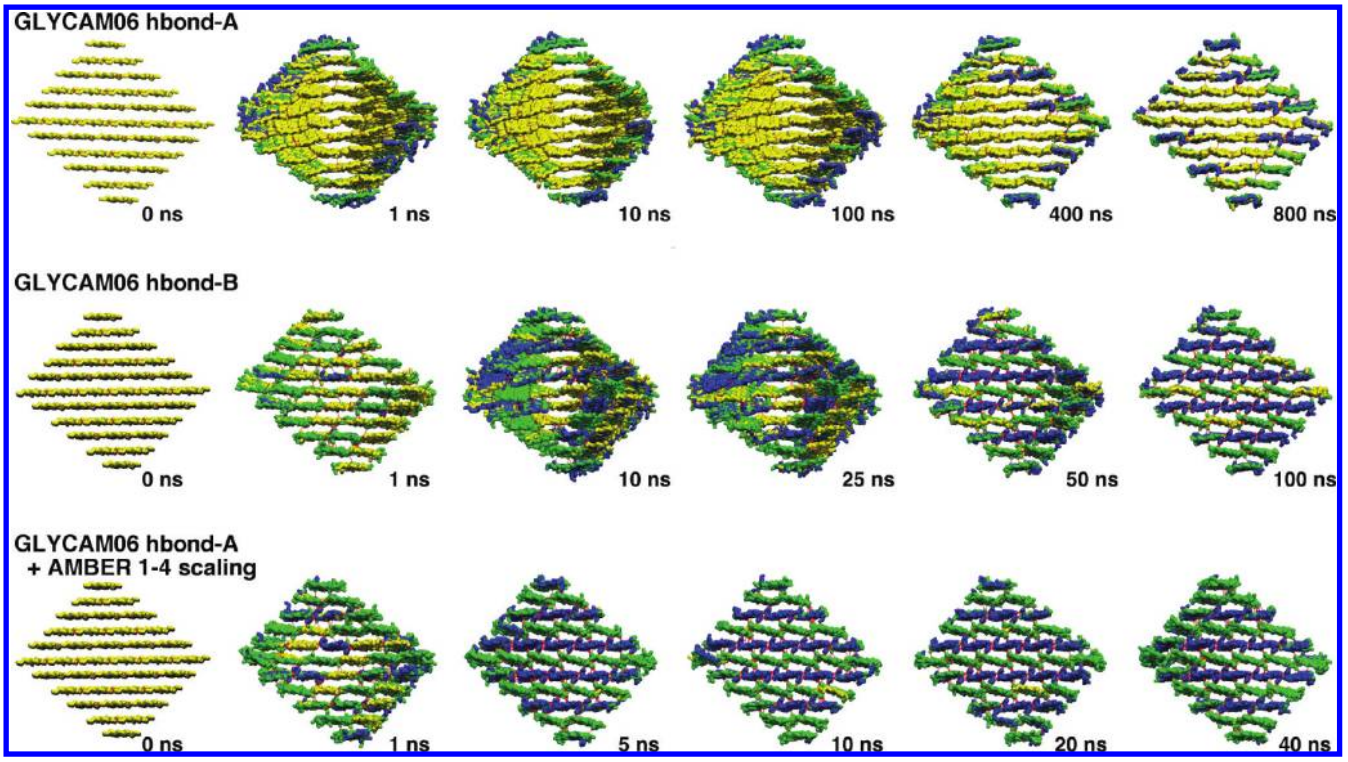

Figure 8. Simulations with the GLYCAM06 force field are sensitive to initial conformations and choice of electrostatic 1,4 scaling factor. As shown at top and in Figures 3−6, starting with the hbond-**A** pattern and using a scaling factor of 1 (appropriate for GLYCAM06) results in very few transitions from initial conformations in the interior of the microfibril for hundreds of nanoseconds. After several hundred nanoseconds at 300 K, nucleation events at surfaces and chain ends result in conformational changes similar to those observed in simulations with this force field at 500 K. Starting with the entire microfibril in the proposed hbond-**B** pattern (middle) results in rapid conformational changes away from the Iβ structure and toward the I-HT structure within a few nanoseconds. Using the AMBER protein electrostatic 1,4 scaling factor (bottom) and starting with the hbond-**A** pattern results in rapid conformational changes to the I-HT structure throughout the microfibril. Interlayer hydrogen bonds are shown in red.

examine the effect of the scaling factor. Figure 8 (bottom) shows the results from a 40 ns MD simulation starting from the **A** pattern with the scaling factor of 5/6 for 1,4 electrostatic interactions. The simulation quickly equilibrates to alternating layers of GG and GT with a three-dimensional hydrogen-bond pattern as in I-HT. Figure 8 shows these interlayer hydrogen bonds in red.

DISCUSSION

Here we have examined three carbohydrate force fields applied to cellulose Iβ at 300 K with a diameter on the same order as cellulose microfibrils found in the plant cell wall. From near-microsecond MD simulations with all three force fields, all three structures diverge from the crystal structure of cellulose Iβ solved using much larger diameter microfibrils from tunicates. Below we discuss the results and the potential reasons for the structural divergence related to the development of each force field and conclude with a discussion related to how our observations compare to experimental studies on plant cellulose.

Observations from C35 and GLYCAM06 Simulations. With both the C35 and GLYCAM06 force fields, we observe similar behavior to the results found at 500 K using the same simulation protocol.[55,73] The MD simulations started in the cellulose Iβ hbond-**A** crystal structure both develop a right-handed twist during the equilibration stage while retaining the TG conformation in the interior. During the subsequent 800 ns MD simulations with the C35 and GLYCAM06 force fields, both microfibrils untwist as conformations diverge from the cellulose Iβ structure. A three-dimensional hydrogen-bond pattern develops due to shifts in the primary alcohol conformations away from the TG conformation, which is consistent with the I-HT structure found in earlier simulations.

