

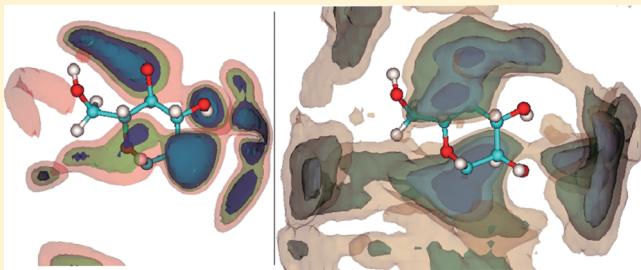
# Thermodynamics of Cellulose Solvation in Water and the Ionic Liquid 1-Butyl-3-Methylimidazolium Chloride

Adam S. Gross, Alexis T. Bell, and Jhih-Wei Chu\*

Department of Chemical and Biomolecular Engineering, University of California, Berkeley, Berkeley, California 94720, United States

Supporting Information

**ABSTRACT:** Cellulose is present in biomass as crystalline microfibrils held together by a complex network of intermolecular interactions making it difficult to initiate its hydrolysis and conversion to fuels. While cellulose is insoluble in water and most organic solvents, complete dissolution of cellulose can be achieved in certain classes of ionic liquids (ILs). The present study was undertaken to analyze the thermodynamic driving forces of this process and to understand how the anions and cations comprising an IL interact with the different moieties of glucose residues to cause dissolution. All-atom molecular dynamics (MD) simulations were performed at two extreme states of cellulose dissolution: a crystalline microfibril and a dissociated state in which all the glucan chains of the microfibril are fully separated from each other by at least four solvation shells. MD simulations of the two states were carried out in water and in the IL 1-butyl-3-methylimidazolium chloride (BmimCl) to provide a comprehensive analysis of solvent effects on cellulose dissolution. The results reveal two important molecular aspects of the mechanism of cellulose dissolution. The first is that the perturbation of solvent structures by the dissolved glucan chains can be a crucial factor in determining solubility, particularly for the insolubility of cellulose in water at 300 K. Second, both the  $\text{Cl}^-$  and the  $\text{Bmim}^+$  ions of BmimCl interact with the moieties of glucan residues that form intersheet contacts, the most robust component of the interaction network of crystalline cellulose.  $\text{Cl}^-$  anions can form hydrogen bonds (HBs) with the hydroxyl groups of glucan chains from either the equatorial or the axial directions. For  $\text{Bmim}^+$  cations, the calculated density profiles reveal that the contacts with glucan chains along the axial directions are closer than those along the equatorial directions. On the basis of the results of atomistic MD simulations, we propose that interacting with glucan chains along axial directions and disrupting the intersheet contacts of cellulose is an important ability of cellulose pretreatment solvents.



## INTRODUCTION

Lignocellulosic biomass has the potential to serve as a renewable feedstock for the production of transportation fuels.<sup>1–8</sup> One of the primary routes to fuels under consideration involves the isolation of the cellulosic and hemicellulosic components of biomass, hydrolysis of these biopolymers to produce glucose and xylose, and subsequent fermentation of these sugars to produce ethanol or butanol. Alternatively, the sugars can be processed with chemical catalysis to produce a broad spectrum of products compatible with gasoline and diesel fuels. Releasing the carbohydrate fraction of biomass remains a major challenge since the cellulosic component is present in the form of microfibrils surrounded by amorphous hemicellulose, which in turn is encased in a waxlike lignin sheath.<sup>9</sup> While hemicelluloses and lignin are amorphous materials, cellulose is crystalline, adding to the difficulty of processing this component. As a result, all current approaches for biomass conversion require harsh and expensive physical or chemical pretreatment methods,<sup>5,8,10,11</sup> and consequently, there is considerable incentive to find alternative pretreatment technologies that could reduce the cost of converting lignocellulosic biomass to fuels.<sup>10,12,13</sup>

Recent studies have shown that all of the components of biomass can be dissolved in ionic liquids (ILs),<sup>14–16</sup> salts that

have melting points below 373 K. Of particular interest is the fact that crystalline cellulose can be dissolved in ILs since the glucan strands of cellulose are held together by a robust interaction network that includes specific interactions consisting of classical O–H···O hydrogen bonds (HBs) and nonconventional C–H···O contacts that together help to render the material insoluble in water and most organic solvents.<sup>17,18</sup> The intermolecular O–H···O HBs, which are part of cellulose's interchain interactions, act in the glucans' equatorial directions and serve to bind neighboring sugar chains together to form flat sheets.<sup>18–28</sup> The C–H···O contacts form along the axial directions of glucan residues and contribute to the intersheet interactions holding glucan sheets together in crystalline cellulose microfibrils. In molecular dynamics (MD) simulations of cellulose microfibrils, the C–H···O contacts of intersheet interactions were observed to be highly intransigent to surface exposure<sup>18</sup> or microfibril twist,<sup>29</sup> which is opposite the behaviors of the O–H···O HB components of interchain interactions. The calculated glucan–glucan interaction energies in cellulose microfibrils

**Received:** March 14, 2011

**Revised:** September 21, 2011

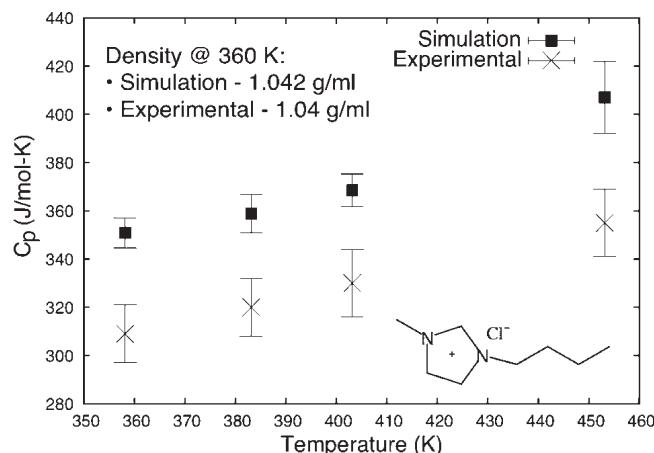
**Published:** September 28, 2011

are stronger for intersheet interactions than for interchain interactions,<sup>18</sup> and path optimization and free-energy simulations of cellulose deconstruction further affirm that intersheet interactions are the main cause of recalcitrance.<sup>30</sup> The signature of stronger intersheet than interchain interactions is also revealed in restrained dissociations of small, soluble model carbohydrates<sup>31–33</sup> which showed that association energy or free energy via axial stacking is greater than that via equatorial pairing.

