THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Degree of Polymerization of Glucan Chains Shapes the Structure Fluctuations and Melting Thermodynamics of a Cellulose Microfibril

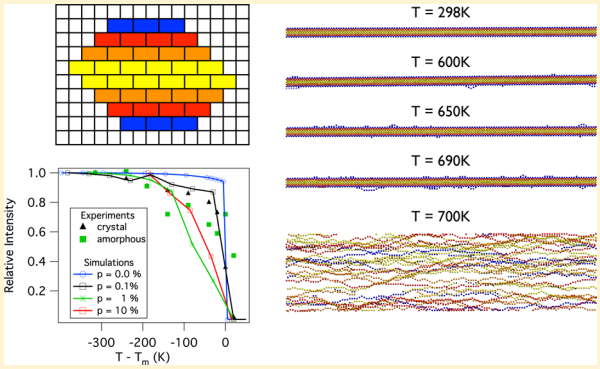
Rakwoo Chang,*[,†] Adam S. Gross,[‡] and Jhih-Wei Chu*[,‡]

[†]Department of Chemistry, Kwangwoon University, Seoul 139-701, Republic of Korea

[‡]Department of Chemical and Biomolecular Engineering and Energy Biosciences Institute, University of California, Berkeley, 94720, United States

Supporting Information

ABSTRACT: A Staggered LATtice (SLAT) model is developed for modeling cellulose microfibrils. The simple representation of molecular packing and interactions employed in SLAT allows simulations of structure fluctuations and phase transition of cellulose microfibrils at sufficiently long and large scales for comparison with experiments. Glucan chains in the microfibril are modeled as connected monomers, each corresponding to a cellobiose subunit, and the surrounding space around the cellulose is composed of solvent cells. Interaction parameters of monomer−monomer interactions were parametrized based on the results of atomistic molecular dynamics simulations. The monomer−solvent interaction was optimized to give a melting temperature of ∼695 K for the 36-glucan chain model cellulose microfibril, which is consistent with the estimation based on experimental data. Monte Carlo simulations of the SLAT model also capture experimentally measured X-ray diffraction patterns of cellulose as a function of temperature, including the region of melting transition, as well as predict the highly flexible regions in the microfibril. Beyond the diameter of ∼3 nm, we found that melting temperature of the cellulose microfibril is not significantly shifted by changing the thickness. On the other hand, a slight decrease in the degree of polymerization of glucan chains is shown to enhance structure fluctuations through the ends of glucan chains, i.e., the defect sites, and thereby significantly reduce the melting temperature. Analysis of the sizes, densities, and lifetimes of defect structures in the microfibril indicates a significant extent of fluctuations on the surfaces even at room temperature and that defect statistics are strong but distinct functions of temperature and solvent quality. The SLAT model is the first of its kind for simulating cellulosic materials, and this work shows that it can be used to incorporate information obtained from atomistic simulations and experimental data to enable the aforementioned findings through computation.



INTRODUCTION

As a result of photosynthetic carbon fixation, cellulose is the main form of organic carbon on earth.[1,2] In addition to its uses in the pulp and paper industries, the abundance of cellulose in inedible biomass also makes it an attractive feedstock for renewable production of fuels and biomaterials.[3−9] If economic production of transportation fuels from cellulosic biomass on commercial scales could be achieved, reliance on fossil fuels as an energy source can be drastically reduced.[10−14] To realize this objective, though, one significant challenge that must be overcome is the elimination of biomass recalcitrance through enhanced pretreatment options and engineering of more effective degradation enzymes.

Toward this end, a key mechanistic understanding that needs to be developed is the statistics of structure fluctuations in cellulose at different thermophysical conditions. This knowledge will inform the structural features that can be acted upon by solvent molecules or cellulose degrading enzymes (cellulases) to deconstruct the material. For example, the active sites of many cellulases in the catalytic domain (CD) have high similarity in shape and size and can only incorporate a single glucan chain. A variety of cellulose-binding modules (CBMs) connected to the CD, however, are found in different cellulases, revealing the need for affixing to diverse structural features on cellulose surfaces. CBMs have thus been applied as probes to characterize cellulose structures.[15] Even for cellulase CDs, strong evidence indicates that peripheral structural elements around the active site recognize specific structures on cellulose surfaces, and such coupling likely plays important roles for the CD to complex with glucan chains from a microfibril.[16] It has also been shown that the interactions between solvent molecules and glucose residues strongly depend on the structural state of a glucan chain in cellulose.[17] Knowledge of structure fluctuations in cellulose is thus essential

Received: March 28, 2012
Revised: June 19, 2012
Published: June 22, 2012

for molecular engineering toward overcoming biomass recalcitrance.

Microfibrils are the elementary structural form of cellulose in plant cell walls. A wide range of diameters (2−20 nm) and lengths (0.1−100 μm) of the material can be found in biomass.[18,19] The crystalline allomorphs of cellulose observed in nature (celluloses I_α and I_β) are collectively called cellulose *I*. In a cellulose microfibril, linear glucan chains of glucose monomers connected by β-1,4-glycosidic bonds arrange laterally in flat sheets, and the sheets stack axially to form a closely packed network that includes intra- and interpolymeric hydrogen bonds as well as intersheet contacts.[20,21] This crystalline interaction network renders cellulose a robust and stable material with a melting temperature as high as 500∼700 K at 1 atm.[22−24] Therefore, harsh conditions (high temperature and/or high pressure) are required to deconstruct cellulose in aqueous solutions and most organic solvents.[25] Special solvents such as certain types of ionic liquids that have molecular moieties compatible with cellulose structures[17] can achieve dissolution under milder conditions.[26−34]

At the atomistic scale, molecular dynamics (MD) simulations are indispensable in resolving the details of structure fluctuations in cellulose needed to complement experimental techniques such as NMR, FTIR, and scattering.[35−48] All-atom molecular mechanical force fields of carbohydrates have allowed MD simulations to capture key experimental observables such as the density, unit cell parameters, and distributions of side chain dihedral angles of cellulose.[41−43] A key finding is that the intramolecular hydrogen bonding between the glucan chains in cellulose depends on specific interactions with solvent molecules.[36−42] Distinct patterns of solvent arrangement around different surfaces of a microfibril have been observed. Furthermore, fluctuations of glucose residues in cellulose were found to be sensitive to their solvent exposure.[40] Since interchain OH−O hydrogen bonds on cellulose can be readily disrupted by hydrogen bonding with water, surface-exposed glucan chains are more flexible. Interactions along the axial directions of glucose residues (mainly in the form of pseudo CH−O HBs), though, are robust, and the strengths are insensitive to solvent exposure and microfibril twist.[38,40,45,46] Therefore, intersheet interactions were identified as the main contributor to the stability of cellulose microfibrils.[17,40] Simulations of cellulose deconstruction using atomistic MD with free-energy methods further illustrate the specific interactions between glucose residues and solvent molecules.[17,47,48] The mean forces encountered during deconstructing a glucan chain from a microfibril affirmed that intersheet contacts are indeed the relevant order parameter that dictates the bottleneck of overcoming cellulose recalcitrance.[17]

Based on structure fluctuations of cellulose observed in atomistic MD simulations, a great deal of understanding on the structures and interaction networks of cellulose has been developed.[35−48] The MD data can also be used to quantify the strengths of glucose−solvent interactions.[17] The accessible length scales and time scales using atomistic models, however, are severely limited. Coupling the MD engine with reaction path and free-energy simulation methods[49,50] can overcome this limitation to some extent, but the resulting calculations quickly become exceedingly demanding for computational resources. In extending the spatial and temporal scales of computer simulation of cellulose, several coarse-grained (CG) models have been developed.[51−55] Despite the differences in parametrization protocols and specific applications, the scope of previous CG modeling mainly focused on the local behavior of the cellulose interaction network.