For the C35 force field, we observe little difference in the final microfibril structure as a function of the initial hydrogen-bond pattern. In contrast, with GLYCAM06 we observe significant variation in the time scale of divergence from the cellulose Iβ structure depending on initial conditions, such as starting either entirely or in part with the hbond-**B** pattern hydroxyl conformation or by using the AMBER 1,4 electrostatics scaling factor. Starting in the **A** pattern with GLYCAM06, the twist goes away as the GG conformation appears within parts of the interior, but even after 800 ns, most of the interior of the microfibril retains the TG conformation as in the initial cellulose Iβ structure.

With GLYCAM06 hbond-**B**, we observe a much faster convergence to the I-HT structure (~50 ns). The difference between the **A** and **B** pattern is the orientation of HO2 and HO6 hydrogen atoms, which results in fewer interchain hydrogen bonds overall in the **B** pattern. Based on this difference, the **B** pattern may allow a higher degree of conformational freedom and reduced stabilizing interchain interactions. Accordingly, the microfibril started in the **B** pattern will be able to converge more quickly to the I-HT structure. Additionally, hydroxyl rotations to convert between the **A** and **B** patterns may allow hydrogen-bond interactions with neighboring layers. These interlayer hydrogen bonds

dx.doi.org/10.1021/ct2007692 | J. Chem. TheoryComput. 2012, 8, 735−748

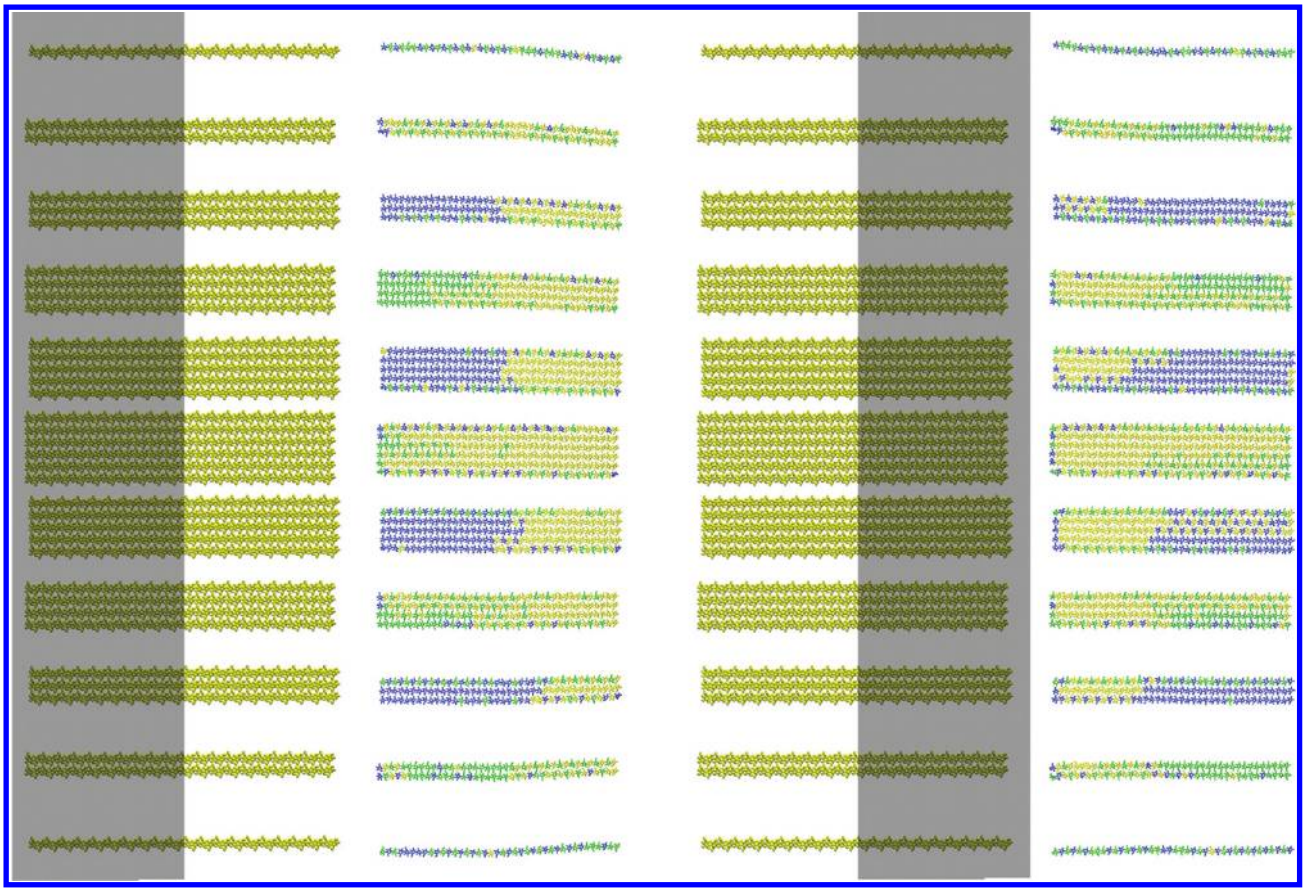

Figure 9. Initial configuration and 100 ns snapshot of GLYCAM06 microfibrils with the nonreducing (left) or reducing (right) end half started in the hbond-**B** pattern. Dark shaded regions indicate the **B** pattern in the initial structures. Views in cross-section for each of the 40 monomers in these microfibrils at 100 ns are shown in Figure S6, Supporting Information, where the correlation between twist and primary alcohol conformations in the microfibril interior is more evident.

