JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

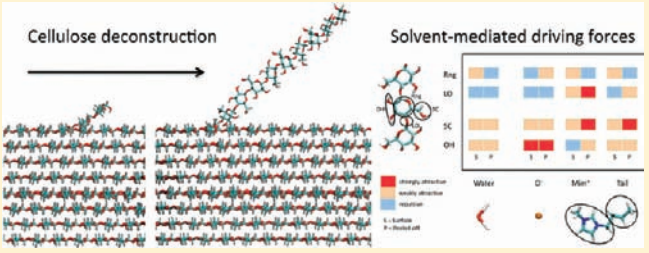
Dissecting Force Interactions in Cellulose Deconstruction Reveals the Required Solvent Versatility for Overcoming Biomass Recalcitrance

Hyung Min Cho, Adam S. Gross, and Jhih-Wei Chu*

Department of Chemical and Biomolecular Engineering, Energy Biosciences Institute, University of California, Berkeley, Berkeley, California, United States

Supporting Information

ABSTRACT: Pretreatment for deconstructing the multifaceted interaction network in crystalline cellulose is a limiting step in making fuels from lignocellulosic biomass. Not soluble in water and most organic solvents, cellulose was found to dissolve in certain classes of ionic liquids (ILs). To elucidate the underlying mechanisms, we simulated cellulose deconstruction by peeling off an 11-residue glucan chain from a cellulose microfibril and computed the free-energy profile in water and in 1-butyl-3-methylimidazolium chloride (BmimCl) IL. For this deconstruction process, the calculated free-energy cost/reduction in water/BmimCl is ∼2 kcal/mol per glucose residue, respectively. To unravel the molecular origin of solvent-induced differences, we devised a coarse graining scheme to dissect force interactions in simulation models by a force-matching method. The results establish that solvent−glucan interactions are dependent on the deconstruction state of cellulose. Water couples to the hydroxyl and side-chain groups of glucose residues more strongly in the peeled-off state but lacks driving forces to interact with sugar rings and linker oxygens. Conversely, BmimCl demonstrates versatility in targeting glucose residues in cellulose. Anions strongly interact with hydroxyl groups, and the coupling of cations to side chains and linker oxygens is stronger in the peeled-off state. Other than enhancing anion−hydroxyl group coupling, coarse-grain analysis of force interactions identifies configuring cations to target side chains and linker oxygens as a useful design strategy for pretreatment ILs. Furthermore, the state dependence of solvent−glucan interactions highlights specific stabilization and/or frustration of the different structure states of cellulose as important design parameters for pretreatment solvents.



1. INTRODUCTION

Lignocellulosic biomass is an abundant renewable energy source, but the complex interaction network of cellulose microfibrils in plant cell walls causes vast recalcitrance toward decomposition.[1−4] Pretreatment for exposing the buried glycosidic bonds to enable catalytic or enzymatic breakdown currently involves high costs, limiting the viability of biomass as a feedstock for fuel synethesis.[5,6] Significant advancements in pretreatment technology are thus urgently needed. Cellulose is not soluble in water and most organic solvents but can be dissolved by certain classes of ionic liquids (ILs) to an appreciable extent; the molecular origins however, remain unclear.[7−17] Although the use of IL as a preteatment solvent still faces serious challenges in cost, handling, and recycling,[18,19] elucidating how the molecular forces provided by IL molecules deconstruct crystalline cellulose can potentially inform novel engineering strategies or designs of pretreatment solvents.

Cellulose microfibrils in plant cell walls are slender aggregates of glucan chains with a wide range of diameters (2−20 nm) and lengths (0.1−100 μm).[20−22] Individual glucan chains are linear polymers composed of glucose residues linked together by β-1,4-glycosidic bonds. In a cellulose microfibril, glucan chains are aligned parallel to form flat sheets, and the sheets are stacked together to form the rodlike material.[23−25] The hydrogen-bonding (HB) network in crystalline cellulose microfibril involves three types of specific molecular interactions: intrachain OH···O hydrogen bonds (HBs) along a single chain, interchain OH···O HBs within a flat sheet, and intersheet CH···O interactions between staggered glucan chains in two different sheets (Figure 1A).[23−25] The donors and acceptors of OH···O HBs are from the equatorial hydroxyl groups and side chains of glucose residues. Interchain OH···O HBs had been considered as the main cause of cellulose recalcitrance.[7−10,26−29]

CH···O interactions in crystalline cellulose have high structural similarity with classical hydrogen bonds;[30,31] donors are from the axial CH groups in sugar rings, and acceptors are the oxygen atoms in ether linkers, hydroxyl groups, and side chains. In addition to their abundance, CH···O HBs are the most intransigent type of HBs as inferred from the fluctuations in their distances and occupancies.[32] On the hydrophobic, CH-exposing surfaces of a cellulose microfibril, all-atom molecular dynamics (MD) simulations illustrate that competitive hydrogen bonding with water can disrupt in-plane OH···O HBs but not the axial CH···O HBs.[32] Therefore, structural fluctuations observed in the MD simulations of cellulose in water highlight intersheet interactions as a key contributor to recalcitrance. As shown later, the energetics of cellulose deconstruction and water−glucan