An important concern of the structure fluctuations in cellulose is the polymeric nature of individual glucan chains. Considering thermal energy and conformational entropy, the polymeric character would inscribe universal structural and dynamical modes in cellulose fluctuations on both local and global scales.[56] For example, a freely jointed chain model is shown to fit very well with the distributions of the end-to-end distances of glucose oligomers observed in atomistic MD simulations.[37] Conventional polymer theories and phenomenological parameters representing the mean-field statistics of structure fluctuations can also be used to model the phase diagram of cellulose semiquantitatively.[57] A remaining challenge is elucidating how the interplay between the polymeric nature of glucan chains and the constriction imposed by the crystalline interaction network in a cellulose microfibril affects structure fluctuations.

In this work, we developed a SLAT (Staggered LATtice) model of structure fluctuations in cellulose to significantly simplify molecular complexity while retaining resolution and architecture at the level of cellulobiose repeat units to represent glucan chain packing in the crystalline microfibril. The computational efficiency of SLAT allows systematic characterization of the structure fluctuations of glucan chains during thermophysical processes such as melting. A key finding of SLAT simulations is that degree of polymerization (DOP) of glucan chains is an important parameter in shaping the structure fluctuations observed in fiber diffraction data of cellulose at different temperatures.

METHODS

In the SLAT model of a solvated cellulose microfibril, the three-dimensional space is partitioned into cubic cells of size σ, which is also the step size of the Monte Carlo (MC) moves. Each monomer in the SLAT model corresponds to an anhydrocellobiose repeat unit (two anhydroglucose molecules linked by a $\beta(1 \rightarrow 4)$ bond, $[OCHCHOCH_2OH(CHOH)_2CH]_2$), and has a size of 2σ in the *x* direction, 1σ in the *y* direction, and 2σ in the *z* direction (four cubic cells total). The *z* axis is taken as the polymerization direction of glucan chains. The monomer units cannot interpenetrate with each other because of the excluded volume interaction.

Covalent connectivity between bonded monomers in a glucan chain is imposed by the following conditions during simulation:

$$|x_{ij}| \le \sigma,\ |y_{ij}| \le \sigma, \quad \text{and} \quad l_{B0} \le |z_{ij}| \le l_{B0} + \sigma, \quad \text{for all} \quad i \quad \text{and} \quad j = i + 1 \tag{1}$$

In this equation, x_{ij}, y_{ij}, and z_{ij} are components of the bond vector between the centers of mass of monomers *i* and *j* in the same chain, and l_{B0} is the equilibrium length between cellobiose repeats, which is 2σ. The current SLAT model does not allow for spontaneous bond breaking.

The cellulose microfibril shown in Figure 1a has 36 glucan chains and is used as a model system.[4,58] The sheets in the microfibril are colored differently according to the levels of solvent exposure: blue (SI), red (SII), orange (SIII), and yellow (SIV). The outermost layers of the microfibril are SI, and the innermost layers are SIV. In representing intersheet contacts, glucan chains in different layers are stacked in a staggered

dx.doi.org/10.1021/jp302974x | J. Phys. Chem. B 2012, 116, 8074−8083

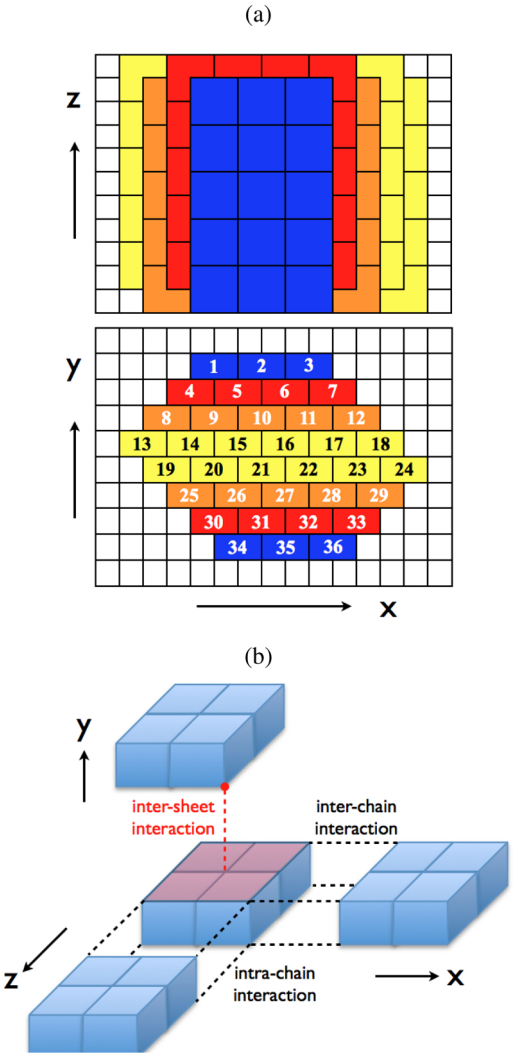


Figure 1. Lattice model of a crystalline cellulose microfibril developed in this study. (a) Four distinct layers are shown in blue, red, orange, and yellow, starting from the outermost layer. Chain index starts from the top left corner. The layers are stacked in a staggered way with an edge-to-face alignment of glucose residues in different layers. (b) Three types of monomer−monomer interactions: intrachain, interchain, and intersheet interactions. The face-to-edge contact used for the intersheet interaction was visualized in red color.

manner by aligning the cell edges of monomers in one sheet with the cell centers of monomers in neighboring sheets. Glucan chains in a sheet are thus shifted by σ in both the x and z directions relative to those in the neighboring layers. The crystal structures of cellulose I_α (or I_β) revealed via X-ray diffraction indicate that the distances between cellobiose repeats along a glucan chain, between neighboring chains in the same sheet, and between contacting sheets are about 1, 0.8, and 0.4 nm, respectively.[20,21] Therefore, the size of the cubic cell, σ, is set to 0.5 nm. Because of its simplicity, the SLAT model does not distinguish between the allomorphs of cellulose *I*.

Interactions between glucose residues in cellulose include intrachain, interchain, and intersheet components as shown in Figure 1b. A favorable intrachain interaction is added to the SLAT energy when two connected monomers along a glucan chain are in a face-to-face alignment ($-\varepsilon_{intra}$ per unit area σ^2). If nonbonded monomer pairs between glucan chains in the same sheet have a face-to-face contact, a favorable interchain interaction of $-\varepsilon_{inter}/\sigma^2$ is added. To represent intersheet coupling, instead of awarding face-to-face contacts a stabilizing interaction energy, a favorable term ($-\varepsilon_{sheet}$ per contact) is added when two monomers in different sheets are in contact in a face-to-edge alignment.

Solvent−glucose interactions are explicitly represented in the SLAT model. Monomer−solvent interactions are based on the contacts between solvent cells and glucose cells. Each contact with an area of σ^2 contributes $-\varepsilon_{ms}$ to the total energy. Following this convention, an isolated monomer in bulk solution would have a solvent interaction of $-16\varepsilon_{ms}$, whereas a monomer in the innermost layer of the cellulose microfibril would have an interaction energy of $-4\varepsilon_{intra} - 4\varepsilon_{inter} - 16\varepsilon_{sheet}$.