While there are innumerable anion–cation pairs that might be considered as solvents for cellulose, the best ILs are those that contain a heterocyclic ring (imidazole or pyridine) substituted with alkyl side chains as the cation (e.g., 1-butyl-3-methylimidazolium) and a conjugate acid anion (e.g., chloride, acetate).<sup>14,34,35</sup>

With such ILs, cellulose solubilities up to 25 wt % have been achieved at moderate temperatures.<sup>14</sup> However, little is known about the means by which ILs promote the dissolution of cellulose. The predominant explanation is that the anions of the IL form hydrogen bonds with the hydroxyl groups of the glucan units thereby weakening the hydrogen bonding between glucan strands. This theory is supported by several experimental and computational investigations that have used water-soluble substrates as model solutes. NMR relaxation experiments carried out with solutions of glucose or cellobiose in 1-butyl-3-methylimidazolium chloride (BmimCl) show a strong stoichiometric correspondence between IL anions and glucan hydroxyl groups.<sup>36,37</sup> The formation of O–H···anion HBs has also been observed in MD simulation studies,<sup>30,38–40</sup> and the interaction energy between a glucan residue and an IL anion has been found to be stronger than that between a glucan residue and a water or methanol molecule in the liquid phase.<sup>41</sup> While a strong dependence of cellulose solubility on the hydrogen-bonding properties of anions is observed,<sup>14,35</sup> O–H···anion HBs alone are not sufficient to explain the solubility of cellulose in different ILs. The cation identity also affects the solubility. The length of the alkyl chains and the structure of the charge-bearing ring of the cation (e.g., imidazolium, pyridinium) both change cellulose solubility.<sup>14,34</sup> Since a definite set of cation-mediated modes of interaction has not yet been identified, the roles of cations are considered mostly nonspecific or indirect.<sup>36,37,39</sup> Furthermore, the strong electrostatic coupling between anions and cations indicates that their interactions with cellulose are likely collective. The multifaceted nature of IL molecules complicates the characterization of their interactions with cellulose, and an integrated approach is thus needed.

The present investigation was undertaken in order to elucidate the IL-mediated interactions leading to cellulose dissolution to establish a molecular basis for improving IL-based pretreatment strategies or for designing novel pretreatment solvents. All-atom MD simulations of cellulose were carried out for two different states, a crystalline microfibril and a state in which the glucan chains of the microfibril are fully dissociated. These two extremes were taken to represent crystalline and dissolved cellulose and are used to understand the modes by which the anions and cations of the IL interact with the glucan strands. Each state was solvated in water or BmimCl to compare and contrast how the two solvents modulate the relative energies of the microfibril and dissociated states. The solvent structures around glucan chains and on the surfaces of the microfibril were examined to study the specific interactions mediated by the cations and anions of the IL.

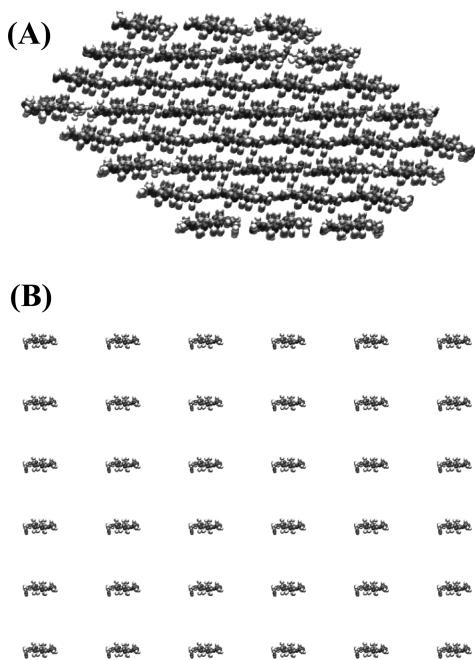


**Figure 1.** Comparison of the calculated heat capacity and density of BmimCl (shown at bottom right) from the newly developed atomistic force field with the experimental data.<sup>48,49</sup>

## METHODS

An all-atom force field of BmimCl was developed using a standardized parametrization procedure<sup>42–44</sup> for compatibility with the existing CHARMM<sup>45</sup> carbohydrate force fields,<sup>46,47</sup> which have been validated against thermodynamic and structural data for several monosaccharides<sup>46</sup> and cellulose.<sup>18</sup> In short, atomic charges and intramolecular potentials were obtained by fitting against the interaction energies and geometries obtained from ab initio calculations at the HF/6-31G\* level. The resulting force field was then validated by comparing the calculated thermodynamic properties with experimental measurements. The reported thermodynamic data of BmimCl include liquid density at 360 K and 1 atm<sup>48</sup> and heat capacities at different temperatures<sup>49</sup> all of which were used for force field development. The atomic and intramolecular parameters were adjusted iteratively to ensure accuracy of both molecular structures and thermodynamic properties. A comparison of the results obtained using our optimized force field with the available experimental data is shown in Figure 1. Although the density is predicted accurately by our BmimCl force field, the heat capacities are consistently overestimated by ~10%. Classical force fields often overestimate heat capacity compared to ab initio calculations because of the use of harmonic potentials.<sup>50</sup> Additionally, since the anion and cation retain their formal charges in our model, the lack of polarizability may lead to enhanced cage effects and energy variations causing the small overestimation of heat capacity.<sup>51,52</sup> However, in this work, MD simulation results are used only for energy comparisons between states of cellulose dissolution, which mitigates the biases of any single trajectory caused by the use of fixed-charge force fields, or for structure characterizations. Previous work has shown that the use of fixed-charge models in MD simulations is able to capture IL structures in the liquid state accurately as compared to those measured by neutron diffraction.<sup>53,54</sup> More details of the development and validation of the force field are provided in the Supporting Information (SI).

Two model states of cellulose were constructed to contrast the thermodynamic and solvation properties of crystalline and dissolved cellulose. Both states contained 36 glucan chains each with a length of 16 glucan units. The crystalline state was taken to be a cellulose I<sub>B</sub><sup>55</sup> microfibril.<sup>9,18,56</sup> The I<sub>B</sub> allomorph is the most



**Figure 2.** Cross sections of the two atomistic models of cellulose simulated in this work: the cellulose microfibril (A) and the dissociated glucan chains (B). Both contain 36 glucan chains, each 16 residues long.