Received: May 19, 2011
Published: July 28, 2011

dx.doi.org/10.1021/ja2046155 | J. Am. Chem. Soc. 2011, 133, 14033–14041

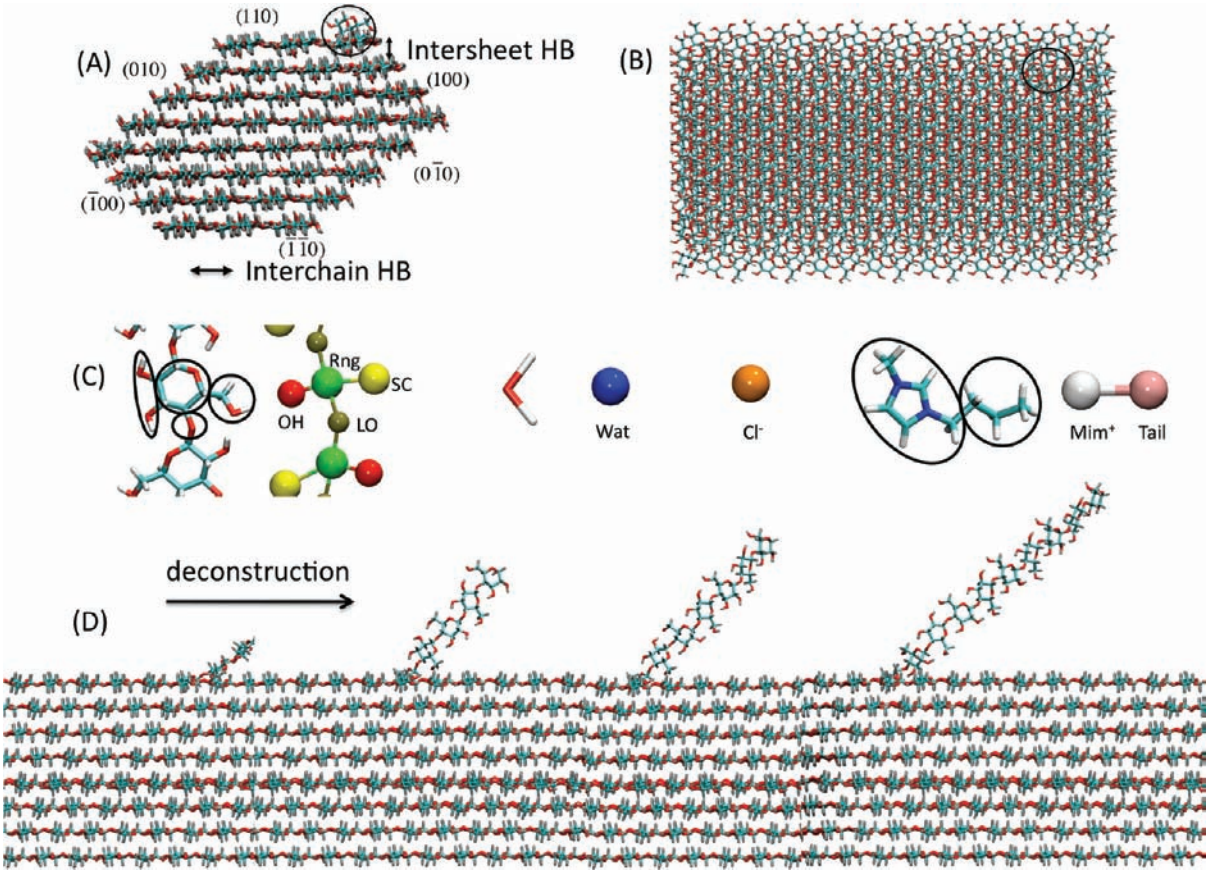


Figure 1. Simulation model of cellulose deconstruction. (A) Cross section of the model cellulose microfibril with 36 glucan chains. The corner chain being peeled off is circled, the different surfaces of cellulose microfibril are also labeled, and the directions of interchain and intersheet interactions are indicated. (B) Side view of the model cellulose microfibril. Each glucan chain has 16 glucose residues. (C) Coarse-graining scheme for computing the radial mean forces and radial distribution functions between glucose and solvent molecules. (D) Snapshots of the energy-minimized reaction path for peeling off the corner glucan chain labeled in panel A.

interactions calculated from free-energy simulations provide further evidence for this result.

While water and most organic solvents cannot solubulize cellulose, certain ionic liquids such as 1-butyl-3-methylimidazolium chloride (BmimCl), were found to dissolve crystalline cellulose to an appreciable extent.[7−10] The predominant explanation is that anions form hydrogen bonds with the equatorial hydroxyl groups and side chains of glucose residues. In model systems of water-soluble glucose or cellulobiose dissolved in IL, NMR measurements of relaxation times show that the stoichiometry of the affected anions is comparable to that of the hydroxyl groups.[11−13] MD simulations of solvated oligosaccharides and single glucan chains of different lengths also show strong HBs between IL anions and the hydroxyl groups.[14−16]

Since intersheet CH···O HBs are the most abundant and robust component in the interaction network,[32] IL anions disrupting interchain OH···O HBs alone is not sufficient for explaining IL's ability to dissolve cellulose.[17] The amount of crystalline cellulose that can be dissolved in IL is sensitive to the properties of the charge-carrying moiety of the cation as well as the length of its alkyl chains.[7−10,17] However, the molecular forces provided by IL molecules in cellulose deconstruction and the specific roles of cations are largely unknown.

To map out the molecular origin of IL-induced cellulose dissolution, we simulated cellulose deconstruction in water and in BmimCl using all-atom force fields and explicit solvent models. In both solvents, we computed the potential of mean force (PMF) of peeling off a glucan chain of 11 glucose residues from the most exposed corner of a model cellulose microfibril (Figure 1) by reaction path optimization and free-energy simulations.[33−35] A key challenge is dissecting how solvent components contribute to sculpting the apparent PMFs of cellulose deconstruction. Although the MD simulations performed in PMF calculatons contain the desired information, atomistic complexity often blurs physical mechanisms. Analysis of structural properties can reveal useful insights but it only provides indirect, convoluted information on interaction forces. For example, if proximity between IL cations and glucose residues was observed, cation−glucan interactions could be the cause but anion−glucan interactions in conjunction with the electrostatic coupling between anions and cations may also lead to the same phenomenon. Therefore, unraveling the specific roles of solvent components in cellulose deconstruction solely on the basis of structure analysis is difficult.

To devise a physically intuitive, interaction-based approach to elucidate the specific roles of solvent moieties in facilitating cellulose deconstruction, we devised a coarse-graining (CG) scheme (Figure 1C) to cluster atoms together according to their functional groups. The positions of atoms recorded in an MD trajectory are mapped to CG configurations and the atomistic

forces are summed to determine the net forces on CG sites. We then use a force matching method to decompose the net forces on CG sites according to the configurations sampled in the same all-atom MD trajectory to determine the radial mean forces between site pairs.[36] Integrals of radial mean forces define the CG potentials for representing the strengths of site−site interactions in the simulation model, including solvent−glucan pairs.

Free-energy simulations and coarse-grain analysis of force interactions illustrate that versatility in coupling to glucose residues from both equatorial and axial sides with sufficient strength is desired for a pretreatment solvent to dissolve crystalline cellulose. Molecular-level elucidation of the specific roles of IL anions and cations in cellulose deconstruction has been difficult and not yet established via other means. Principles such as optimizing cations to target the region around the side chain and linker oxygen in a glucose residue can also be extracted from simulation results for IL design. Deconstruction-state dependence of the calculated solvent−glucan CG potentials highlights specific stabilization and/or frustration of the different structure states of cellulose as important design parameters for pretreatment solvents.