In addition, an angle potential (U_{angle}) is imposed on every three consecutive monomers to represent the bending stiffness of the polymer. The lateral and axial bending modes of a glucan chain are treated separately as

$$U_{angle} = U_{angle,xz} + U_{angle,yz} \quad (2)$$

$$U_{angle,\alpha z} = k_{angle,\alpha z}(\cos\theta_{\alpha z} + 1) \quad (\alpha = x, y) \quad (3)$$

where $k_{angle,\alpha z}$ ($\alpha = x, y$) is the force constant for each plane and $\theta_{\alpha z}$ is the angle between three successively connected monomers projected on the αz plane. It is noted that the above expression is similar to the conventional harmonic potential near $\theta_{\alpha z} \approx \pi$. The SLAT model of cellulose thus contains six parameters for computing the total energy of a solvated cellulose microfibril: ε_{intra}, ε_{inter}, ε_{sheet}, ε_{ms}, $k_{angle,xz}$, and $k_{angle,yz}$. Parametrization of the SLAT model is discussed later.

The simulation model constructed for the cellulose microfibril has a size of 32 nm × 32 nm × 128 nm (1 048 576 unit cells). For the 36-chain model shown in Figure 1a, the total number of monomers (N_{mon}) of cellulose is 4 608 (18 432 unit cells). Cells not taken up by cellulose are occupied by the solvent. The length of the microfibril spans the entire simulation cell along the z direction (128 nm), and the micrifibril is made effectively infinitely long through the periodic boundary condition. This setup corresponds to cellulose concentration of 5.3 g/L.

Using the SLAT model, we performed canonical ensemble MC simulations to sample the equilibrium configurations of the CG cellulose microfibril at different temperatures. Simple trial moves such as translation of monomers were employed and accepted according to the Metropolis algorithm.[59,60] If not specified otherwise, the initial configuration corresponds to the crystalline structure shown in Figure 1a. Equilibration MC simulations continue until both the system potential energy and the structural properties do not drift. After the equilibration period, the production MC run contains a total of 10^6∼10^7 steps per monomer with a snapshot saved every 100 × N_{mon} steps to generate 10^4∼10^5 configurations for analysis. The results reported in this work are averages over the production-run configurations with error bars estimated as 1 standard deviation of block averages (3−5 blocks).

A fundamental difficulty of using atomistic MD simulations to characterize complex molecular systems such as cellulose microfibrils is that longer-range modes require power-law growing time scales to sample, but atomistic-scale motions set the system-independent upper limit of the time step for stable dynamics propagation. In the SLAT model, the molecular details of glucose residues are discarded and only the architecture of glucan chains is retained. As shown later, this

dx.doi.org/10.1021/jp302974x | J. Phys. Chem. B 2012, 116, 8074−8083

reduction in complexity enables access of longer-range fluctuations in cellulose via computer simulation. However, model simplicity also reduces the extent of structural details that can be captured. For example, the on-lattice construction and broken rotational symmetry for mimicking the interaction network in cellulose prevent explicit representation of glucan chain tumbling with SLAT. Based on the principles of statistical mechanics, effects of missing details can be incorporated as they are averaged out through the model parameters in a mean-field manner.[61] In the SLAT model, thermodynamic behaviors of cellulose are captured by the aforementioned parameters: ε_{intra}, ε_{inter}, ε_{sheet}, ε_{ms}, $k_{angle,xz}$, and $k_{angle,yz}$. In parametrizing the SLAT model, our strategy was to use the results of atomistic MD simulations to estimate the interaction parameters between glucan chains and employ the experimental data of melting thermodynamics to tune the cellulose−solvent interactions.

The interchain interaction energies between two glucan chains on a cellulose microfibril have been computed to be around −2.0 kcal/mol-glucose.[40] Since each monomer has two interaction sites per contact surface for interchain interactions in the microfibril, ε_{inter} is set to 1.0 kcal/mol. Because intrachain interactions arise from the same molecular origin as interchain interactions, OH−O HBs, ε_{intra} was also set to 1.0 kcal/mol. Since the interaction energies between two adjacent sheets are calculated to be around −5.0 kcal/mol-glucose in atomistic MD simulations,[40] counting up and normalizing the intersheet interactions sets ε_{sheet} at a value of 0.5 kcal/mol in the SLAT model. Therefore, ε_{inter} is employed as the basic energy scale for glucose−glucose interactions in cellulose, and the 1:1:0.5 ratio observed in atomistic MD simulations is used for ε_{intra}, ε_{inter}, and ε_{sheet}. The angle force constant along a glucan chain, k_{angle}, is directly related to the persistence length, P. Following a wormlike chain (WLC) model,[62] k_{angle} is related to P as

$$k_{angle} = k_B T \frac{P}{l} \tag{4}$$

where k_B, T, and l are the Boltzmann constant, absolute temperature, and the bond length of polymer chains, respectively. The calculated persistence length from the atomistic MD simulations of glucose oligomers is about 10.6 nm.[37] With this result, thermal energy $k_BT \approx 0.6$ kcal/mol at room temperature, and $l \approx 1$ nm, $k_{angle,xz}$ and $k_{angle,yz}$ are set to 5 kcal/mol.

The remaining parameter ε_{ms} is the effective interaction energy for the contact interface between a solvent cell and a cellulose cell, which is a measure of the solvophilicity of glucan chains and controls the phase behavior of cellulose microfibrils.[17,63,64] Although the available data on thermophysical characterization of plant materials are limited, the phase behaviors of raw cellulose and that in woody biomass in dried and wet conditions have been studied systematically.[65,66] Upon heat treatment, the crystallinity of wood and raw celluloses shows a slight increase in the initial stage followed by an abrupt decrease.[65] This two-stage behavior is attributed to the crystallization of the quasicrystalline region in the first stage and thermally induced deconstruction in the second stage. Reduction in degree of polymerization of glucan chains in cellulose has also been observed alongside the reduction in crystallinity during the heat treatment.[66] That the temperature of chemical degradation overlaps with that of thermophysical melting is a reflection of the strong interaction network in cellulose.

The temperature of cellulose melting/decomposition determined by X-ray diffraction experiments varies from 200 to 400 °C, depending on the origin of the material, thickness of microfibrils, crystallinity, and the treatment conditions such as heat rate, humidity, and the form of the sample (powder or solid).[22,64,67] Therefore, in developing the SLAT model for cellulose, the targeted temperature of melting (deconstruction) is set to 500∼700 K. In this work, chemical degradation of glucan chains is not considered, although it is straightforward to include in the framework of the SLAT model. This study focuses on the physical deconstruction of cellulose and the structure fluctuations of glucan chains at different temperatures. The ε_{ms} parameter is thus scanned to capture the observed melting thermodynamics of cellulose.

The extent of crystallinity of the cellulose microfibrils in MC simulation is quantified by the root of mean squared deviation (rmsd) of cellulose monomers with respect to the positions in the reference crystalline structure shown in Figure 1a. The rmsd is calculated as

$$\text{rmsd} \equiv \sqrt{\left\langle \frac{1}{N_{mon}} \sum_{i=1}^{N_{mon}} [(x_i - x_{i0})^2 + (y_i - y_{i0})^2] \right\rangle} \tag{5}$$

where x_i and y_i are the components of the position vector of monomer i in the directions normal to the microfibril axis and x_{i0} and y_{i0} are the positions in the reference structure. Net translation of infinitely long glucan chains across the periodic boundary along the z direction is prevented via the bond restraints (eq 1).