(from O6 in center layers to O2 in origin layers) may initiate the transition to the I-HT structure.[101]

The GLYCAM06 simulation using the AMBER scaling factor also converges to the I-HT structure much more rapidly (<40 ns) than in the 800 ns simulation (which has not fully converged). We speculate that using the AMBER 1,4 electrostatic scaling factor of 5/6 reduces the barriers to rotation away from the TG conformation, which accounts for the rapid structural convergence in the simulation using this scaling factor. Changes to the electrostatic scaling factor for GLYCAM06 can affect both the thermodynamics and the dynamics, both of which have significant implications for cellulose structure studied with MD simulation.

While we do not dispute that cellulose I contains the TG conformation, it is common for nanometer-scale crystals to differ in structure from larger particles of the same material.[33,34,122,123] We speculate that in cellulose Iβ simulations with the GLYCAM06 force field, the apparent stability of the TG conformation may be due to two reasons: (1) overstabilization of the TG conformation for glucose in vacuum with this force field (incorrect potential of mean force (PMF) for hydroxymethyl rotation in vacuum)[124−129] and (2) more stable hydrogen bonds as compared to the C35 and Gromos 45a4 force fields, due to differences in partial charges on oxygen atoms (Table S1, Supporting Information). Conditional potential energy surfaces for hydroxymethyl rotation with GLYCAM06 do match the reference quantum surface when hydrogen bonding is disallowed.[130]

We suspect that longer MD simulations with GLYCAM06 hbond-**A** will eventually result in convergence to the I-HT structure. Nucleation and propagation of the I-HT structure with the GLYCAM06 force field may be slow due to this incorrect hydroxymethyl PMF for an isolated monomer. This hypothesis is directly supported by the results shown in Figure 9, wherein the portion of the structure started in the **B** hydrogen-bond pattern forms the I-HT structure, which then grows into the portion of the microfibril started in the Iβ-like structure. This result qualitatively implies that the I-HT structure is more thermodynamically stable than the Iβ structure at both 300 and 500 K, for the system studied here with the GLYCAM06 force field.

The *c* lattice parameter experimental value is 10.38 Å, but for GLYCAM06 this lattice parameter is 10.77 Å, for Gromos 45a4 10.57 Å, and for C35 10.42 Å. The extension in the *c* lattice parameter for the GLYCAM06 and Gromos 45a4 force fields will not likely significantly impact structural properties at the length scale of a monomer. However, we note that over extended regions of contact, this mismatch can lead to inaccurate interactions between, for example, enzymes and the cellulose surface. In simulations of infinite microfibrils, where the cellulose chains in the primary cell are covalently linked across the periodic boundary, enforcing the experimental *c* lattice parameter value leads to an effective compression of the cellulose chains. Deviations from *c* parameter repeat distances of 10.42 Å with C35, 10.77 Å with GLYCAM06, and 10.57 Å with Gromos 45a4 indicate the presence of an