2. COMPUTATIONAL METHODS

2.1. Molecular Dynamics Simulation. The recently developed CHARMM carbohydrate force fields[37,38] are used for all simulations performed in this work. The TIP3P model[39] and an atomic force field of BmimCl were used for explicit solvent simulations. We developed an all-atom force field for the BmimCl following the self-consistent protocol of CHARMM force field development.[40] The reported thermodynamic data of BmimCl (densities and heat capacities)[41,42] were employed in the iterative validation against the intramolecular potential energy surfaces and the interaction energies with water obtained from ab initio calculations.[40,43] A similar strategy of force-field development based on CHARMM parameters has also been applied to different ILs and demonstrated predictability of an array of experimental observables.[44−46] Validation and improvement of IL force field upon the availability of new experimental data is a continuous effort of ours. The proposed framework of dissecting force interactions in a simulation model can also be utilized to analyze the emergent properties of different force fields.

The model cellulose microfibril is adopted from those previously published,[32] which contains 36 glucan chains each with 16 glucose residues. The microfibril alone has 12 207 atoms, a diameter of 40 Å, and a length of 86 Å. In both water and BmimCl, at least four solvent shells were added to each side of the microfibril for solvation, and 18 931 water molecules were involved in the water simulations, for a total of 69 000 atoms. For BmimCl, 3661 IL pairs were used to solvate the microfibril, resulting in 107 393 atoms in the BmimCl simulations. To mimic crystalline cellulose microfibrils with a finite length system, harmonic potentials with a force constant of 100 kcal·mol^{-1}·Å^{-2} were added to restrain the C4 atoms at chain ends according to their positions in a crystalline cellulose, except for those on the solvent-exposing surfaces of the microfibril. The statistics of interchain and intersheet HBs in the simulations of the finite microfibril are in quantitative agreement with those observed in simulations in which the covalent bonds along a glucan chain are linked across the longitudinal periodic boundary to mimic an infinitely long cellulose microfibril.[32] Before heating to the targeted thermodynamic states, 50 000 conjugated energy-minimization steps are performed. Heating was done by random velocity assignment with a 4 K/ps heating rate over 100 ps followed by a 4 ns equilibration at the targeted thermodynamic state using Langevin dynamics[47] and a Langevin piston barostat.[48] The protocols of nonbound cutoffs, particle-mesh Ewald parameters, and other details in MD simulation protocol are the same as those previously published.[32] All stimulations were performed with the NAMD program;[47] system setup and analyses, force matching, and the calculations of RDFs were performed with the CHARMM program[49] and in-house codes.

2.2. Reaction Path Optimization and Free-Energy Simulations. A model cellulose microfibril composed of 36 glucan chains each with 16 glucose residues[32] was developed based on the I_α form of crystalline cellulose (Figure 1). The most solvent-exposed glucan chain on the (110) surface of microfibril is peeled off from the reducing end (Figure 1). In the final replica, 11 glucose residues are peeled off away from the microfibril and three glucose residues in the nonreducing end of the targeted glucan chain are left on the surface. The peeled-off structure was first stabilized by harmonic restraint potentials for reaction path optimization[33,34] with an implicit solvent model based on the generalized Born theory.[50] In the reaction path optimization using a chain of replicas, the distances between neighboring replicas along the path are defined by the Cartesian coordinates of all the atoms of the peeled-off glucan chain and were all maintained at the same value.[33,34] By use of the intact and peeled-off states of the microfibril as end replicas, several rounds of reaction path optimization were performed to determine the total number of replicas in the path. Starting from a three-replica path with a single intermediate state, we double the resolution of the path by inserting a new replica in between two neighboring replicas in the previous round.[33] The increase in resolution stops when doubling the resolution results in only insignificant changes (±1 kcal/mol) in the profile of the accumulated work. The smooth, energy-minimized 32-replica path with the rmsd (root-mean-squared difference) between neighboring replicas being 0.67 Å was used as the initial path for finite-temperature path optimization in explicit water.[35]

In the energy-minimized path, the total contact number associated with the peeled-off glucan gradually decreases during deconstruction. The total contact number[51] of the heavy atoms that form interchain OH···O and CH···O HBs with the peeled-off glucan chain was used as the collective variable for path optimization in explicit water at 300 K and 1 atm. Restrained MD simulations with a force constant of 50 kcal·mol^{-1}·Å^{-2} on the collective variable were performed for 1 ns to calculated mean forces.[35] The averaged displacements of OH···O and CH···O contacts observed in restrained MD simulations showed that reforming CH···O contacts is the dominant mechanism for the peeled-off chain to mend itself back to the surface of the microfibril. Therefore, the weighting of OH···O contacts in describing deconstruction was gradually reduced to zero. This result validates the relative robustness of interchain and intersheet interactions in water as inferred from the structural fluctuations observed in MD simulations[32] and provides direct evidence that disrupting CH···O HBs dominates the free-energy cost of cellulose deconstruction in water.

The CH···O contact number was thus used as the collective variable for PMF calculations. In water, the free-energy simulation was carried out at 300 K and 1 atm. The mean forces were calculated with 6 ns restrained MD simulations (after 100 ps of

heating and 3 ns of equilibration) for 32 equally spaced values of CH···O contacts. In BmimCl, the free-energy simulation was conducted at 450 K and 1 atm for accelerated dynamics (experimental melting point 340 K for BmimCl) and is within the temperature range of the experimental data used for force field validation.[41−43] Production runs (10 ns) after 100-ps heating and 3-ns equilibration were collected for computing the mean forces at the 32 equally spaced values of CH···O contacts (the same as in the water case) to compute the PMF of cellulose deconstruction in BmimCl.

2.3. Dissecting Force Interactions by Coarse Graining. To compute the interaction potentials between solvent and glucose CG sites, configurations in all-atom MD simulations were saved every 1.0 ps for force-matching calculations.[36] The radial force approximation was applied to the nonbonded interactions between solvent−solvent and solvent−glucan CG sites with force tables of 0.1 Å resolution. Between CG sites with covalent bonds, harmonic bond potentials were used to describe their interactions with the equilibrium lengths calculated from the ensemble averages of the all-atom simulation data. Between glucose CG sites in the cellulose microfibril, each site was connected to the neighboring sites within a cutoff radius of 10 Å to form an initial bond network; the actual number of elastic bonds attached to a glucose CG site was then adjusted according to bond lengths to reproduce the thermal fluctuations of CG sites calculated from the all-atom MD simulation. The force constants were determined along with the force tables of nonbound interactions by the force-matching method.[36]

In water, CG potentials calculated from the 6-ns replica-32 simulation quantitatively agree with those calculated from the 6-ns replica-1 simulation for glucose residues in the microfibril that were not peeled off. If the 6-ns simulation data of all 32 replicas were combined with separate force tables for glucose residues in the peeled-off state and for those in the microfibril state, the resulting CG potentials also agree with those calculated from the replica-32 simulation, demonstrating reproducibility in the results of force matching. In BmimCl, the replica-32 simulation was extended to 100 ns to obtain smooth CG potentials. For the solvent−glucan CG potentials of glucose residues in the microfibril state, quantitative agreement is achieved between the results calculated from 100 ns replica-32 simulation and those from 100 ns replica-1 simulation. For the solvent−glucan CG potentials in both the peeled-off and the microfibril states, quantitative agreement is also achieved between the results from the 100 ns replica-32 simulation and those from combining the 10-ns MD data of all 32 replicas. All glucose residues in the system were involved in calculating CG potentials, but those on surfaces have the dominant contribution to solvent interactions. If only surface residues were used for force matching, the resulting CG potentials are essentially the same. To examine the solvation structures around the peeled-off chain at different deconstruction states, we also computed the radial distribution functions (RDFs) between the CG sites of solvent and glucose residues.