Figure 2a shows the averaged rmsd of the MC simulations as a function of ε_{ms} at T = 298, 500, 600, and 700 K. Representative snapshots at several data points are also shown to highlight the nature of microfibril fluctuations. Each solid sphere in the snapshots corresponds to the center-of-mass coordinate of a monomer projected on the plane normal to the microfibril axis, and the coloring scheme shown in Figure 1a was used. It is clear that increasing the strength of monomer−solvent interaction, ε_{ms}, can lead to a sharp melting transition at a given temperature. At T = 298 K, ε_{ms} of 0.44 kcal/mol is required to initiate melting, while at T = 700 K, only a value of ε_{ms} = 0.34 kcal/mol is necessary. Thermophysical melting of the model cellulose microfibril represented by rmsd as a function of temperature is shown in Figure 2b with different ε_{ms} values. A clear signature of melting transition is seen with the sharp increase of rmsd with T. Therefore, ε_{ms} is set to 0.35 kcal/mol to give a melting temperature in the experimentally observed range of 500∼700 K. Simulation studies with three sizes of the simulation box (L_z = 64, 128, 256 nm) indicate that the melting temperature is not affected by the finite length effect (data not shown). The 0.35 kcal/mol value of ε_{ms} also limits the solubility of cellodextrins to oligomers with a length ≤6, the same limit as that found experimentally.[68,69] The results of perturbing SLAT parameters from those shown in Table 1 are shown in the Supporting Information. Although such changes do not alter the qualitative behaviors of the observed phenomena, the calculated values of the thermodynamic and mechanical properties of our cellulose model are sensitive to these parameters. Therefore, the set of model parameters can be identified unambiguously.

Another property of the SLAT model that can be compared with the result of atomistic MD simulation is the spatial dependence of the structure fluctuations of glucan chains in the microfibril. Figure 3 displays the averaged rmsd of each glucan

dx.doi.org/10.1021/jp302974x | J. Phys. Chem. B 2012, 116, 8074−8083

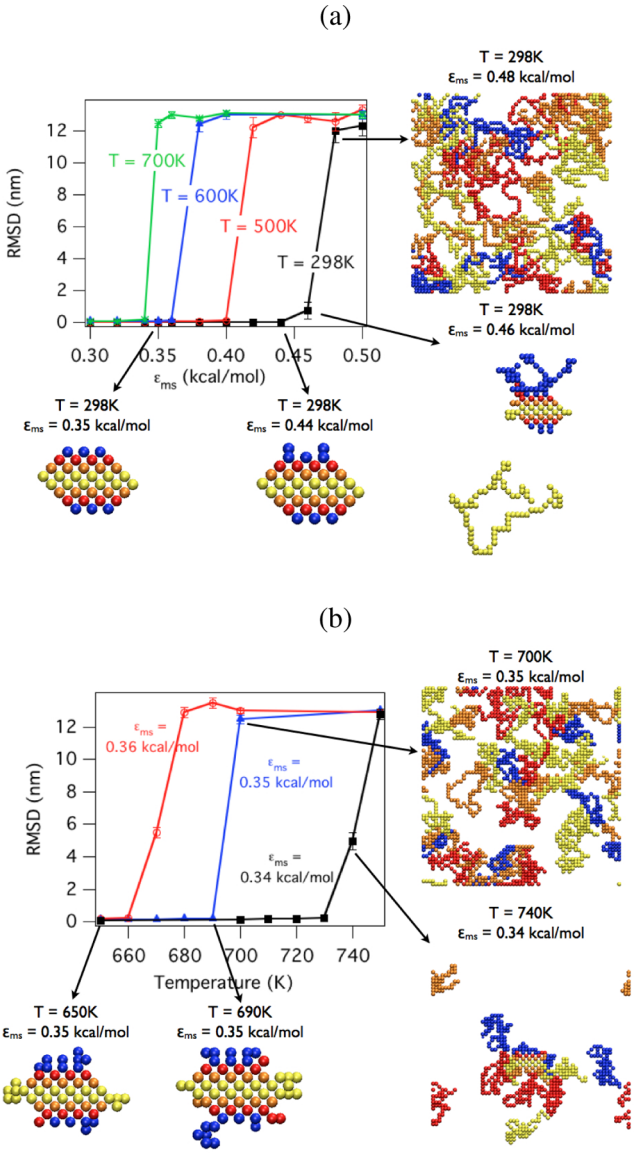


Figure 2. Melting transition of the model cellulose microfibril shown in Figure 1. (a) Averaged root-mean-square deviation (rmsd) of monomer positions from the initial crystalline structure as a function of the monomer−solvent interaction, ε_{ms}, at various temperatures and (b) rmsd as a function of temperature for ε_{ms} = 0.34, 0.35, and 0.36 kcal/mol. Representative snapshots at several data points are also shown to illustrate the structural state of the cellulose microfibril. Each solid sphere in the snapshots corresponds to the center-of-mass coordinate of a monomer projected on the plane normal to the microfibril axis, and the coloring scheme follows Figure 1a.

Table 1. Optimized Parameters of the SLAT Model for a Cellulose Microfibril

parameter	ε_{intra}	ε_{inter}	ε_{sheet}	ε_{ms}	$k_{angle,xz}$	$k_{angle,yz}$
energy (kcal/mol)	1.0	1.0	0.5	0.35	5.0	5.0

chain as a function of chain index at T = 298 K. The chain index is shown in Figure 1a. In comparison to an atomistic MD simulation of a cellulose microfibril with the same number and arrangement of glucan chains,[40] the SLAT model successfully identifies the most structurally fluctuating chains, which are the outermost chains in the middle layer (chains 13 and 24). The relative structure fluctuations in the second most flexible chains

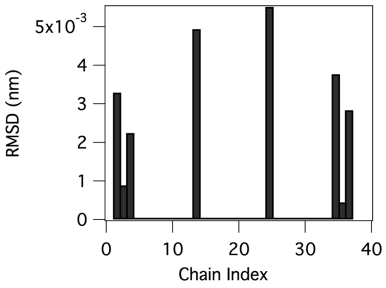


Figure 3. Rmsd of each cellulose chain as a function of chain index at T = 298 K. The chain index follows that in Figure 1a.

(the outermost layers) are also consistent with those in the atomistic MD simulation. Such ability in discerning the spatial distribution of chain flexibility by using a simple lattice model like SLAT is encouraging.

RESULTS AND DISCUSSION

Structural Properties of a Cellulose Microfibril at Different Temperatures. Figure 4a−e displays snapshots that

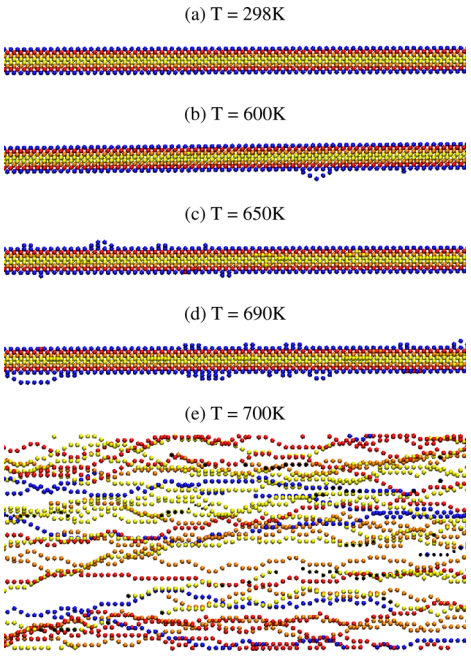


Figure 4. Representative snapshots of the cellulose microfibril viewed from along the x axis at T = (a) 298, (b) 600, (c) 650, (d) 690, and (e) 700 K.

show the shifting of glucan chain fluctuations from local to global length scales with increasing temperature; Figure 2 shows these snapshots along the axial view. At room temperature, local dislocations of monomers quickly anneal back to the most stable configuration and do not lead to global structural changes. Local deformation becomes more prominent and persistent as the temperature approaches the melting point of 695 K, but the microfibril structure still remains stable up to 690 K. Typically observed structure fluctuations in the intermediate temperature range include dislocated loops migrating along the microfibril axis, disruption and annealing of monomers at the ends of a loop, and coalescing of neighboring loops into larger ones. When the temperature is raised to the melting point and beyond, the deformation of

dx.doi.org/10.1021/jp302974x | J. Phys. Chem. B 2012, 116, 8074−8083

glucan chains on cellulose surfaces becomes long-ranged and leads to dissolution, finally resulting in complete disaggregation of the cellulose microfibril. The periodic boundary condition is imposed in the MC simulations shown in Figure 4. We have also examined the finite size effect of the model system by increasing the simulation box size twice in the z direction and found no significant change except the melting curve becomes sharper. This result suggests that the structural fluctuations observed at ambient temperature are mostly local.