dx.doi.org/10.1021/ct2007692 | J. Chem. TheoryComput. 2012, 8, 735−748

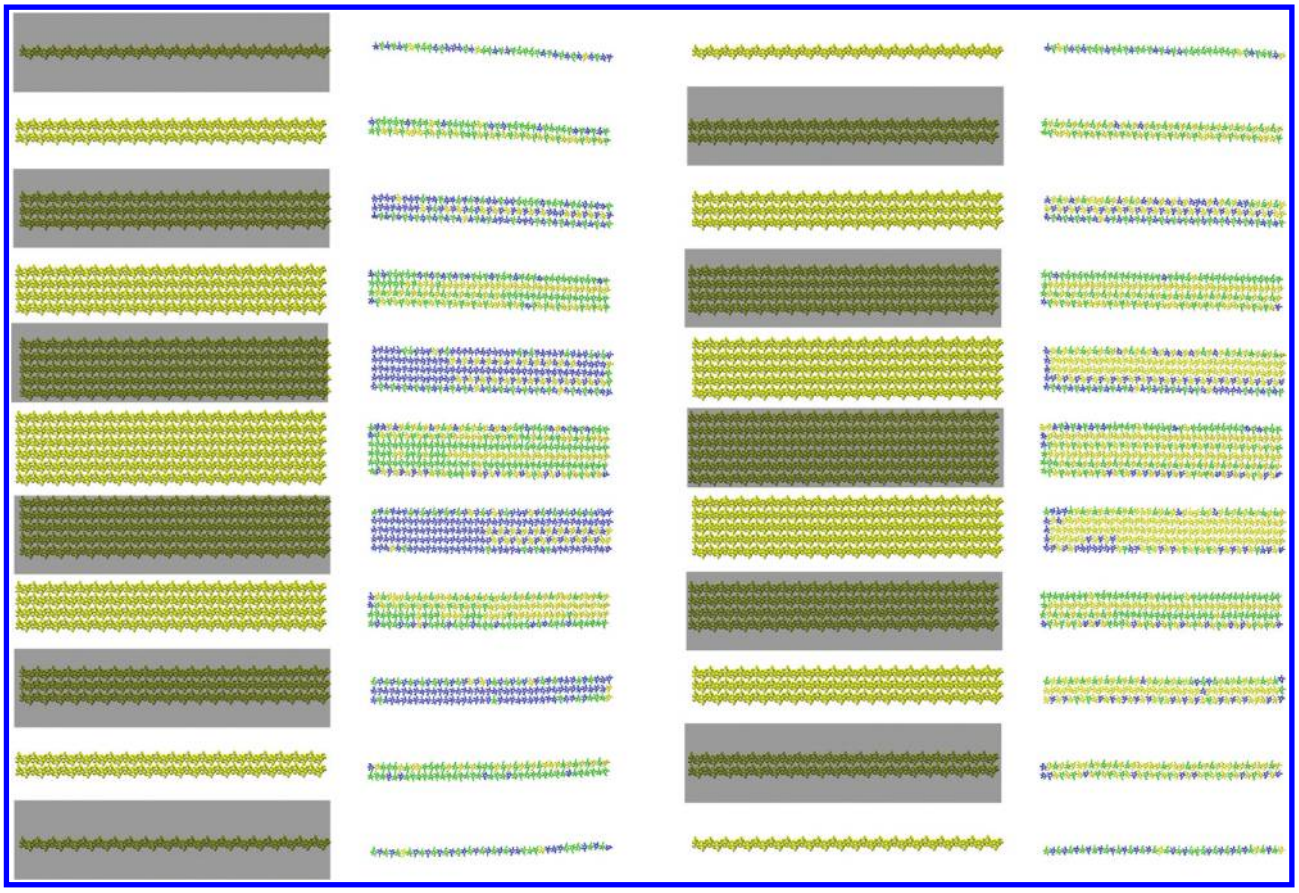

Figure 10. Initial configuration and 100 ns snapshot of GLYCAM06 microfibrils with the center (left) or origin (right) layers started in the hbond-**B** pattern. Dark shaded region indicates the **B** pattern in the initial structure.

artifact due to simulation procedures and care should be taken in periodic cellulose simulations to avoid this pitfall.

Observations from Gromos 45a4 Simulations. The Gromos 45a4 simulation exhibits significant deviations from the cellulose Iβ structure within the first nanosecond, most dramatically illustrated in the change of the γ angle from 96.5° to approximately 70°. Microfibril twist is present throughout the entire simulation. In origin layers, TG conformations are retained throughout the 800 ns MD simulation. The resulting Gromos 45a4 microfibril is significantly different from the structures obtained with GLYCAM06 and C35 and differs from previous simulations with this force field using constrained boundary conditions.

Implications for Computational and Experimental Studies of Cellulose. As mentioned previously, cellulose MD simulations have been used for a wide variety of computational studies. Reliable application of classical force fields for cellulose requires convergence in the properties of interest for a given question. For example, interactions of enzymes with cellulose surfaces do not likely require that the microfibril interior be completely converged but will certainly depend strongly on the surface structure. Conversely, direct comparison of MD simulation data to NMR and vibrational spectroscopy, which are ensemble experimental measurements, will require convergence of the entire microfibril. From our results with GLYCAM06 hbond-**A**, we note that the convergence time is likely on the time scale microseconds or longer for the interior of the microfibril. Surface chain conformations in the GLYCAM06 microfibril converge in tens of nanoseconds. Microfibril twist, which appeared during equilibration, goes away after hundreds of nanoseconds as the GG conformation in center layers appears. The results for C35 are similar to the GLYCAM06 results in that the surface chains equilibrate within the first few tens of nanoseconds, and the twist disappears within several hundred nanoseconds as regular interlayer hydrogen bonds appear. The interior of the C35 microfibril, as shown in Figure 6, is likely not converged because populations of primary alcohol conformers are changing and significant positional differences remain. In the Gromos 45a4 simulations, neither the twist nor the interior conformations of the microfibril converge within 800 ns. However, surface chain conformations converge within tens of nanoseconds, and there is a rapid change in unit cell parameters toward a structure unlike any determined from experiment to our knowledge.

From these comparison simulations, the time scale achieved is likely sufficient to converge interactions of interest which occur at surfaces, such as solvent−cellulose or catalyst−cellulose interactions. High confidence in the comparison of bulk cellulose properties to spectroscopic measurements will require better sampling, which has not yet been achieved. We note that the computational resources required for the simulations conducted here represent the current practical limit of computing capabilities for classical, all-atom MD simulations with typical supercomputers.

Additionally, we cannot definitively state which force field yields the "correct" behavior for particles of this size without further experimental data. It is not known from experiment if

particles of this width and length have the same structure as in larger, better-characterized cellulose microfibrils. Simulations with all three force fields diverge from the initial Iβ structure. Further experimental characterization of small diameter cellulose is required to obtain a better reference for comparison. Moreover, the number of chains and cross-sectional arrangement of small diameter microfibrils in plants is a significant detail yet to be elucidated.[3,12]

One method to estimate the width of cellulose microfibrils uses the ratio of NMR peaks in the C4 region of 90 to 80 ppm.[131] The chemical shift at C4 in glucose is sensitive to hydroxymethyl orientation and is also sensitive to the conformation of the glycosidic linkage in cellulose.[28] Chemical shift peaks at C4 and C6 have been decomposed into several components based on T_1 relaxation times, with faster relaxations tentatively assigned to solvent accessible surfaces and "amorphous" regions.[132] This correspondence is not completely understood. Peaks at 90 ppm and 88 ppm have been assigned to C4 in the interior of both cellulose I (TG) and of cellulose II and III_I (GT) crystals. Several C4 peaks near 84 ppm have been assigned to separate crystallographic surfaces,[133] but it is not clear whether this difference in chemical shift from those peaks assigned to the interior is due to differences in glycosidic linkage conformations or to differences in hydroxymethyl conformations and hydrogen bonding.[133,134]

This NMR approach to calculate microfibril width assumes surface and interior contribution in the C4 region can be decomposed, which may be susceptible to error if the structure of small diameter microfibrils differs from the structure of larger microfibrils. The interior of I-HT-like microfibrils, which have rapid hydroxyl reorientation, may give rise to components of the NMR spectra with fast T_1-relaxation times currently assigned to surfaces. For Iβ-like microfibril structure, monomers in surface chains with interior-facing primary alcohol groups can retain the TG conformation.