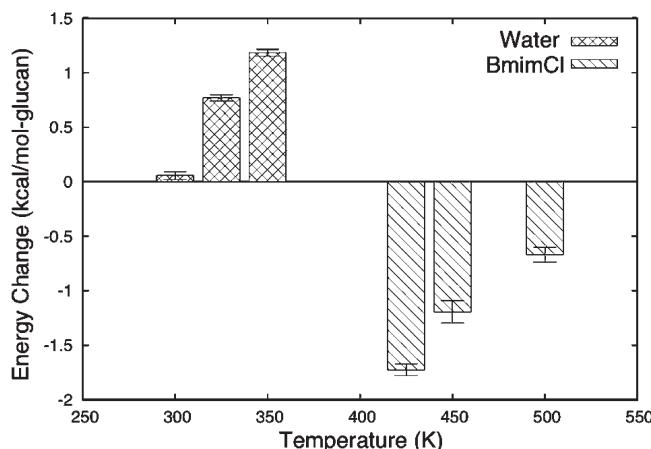
stable form of naturally occurring cellulose.<sup>57,58</sup> In the fully dissociated state, the 36 chains were arranged into a six-by-six grid with a spacing of at least four solvation shells between chains (16 Å in water and 30 Å in IL). Cross sections of the two model states are shown in Figure 2A and B. Both states were solvated in either TIP3P<sup>59</sup> water or BmimCl and have the same number of total atoms and composition for a given solvent choice. The cellulose/water simulations contained 28 443 waters, and the cellulose/IL simulations contained 12 180 BmimCl pairs. These correspond to a mass fraction of cellulose of 0.15 in water and 0.04 in BmimCl, which are well within the solubility range of the latter system.<sup>14,35</sup>

Simulations were performed with the NAMD software<sup>60</sup> at constant temperature and pressure (1 atm) conditions using the Langevin thermostat and Nose-Hoover Langevin barostat.<sup>61,62</sup> The SHAKE algorithm was used to constrain the length of all covalent bonds to hydrogen.<sup>63</sup> Cellulose/water trajectories were generated at 300, 325, and 350 K, and cellulose/IL trajectories were generated at 425, 450, and 500 K. We simulated the BmimCl systems at temperatures elevated compared to the melting point (340 K<sup>64</sup>) for accelerated dynamics. At the start of all simulations, harmonic restraint potentials were placed on all C1 sugar carbons (force constant 5 kcal/mol/Å<sup>2</sup> for water systems and 10 kcal/mol/Å<sup>2</sup> for IL systems) to keep the glucans around their initial positions. After 15 000 steps of energy minimization, the systems were heated gradually to their target temperatures and were equilibrated for at least 5 ns for the cellulose/water systems and 15 ns for the cellulose/IL systems. After equilibration, the harmonic restraints were removed except for those on the terminal glucans of each chain to retain the system in the microfibril or fully dissociated state. Since cellobextrins in water exist in extended chain conformations,<sup>32,65,66</sup> the effects of the remaining restraints on biasing the conformations of the individual glucan chains are likely insignificant.

A comparison of the glycosidic linkage torsion angles sampled in our trajectories with those from the simulation of free-short-chain cellooligomers in water<sup>65</sup> shows a similar distribution (see the SI for further details). We also conducted additional simulations in which all harmonic restraints were removed after equilibration, that is, free simulations, to examine chain collapse behaviors in water and in BmimCl. For all simulations, a time step of 2 fs was employed, and the particle mesh Ewald (PME) sum was used to calculate long-range electrostatic interactions. Data was taken from production runs of at least 15 ns with frames saved every 1 ps for analysis from the cellulose/water trajectories and from runs of at least 30 ns with frames saved every 2 ps from the cellulose/IL trajectories. The simulation lengths of both system types were examined to ensure the convergence of solvent structures by comparing the solvent density profiles calculated from subportions of the trajectories (see the SI for further details).

Three-dimensional plots for the number density of solvent molecules around dissociated glucan chains were generated in the following manner. First, a local Cartesian coordinate system was created for each glucan in a trajectory frame. For a given glucan, the vector that connects the two O4 ether oxygens that bound it defines the z-axis. A pseudo-x-axis was defined as the average directions of the C2—O2 and C3—O3 bonds. The cross product of the z- and pseudo-x-axes then defines the y-axis, and the cross product of the y- and z-axes defines the x-axis. This methodology assigns the x-axis along the equatorial direction and the y-axis along the axial direction of the selected glucan. The origin of the coordinate system was chosen to be the center of mass of the glucan ring. A schematic of the process for creating the glucan-centered coordinate system is shown in the SI (Figure S10A). After creating a local Cartesian coordinate system for a glucan, the space around the glucan was discretized into small cubic volumes of side length 0.25 Å, and the occupancy of each grid point was calculated. This process was repeated and was averaged over all glucans and frames in a trajectory to determine the three-dimensional number density distributions, which are normalized by the density of pure solvent at the same temperature and pressure. To determine locations of the non-monatomic solvent molecules, the center of mass of the whole molecule was used for water, while the center of mass of the imidazolium ring, the solitary methyl group, and the first methyl group off the imidazolium ring in the butyl group was used to represent the charge-carrying portion of the Bmim<sup>+</sup> cation, which we name the ring site. The center of mass of the remaining atoms of the Bmim<sup>+</sup> cation is then the tail portion. The distribution of Bmim<sup>+</sup> tails is presented in the SI (Figure S13).

To calculate the solvent density profile as a function of cylindrical radius around a glucan chain, that is, the cylindrical radial distribution function (RDF), the center of mass of each glucan ring was determined. The vector connecting the centers of masses of sequential glucans is defined as the z-axis of the cylinder associated with a pair of neighboring sugar units. The length of the inter-center-of-mass vector defines the length of the cylinder within which the density profile along the radial direction was calculated. The distance from each solvent molecule to the local z-vector was then calculated, binned, and averaged over all neighboring glucan—glucan pairs and frames to determine the cylindrical RDF. A schematic representation of the construction of a cylindrical coordinate system for a glucan pair is shown in the SI (Figure S10B).