3. RESULTS AND DISCUSSION

3.1. Potential of Mean Force of Peeling Off a Glucan Chain from Cellulose Microfibril. At 300 K and 1 atm in water, the calculated PMF of the cellulose deconstruction is shown in Figure 2. The PMF increases with the progression of peeling off the targeted glucan chain (Figure 1A) and reaches a value of 19.3 ± 1.2 kcal/mol at the final replica. The free-energy cost of deconstructing a glucan from crystalline cellulose in water is thus ∼1.9 kcal/mol per glucose residue. This value is also close to those reported in a recent simulation work.[52]

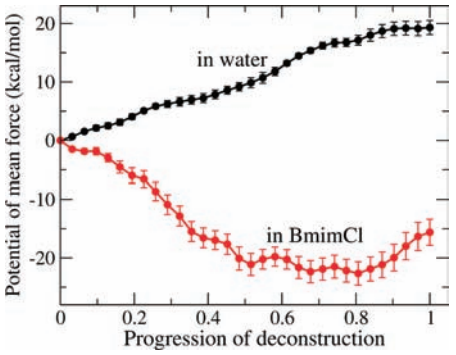


Figure 2. Potential of mean force, in kilocalories per mole, for peeling off the corner glucan chain from the microfibril shown in Figure 1 in water and in BmimCl.

The PMF of peeling off the same corner glucan chain from the microfibril in BmimCl is calculated at 450 K and 1 atm via 10 ns production-stage restrained MD simulations for all 32 replicas. BmimCl is a potent IL for dissolving cellulose,[7−13] and the calculated PMF of deconstruction is indeed downhill, as shown in Figure 2, and is clearly distinct from the PMF in water. Quantitative measurements of the solubility of crystalline cellulose in IL and the transfer free energy of a glucan chain from the crystalline phase to BmimCl are not available to the best of our knowledge. This work thus provides a simulation-based prediction to motivate such measurements for further validation and improvement of the simulation model and IL force fields. We also recognize that IL molecules would affect all accessible glucan chains, and the evolution in chain−chain packing and microfibril conformation during deconstruction would be more complicated than the process simulated here, which focuses on peeling off a single glucan chain while keeping the rest of the microfibril in its crystalline arrangement (details are discussed in section 2.2). We adopt this simple model process to compare the specific roles of water and BmimCl in cellulose deconstruction under the same procedure. Toward the end of peeling off the targeted chain after the 0.8 value of deconstruction progression, an increase in the BmimCl PMF followed by plateauing is observed, suggesting that barrier crossing may be involved in deconstruction. Although the calculated PMF in BmimCl is specific to the simulated model process, kinetically trapped dissolution of cellulose in IL is indeed often observed experimentally.[9] Based on the PMF calculated in BmimCl, the free-energy reduction of peeling off the glucan chain is estimated to be around 2.0 kcal/mol per glucose residue.

The calculated PMFs of cellulose deconstruction agree with the observations that cellulose can be dissolved in BmimCl but not in water.[7−13] To elucidate the molecular origin, we computed the effective interactions between the CG sites of glucan and solvent molecules. As shown in Figure 1C, the OH site contains O2 and O3 hydroxyl groups; the SC site contains side-chain atoms, the LO site contains only the linker oxygen, and the Rng site contains the remaining atoms in the sugar ring. The donors and acceptors of interchain OH···O HBs are in OH and SC sites, while the Rng sites contain CH···O donors and the OH, SC, and LO sites contain CH···O acceptors. Each water molecule is coarse-grained into a single site, and each BmimCl molecule is coarse-grained into three sites: the Cl^- single-atom

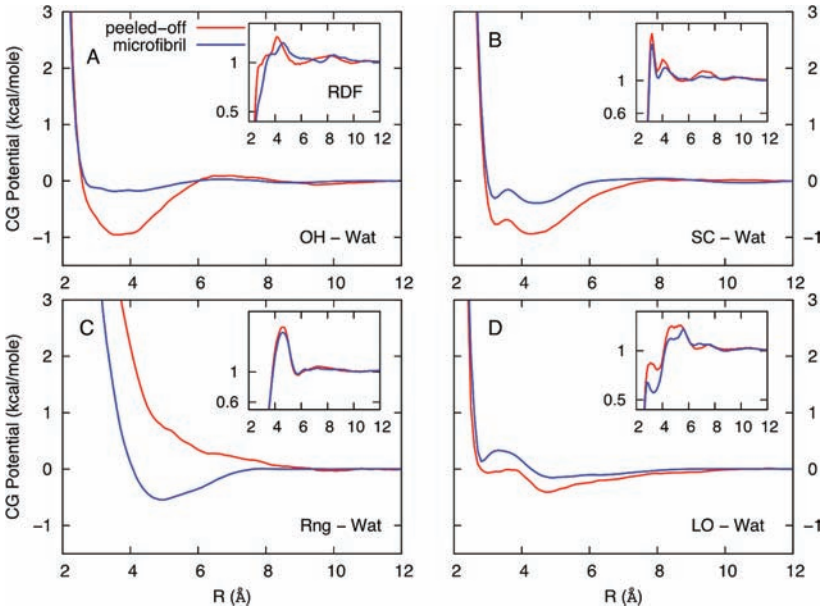


Figure 3. CG potentials and RDFs between CG sites of water and glucose residues defined in Figure 1C: (A) OH−Wat, (B) SC−Wat, (C) Rng−Wat, and (D) LO−Wat.

site, the Mim^+ site containing the charge carrying portion of the cation, and the Tail site containing the remaining aliphatic groups. In this way, how solvent moieties target glucose residues to disrupt the interaction network in cellulose can be analyzed systematically. For example, the well depth in the CG potential between OH and Cl^- sites represents the anion's ability to disrupt interchain OH···O HBs.