It is generally considered that the decomposition temperature of cellulose microfibrils is affected by their diameter.[22,64,67] Therefore, we studied the effects of microfibril thickness on the thermal behavior of crystalline cellulose by comparing three microfibril systems with different numbers of cellulose chains N_p = 6, 18, and 36, which correspond to approximately 2, 3, and 5 nm in diameters, respectively. The diameter of each microfibril was estimated from the arrangement shown in Figure 1a. For example, the 36-chain system has 8 layers and 6 chains in the innermost layers, which correspond to 4 and 6 nm in the x and y directions, respectively. Hence, the diameter in the 36-chain system is approximated as 5 nm on average.

The microfibril system with N_p = 18 was constructed by removing 18 solvent-exposed glucan chains from the original 36-chain system. Similarly, the system with N_p = 6 was constructed by removing 12 solvent-exposed glucan chains from the 18-chain system. Figure 5 shows rmsd as a function of

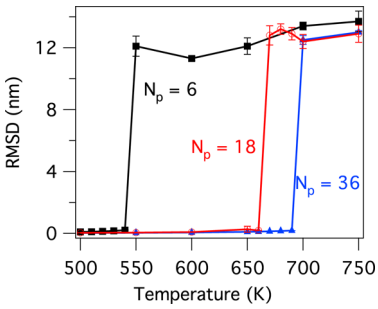


Figure 5. Effect of microfibril thickness on structure fluctuations of cellulose microfibrils. The averaged rmsd values as a function of temperature for microfibrils with different radii are plotted.

temperature for three cellulose microfibril systems (N_p = 6, 18, and 36). Although all systems investigated are stable at room temperature, the 6-chain system is decomposed completely at T = 550 K. However, the 18-chain microfibril system shows similar melting behavior to the 36-chain system, with its melting temperature slightly shifted to 670 K. The stability of the microfibril thus depends on the level of coordination between monomers. Glucan chains (chains 10 and 27) in top or bottom layers of the microfibril with N_p = 6 (a microfibril consisting of chains 10, 15, 16, 21, 22, and 27 in Figure 1a) have two fewer sites of interchain interactions and two more solvent interaction sites per monomer than the corresponding chains of the microfibril with N_p = 18 and 36. This counting corresponds to a $2\varepsilon_{inter} - 2\varepsilon_{ms}$ = 1.3 kcal/mol-monomer reduction in the energy increase for transferring surface-exposed glucan chains to the bulk and hence the lower melting temperature of the N_p = 6 microfibril.

Comparison of Simulated and Measured Patterns of X-ray Diffraction. X-ray diffraction is a commonly used experimental technique for probing the structures of complex materials and has been used to determine the decomposition temperature of crystalline cellulose. Since X-ray diffraction data was not used in parametrizing the SLAT model, comparing simulation results with the measured patterns will inform the accuracy and predictability of the SLAT model. For each cellulose configuration sampled in the MC simulation, we calculated the intensity, $I(R, Z)$, of fiber diffraction using the following equation[70]

$$I(R, Z) = \sum_n |G_n(R, Z)|^2 \tag{6}$$

and

$$G_n(R, Z) = \sum_j f_j J_n(2\pi R r_j) \exp i(-n\phi_j + 2\pi Z z_j) \tag{7}$$

where R and Z are the cylindrical coordinates in the reciprocal space, whereas r_j, ϕ_j, and z_j are the real space coordinates of monomer j in the simulation model and f_j is the scattering factor of the monomer, set to unity in the SLAT model. J_n is the Bessel function of the first kind of order n. In eq 7, cylindrical averaging along the fibril axis has been taken into account. The reciprocal coordinates R and Z are related to the Miller index (h k l) in the cubic lattice system

$$R = \sqrt{(h/a)^2 + (k/b)^2}, \quad Z = l/c \tag{8}$$

where a, b, and c are unit cell lengths in the real cubic lattice and are set to 2σ following the crystal lattice of the I_β allomorph.

Figure 6a shows relative diffraction intensities of Miller index (1 1 0) as a function of temperature for different simulation models and the experimental measurement of wood cellulose. The experimental values of the diffraction patterns of cottonwood, in which the cellulose is naturally highly orientated and crystalline,[22] are used to compare with the simulated values. The cottonwood material used in this experiment has been conditioned at 20 °C and stored at 65% relative humidity for over 1 year. The abscissa of Figure 6a is not absolute temperature but the temperature difference relative to the melting temperature, T_m, of the cellulose. This shift allows comparison of the diffraction patterns of cellulose models with different melting temperatures. To elucidate the effects of finite chain lengths in cellulose, i.e., finite degrees of polymerization (DOP), free ends were introduced to the SLAT simulation model by randomly disconnecting intrachain bonds. The p parameter is the density of the disconnected bonds, which is varied from 0 to 10% in this work, and is directly related to the degree of polymerization (DOP) of glucan chains. On average, DOP = $2/p$ for $p > 1/N_m$ (N_m = the number of monomers in a glucan chain). The factor of 2 in relating N_m to p comes from the fact that the monomer in the SLAT model corresponds to a cellobiose unit (two glucose units). With the current model of the cellulose microfibril with 36 glucan chains and N_m = 128 for each chain ($1/N_m$ = 0.78%), DOP is about 20 for p = 10% and 200 for p = 1%. The inequality criterion of $p > 1/N_m$ for DOP = $2/p$ is due to the fact that since the periodic boundary condition is applied, at least one bond disconnection is needed to convert a chain with an effectively infinite length into a finite-length chain. For example, in the case of p = 0.1% for the 36-chain microfibril system, there are about four or five disconnected bonds out of 4608 intrachain bonds in the system: i.e., 31 or 32 glucan chains effectively have an infinite length, and only 4 or 5 chains have a

dx.doi.org/10.1021/jp302974x | J. Phys. Chem. B 2012, 116, 8074−8083

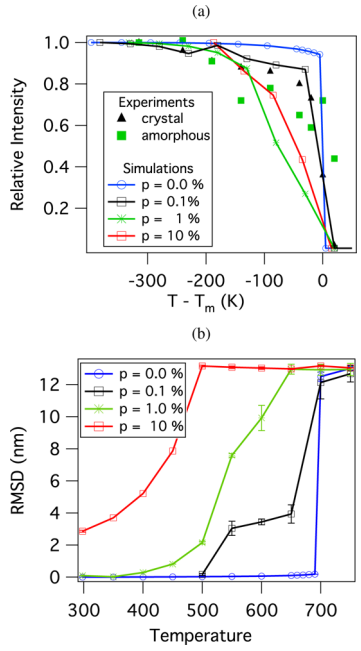


Figure 6. Comparison of SLAT simulations with the experimental X-ray diffraction results of cellulose. (a) Temperature dependence of relative diffraction intensities of the cellulose Miller index (1 1 0) plane from both experiments and simulations. Experimental data were taken from Kim et al.[22] The melting temperatures, T_m, for the simulation systems were 695, 680, 630, and 485 K for p = 0.0, 0.1, 1.0, and 10%, respectively, where p is the fraction of free ends of glucan chains in the system. (b) Effects of temperature on the averaged rmsd of the cellulose microfibril at different values of p.

DOP = 256 on average. In this work, the p value is used to label simulation models with different DOP.