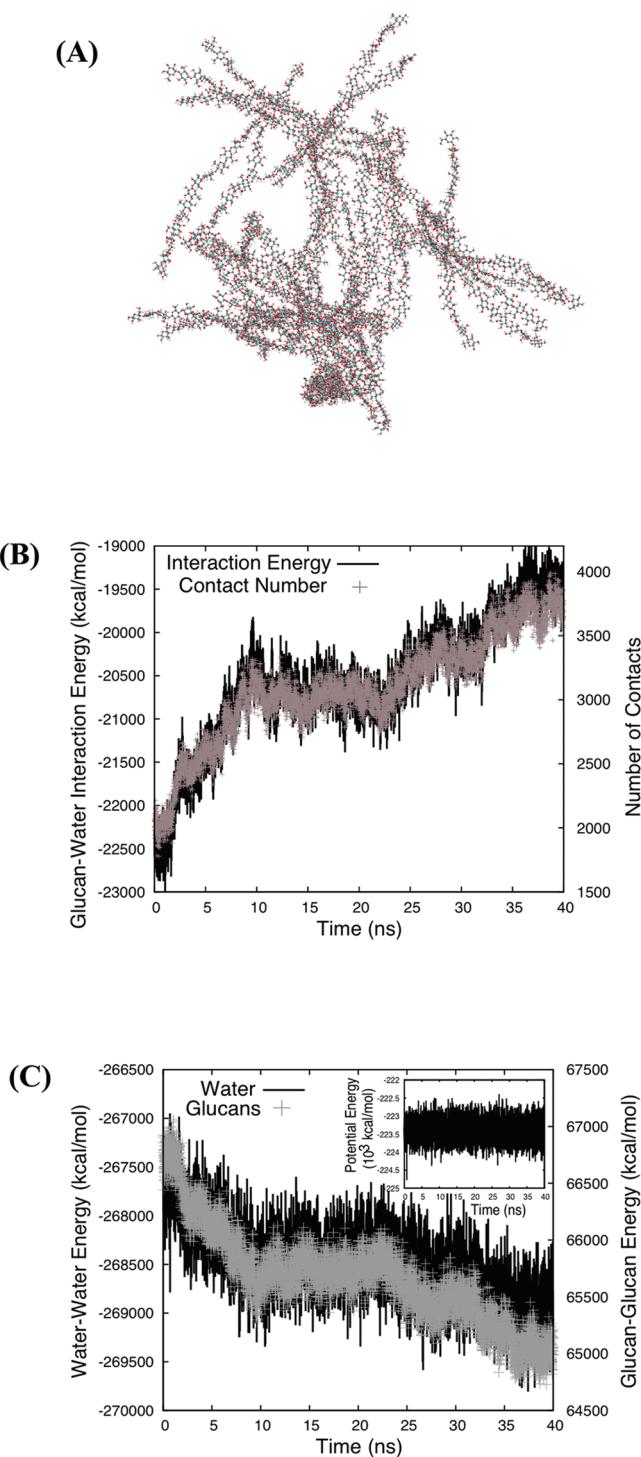


**Figure 3.** Difference in energy between the dissociated and the microfibril states of cellulose (dissociated–microfibril) in both water and BmimCl, normalized per glucan unit, from the all-atom MD simulations.

## RESULTS AND DISCUSSION

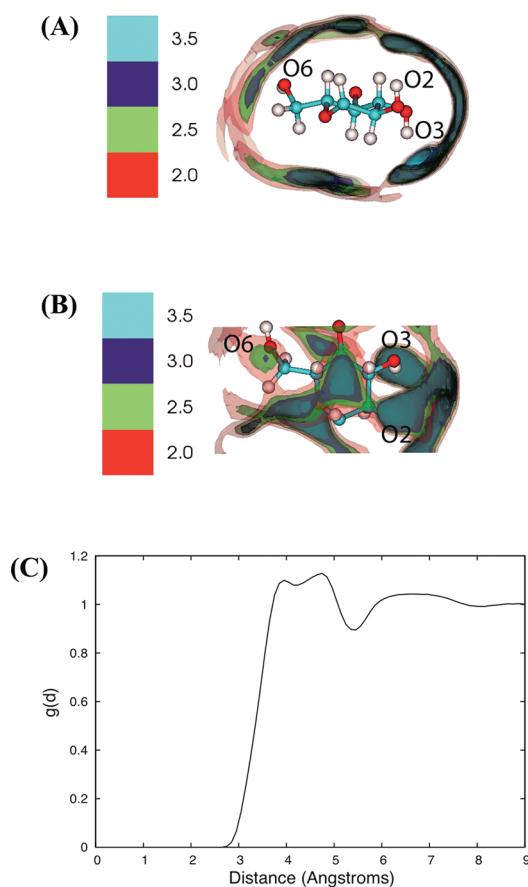
**Insolubility of Cellulose in Water.** Figure 3 shows the differences in the average total potential energies between the dissociated and the microfibril states of cellulose in water and BmimCl at several temperatures normalized by the number of glucans. In water, the two states are nearly isoenergetic at 300 K. At higher temperatures, the dissociated state has increasingly higher potential energies in water as the heat capacity of the more disordered state is larger.<sup>67</sup> The calculated energy differences from simulating these polymeric materials agree semiquantitatively with the measured solvation enthalpies of monosaccharides, ~1 kcal/mol, over the investigated temperature range.<sup>68,69</sup> The increase with temperature of the potential energy of the dissociated state relative to the microfibril state is also in line with the temperature dependence of the solvation enthalpy of glucose.<sup>68</sup> Higher potential energies of the dissociated state at elevated temperatures represent part of the thermodynamic penalty of dissolving cellulose in water.

Although the microfibril and the dissociated states of cellulose are isoenergetic at 300 K in water, collapse of glucan chains was observed in the free simulation after removing the restraint potentials at the chain ends used for keeping the glucan chains in a dissociated state. The amorphous aggregate at the end of the free simulation is shown in Figure 4A. The number of glucan–glucan contacts, defined as the number of glucan atoms within 3 Å of any other glucan atom in any other chain, increased by a factor of 2 over the length of the free simulation (Figure 4B). Analyzing the different components of interaction energies in the system during chain collapse revealed that the loss of glucan–water contacts increased the glucan–water interaction energy (Figure 4B) but was compensated by enhanced water–water and glucan–glucan interactions (Figure 4C). As a result, the total potential energy does not show any significant drift during the free simulation (inset, Figure 4C). The magnitude of the change in energy of the system, ~40 kcal/mol, as calculated from the differences in the potential energies of the first and last 10 ns of the trajectory, is small compared to the size of a typical potential energy fluctuation, 270 kcal/mol (the running average of the potential energy of the system is included in the SI, Figure S15). Thus, chain collapse is considered nearly isoenergetic at 300 K. This result is in line with the data in Figure 3, which show that the



**Figure 4.** Results of the collapse of glucan chains in water at 300 K during the free simulation. (A) Structure of the collapsed aggregate of glucan chains at the end of the simulation. (B) The temporal evolution of the total glucan–water interaction energy and the number of glucan–glucan contacts within 3 Å during the free simulation in water. Only the contacts between glucan atoms of different chains are included. (C) The temporal evolution of the total water–water and glucan–glucan interaction energies during the free simulation. The inset shows the temporal evolution of the total potential energy of the entire system.