We calculated the radial mean forces between CG sites from the positions and forces mapped from all-atom MD simulation by a force matching method.[36] The integrals of radial mean forces define CG potentials between different site types in the simulation model. Since the net forces acting on CG sites were decomposed according to the configurations sampled in the same MD trajectory, CG potentials provide a deconvoluted view of the specific molecular interactions in a multicomponent, condensed-phase system.[36,53] They also inform how solvent components in the simulation model favor or disfavor the peeled-off state, at which the targeted glucan chain has a higher number of solvent contacts. A critical question is if solvent−glucan coupling exhibits different stabilization effects depending on the structure state of cellulose, especially when molecular solvents such as water and IL are used. To learn this information from molecular simulation, we compare the CG potentials of solvent−glucose coupling at different states of cellulose deconstruction.

We categorize glucan−solvent CG potentials based on their depths in the nearest-neighbor well. Since the free-energy cost or reduction in peeling off the corner glucan chain in water and BmimCl is around ±2 kcal/mol, CG potentials with the nearest-neighbor well lower than −2 kcal/mol are considered strongly attractive, those between 0 and −2 kcal/mol are considered weakly attractive, and others are considered repulsive.

3.2. Water CG Potentials Are State-Dependent and Lack Driving Forces to Couple to Sugar Rings and Linker Oxygens. The CG potentials between water and glucan sites calculated from the simulation of replica-32 are shown in Figure 3 along with their radial distribution functions (RDFs). The weakly attractive CG potential and RDF of the Wat−Wat pair are shown in Figure S1 (Supporting Information) for comparison. Figure 3A and 3B indicate that water molecules exhibit weakly attractive interactions with the OH and SC sites. Water−OH and water−SC potentials have clear state dependence and become more attractive in the peeled-off state, illustrating the driving forces of water molecules to interact with the donors and acceptors of interchain OH···O HBs. The deconstruction-state dependence of CG potentials also indicates that specific stabilization of the different structure states of cellulose is indeed an important property of pretreatment solvents.

The CG potentials between water and the Rng and LO sites (Figure 1C,D), however, are repulsive or nearly repulsive. Figure 3C shows that, for glucose residues on the microfibril, Rng−Wat interactions are weakly attractive and become repulsive in the peeled-off state. This qualitative change illustrates the lack of driving forces for exposing the aliphatic CH groups of glucose residues in water. Furthermore, water molecules have repulsive interactions with LO sites, another component of CH···O acceptors (Figure 3D). In addition to the results of path optimization that show that CH···O contacts are the dominant mechanism for the peeled-off glucose residues to heal themselves back, coarse-grain analysis of force interactions provides further evidence that CH···O contacts are the major bottleneck for cellulose deconstruction in water.

3.3. BmimCl Anions Strongly Interact with the Hydroxyl Groups of Glucose Residues. For OH and SC sites, the CG potentials and RDFs with the Cl^- and Mim^+ sites of BmimCl are shown in Figure 4. The interactions between OH and Cl^- sites are strongly attractive, as expected (Figure 4A).[11−16] In comparison with glucose residues on the microfibril, the OH−Cl^- pair in the peeled-off state has a very similar CG potential. Although the OH−Cl^- potential is transferrable, distinct solvent structures around glucan chains are observed at different structure states. The first-peak heights of OH−Cl^- RDFs in the peeled-off state and the microfibril states are very close, but the second peak is much higher in the peeled-off state. Further analyses show that this height increase in the second peak of OH−Cl^- RDF comes

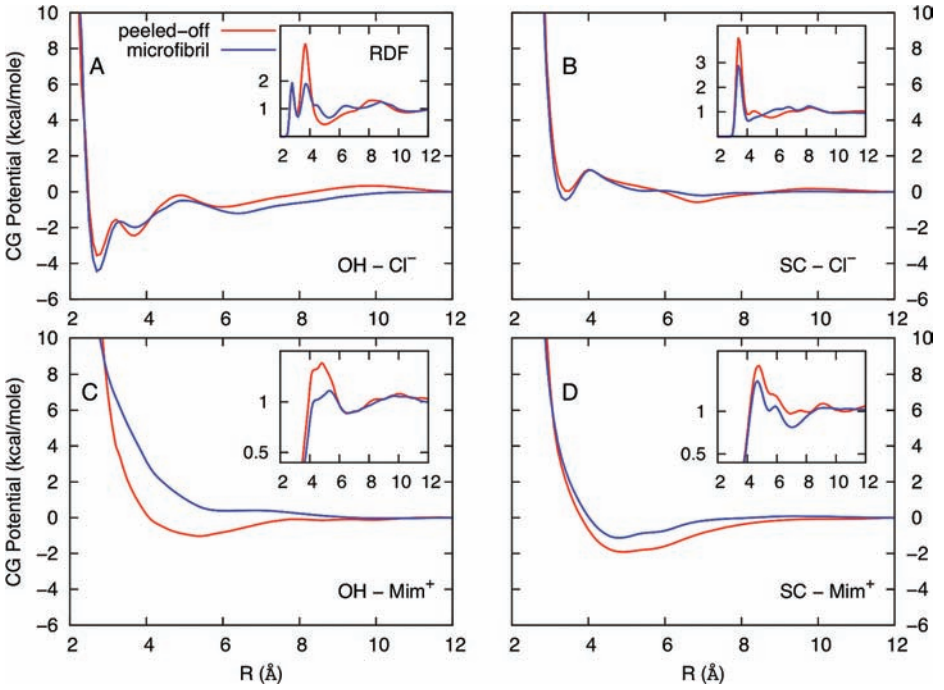


Figure 4. CG potentials and RDFs between CG sites of BmimCl and hydroxyl (OH) and side-chain (SC) sites of glucose residues defined in Figure 1C: (A) OH−Cl^-, (B) SC−Cl^-, (C) OH−Mim^+, and (D) SC−Mim^+.

from Cl^- ions interacting with the hydroxyl groups from the axial directions after CH···O contacts in the microfibril state are removed in the peeled-off state. The OH−Cl^- pair thus also provides driving forces for disrupting intersheet interactions in cellulose in addition to intrachain interactions. To side chains, SC−Cl^- interactions are only weakly attractive, and glucose residues in the peeled-off and microfibril states also have very similar CG potentials (Figure 4B). The CG potentials of OH−Cl^- and SC−Cl^- pairs establish that Cl^- anions disrupt the interchain and intersheet interactions in cellulose mainly via interactions with the OH sites of glucose residues.