Uncertainty in the origin of these chemical shift components of the C4 region warrants caution to not overinterpret these data. If we assume peaks in the region near 84 ppm correspond to non-TG conformers, our simulation structures yield vastly different estimates for microfibril width. For idealized square cross-section microfibrils, the assumption that surface chains are not in the TG conformation may overestimate the width of Iβ-like microfibrils. For example, in the 36-chain GLYCAM06 hbond-**A** Iβ-like microfibril at 400 ns reported here, the 16 interior chains exhibit the TG conformation (44%). Of the 20 surface chains (66%), 18 of these chains contain approximately half TG conformations (monomers where C6 faces the interior), while chains number 11 and 66 have almost no TG conformers. Incorrectly assuming all TG conformations occur exclusively in the interior gives a ratio of 69% "interior" chains, which for a microfibril with square cross-section would be an estimated 7 Å wide and contain approximately 12 chains per side.[131] Conversely, in the I-HT-like C35 microfibril at 400 ns reported here, approximately 24% exhibit the TG conformation. Assuming 24% "interior" chains and a square cross section yields an estimated 2.5 Å microfibril containing 4 chains per side. More information about the distribution of surface and interior contributions to the NMR spectra of small diameter cellulose microfibrils is needed to address these uncertainties, such has been determined for bacterial cellulose ribbons.[27]

Finally, we note that single cellulose synthase terminal complexes may produce microfibrils with fewer than the 36 chains simulated here, but aggregation of neighboring microfibrils into microfibril bundles allows a variety of widths to exist within a single plant cell wall. Microfibril size and surface morphology may have significant effects on the convergence of observables obtained via MD simulation, but this is outside of the scope of the current study. Further experimental work is required to understand small diameter cellulose microfibril morphology and will be of significant value for more accurate cellulose simulations.

CONCLUSIONS

Here we have used three carbohydrate force fields in near-microsecond MD simulations of cellulose Iβ microfibrils: C35, GLYCAM06, and Gromos 45a4. Our results suggest the following:

- Simulations with all force fields examined diverge from the Iβ structure.
- The internal structure of all three microfibrils started in the hbond-**A** pattern do not converge in 800 ns.
- Solvent-exposed primary alcohol conformations rapidly equilibrate in simulations with all three force fields.
- With C35 for tens of nanoseconds and with GLYCAM06 for hundreds of nanoseconds, interior primary alcohol conformations retain the initial TG conformation as in the cellulose Iβ crystal structure.
- The C35 and GLYCAM06 structures resemble the I-HT structures found at 500 K in previous studies.[55,73]
- The C35 and GLYCAM06 microfibrils both initially twist during equilibration while retaining the Iβ-like TG conformation in the interior, and the microfibrils subsequently untwist after several hundred nanoseconds as a regular three-dimensional hydrogen-bond pattern forms, as is observed at 500 K with these force fields.[55]
- The Gromos 45a4 simulation results exhibit a dramatic shift in the γ angle from the initial experimental value of 96.5 to 70°, and this structure is not a good match to I-HT or to any other cellulose structure from experiment.
- Simulations with the GLYCAM06 and Gromos 45a4 force fields produce a *c*-lattice parameter (covalently bonded direction along the molecular chain axis) of 10.77 or 10.57 Å respectively, which are significantly larger than the experimental value of 10.38 Å.
- Cellulose simulations with the GLYCAM06 force field are sensitive to initial conformations and choice of 1,4 electrostatic scaling factor.
- The convergence time scales for properties of interest in cellulose MD simulations vary by orders of magnitude.

Further experiments are needed to ascertain the structure of small diameter cellulose microfibrils, which will enable more realistic studies of this important structural carbohydrate polymer. More generally, our study highlights how differences in carbohydrate force field parametrization can yield dramatically different computational observables in complex systems of interest to the computational chemistry community.

ASSOCIATED CONTENT

Supporting Information

Additional figures for primary alcohol conformations in $1\bar{1}0$ layers of each microfibril, the unit cell parameters and angles for each chain; partial atomic charges for the force fields studied here. This information is available free of charge via the Internet at http://pubs.acs.org

AUTHOR INFORMATION

Corresponding Author

*E-mail: james.matthews@nrel.gov; michael.crowley@nrel.gov.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by an award (DE-AC36-08GO28308) from the Office of Science's Office of Biological and Environmental Research, U.S. Department of Energy. J.F.M., M.E.H., and M.F.C. also were partially supported by the Center for Direct Catalytic Conversion of Biomass to Biofuels (C3Bio), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under award no. DE-SC0000997. M.B.-W. thanks Alf de Ruvo Memorial Foundation of SCA AB and Sweden-America Foundation for financial support. Computer time was provided by the TACC Ranger cluster under the National Science Foundation Teragrid grant no. MCB-090159. Some figures were made with VMD.[135] We thank Chris Pelkie for visualization expertise, Miriam Estin for editorial assistance, and Robert Woods and his group for helpful comments.