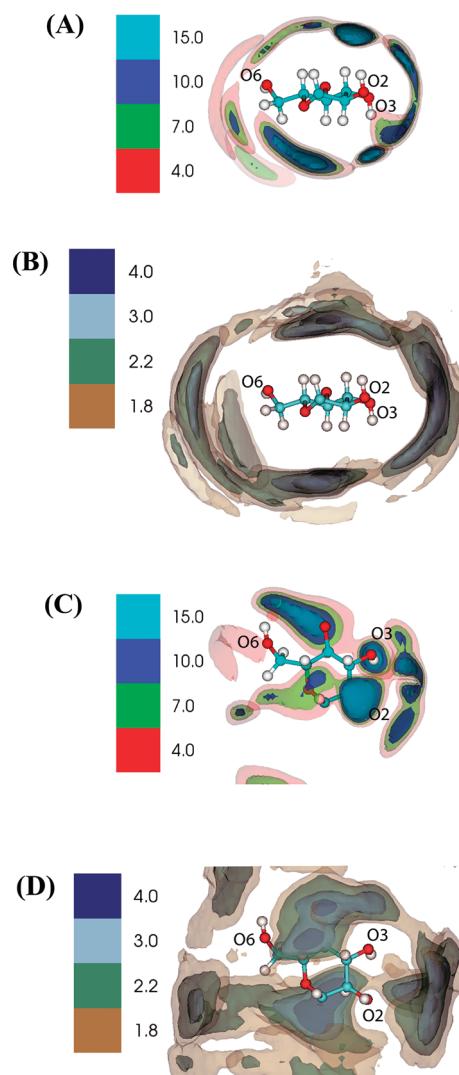
microfibril and the dissociated states of cellulose are also nearly isoenergetic at 300 K. Therefore, the chain collapse in water near



**Figure 5.** The calculated profile of the temporally and spatially averaged number density of water around a glucan chain normalized by the bulk value viewed from along the chain axis (A) and above the plane of the glucan ring (B) at 300 K and 1 atm. (C) The cylindrical radial distribution function of water, by center of mass, around the dissociated chains.

room temperature observed in our atomistic MD simulations is considered entropy driven, and solvent structures near different states and conformations of cellulose are expected to play an essential role in the insolubility of cellulose in water.<sup>18</sup> A similar free simulation performed at 300 K with a smaller system, 18 10-mer glucan chains in water, showed identical results, that is, a collapse of the chains with no drift in total potential energy (see the SI for further information). To analyze how glucan chains affect solvent structures, we discuss next the number density of water molecules around the dissociated glucan chains and on the surfaces of a microfibril.

In Figure 5A and B, the three-dimensional distribution of water density around the dissolved glucan chains is plotted as a contour map. Regions of high water density can be observed around the interchain O—H $\cdots$ O HB forming OH groups on both the equatorial and the axial sides of the glucan ring.<sup>18,70</sup> In addition to regions near the HB donors and acceptors of the glucan units, however, high water density was also observed above and below the planes of the glucan rings near the CH groups protruding outward in the axial direction (Figure 5A, B). Inspection of water structures in this region reveals that the lack of hydrogen-bonding ability with the aliphatic CH groups gives rise to a denser water–water interaction network nearby, which is a signature of small-scale hydrophobic effects.<sup>71</sup> The



**Figure 6.** The calculated profiles of the temporally and spatially averaged number densities of the centers of masses of anions and cation rings around a glucan chain normalized by bulk values at 450 K and 1 atm. The density profiles viewed from along the chain axis: Cl<sup>−</sup> (A) and Bmim<sup>+</sup> (B). The density profiles viewed from the top of the glucan ring: Cl<sup>−</sup> (C) and Bmim<sup>+</sup> (D).

amphiphilicity of glucan chains thus results in a peculiar ordering of water around the dissociated cellulose. We also calculated the average density profiles of water as a function of the cylindrical radius around a dissociated chain, and the results are shown in Figure 5C. The two peaks in the first shell at 3.8 and 4.8 Å correspond to the high water density regions in the axial and equatorial directions of a glucan chain, respectively.

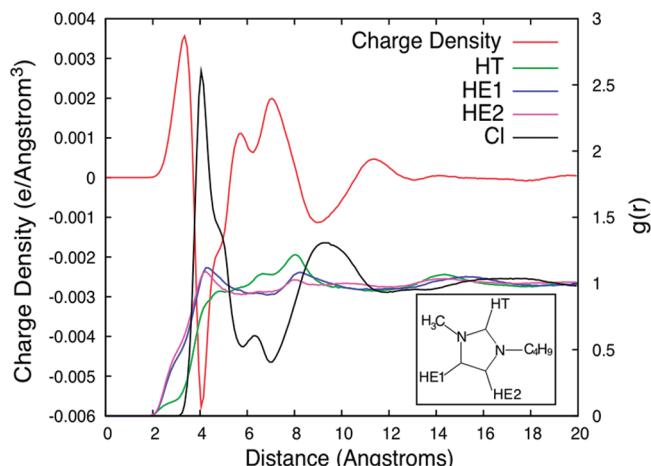
The amphiphilicity of glucan units and the distribution pattern of exposed groups on the surfaces of a cellulose microfibril give rise to distinctive water structures at different solid–liquid interfaces. On the hydrophobic surfaces that expose CH groups, significant water ordering was observed, whereas on the hydrophilic surfaces of a microfibril, which expose mostly hydroxyl groups, solvent ordering was much less prominent.<sup>18</sup> Since the water-ordering CH groups in the dissolved state of cellulose are solvent accessible, chain collapse occurred spontaneously in the free simulation to reduce their exposure. This entropy-based mechanism is supported by the observation that the total potential

energy does not drift during chain collapse (Figure 3). Our results indicate that the amphiphilicity of glucan chains and their tendency to restrict available water network structures are dominant causes of the insolubility of cellulose in water.

**Dissolution of Cellulose in BmimCl.** Figure 3 shows that the potential energy of the dissociated state of cellulose in BmimCl is lower than that of the microfibril state, indicating more favorable interactions of BmimCl molecules with dissociated glucan chains than with the microfibril. Increasing temperature reduces the difference in potential energy since the more disordered state has a higher heat capacity, but for all investigated temperatures, the dissociated state has a lower potential energy in BmimCl indicating the enthalpic driving force for cellulose dissolution in the IL. In the free simulation in BmimCl at 450 K, the dissociated glucan chains showed no sign of collapse and remained separated over a 45 ns MD simulation. These results are consistent with the observation that BmimCl can dissolve cellulose at the concentration (4 wt %) used in the present simulations.