To OH sites, the covalently linked Mim^+ and Tail sites of cations exhibit weakly attractive or repulsive interactions. Figure 4C shows the CG potentials and RDFs of the OH−Mim^+ pair, and Figure S2A (Supporting Information) shows those of the OH−Tail pair. The OH−Mim^+ interaction is repulsive for glucose residues in the microfibril state and weakly attractive for those in the peeled-off state; the OH−Tail pair follows a similar trend. These results indicate that the limited space around glucose residues on the microfibril hindered the formation of OH−cation contacts and their interactions become more attractive after the glucan chain is peeled off. In the peeled-off state, the OH−Mim^+ and OH−Tail RDFs clearly have higher peaks; see Figure 4C and Figure S2A (Supporting Information).

To equatorial SC sites, the covalently linked Mim^+ and Tail sites of BmimCl cations exhibit strongly attractive interactions with glucose residues in the peeled-off state and weakly attractive interactions for those in the microfibril state; see Figure 4D and Figure S2B (Supporting Information). The SC−Mim^+ and SC−Tail RDFs shown in the same figures indicate higher concentrations of cation moieties around the glucan chain after it is peeled off. Coarse-grain analysis establishes that the amphiphilic glucose side chains and IL cations better interact with each other in the peeled-off state. As SC sites are involved in forming OH···O HBs as well as CH···O HBs, cations also provide strong driving forces to disrupt the interchain as well as intersheet interactions in cellulose.

3.4. BmimCl Cations Strongly Interact with Side Chains and Linker Oxygens in Glucose Residues. The above analysis of IL−glucan coupling reveals that both anions and cations participate in targeting the equatorial groups of glucose residues for cellulose deconstruction. Anions dominate the coupling to OH sites, whereas cations interact with SC sites more favorably. Both equatorial sites participate in forming interchain OH···O HBs and intersheet CH···O HBs, and IL anions and cations are thus both involved in disrupting these interactions in cellulose.

The CH groups protruding outward along the axial directions of sugar ring are donors of the intersheet CH···O HBs in cellulose[32] and are coarse-grained into the Rng sites as shown in Figure 1C. CH···O acceptors include the OH site and linker oxygens that are assigned a stand-alone CG site. Figure 5A shows that the Rng−Cl^- pair has a repulsive CG potential for glucose residues in the microfibril state but a weakly attractive CG potential for residues in the peeled-off state. Therefore, Cl^- anions contribute to disrupting intersheet interactions also by interacting with CH donors in addition to their strong coupling to OH sites. The LO−Cl^- pair is repulsive in both states (Figure 5B).

The Rng−Mim^+ CG potentials have a state dependence opposite to those of the Rng−Cl^- pair. The Rng−Mim^+ pair is weakly attractive for glucose residues in the microfibril state but repulsive for those in the peeled-off state (Figure 5C), and the Rng−Tail pair follows a similar trend as shown in Figure S2C (Supporting Information). Cl^- and Mim^+ both contain acceptors to interact with the axial CH donors in Rng sites, and the opposite dependence of Rng−Mim^+ and Rng−Cl^- pairs on deconstruction state indicates the complementarity of cations and anions in coupling to the Rng sites. For glucose residues in

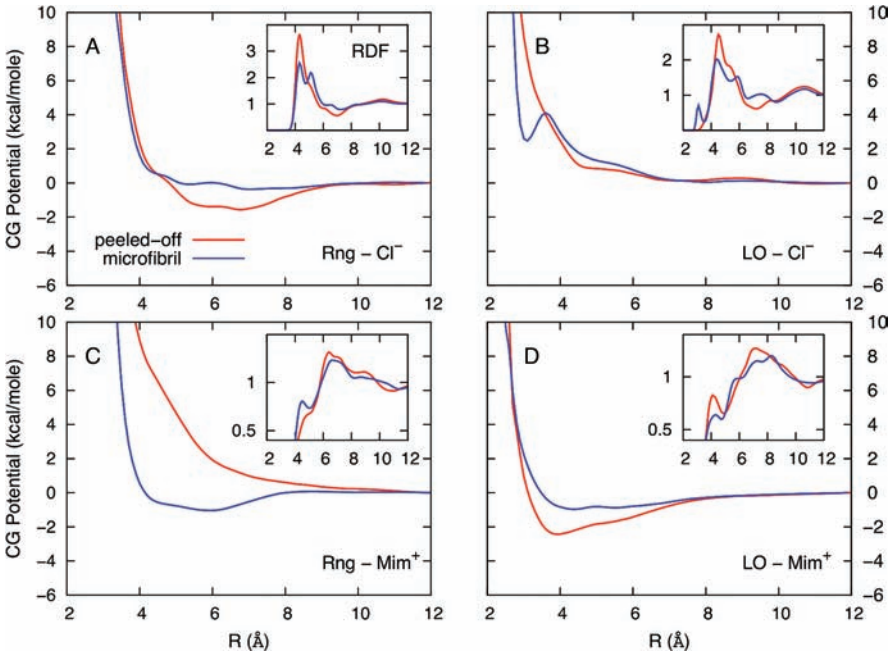


Figure 5. CG potentials and RDFs between CG sites of BmimCl and CG sites of sugar rings and linker oxygens of glucose residues defined in Figure 1C: (A) Rng−Cl^-, (B) LO−Cl^-, (C) Rng−Mim^+, and (D) LO−Mim^+.

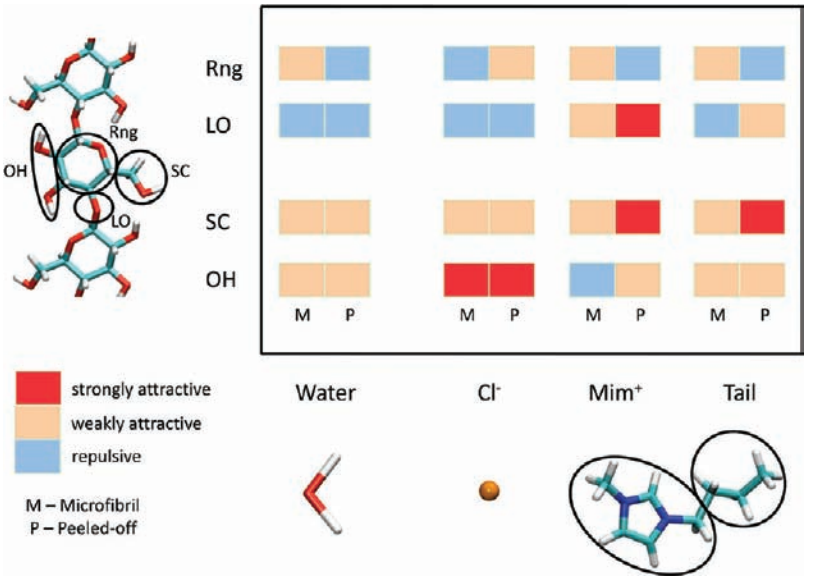


Figure 6. Representation of the solvent versatility provided by water and BmimCl molecules in interacting with glucose residues in the microfibril and the peeled-off glucan chain.

the microfibril state, the surface topology allows favorable Rng−Mim^+ coupling, but not in the peeled-off state (Figure 5C). The aforementioned access of Cl^- anions to OH sites from the axial directions of glucose residues in the peeled-off state facilitates Rng−Cl^- coupling but also hinders Rng−Mim^+ interactions. Furthermore, as SC−Mim^+ and SC−Tail CG potentials both become more attractive in the peeled-off state via axial access, Rng−Mim^+ and Rng−Tail coupling is further inhibited. The responsive state dependence of the CG potentials of Rng−cation pairs reflects the weak nature of their interactions.