The melting temperatures, T_m's, of different simulation models are 680, 630, and 485 K for p = 0.1%, 1.0%, and 10%, respectively. This range covers the experimentally measured T_m of wood cellulose (613 K).[22] With p = 0.0%, the SLAT model of the microfibril effectively has an infinite length with the periodic boundary condition. The higher value of T_m than that of wood cellulose with a finite value of DOP is thus expected. In addition to a higher melting point for the case of p = 0.0%, the melting transition of the SLAT model is also significantly steeper than that of the experiment and SLAT models with nonzero p values (see Figure 6a). As shown later, a key finding is that DOP is a crucial factor that provides irregularities in the crystalline structure of cellulose, shifts the melting temperature, and shapes the melting transition of the cellulose microfibril.

The width of the polymer melting transition regime scales as $1/(N_m)^{1/2}$.[56] A lower DOP in cellulose increases the density of free ends, and more pronounced free end fluctuations may affect the thermophysical properties of the material. Since DOP is an important parameter that distinguishes cellulose in different plants, how DOP modulates melting is a crucial concern of plant biophysics. It is clearly seen in Figure 6a that the transition regime becomes broader with increasing p and the microfibril with p = 0.1% shows an almost quantitative agreement with the measured X-ray diffraction pattern of wood cellulose. The presence of disconnected bonds thus has a significant effect on the stability of cellulose microfibrils. This behavior can be seen clearly in Figure 6b, which shows the averaged rmsd valuess from the crystalline structure of SLAT microfibrils with different p values. For p = 0.0%, the rmsd melting curve in Figure 6b changes sharply at the transition,

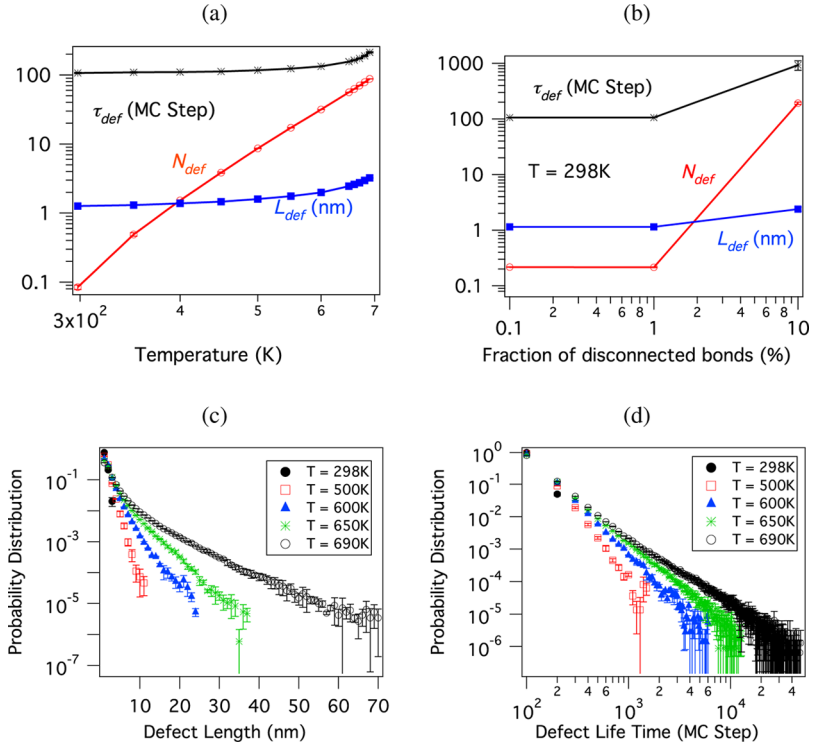


Figure 7. Statistics of defect structures in a cellulose microfibril. The averaged number (N_{def}), length (L_{def}), and lifetime (τ_{def}) of site defects as a function of (a) temperature and (b) p, the fraction of free ends of glucan chains in the system. Probability distributions of (c) defect length (L_{def}) and (d) defect lifetime (τ_{def}) at different temperatures.

dx.doi.org/10.1021/jp302974x | J. Phys. Chem. B 2012, 116, 8074−8083

just as in the diffraction curve shown in Figure 6a. The microfibril with slightly reduced DOP (p = 0.1%) shows an interesting double transition behavior. The initial transition occurs between 500 and 550 K, followed by a second transition between 650 and 700 K. Inspection of structural details indicates that the initial transition for the p = 0.1% model is due to the desorption of a disconnected chain end on the microfibril surface and the second transition is due to global melting. At a higher p value of 1%, the microfibril still retains a crystalline-like conformation at room temperature (see Figure 6b). However, for p = 1.0%, a lower DOP leads to significant structure fluctuations even at T = 500 K, and the melting transition regime is further broadened compared to that of the p = 0.1% model. The microfibril with the lowest DOP (p = 10%) has a high rmsd from the crystalline structure even at room temperature (Figure 6b). This behavior is mainly due to cellulose oligomers desorbed from the main structure; the overall microfibril remains intact at room temperature. The melting transition of the p = 10% model becomes less steep than that of the p = 1% case. Therefore, the results of SLAT simulations indicate that in addition to solvent quality and microfibril thickness, DOP is also an important factor affecting the thermodynamics of cellulose microfibrils.

Statistics of Site Defects. Simulation results using the SLAT model indicate that glucan chains in a cellulose microfibril exhibit various levels of structure fluctuations. Considering the crystalline form as the reference state, structure fluctuations of individual chains cause defects on the microfibril. To resolve structure fluctuations of glucan chains on cellulose, a defect site is defined as an empty cubic cell which was initially occupied by a crystalline cellulose monomer. In this definition, a defect occurs when any monomer moves away from its original site in the crystalline form and leaves the site unoccupied. From the snapshots sampled in MC simulations, the averaged number (N_{def}), length (L_{def}), and lifetime (τ_{def}) of defects are calculated and plotted as a function of temperature at p = 0 (Figure 7a) or as a function the fraction of disconnected bonds (p) at T = 298 K (Figure 7b). A logarithmic scale is used in Figure 7. The L_{def} value of a defect region measures the length along the glucan chain; the lateral width does not contribute to the value. In calculating the lifetime (τ_{def}) of a defect site, the time scale ($\tau_{MC} \equiv$ MC Step) is employed to approximately relate a MC step to physical time by the Einstein relation: $\tau_{MC} \approx (\Delta R^2)/(6D_{mon}) = (0.5\ nm^2)/(6 \times 10^{-6}\ cm^2/s) \approx 1$ ns, where D_{mon} is the translational diffusion coefficient of cellobiose in water.[71] Since MC moves of monomers give rise to the statistics of Brownian diffusion at long times, we utilize this approximate approach to provide a rough estimate of the time scale associated with the structural fluctuations of the cellulose microfibril.

A prominent feature in Figure 7a is that N_{def} increases exponentially with increasing temperature, whereas L_{def} and τ_{def} show a much weaker dependence; such trends are also observed as a function of solvent quality (data not shown). The introduction of free ends does not have significant effects on the defect dynamics up to p = 1% at room temperature, where defect sites are mostly localized, Figure 7b. When p is increased to 10%, the number of defect sites becomes almost 10^3 times that of the p = 1% case, indicating that the microfibril surface becomes highly susceptible to thermal fluctuation. Snapshots of p = 10% simulations indeed show the dissolution of small oligomers into bulk phase even at room temperature. Another behavior in Figure 7b distinct from the temperature dependence of defects (Figure 7a) is that L_{def} and τ_{def} have the same qualitative dependence on p as that of N_{def}.