To analyze the interactions of BmimCl molecules with glucan chains in the dissociated state, we calculated the three-dimensional density distribution of  $\text{Cl}^-$  anions and  $\text{Bmim}^+$  cations around the dissolved glucan chains. The average  $\text{Cl}^-$  density around a dissolved glucan unit, normalized by the bulk value, is shown in Figure 6A, C. As expected from NMR measurements and MD simulations of glucose or cellobiose molecules dissolved in IL,<sup>36–39</sup>  $\text{Cl}^-$  anions form HBs with the OH groups of dissolved glucan chains, and regions of high  $\text{Cl}^-$  density are evident around the hydroxyl groups. Well-defined regions of high  $\text{Cl}^-$  density can be observed around the O2–H and O3–H groups as well as in the space between them (Figure 6A, C). Near the C6 hydroxymethyl group, which samples different rotameric states in solution, discrete regions of high  $\text{Cl}^-$  density can also be found about the circumference of the C6–O6–H rotation. We found that  $\text{Cl}^-$  anions can form individual O–H $\cdots$  $\text{Cl}^-$  HBs with each hydroxyl group in a glucan unit as well as form shared HBs between neighboring OHs in the same residue or across two adjacent residues along a glucan chain. Calculation of the coordination number of  $\text{Cl}^-$  up to the first minimum in the sugar anion radial distribution function gives two anions per glucan and an average anion to OH ratio of 2:3. This value is similar to the reported 4:5 anion to OH ratio from MD simulations of glucose in 1,3-dimethylimidazolium chloride<sup>39</sup> and the 1:1 ratio derived from NMR measurements of cellobiose solvated in BmimCl<sup>36</sup> in that it shows a strong stoichiometric association between the sugar OH groups and the  $\text{Cl}^-$  anion. The anion to OH ratio is lower here than for those with glucose or cellobiose because of the use of polymeric solutes in our simulations (more discussion is provided in the SI).

While the density distribution of  $\text{Cl}^-$  anions establishes their formation of HBs with the equatorial hydroxyl groups of glucose residues, the density distribution of the centers of masses of  $\text{Bmim}^+$  cation rings shown in Figure 6B, D reveals closer contacts to the axial moieties of glucan chains than to the equatorial ones. In Figure 6D, large volumes of high-density regions of  $\text{Bmim}^+$  rings are observed over the glucan rings of the chains. Next to the axial moieties of glucose residues, the high-density regions of  $\text{Cl}^-$  and  $\text{Bmim}^+$  rings have similar distances to the glucans, whereas near the equatorial groups, the high-density zones of the cations are further away from the sugars than those of the anions (more details are provided in the SI). Therefore, both  $\text{Cl}^-$  and  $\text{Bmim}^+$  rings interact with glucan chains along axial directions, but close



**Figure 7.** The calculated cylindrical radial distribution profile of charge density of BmimCl on the basis of the partial charges of the IL atoms and the cylindrical radial density profiles of  $\text{Cl}^-$  and three electrophilic protons of the imidazolium ring, around the dissociated glucan chains, at 450 K and 1 atm. The inset shows the labeling of the three hydrogens in the imidazolium ring.

equatorial contacts are dominated by  $\text{Cl}^-$  anions. In the MD simulations of the dissociated and microfibril states, the observed signature of ring–ring stacking between the  $\text{Bmim}^+$  cations and the sugar rings in dissociated glucan chains or on cellulose microfibril surfaces are weak, in agreement with earlier simulations of dissolved glucose molecules in BmimCl<sup>39</sup> (see Supporting Information for further discussion). It is also clear from Figure 6 that the  $\text{Bmim}^+$  and  $\text{Cl}^-$  densities are arranged in a highly complementary manner, with each ion occupying specific regions near a glucan chain while maintaining local charge neutrality. This result, combined with the lower potential energy of the dissociated state of cellulose compared to the microfibril state in BmimCl, indicates that the specific interactions between BmimCl and the amphiphilic glucan chains are compatible with the electrostatic coupling between cations and anions.

The liquid-state structures of BmimCl around a glucan chain can be further illustrated by cylindrical density profiles. In Figure 7, we plot the cylindrical density profiles of the  $\text{Cl}^-$  anion and the three electrophilic hydrogen atoms of the  $\text{Bmim}^+$  ring. It is clear that the density distribution around a glucan chain of BmimCl atoms with opposite partial charges oscillates and plateaus after around  $\sim 14 \text{ \AA}$  from the chain. Converting the structural distributions of BmimCl atoms around a glucan chain into a charge distribution based on their partial charges clearly shows that electrostatic coupling strongly influences the liquid-state structures of BmimCl. The first layer of high positive charge density around a glucan chain in Figure 7 results from the three electrophilic hydrogen atoms on the  $\text{Bmim}^+$  ring. This result is also in line with the observation shown in Figure 6B and D that the centers of mass of the  $\text{Bmim}^+$  rings are positioned near the ether oxygen atoms, which carry a negative partial charge. The second layer of high charge density around a glucan chain is negative and comes from  $\text{Cl}^-$  anions as expected. Further away from the glucan chains, alternative layers of positive and negative charge can be observed and are coherent with the density profile of  $\text{Cl}^-$ .

In both water and BmimCl, the molecular structures of the amphiphilic glucan chains modulate the nearby solvent structures.

The structure of the hydrogen-bonding network of water molecules depends on their proximity to the polar hydroxyl groups or the aliphatic CH groups of the glucan chains. In BmimCl, anions form HBs with hydroxyl groups whereas the Bmim<sup>+</sup> rings of the cations have closer contacts with the ether oxygen atoms and the CH groups along the axial directions of glucose residues than with groups in equatorial positions. The tail portions of the cations have a similar pattern of contacts with glucan chains as the Bmim<sup>+</sup> rings. The electrostatic coupling of BmimCl also leads to oscillatory charge density profiles around the glucan chains. Such spatial distributions of solvent molecules are expected to affect the diffusion of catalyst molecules to the ether linkages as well as the complexation of cellulase enzymes with glucan chains. Therefore, the reactivity and stability of molecular catalysts or enzymes with glucan strands can be expected to depend on the solvation structures and charge distributions around the dissolved chains. Thus, in addition to the dissolution of cellulose, the amphiphilicity of glucan chains and the surrounding arrangement of solvation structures also have significant implications on the hydrolysis of the dissolved polymer chains.

## CONCLUSION

We have performed all-atom MD simulations with the aim of understanding the molecular origin of the insolubility of cellulose in water and its solubility in BmimCl. Two extreme states of cellulose were considered: a crystalline microfibril state and a dissociated state in which the glucan chains are fully detached from each other (Figure 2). Comparing the average potential energies of the two states reveals a thermodynamic driving force for the dissolution process: the dissociated state has an equal or higher potential energy in water and a lower potential energy in BmimCl (Figure 3). Furthermore, starting from the dissociated state, chain collapse occurs spontaneously in water at 300 K during which the potential energy does not drift away from its initial value. This result, combined with the analysis of water structures near the glucan chains (Figure 5), indicates that the insolubility of cellulose in water originates mostly from reduction in solvent entropy.