On the other hand, the LO−Mim^+ pair exhibits strongly attractive interactions for the peeled-off glucose residues and weakly attractive interactions for those on the microfibril (Figure 5D). The LO−Tail pair follows a similar qualitative trend and is weakly attractive in the peeled-off state and repulsive for glucose residues in the microfibril state (Figure S2D, Supporting Information). The LO−Mim^+ and LO−Tail CG potentials and their state dependence indicate that Bmim cations provide strong driving forces to target linker oxygens in addition to side chains for deconstructing cellulose.

dx.doi.org/10.1021/ja2046155 |*J. Am. Chem. Soc.* 2011, 133, 14033–14041

The CG potentials and RDFs between BmimCl sites are shown in Figure S3 (Supporting Information). The strong electrostatic coupling is clear in the deep well of the Cl^-—Mim^+ potential, and the Cl^-—Cl^- and Mim^+—Mim^+ pairs have strongly repulsive CG potentials as expected. The other IL—IL pairs have only weakly attractive CG potentials. Even though the well depth of the most attractive BmimCl—glucan potential, the OH—Cl^- pair, is only −4 kcal/mol (Figure 4A), much less than that of the Cl^-—Mim^+ pair, −14 kcal/mol (Figure S3C, Supporting Information), an overall reduction in free energy of cellulose deconstruction can still be achieved as Cl^- and Mim^+ sites can exhibit strongly attractive or weakly attractive interactions with glucose sites (Figures 4 and 5) while still coupling to each other. The strong electrostatic coupling between anions and cations thus allows colocalization of distinct molecular moieties in BmimCl to interact with glucose residues for overcoming the multifaceted interaction network in cellulose. Using coarse-grain analysis to dissect force interactions reveals both the structures and strengths of solvent—glucan coupling, thus enabling the deduction of an unprecedented mechanistic understanding of solvent-induced cellulose deconstruction.

3.5. Required Solvent Versatility for Overcoming Biomass Recalcitrance. With a complex interaction network in crystalline cellulose, an effective pretreatment solvent needs to disrupt both the interchain and intersheet interactions therein. Dissecting force interactions in simulation models reveals that water lacks the driving forces to interact with Rng and LO sites, and the coupling to equatorial OH and SC sites is only weakly attractive. Therefore, water is a poor solvent for deconstructing cellulose. Conversely, BmimCl has the versatility that Cl^- anions have strongly attractive interactions with hydroxyl groups and weakly attractive interactions with side chains and sugar rings, and cations have strongly attractive interactions with side chains as well as linker oxygens. These results are summarized in Figure 6 to contrast the two solvents in cellulose deconstruction.

As neither anions nor cations alone cover all of the strongly attractive interactions with glucose residues, their coupled actions are necessary, and the strong electrostatic coupling in IL enables colocalization of different molecular moieties to target glucose residues. For interchain interactions in cellulose, OH and SC sites contain OH···O donors and acceptors, and both the anions and cations of BmimCl couple to these moieties. For intersheet interactions, anions block CH···O contacts by interacting with equatorial OH and SC sites as well as coupling to the Rng sites that contain CH···O donors, and Bmim cations exhibit strongly attractive interactions with linker oxygens and side chains. Versatility in interacting with glucose residues from both the equatorial and axial sides with sufficient strength thus emerges as the desired property in a pretreatment solvent to deconstruct cellulose.

4. CONCLUSIONS

Converting lignocellulosic biomass into liquid fuels would be a far more productive route for producing fuels if the strongly interconnected structures of cellulose microfibrils could be deconstructed more effectively. The recalcitrant interaction network in cellulose originates from coupling between the equatorial and axial moieties of glucose residues. Both OH and SC groups participate in forming interchain and intersheet interactions with their hydroxyl groups, and sugar rings (donors) and linker oxygens (acceptors) contribute to forming intersheet interactions. To make the buried glycosidic bonds in cellulose accessible for catalytic or enzymatic breakdown, a pretreatment solvent needs to disrupt these interactions. Although certain ILs were found to dissolve cellulose, the molecular forces that are involved, especially the specific roles of anions and cations, are largely unclear.[7−17] The lack of mechanistic understanding prevents breakthroughs and rational design in advancing pretreatment technologies.

By using coarse-grain analysis to acquire both the structures and strengths of solvent—glucan interactions, we developed an unprecedented mechanistic understanding of solvent-induced cellulose deconstruction. The calculated water—glucan and cation—glucan CG potentials depend on the state of cellulose deconstruction, while the interactions associated with monoatomic Cl^- anions are transferrable. Water molecules favorably interact with the hydroxyl and side-chain groups of glucose residues in the peeled-off state but lack driving forces to interact with the sugar rings and linker oxygens. Conversely, the BmimCl IL demonstrates versatility in targeting glucose residues in cellulose. Anions strongly interact with hydroxyl groups, and the coupling of cations to side chains and linker oxygens is stronger in the peeled-off state. In addition to enhancing anion—hydroxyl group coupling, optimizing cations to target the side chains and linker oxygens of glucose residues is also identified as a useful strategy to optimize ILs for pretreating cellulose. Furthermore, the state dependence of solvent—glucan CG potentials indicates that specific stabilization and/or frustration of the different structure states of cellulose are also important design parameters in engineering pretreatment solvents. Solvent components with complex molecular structures such as water and BmimCl cation tend to have more responsive CG potentials to the deconstruction state than the simple, monoatomic Cl^- anion.

■ ASSOCIATED CONTENT

Supporting Information. Three additional figures and complete ref 49. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
jwchu@berkeley.edu

■ ACKNOWLEDGMENT

We thank the Energy Biosciences Institute (Grant OO0J04) and the University of California, Berkeley, for supporting this research. The research was performed by use of the computational resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC02-05CH11231, the NISC Advanced Computing Center Kraken cluster under the National Science Foundation Teragrid Grants TG-MCB100106 and TG-MCB100104, and the UIUC Advanced Computing Center Lincoln cluster under the National Science Foundation Teragrid Grant TG-CTS100070.