REFERENCES

(1) Carpita, N. C. *Plant Physiol.* **2011**, *155*, 171.
(2) Nishiyama, Y. *J. Wood Sci.* **2009**, *55*, 241.
(3) Chundawat, S. P. S; Beckham, G. T.; Himmel, M. E.; Dale, B. E. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 121.
(4) 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.
(5) Moon, R.; Martini, A.; Nairn, J.; Simonsen, J.; Youngblood, J. *Chem. Soc. Rev.* **2011**, *40*, 3941.
(6) Atalla, R. H.; Vanderhart, D. L. *Science* **1984**, *223*, 283.
(7) Vanderhart, D. L.; Atalla, R. H. *Macromolecules* **1984**, *17*, 1465.
(8) Sugiyama, J.; Persson, J.; Chanzy, H. *Macromolecules* **1991**, *24*, 2461.
(9) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074.
(10) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300.
(11) Gross, A. S.; Chu, J. *J. Phys. Chem. B* **2010**, *114*, 13333.
(12) Beckham, G. T.; Matthews, J. F.; Peters, B.; Bomble, Y. J.; Himmel, M. E.; Crowley, M. F. *J. Phys. Chem. B* **2011**, *115*, 4118.
(13) Langan, P.; Nishiyama, Y.; Chanzy, H. *Biomacromolecules* **2001**, *2*, 410.
(14) Wada, M.; Chanzy, H.; Nishiyama, Y.; Langan, P. *Macromolecules* **2004**, *37*, 8548.
(15) Langan, P.; Sukumar, N.; Nishiyama, Y.; Chanzy, H. *Cellulose* **2005**, *12*, 551.
(16) Nishiyama, Y.; Langan, P.; Wada, M.; Forsyth, V. T. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 1172.
(17) Kennedy, C. J.; Cameron, G. J.; Sturcova, A.; Apperley, D. C.; Altaner, C.; Wess, T. J.; Jarvis, M. C. *Cellulose* **2007**, *14*, 235.
(18) Ha, M. A.; Apperley, D. C.; Evans, B. W.; Huxham, M.; Jardine, W. G.; Vietor, R. J.; Reis, D.; Vian, B.; Jarvis, M. C. *Plant J.* **1998**, *16*, 183.
(19) Lahiji, R. R.; Xu, X.; Reifenberger, R.; Raman, A.; Rudie, A.; Moon, R. J. *Langmuir* **2010**, *26*, 4480.
(20) Niimura, H.; Yokoyama, T.; Kimura, S.; Matsumoto, Y.; Kuga, S. *Cellulose* **2010**, *17*, 13.
(21) Leppanen, K.; Andersson, S.; Torkkeli, M.; Knaapila, M.; Kotelnikova, N.; Serimaa, R. *Cellulose* **2009**, *16*, 999.
(22) Sturcova, A.; His, I.; Apperley, D. C.; Sugiyama, J.; Jarvis, M. C. *Biomacromolecules* **2004**, *5*, 1333.
(23) Heux, L.; Dinand, E.; Vignon, M. R. *Carbohydr. Polym.* **1999**, *40*, 115.
(24) Malm, E.; Bulone, V.; Wickholm, K.; Larsson, P. T.; Iversen, T. *Carbohydr. Res.* **2010**, *345*, 97.
(25) Newman, R.; Davidson, T. *Cellulose* **2004**, *11*, 23.
(26) Vietor, R. J.; Newman, R. H.; Ha, M. A.; Apperley, D. C.; Jarvis, M. C. *Plant J.* **2002**, *30*, 721.
(27) Yamamoto, H.; Horii, F.; Hirai, A. *Cellulose* **2006**, *13*, 327.
(28) Suzuki, S.; Horii, F.; Kurosu, H. *J. Mol. Struct.* **2009**, *921*, 219.
(29) Santa-Maria, M.; Jeoh, T. *Biomacromolecules* **2010**, *11*, 2000.
(30) Li, Q.; Renneckar, S. *Cellulose* **2009**, *16*, 1025.
(31) Driemeier, C.; Pimenta, M. T. B.; Rocha, G. J. M.; Oliveira, M. M.; Mello, D. B.; Maziero, P.; Gonçalves, A. R. *Cellulose* **2011**, *18*, 1509−1519.
(32) Fink, H. P.; Hofmann, D.; Philipp, B. *Cellulose* **1995**, *2*, 51.
(33) Jacobs, K.; Zaziski, D.; Scher, E. C.; Herhold, A. B.; Alivisatos, A. P. *Science* **2001**, *293*, 1803.
(34) Zheng, H.; Rivest, J. B.; Miller, T. A.; Sadtler, B.; Lindenberg, A.; Toney, M. F.; Wang, L.; Kisielowski, C.; Alivisatos, A. P. *Science* **2011**, *333*, 206.
(35) Marechal, Y.; Chanzy, H. *J. Mol. Struct.* **2000**, *523*, 183.
(36) Barsberg, S. *J. Phys. Chem. B* **2010**, *114*, 11703.
(37) Barsberg, S.; Sanadi, A. R.; Jorgensen, H. *Carbohydr. Polym.* **2011**, *85*, 457.
(38) Barnette, A. L.; Bradley, L. C.; Veres, B. D.; Schreiner, E. P.; Park, Y. B.; Park, J.; Park, S.; Kim, S. H. *Biomacromolecules* **2011**, *12*, 2434.
(39) Baker.; Helbert, W.; Sugiyama, J.; Miles. *Biophys. J.* **2000**, *79*, 1139.
(40) Hult, E. L.; Liitia, T.; Maunu, S. L.; Hortling, B.; Iversen, T. *Carbohydr. Polym.* **2002**, *49*, 231.
(41) Yamamoto, H.; Horii, F.; Hirai, A. *Cellulose* **2006**, *13*, 327.
(42) Wada, M.; Heux, L.; Sugiyama, J. *Biomacromolecules* **2004**, *5*, 1385.
(43) Newman, R. *Cellulose* **2008**, *15*, 769.
(44) Sueoka, A.; Hayashi, J.; Watanabe, S. *Nippon Kagaku Kaishi* **1973**, 1345.
(45) Atalla, R. H.; Whitmore, R. E. *J. Polym. Sci., Part C: Polym. Lett.* **1978**, *16*, 601.
(46) Atalla, R. H.; Ellis, J. D.; Schroeder, L. R. *J. Wood Chem. Technol.* **1984**, *4*, 465.
(47) Takahashi, M.; Ookubo, M. *Kobunshi Ronbunshu* **1994**, *51*, 107.
(48) Lai-Kee-Him, J. *J. Biol. Chem.* **2002**, *277*, 36931.
(49) Isogai, A.; Usuda, M.; Kato, T.; Uryu, T.; Atalla, R. H. *Macromolecules* **1989**, *22*, 3168.
(50) Dudley, R. L.; Fyfe, C. A.; Stephenson, P. J.; Deslandes, Y.; Hamer, G. K.; Marchessault, R. H. *J. Am. Chem. Soc.* **1983**, *105*, 2469.
(51) Li, Y.; Lin, M. L.; Davenport, J. W. *J. Phys. Chem. C* **2011**, *115*, 11533.
(52) Nishiyama, Y.; Johnson, G. P.; French, A. D.; Forsyth, V. T.; Langan, P. *Biomacromolecules* **2008**, *9*, 3133.
(53) French, A. D.; Johnson, G. P. *Can. J. Chem.* **2006**, *84*, 603.
(54) 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.
(55) Matthews, J. F.; Bergenstrahle, M.; Beckham, G. T.; Himmel, M. E.; Nimlos, M. R.; Brady, J. W.; Crowley, M. F. *J. Phys. Chem. B* **2011**, *115*, 2155.
(56) Beckham, G. T.; Bomble, Y. J.; Bayer, E. A.; Himmel, M. E.; Crowley, M. F. *Curr. Opin. Biotechnol.* **2011**, *22*, 231.
(57) Bellesia, G.; Asztalos, A.; Shen, T. Y.; Langan, P.; Redondo, A.; Gnanakaran, S. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 1184.
(58) Paavilainen, S.; Rog, T.; Vattulainen, I. *J. Phys. Chem. B* **2011**, *115*, 3747.
(59) Mazeau, K.; Heux, L. *J. Phys. Chem. B* **2003**, *107*, 2394.
(60) Bu, L. T.; 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.
(61) Hynninen, A. P.; Matthews, J. F.; Beckham, G. T.; Crowley, M. F.; Nimlos, M. R. *J. Chem. Theory Comput.* **2011**, *7*, 2137.