The Cl<sup>-</sup> anions of BmimCl form strong HBs with the hydroxyl groups of the glucan chains of cellulose in both the equatorial and the axial orientations, which would help to replace the interchain as well as the intersheet interactions of cellulose lost upon dissolution. The Bmim<sup>+</sup> cation rings also interact with glucan chains, with the contact distances to glucose residues along axial directions smaller than those along equatorial directions. Therefore, cation–glucan interactions also appear to be an important compensator for lost cellulose intersheet interactions. Furthermore, a clear complementarity in the density profiles of Bmim<sup>+</sup> and Cl<sup>-</sup> is observed around glucan chains. This result and the observation that the dissociated state of cellulose has a lower potential energy than the microfibril state in BmimCl indicate that the electrostatic coupling between Bmim<sup>+</sup> and Cl<sup>-</sup> is compatible with the presence of aliphatic glucan chains.

Together, the results presented in this work highlight two mechanistic insights for the dissolution of cellulose by solvent-mediated interactions. First, in water, the perturbation of solvent structures by dissolved glucan chains is a dominant factor in limiting solubility by leading to the entropic penalty of the free-energy cost of dissolution. Second, in BmimCl, in addition to interacting with the equatorial moieties of glucose residues that form interchain interactions in cellulose, both the Cl<sup>-</sup> and the

Bmim<sup>+</sup> ions were found to interact with glucose residues along their axial directions along which the intersheet interactions in cellulose are formed. Cl<sup>-</sup> anions can form HBs with hydroxyl groups from the axial directions of sugar rings while Bmim<sup>+</sup> cations make contacts with the donors and acceptors of intersheet C–H···O HBs. Solvent molecules contacting glucose moieties along axial directions is important since intersheet interactions are the most intransigent component in the interaction network of cellulose.<sup>18</sup> Therefore, we propose that the formation of solvation environments that are effective in interacting with glucan chains along both equatorial and axial directions, as in BmimCl, is likely an indicative signature of an effective pretreatment solvent.

## ASSOCIATED CONTENT

**S Supporting Information.** In depth discussion of the BmimCl CHARMM force field development and force field parameters, Cl<sup>-</sup> coordination around and hydrogen bonding to the dissociated glucan chains, Bmim<sup>+</sup> ring stacking on the glucan rings of the dissociated chains and microfibril surfaces, the effects of end restraints on glucan chain conformation, schematics of the local coordinate systems for both the three-dimensional density plots and the cylindrical RDFs, collapse behavior of the 18 10-mer glucan chain system in water at room temperature, convergence of solvent structures, further information on the three-dimensional BmimCl solvent structures, and running average of the potential energy of the unconstrained dissociated cellulose/water system during collapse are included. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: jwchu@berkeley.edu. Phone: (510) 642-4398. Fax: (510) 642-4778.

## ACKNOWLEDGMENT

We thank Prabhat from the NERSC (National Energy Research Scientific Computing Center) Visualization Group for his help to produce several of the manuscript's graphics. This project was supported by the Energy Biosciences Institute (grant numbers OO7G03 and OO0J04) and the University of California, Berkeley. We also thank the computational resources provided by NERSC, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

## REFERENCES

- (1) Schubert, C. *Nat. Biotechnol.* **2006**, *24*, 777–784.
- (2) Huber, G. W.; Iborra, S.; Corma, A. *Chem. Rev.* **2006**, *106*, 4044–4098.
- (3) Farrell, A. E.; Plevin, R. J.; Turner, B. T.; Jones, A. D.; O'Hare, M.; Kammen, D. M. *Science* **2006**, *311*, 506–508.
- (4) Wyman, C. E. *Appl. Biochem. Biotechnol.* **2001**, *91*–*3*, 5–21.
- (5) Ragauskas, A. J.; Williams, C. K.; Davison, B. H.; Britovsek, G.; Cairney, J.; Eckert, C. A.; Frederick, W. J.; Hallett, J. P.; Leak, D. J.; Liotta, C. L.; et al. *Science* **2006**, *311*, 484–489.
- (6) Demirbas, A. *Energy Convers. Manage.* **2008**, *49*, 2106–2116.
- (7) Perlack, R. D. *Biomass as Feedstock for a Bioenergy and Bio-products Industry: The Technical Feasibility of a Billion-Ton Annual Supply*;