■ REFERENCES

(1) Himmel, M. E.; Ding, S.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. *Science* **2007**, *315*, 804.
(2) Carroll, A.; Somerville, C. *Annu. Rev. Plant Biol.* **2009**, *60*, 165.

(3) Galbe, M.; Zacchi, G. *Adv. Biochem. Eng. Biotechnol.* **2007**, *108*, 41.
(4) Chandra, R. P.; Bura, R.; Mabee, W. E.; Berlin, A.; Pan, X.; Saddler, J. N. *Adv Biochem Eng Biot* **2007**, *108*, 67.
(5) Eggeman, T.; Elander, R. *Bioresour. Technol.* **2005**, *96*, 2019.
(6) Yang, B.; Wyman, C. E. *Biofuels, Bioproducts Biorefining* **2008**, *2*, 26.
(7) Swatloski, R.; Spear, S.; Holbrey, J.; Rogers, R. *J. Am. Chem. Soc.* **2002**, *124*, 4974.
(8) Zhu, S.; Wu, Y.; Chen, Q.; Yu, Z.; Wang, C.; Jin, S.; Ding, Y.; Wu, G. *Green Chem.* **2006**, *8*, 325.
(9) Pinkert, A.; Marsh, K. N.; Pang, S.; Staiger, M. P. *Chem. Rev.* **2009**, *109*, 6712.
(10) Zakrzewska, M. E.; Bogel-Lukasik, E.; Bogel-Lukasik, R. *Energy Fuels* **2010**, *24*, 737.
(11) Moulthrop, J.; Swatloski, R.; Moyna, G.; Rogers, R. *Chem. Commun.* **2005**, 1557.
(12) Remsing, R.; Swatloski, R.; Rogers, R.; Moyna, G. *Chem. Commun.* **2006**, 1271.
(13) Remsing, R. C.; Hernandez, G.; Swatloski, R. P.; Massefski, W. W.; Rogers, R. D.; Moyna, G. *J. Phys. Chem. B* **2008**, *112*, 11071.
(14) Youngs, T. G. A.; Hardacre, C.; Holbrey, J. D. *J. Phys. Chem. B* **2007**, *111*, 13765.
(15) Remsing, R. C.; Liu, Z.; Sergeyev, I.; Moyna, G. *J. Phys. Chem. B* **2008**, *112*, 7363.
(16) Liu, H.; Sale, K. L.; Holmes, B. M.; Simmons, B. A.; Singh, S. *J. Phys. Chem. B* **2010**, *114*, 4293.
(17) Lindman, B.; Karlstrom, G.; Stigsson, L. *J. Mol. Liq.* **2010**, *156*, 76.
(18) Simmons, B. A.; Singh, S.; Holmes, B. M.; Blanch, H. W. *Chem. Eng. Prog.* **2010**, *106*, 50.
(19) Shill, K.; Padmanabhan, S.; Xin, Q.; Prausnitz, J. M.; Clark, D. S.; Blanch, H. W. *Biotechnol. Bioeng.* **2011**, *108*, 511.
(20) Atalla, R.; Vanderhart, D. *Science* **1984**, *223*, 283.
(21) Somerville, C. *Annu. Rev. Cell. Dev. Biol.* **2006**, *22*, 53.
(22) Ding, S.; Himmel, M. *J. Agric. Food Chem.* **2006**, *54*, 597.
(23) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074.
(24) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300.
(25) Nishiyama, Y.; Johnson, G. P.; French, A. D.; Forsyth, V. T.; Langan, P. *Biomacromolecules* **2008**, *9*, 3133.
(26) Watanabe, A.; Morita, S.; Ozaki, Y. *Biomacromolecules* **2006**, *7*, 3164.
(27) Tashiro, K.; Kobayashi, M. *Polymer* **1991**, *32*, 1516.
(28) Marechal, Y.; Chanzy, H. *J. Mol. Struct.* **2000**, *523*, 183.
(29) Kondo, T. In *Polysaccharides: Structural diversity and functional versatility*, 2nd ed.; Dumitriu, S., Ed.; Marcel Dekker: New York, 2005; p 69.
(30) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 5063.
(31) Steiner, T.; Saenger, W. *J. Am. Chem. Soc.* **1992**, *114*, 10146.
(32) Gross, A. S.; Chu, J.-W. *J. Phys. Chem. B* **2010**, *114*, 13333.
(33) Brokaw, J. B.; Haas, K. R.; Chu, J.-W. *J. Chem. Theory Comput.* **2009**, *5*, 2050.
(34) Haas, R. K.; Chu, J.-W. *J. Chem. Phys.* **2009**, *131*, No. 144105.
(35) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. *J. Chem. Phys.* **2006**, *125*, No. 024106.
(36) Noid, W. G.; Chu, J.-W.; Ayton, G. S.; Krishna, V.; Izvekov, S.; Voth, G. A.; Das, A.; Andersen, H. C. *J. Chem. Phys.* **2008**, *128*, No. 244114.
(37) 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.
(38) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. *J. Chem. Theory Comput.* **2009**, *5*, 2353.
(39) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926.
(40) Mackerell, A. D. *J. Comput. Chem.* **2004**, *25*, 1584.
(41) Holbrey, J.; Reichert, W.; Reddy, R.; Rogers, R. *ACS Symp. Ser.* **2003**, *856*, 121.
(42) Pomelli, C. S.; Chiappe, C.; Vidis, A.; Laurenczy, G.; Dyson, P. J. *J. Phys. Chem. B* **2007**, *111*, 13014.
(43) Gross, S. A.; Bell, A. T.; Chu, J.-W. *J. Phys. Chem. B* **2011**, under review.
(44) Morrow, T.; Maginn, E. *J. Phys. Chem. B* **2002**, *106*, 12807.
(45) Lee, S.; Jung, J.; Han, Y. *Chem. Phys. Lett.* **2005**, *406*, 332.
(46) Maginn, E. J. *J. Phys.: Condens. Matter* **2009**, *21*, 373101.
(47) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
(48) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613.
(49) Brooks, B. R.; et al. *J. Comput. Chem.* **2009**, *30*, 1545.
(50) Haberthuer, U.; Caflisch, A. *J. Comput. Chem.* **2008**, *29*, 701.
(51) Iannuzzi, M.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2003**, *90*, No. 238302.
(52) Beckham, G. T.; Matthews, J. F.; Peters, B.; Bomble, Y. J.; Himmel, M. E.; Crowley, M. F. *J. Phys. Chem. B* **2011**, *115*, 4118.
(53) Cho, H.; Chu, J.-W. *J. Chem. Phys.* **2009**, *131*, No. 134107.