(62) Shen, T. Y.; Gnanakaran, S. *Biophys. J.* **2009**, *96*, 3032.
(63) Wohlert, J.; Berglund, L. A. *J. Chem. Theory Comput.* **2011**, *7*, 753.
(64) 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.
(65) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. *J. Chem. Theory Comput.* **2009**, *5*, 2353.
(66) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622.
(67) Lins, R.; Hunenberger, P. *J. Comput. Chem.* **2005**, *26*, 1400.
(68) Damm, W.; Frontera, A.; TiradoRives, J.; Jorgensen, W. L. *J. Comput. Chem.* **1997**, *18*, 1955.
(69) Kony, D.; Damm, W.; Stoll, S.; van Gunsteren, W. F. *J. Comput. Chem.* **2002**, *23*, 1416.
(70) Hansen, H.; Hunenberger, P. *J. Comput. Chem.* **2010**, *32*, 998.
(71) Payne, C. M.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. *J. Phys. Chem. Lett.* **2011**, *2*, 1546.
(72) Beckham, G. T.; Crowley, M. F. *J. Phys. Chem. B* **2011**, *115*, 4516.
(73) Zhang, Q. O.; Bulone, V.; Agren, H.; Tu, Y. Q. *Cellulose* **2011**, *18*, 207.
(74) Bergenstrahle, M.; Berglund, L. A.; Mazeau, K. *J. Phys. Chem. B* **2007**, *111*, 9138.
(75) Yui, T.; Hayashi, S. *Biomacromolecules* **2007**, *8*, 817.
(76) Yui, T.; Hayashi, S. *Cellulose* **2009**, *16*, 151.
(77) Yui, T.; Nishimura, S.; Akiba, S.; Hayashi, S. *Carbohydr. Res.* **2006**, *341*, 2521.
(78) Yui, T.; Okayama, N.; Hayashi, S. *Cellulose* **2010**, *17*, 679.
(79) 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.
(80) Zhong, L.; Matthews, J. F.; Crowley, M. F.; Rignall, T.; Talon, C.; Cleary, J. M.; Walker, R. C.; Chukkapalli, G.; McCabe, C.; Nimlos, M. R.; Brooks, C. L.; Himmel, M. E.; Brady, J. W. *Cellulose* **2008**, *15*, 261.
(81) 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.
(82) Lin, Y.; Silverstre-Ryan, J.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T.; Chu, J. W. *J. Am. Chem. Soc.* **2011**, *133* (41), 16617–16624.
(83) Bellesia, G.; Chundawat, S. P. S; Langan, P.; Dale, B. E.; Gnanakaran, S. *J. Phys. Chem. B* **2011**, *115*, 9782.
(84) Liu, H.; Sale, K. L.; Simmons, B. A.; Singh, S. *J. Phys. Chem. B* **2011**, *115*, 10251.
(85) Gross, A. S.; Bell, A. T.; Chu, J. *J. Phys. Chem. B* **2011**, *115*, 13433.
(86) Cho, H. M.; Gross, A. S.; Chu, J. *J. The Am. Chem. Soc.* **2011**, *133*, 14033.
(87) Wada, M.; Nishiyama, Y.; Bellesia, G.; Forsyth, T.; Gnanakaran, S.; Langan, P. *Cellulose* **2011**, *18*, 191.
(88) Nishiyama, Y.; Johnson, G.; French, A.; Forsyth, V. T.; Langan, P. *Biomacromolecules* **2008**, *9*, 3133.
(89) Xu, P.; Donaldson, L. A.; Gergely, Z. R.; Staehelin, L. A. *Wood Sci. Technol.* **2007**, *41*, 101.
(90) Brown, R.; Willison, J.; Richardson, C. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 4565.
(91) Hanley, S. J.; Revol, J. F.; Godbout, L.; Gray, D. G. *Cellulose* **1997**, *4*, 209.
(92) Elazzouzi-Hafraoui, S.; Nishiyama, Y.; Putaux, J.; Heux, L.; Dubreuil, F.; Rochas, C. *Biomacromolecules* **2008**, *9*, 57.
(93) Horikawa, Y.; Itoh, T.; Sugiyama, J. *Cellulose* **2006**, *13*, 309.
(94) Iwamoto, S.; Kai, W.; Isogai, A.; Iwata, T. *Biomacromolecules* **2009**, *10*, 2571.
(95) Perez, S.; Imberty, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinosa, J. F.; van Eyck, B. P.; Johnson, G.; French, A. D.; Louise, M.; Kouwijzer, C. E.; Grootenuis, P. D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K. *Carbohydr. Res.* **1998**, *314*, 141.
(96) Stortz, C. A.; Johnson, G. P.; French, A. D.; Csonka, G. I. *Carbohydr. Res.* **2009**, *344*, 2217.