- DOE/GO-102995-2135; Oak Ridge National Laboratory: Oak Ridge, TN, 2005.
- (8) Lynd, L. R.; Wyman, C. E.; Gerngross, T. U. *Biotechnol. Prog.* **1999**, *15*, 777–793.
  - (9) Perez, S.; Mazeau, K. In *Polysaccharides: Structural Diversity and Functional Versatility*, 2nd ed.; Dumitriu, S., Ed.; Marcel Dekker: New York, 2005; pp 41–68.
  - (10) Yang, B.; Wyman, C. E. *Biofuels, Bioprod. Biorefin.* **2008**, *2*, 26–40.
  - (11) Dale, B. J. *Agric. Food. Chem.* **2008**, *56*, 3885–3891.
  - (12) Simmons, B. A.; Loque, D.; Blanch, H. W. *Genome Biol.* **2008**, *9*, 242.
  - (13) Wyman, C. E. *Trends Biotechnol.* **2007**, *25*, 153–157.
  - (14) Swatloski, R. P.; Spear, S. K.; Holbrey, J. D.; Rogers, R. D. *J. Am. Chem. Soc.* **2002**, *124*, 4974–4975.
  - (15) Moulthrop, J. S.; Swatloski, R. P.; Moyna, G.; Rogers, R. D. *Chem. Commun.* **2005**, 1557–1559.
  - (16) Singh, S.; Simmons, B. A.; Vogel, K. P. *Biotechnol. Bioeng.* **2009**, *104*, 68–75.
  - (17) 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.
  - (18) Gross, A. S.; Chu, J. W. *J. Phys. Chem. B* **2010**, *114*, 13333–13341.
  - (19) Kondo, T. In *Polysaccharides: Structural Diversity and Functional Versatility*, 2nd ed.; Dumitriu, S., Ed.; Marcel Dekker: New York, 2005; pp 69–98.
  - (20) Marechal, Y.; Chanzy, H. *J. Mol. Struct.* **2000**, *523*, 183–196.
  - (21) Watanabe, A.; Morita, S.; Ozaki, Y. *Biomacromolecules* **2006**, *7*, 3164–3170.
  - (22) Watanabe, A.; Morita, S.; Ozaki, Y. *Biomacromolecules* **2007**, *8*, 2969–2975.
  - (23) Kondo, T. *J. Polym. Sci., Part B: Polym. Phys.* **1997**, *35*, 717–723.
  - (24) Heiner, A. P.; Sugiyama, J.; Teleman, O. *Carbohydr. Res.* **1995**, *273*, 207–223.
  - (25) Tashiro, K.; Kobayashi, M. *Polymer* **1991**, *32*, 1516–1530.
  - (26) Bergenstrahle, M.; Berglund, L. A.; Mazeau, K. *J. Phys. Chem. B* **2007**, *111*, 9138–9145.
  - (27) Shen, T. Y.; Gnanakaran, S. *Biophys. J.* **2009**, *96*, 3032–3040.
  - (28) Mazeau, K.; Heux, L. *J. Phys. Chem. B* **2003**, *107*, 2394–2403.
  - (29) Paavilainen, S.; Rog, T.; Vattulainen, I. *J. Phys. Chem. B* **2011**, *115*, 3747–3755.
  - (30) Cho, H. M.; Gross, A. S.; Chu, J.-W. *J. Am. Chem. Soc.* **2011**, *133*, 14033–14041.
  - (31) Cousins, S. K.; Brown, R. M. *Polymer* **1995**, *36*, 3885–3888.
  - (32) Bergenstrahle, M.; Wohlert, J.; Himmel, M. E.; Brady, J. W. *Carbohydr. Res.* **2010**, *345*, 2060–2066.
  - (33) Payne, C. M.; Himmel, M. E.; Crowley, M. F.; Beckham, G. T. *J. Phys. Chem. Lett.* **2011**, *2*, 1546–1550.
  - (34) El Seoud, O. A.; Koschella, A.; Fidale, L. C.; Dorn, S.; Heinze, T. *Biomacromolecules* **2007**, *8*, 2629–2647.
  - (35) Pinkert, A.; Marsh, K. N.; Pang, S. S.; Staiger, M. P. *Chem. Rev.* **2009**, *109*, 6712–6728.
  - (36) Remsing, R. C.; Swatloski, R. P.; Rogers, R. D.; Moyna, G. *Chem. Commun.* **2006**, 1271–1273.
  - (37) Remsing, R. C.; Hernandez, G.; Swatloski, R. P.; Massefski, W. W.; Rogers, R. D.; Moyna, G. *J. Phys. Chem. B* **2008**, *112*, 11071–11078.
  - (38) Youngs, T. G. A.; Holbrey, J. D.; Deetlefs, M.; Nieuwenhuyzen, M.; Gomes, M. F. C.; Hardacre, C. *ChemPhysChem* **2006**, *7*, 2279–2281.
  - (39) Youngs, T. G. A.; Hardacre, C.; Holbrey, J. D. *J. Phys. Chem. B* **2007**, *111*, 13765–13774.
  - (40) Liu, H. B.; Singh, S.; Sale, K. L.; Simmons, B. A. *J. Phys. Chem. B* **2011**, *115*, 10251–10258.
  - (41) Liu, H. B.; Sale, K. L.; Holmes, B. M.; Simmons, B. A.; Singh, S. *J. Phys. Chem. B* **2010**, *114*, 4293–4301.
  - (42) Mackerell, A. D.; Wiorkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946–11975.
  - (43) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
  - (44) Foloppe, N.; MacKerell, A. D. *J. Comput. Chem.* **2000**, *21*, 86–104.
  - (45) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
  - (46) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D. *J. Comput. Chem.* **2008**, *29*, 2543–2564.
  - (47) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. *J. Chem. Theory Comput.* **2009**, *5*, 2353–2370.
  - (48) Pomelli, C. S.; Chiappe, C.; Vidis, A.; Laurenczy, G.; Dyson, P. J. *J. Phys. Chem. B* **2007**, *111*, 13014–13019.
  - (49) Holbrey, J. D.; Reichert, W. M.; Reddy, R. G.; Rogers, R. D. *ACS Symp. Ser.* **2003**, *856*, 121–133.
  - (50) Cadena, C.; Zhao, Q.; Snurr, R. Q.; Maginn, E. J. *J. Phys. Chem. B* **2006**, *110*, 2821–2832.
  - (51) Yan, T. Y.; Burnham, C. J.; Del Popolo, M. G.; Voth, G. A. *J. Phys. Chem. B* **2004**, *108*, 11877–11881.
  - (52) Habasaki, J.; Ngai, K. L. *J. Chem. Phys.* **2008**, *129*, 194501.
  - (53) Youngs, T. G. A.; Holbrey, J. D.; Mullan, C. L.; Norman, S. E.; Lagunas, M. C.; D'Agostino, C.; Mantle, M. D.; Gladden, L. F.; Bowron, D. T.; Hardacre, C. *Chem. Sci.* **2011**, *2*, 1594–1605.
  - (54) Bowron, D. T.; D'Agostino, C.; Gladden, L. F.; Hardacre, C.; Holbrey, J. D.; Lagunas, M. C.; McGregor, J.; Mantle, M. D.; Mullan, C. L.; Youngs, T. G. A. *J. Phys. Chem. B* **2010**, *114*, 7760–7768.
  - (55) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074–9082.
  - (56) Ding, S. Y.; Himmel, M. E. *J. Agric. Food Chem.* **2006**, *54*, 597–606.
  - (57) Horii, F.; Yamamoto, H.; Kitamaru, R.; Tanahashi, M.; Higuchi, T. *Macromolecules* **1987**, *20*, 2946–2949.
  - (58) 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.
  - (59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
  - (60) 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*, 1781–1802.
  - (61) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177–4189.
  - (62) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613–4621.
  - (63) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: New York, 1987.
  - (64) Wilkes, J. S.; Levisky, J. A.; Wilson, R. A.; Hussey, C. L. *Inorg. Chem.* **1982**, *21*, 1263–1264.
  - (65) Shen, T. Y.; Langan, P.; French, A. D.; Johnson, G. P.; Gnanakaran, S. *J. Am. Chem. Soc.* **2009**, *131*, 14786–14794.
  - (66) Umemura, M.; Yuguchi, Y.; Hirotsu, T. *J. Phys. Chem. A* **2004**, *108*, 7063–7070.
  - (67) De Podesta, M. *Understanding the Properties of Matter*, 2nd ed.; Taylor & Francis: London, 2001.
  - (68) Schwarz, F. P. *J. Solution Chem.* **1996**, *25*, 471–484.
  - (69) Jasra, R. V.; Ahluwalia, J. C. *J. Solution Chem.* **1982**, *11*, 325–338.
  - (70) Vietor, R. J.; Newman, R. H.; Ha, M. A.; Apperley, D. C.; Jarvis, M. C. *Plant J.* **2002**, *30*, 721–731.
  - (71) Chandler, D. *Nature* **2005**, *437*, 640–647.