The probability distribution of L_{def} for infinitely long chains (p = 0) shown in Figure 7c indicates that defect size is about 1−3 nm at ambient temperature, and most site defects are local, each involving only 1−3 monomers. The probability of site defects with length less than or equal to 3 nm is 98% at T = 298 K and reduces to 81% at T = 650 K as the probability of longer defects increases with temperature. The probability of $L_{def} \geq 10$ nm increases from 0% at T = 298 K to 2% at T = 650 K and 5% at T = 690 K. Similarly, the probability distribution of τ_{def} for infinitely long chains (p = 0) shown in Figure 7d indicates that the lifetime of site defects is weakly dependent on temperature. Therefore, most defects are short-lived and local, and temperature or solvent quality do not significantly alter this behavior. However, the frequency of occurrence of defects increases exponentially with temperature and solvent quality. Therefore, isolated defects with a longer segment of glucan chain are infrequent even at intermediate temperatures. Only when close to the melting point, which is about 695 K, do long-range defects become longer-lived (Figure 7d) and lead to global structural change.

It should be mentioned that the length and time of defects measured in the simulations only take into account the fluctuation along the z direction, and not in the x and y directions. At ambient temperature, this structural fluctuation occurs among 2−3 monomers in a very short time. Hence, the correlation between neighboring defects in the x and y directions would not be significant.

SUMMARY AND CONCLUSIONS

In this work, a primitive lattice model named Staggered LATtice (SLAT) has been developed for modeling structure fluctuations and the melting transition of crystalline cellulose microfibrils. The simple representation of molecular structures and interactions in SLAT gives the computational efficiency necessary for studying these behaviors in sufficiently long time and length scales to compare with experimental observables such as X-ray diffraction at different temperatures. Each glucan chain in a cellulose microfibril is modeled as connected monomers of a cuboidal shape surrounded by solvent cells. Parameters of monomer−monomer interactions and the angular force constants along a glucan chain were estimated based on the results of atomistic MD simulations of cellulose. The monomer−solvent interaction was optimized by calibrating the melting temperature of the 36-chain cellulose microfibril against the experimental value of ∼695 K. Canonical MC simulations with the SLAT model show good agreement with experiments in the temperature dependence of X-ray diffraction patterns near the melting transition region and also predict the highly flexible regions of the microfibril. Analysis of the statistics of defects in a cellulose microfibril indicates significant structure fluctuations due to thermal energy even at room temperature. Defects are shown to be present in the form of loops on microfibril surfaces and are typically local and appear intermittently at lower temperatures or in poorer solvents. Defect loops start to coalesce into larger and longer-lived ones with increasing temperature or improving solvent quality. At the melting point, global structural changes such as surface migration or complete dissolution eventually occur.

Furthermore, structure fluctuations and melting of cellulose are also shown to be affected by the thickness of the microfibril and the degree of polymerization of the glucan chains within.

dx.doi.org/10.1021/jp302974x | J. Phys. Chem. B 2012, 116, 8074−8083

SLAT simulations indicate that the melting temperature of the microfibril model is not reduced significantly by reducing thickness if the diameter is above ~3 nm. Conversely, flexibility of glucan chains and melting of the microfibril are very sensitive to degree of polymerization: increasing the density of free chain ends facilitates formation of more pronounced defects and drastically enhances structure fluctuations.

The SLAT model thus provides an effective route for investigating the long spatial and temporal behaviors of cellulose microfibrils that cannot be readily achieved by employing models with atomistic details. Caution, though, must be exercised when interpreting and utilizing the results of SLAT simulations due to its underlying simplicity. For example, lack of atomistic details does not allow the SLAT model to distinguish the greater details of glucose residues and solvent molecules at different locations in the microfibril, such as the differences between the allomorphs of cellulose I_α and I_β. Potential occurrences of chemical reactions, such as hydrolysis of glycosidic bonds, with the structure fluctuations are not taken into account in the current SLAT model either, although it would to be straightforward to include chemical reaction schemes for more realistic simulations under harsh conditions. Furthermore, the current SLAT model does not yet allow for thermal expansion or helical twisting of cellulose microfibrils, which are observed in both experiments and atomistic computer simulations although to very different extents. Incorporation of these behaviors into the model, as well as modeling enzymatic activities in SLAT for simulating cellulase–cellulose systems[72−74] with explicit representaiton of cellulose structures, is currently in progress.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Rmsd of monomers near a disconnected site, rmsd of single cellulose oligomers, and optimization of interaction parameters. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: rchang@kw.ac.kr (R.C.); jwchu@berkeley.edu (J.-W.C.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Energy Bioscience Institute (grant number OO0J04) and Korea Research Foundation (KRF) grant funded by the Korea government (MEST) (grant number 2010-0003087). R.C. also acknowledges Kwangwoon University for supporting his sabbatical year.

■ REFERENCES

(1) Bassham, J. A.; Benson, A. A.; Calvin, M. *J. Biol. Chem.* **1950**, *185*, 781−787.
(2) Delmer, D. P. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1999**, *50*, 245−276.
(3) Agrawal, R.; Singh, N. R. *Annu. Rev. Chem. Biomol.* **2010**, *1*, 343−364.
(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*, 804−807.
(5) Zinoviev, S.; Muller-Langer, F.; Das, P.; Bertero, N.; Fornasiero, P.; Kaltschmitt, M.; Centi, G.; Miertus, S. *ChemSusChem* **2010**, *3*, 1106−1133.
(6) Sheehan, J. J. *Curr. Opin. Biotechnol.* **2009**, *20*, 318−324.
(7) Alper, H.; Stephanopoulos, G. *Nat. Rev. Microbiol.* **2009**, *7*, 715−723.
(8) Stephanopoulos, G. *Science* **2007**, *315*, 801−804.
(9) Yuan, J. S.; Tiller, K. H.; Al-Ahmad, H.; Stewart, N. R.; Stewart, C. N. *Trends Plant Sci.* **2008**, *13*, 421−429.
(10) Chundawat, S. P. S.; Beckham, G. T.; Himmel, M. E.; Dale, B. E. *Annu. Rev. Chem. Biomol. Eng.* **2010**, *2*, 6.1−6.25.
(11) Carroll, A.; Somerville, C. *Annu. Rev. Plant Biol.* **2009**, *60*, 165−182.
(12) Igarashi, K.; Uchihashi, T.; Koivula, A.; Wada, M.; Kimura, S.; Okamoto, T.; Penttila, M.; Ando, T.; Samejima, M. *Science* **2011**, *333*, 1279−1282.
(13) Baker, E.; Keisler, J. M. *Energy* **2011**, *36*, 595−605.
(14) Regalbuto, J. R. *Science* **2009**, *325*, 822−824.
(15) McLean, B.; Boraston, A.; Brouwer, D.; Sanaie, N.; Fyfe, C.; Warren, R.; Kilburn, D.; Haynes, C. *J. Biol. Chem.* **2002**, *277*, 50245−50254.
(16) Lin, Y.; Silvestre-Ryan, J.; Himmel, M.; Crowley, M.; Beckham, G.; Chu, J.-W. *J. Am. Chem. Soc.* **2011**, *133*, 16617−16624.
(17) Cho, H.; Gross, A.; Chu, J.-W. *J. Am. Chem. Soc.* **2011**, *133*, 14033−14041.
(18) Perez, S.; Mazeau, K. Conformations, Structures, and Morphologies of Celluloses. In *Polysaccharides: Structural Diversity and Functional Versatility*; Dumitriu, S., Ed.; Marcel Dekker: New York, 2005; pp 41−68.
(19) Fernandes, A. N.; Thomas, L. H.; Altaner, C. M.; Callow, P.; Forsyth, V. T.; Apperley, D. C.; Kennedy, C. J.; Jarvis, M. C. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, E1195−E1203.
(20) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074−9082.
(21) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300−14306.
(22) Kim, D.-Y.; Nishiyama, Y.; Wada, M.; Kuga, S.; Okano, T. *Holzforschung* **2001**, *55*, 521−524.
(23) Watanabe, A.; Morita, S.; Ozaki, Y. *Biomacromolecules* **2006**, *7*, 3164−3170.
(24) Watanabe, A.; Morita, S.; Ozaki, Y. *Biomacromolecules* **2007**, *8*, 2969−2975.
(25) JL, W.; O, B.; Mericer, J. *Cellulose Science and Technology*; Lausanne EPFL Press: Lausanne, Switzerland, 2010.
(26) Swatloski, R. P.; Spear, S. K.; Holbrey, J. D.; Rogers, R. D. *J. Am. Chem. Soc.* **2002**, *124*, 4974−4975.
(27) Zhu, S.; Wu, Y.; Chen, Q.; Yu, Z.; Wang, C.; Jin, S.; Ding, Y.; Wu, G. *Green Chem.* **2006**, *8*, 325−327.
(28) Pinkert, A.; Marsh, K. N.; Pang, S.; Staiger, M. P. *Chem. Rev.* **2009**, *109*, 6712−6728.
(29) Zakrzewska, M. E.; Bogel-Łukasik, E.; Bogel-Łukasik, R. *Energy Fuels* **2010**, *24*, 737−745.
(30) Moulthrop, J. S.; Swatloski, R. P.; Moyna, G. *Chem. Commun.* **2005**, 1557−1559.
(31) Remsing, R. C.; Swatloski, R. P.; Rogers, R. D.; Moyna, G. *Chem. Commun.* **2006**, 1271−1273.
(32) Remsing, R. C.; Hernandez, G.; Swatloski, R. P.; Massefski, W. W.; Rogers, R. D.; Moyna, G. *J. Phys. Chem. B* **2008**, *112*, 11071−11078.
(33) Liu, H.; Sale, K. L.; Holmes, B. M.; Simmons, B. A.; Singh, S. *J. Phys. Chem. B* **2010**, *114*, 4293−4301.
(34) Lindman, B.; Karlstroem, G.; Stigsson, L. *J. Mol. Liq.* **2010**, *156*, 76−81.
(35) Matthews, J.; Skopec, C.; Mason, P.; Zuccato, P.; Torget, R.; Sugiyama, J.; Himmel, M.; Brady, J. *Carbohydr. Res.* **2006**, *341*, 138−152.
(36) Beckham, G. T.; Bomble, Y. J.; Bayer, E. A.; Himmel, M. E.; Crowley, M. F. *Curr. Opin. Biotechnol.* **2010**, 1−8.