(97) Spiwok, V.; Kralova, B.; Tvaroska, I. *Carbohydr. Res.* **2011**, *345*, 530.
(98) Krautler, V.; Muller, M.; Hunenberger, P. H. *Carbohydr. Res.* **2007**, *342*, 2097.
(99) Wada, M.; Hori, R.; Kim, U.-J.; Sasaki, S. *Polym. Degrad. Stab.* **2010**, *95*, 1330.
(100) Horikawa, Y.; Clair, B.; Sugiyama, J. *Cellulose* **2008**, *16*, 1.
(101) Matthews, J. F.; Himmel, M. E.; Crowley, M. F. *Cellulose* 2012, *12*, 297–306.
(102) Crowley, M. F.; Williamson, M. J.; Walker, R. C. *Int. J. Quantum Chem.* **2009**, *109*, 3767.
(103) 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.
(104) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.
(105) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926.
(106) Durell, S. R.; Brooks, B. R.; Ben-Naim, A. *J. Phys. Chem.* **1994**, *98*, 2198.
(107) Ryckaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, 23.
(108) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(109) Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D. *J. Chem. Phys.* **2001**, *114*, 2090.
(110) Lins, R. D.; Hunenberger, P. H. *J. Comput. Chem.* **2005**, *26*, 1400.
(111) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. *Intermolecular Forces*; Reidel: Dordrecht, The Netherlands, 1981.
(112) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435.
(113) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 7.
(114) Parrinello, M.; Rahman, A. *Phys. Rev. Lett.* **1980**, *45*, 1196.
(115) Hess, B. *J. Chem. Theory Comput.* **2008**, *4*, 116.
(116) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952.
(117) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8857.
(118) Keller, A.; Hikosaka, M.; Rastogi, S.; Toda, A.; Barham, P. J.; Goldbeckwood, G. *Philos. Trans. R. Soc., A* **1994**, *348*, 3.
(119) Beckham, G. T.; Peters, B.; Starbuck, C.; Variankaval, N.; Trout, B. L. *J. Am. Chem. Soc.* **2007**, *129*, 4714.
(120) Beckham, G. T.; Peters, B.; Trout, B. L. *J. Phys. Chem. B* **2008**, *112*, 7460.
(121) Kirschner, K.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622.
(122) Keller, A.; Hikosaka, M.; Rastogi, S.; Toda, A.; Barham, P.; Goldbeck-Wood, G. *Philos. Trans. R. Soc., A* **1994**, *348*, 3.
(123) Keller, A.; Hikosaka, M.; Rastogi, S.; Toda, A.; Barham, P. J.; Goldbeck-Wood, G. *J. Mater. Sci.* **1994**, *29*, 2579.
(124) Kirschner, K. N.; Woods, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10541.
(125) Thibaudeau, C.; Stenutz, R.; Hertz, B.; Klepach, T.; Zhao, S.; Wu, Q.; Carmichael, I.; Serianni, A. *J. Am. Chem. Soc.* **2004**, *126*, 15668.
(126) Spiwok, V.; Tvaroska, I. *Carbohydr. Res.* **2009**, *344*, 1575.
(127) Barnett, C.; Naidoo, K. *J. Phys. Chem. B* **2008**, *112*, 15450.

(128) Mason, P.; Neilson, G.; Enderby, J.; Saboungi, M.; Cuello, G.; Brady, J. *J. Chem. Phys.* **2006**, *125*, 224505.
(129) Woodcock, H. L.; Brooks, B. R.; Pastor, R. W. *J. Am. Chem. Soc.* **2008**, *130*, 6345.
(130) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.
(131) Heux, L.; Dinand, E.; Vignon, M. *Carbohydr. Polym.* **1999**, *40*, 115.
(132) Wickholm, K.; Larsson, P.; Iversen, T. *Carbohydr. Res.* **1998**, *312*, 123.
(133) Bergenstrahle, M.; Wohlert, J.; Larsson, P. T.; Mazeau, K.; Berglund, L. A. *J. Phys. Chem. B* **2008**, *112*, 2590.
(134) Sternberg, U.; Koch, F.; Prieü, W.; Witter, R. *Cellulose* **2003**, *10*, 189.
(135) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.