(37) Shen, T.; Langan, P.; French, A. D.; Johnson, G. P.; Gnanakaran, S. *J. Am. Chem. Soc.* **2009**, *131*, 14786−14794.
(38) Bergenstrahle, M.; Wohlert, J.; Himmel, M. E.; Brady, J. W. *Carbohydr. Res.* **2010**, *345*, 2060−2066.
(39) Yui, T.; Shiiba, H.; Tsutsumi, Y.; Hayashi, S.; Miyata, T.; Hirata, F. *J. Phys. Chem. B* **2010**, *114*, 49−58.
(40) Gross, A. S.; Chu, J.-W. *J. Phys. Chem. B* **2010**, *114*, 13333−13341.
(41) Mazeau, K.; Heux, L. *J. Phys. Chem. B* **2003**, *107*, 2394−2403.
(42) Guvench, O.; Hatcher, E.; Venable, R.; Pastor, R.; MacKerell, A. *J. Chem. Theory Comput.* **2009**, *5*, 2353−2370.
(43) Matthews, J.; Beckham, G.; Bergenstrahle-Wohlert, M.; Brady, J.; Himmel, M.; Crowley, M. *J. Chem. Theory Comput.* **2012**, *8*, 735−748.
(44) Gross, A. S.; Bell, A. T.; Chu, J.-W. *J. Phys. Chem. B* **2011**, *115*, 13433−13440.
(45) Cousins, S.; Brown, R. *Polymer* **1995**, *36*, 3885−3888.
(46) Paavilainen, S.; Rog, T.; Vattulainen, I. *J. Phys. Chem. B* **2011**, *115*, 3747−3755.
(47) Beckham, G. T.; Matthews, J. F.; Peters, B.; Bomble, Y. J.; Himmel, M. E.; Crowley, M. F. *J. Phys. Chem. B* **2011**, *115*, 4118−4127.
(48) Payne, C. M.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. *J. Phys. Chem. Lett.* **2011**, *2*, 1546−1550.
(49) Brokaw, J. B.; Haas, K. R.; Chu, J.-W. *J. Chem. Theory Comput.* **2009**, *5*, 2050−2061.
(50) Haas, K.; Chu, J.-W. *J. Chem. Phys.* **2009**, *131*, 144105.
(51) Shen, T.; Gnanakaran, S. *Biophys. J.* **2009**, *96*, 3032−3040.
(52) Queyroy, S.; Neyertz, S.; Brown, D.; Muller-Plathe, F. *Macromolecules* **2004**, *37*, 7338−7350.
(53) 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.
(54) Wohlert, J.; Berglund, L. A. *J. Chem. Theory Comput.* **2011**, *7*, 753−760.
(55) Hynninen, A.-P.; Matthews, J. F.; Beckham, G. T.; Crowley, M. F.; Nimlos, M. R. *J. Chem. Theory Comput.* **2011**, *7*, 2137−2150.
(56) de Gennes, P.-G. *Scaling Concepts in Polymer Physics*, 1st ed.; Cornell University Press: Ithaca, NY, 1979.
(57) Frey, M.; Theil, M. *Cellulose* **2004**, *11*, 53−63.
(58) Ding, S.; Himmel, M. *J. Agric. Food Chem.* **2006**, *54*, 597−606.
(59) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*, 2nd ed.; Oxford Science Publications; Clarendon Press: Oxford, U.K., 2006.
(60) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Prentice Hall: New York, 2001.
(61) Cho, H. M.; Chu, J.-W. *J. Chem. Phys.* **2009**, *131*, 134107.
(62) Rubinstein, M.; H Colby, R. *Polymer Physics*, 1st ed.; Oxford University Press: New York, 2003.
(63) Esteves, B. M.; Pereira, H. M. *BioRes.* **2009**, *4*, 341−404.
(64) Taniguchi, T.; Nakato, K. *Bull. Kyoto Univ. For.* **1966**, *38*, 192−1966.
(65) Bhuiyan, M. T. R.; Hirai, N.; Sobue, N. *J. Wood Sci.* **2000**, *46*, 431−436.
(66) Fengel, D.; Wegener, G. *Wood Chemistry, Ultrastructure, Reactions*; Remagen Kessel Verlage: Remagen, Germany, 2003.
(67) Chow, S.; Pickles, K. *Wood Fiber Sci.* **1971**, *3*, 166−178.
(68) Pereira, A. N.; Mobedshahi, M.; R.Ladisch, M. *Methods Enzymol.* **1988**, *160*, 26−38.
(69) Taylor, J. B. *Trans. Faraday Soc.* **1957**, *53*, 1198−1203.
(70) Stubbs, G. *Rep. Prog. Phys.* **2001**, *64*, 1389−1425.
(71) Hardy, B.; Sarko, A. *J. Comput. Chem.* **1993**, *14*, 848−857.
(72) Levine, S. E.; Fox, J. M.; Blanch, H. W.; Clark, D. S. *Biotechnol. Bioeng.* **2010**, *107*, 37−51.
(73) Shao, X.; Lynd, L.; Wyman, C.; Bakker, A. *Biotechnol. Bioeng.* **2009**, *102*, 59−65.
(74) Shao, X.; Lynd, L.; Wyman, C. *Biotechnol. Bioeng.* **2009**, *102*, 66−72.