Local Phase Separation of Co-solvents Enhances Pretreatment of Biomass for Bioenergy Applications

Barmak Mostofian,[†,‡,§] Charles M. Cai,[§,∥,⊥] Nicholas Dean Smith,[†,§,#] Loukas Petridis,[†,§,#] Xiaolin Cheng,[†,§,#] Charles E. Wyman,[§,∥,⊥] and Jeremy C. Smith*,[†,§,#]

[†]UT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, United States

[‡]Joint Institute for Biological Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, United States

[§]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, United States

[∥]Center for Environmental Research and Technology (CE-CERT), Bourns College of Engineering, University of California, Riverside, California 92507, United States

[⊥]Department of Chemical and Environmental Engineering, Bourns College of Engineering, University of California, Riverside, California 92521, United States

[#]Department of Biochemistry, Cellular & Molecular Biology, The University of Tennessee, Knoxville, Tennessee 37996, United States

S *Supporting Information*

ABSTRACT: Pretreatment facilitates more complete deconstruction of plant biomass to enable more economic production of lignocellulosic biofuels and byproducts. Various co-solvent pretreatments have demonstrated advantages relative to aqueous-only methods by enhancing lignin removal to allow unfettered access to cellulose. However, there is a limited mechanistic understanding of the interactions between the co-solvents and cellulose that impedes further improvement of such pretreatment methods. Recently, tetrahydrofuran (THF) has been identified as a highly effective co-solvent for the pretreatment and fractionation of biomass. To elucidate the mechanism of the THF−water interactions with cellulose, we pair simulation and experimental data demonstrating that enhanced solubilization of cellulose can be achieved by the THF−water co-solvent system at equivolume mixtures and moderate temperatures ($\leq$445 K). The simulations show that THF and water spontaneously phase separate on the local surface of a cellulose fiber, owing to hydrogen bonding of water molecules with the hydrophilic cellulose faces and stacking of THF molecules on the hydrophobic faces. Furthermore, a single fully solvated cellulose chain is shown to be preferentially bound by water molecules in the THF−water mixture. In light of these findings, co-solvent reactions were performed on microcrystalline cellulose and maple wood to show that THF significantly enhanced cellulose deconstruction and lignocellulose solubilization at simulation conditions, enabling a highly versatile and efficient biomass pretreatment and fractionation method.
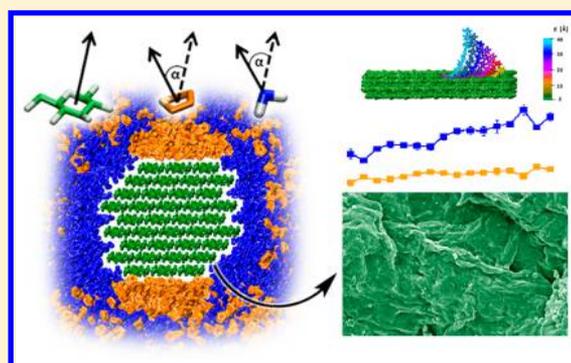
INTRODUCTION

Cellulose, the most abundant biopolymer on Earth, is a major structural component of plant cell walls, in which it is found in both crystalline and amorphous microfibrils and interacts with other organic plant biopolymers, such as lignin.[1,2] With depleting fossil fuels and increasing environmental awareness, lignocellulosic biomass is being recognized as sufficiently abundant and inexpensive to sustainably replace a sizable portion of petroleum resources for the production of fuels and chemicals.[3−6] Moreover, the use of cellulosic biofuels also has the potential to significantly reduce greenhouse gas emissions compared to conventional gasoline.[7] Additionally, unlike corn ethanol, cellulosic ethanol is derived from agricultural and forestry residue crops that do not compete with food resources.[8]

The economic production of cellulosic biofuels has historically necessitated a pretreatment step to overcome the natural recalcitrance of lignocellulosic biomass to its deconstruction. Biomass recalcitrance to deconstruction arises mainly from the tight fibrillar packing of cellulose chains and the heterogeneous matrix of lignin and hemicellulose surrounding the cellulose that limits its accessibility to enzymatic or catalytic breakdown to sugars.[9,10] Pretreatment methods, typically, are divided into three categories: physical, chemical, and biological.[11,12] Physical pretreatments apply mechanical force and/or heat to open up the biomass structure and reduce its particle size, such as by grinding, milling or fiber steam-explosion.[13−16] Chemical pretreatments apply processing

with ionic liquids or dilute acids, bases, or other reagents to dissociate the biomass fractions and alter its composition.[17−19] Finally, biological pretreatment refers to the use of microbes or fungal enzymes with cellulolytic and ligninolytic activity to digest plant biomass.[20,21]

Chemical pretreatment methods have received considerable attention in recent years.[22] Aqueous pretreatment of lignocellulose with alkali or ammonia results in structural changes of cellulose and in the degradation and partial removal of other cell wall components through saponification.[23−25] In contrast, aqueous pretreatment in dilute acid mainly hydrolyzes covalent bonds within the polysaccharides.[26,27] However, these methods have limited effectiveness at high solids loadings and on highly recalcitrant feedstocks, such as wood, which requires the reaction conditions to be optimized.[28,29]

Co-solvents such as alcohols, ketones, ethers, or ionic liquids can be added to aqueous dilute acid in monophasic bulk mixtures to promote cellulose disruption and lignin removal during pretreatment, resulting in greater deconstruction of biomass. However, little is known about the interactions at the co-solvent−biomass interface and the mechanism of co-solvent pretreatment. Consequently, these methods are typically performed under high co-solvent and acid concentrations and/or high reaction temperatures to produce highly digestible pretreated solids at the expense of significant total sugar losses, expensive solvent recovery techniques, or excessive solvent slippage during liquid−solids separations.[30−32]

Recently, a pretreatment method, named Co-solvent Enhanced Lignocellulosic Fractionation (CELF), was developed that augments traditional aqueous-based pretreatment by employing tetrahydrofuran (THF) as a co-solvent in water to significantly improve biomass delignification and deconstruction of sugar polymers.[33] This approach has been shown to effectively reduce biomass recalcitrance, e.g., by requiring 10 times less enzymes in subsequent enzymatic hydrolysis than dilute acid pretreatment and still achieve over 95% total C5 and C6 sugar recovery from corn stover.[34,35] Furthermore, THF is particularly appealing for a sustainable bioindustry because it can be produced from biomass-derived furfural[36] and easily recovered due to its low boiling point. In addition, for a range of reaction conditions, over 90% of the lignin is solubilized in THF−water pretreatment, producing a highly pure lignin product that can be amenable for valorization.[37]

Understanding the physicochemical effects of THF−water solutions on lignin and cellulose should reveal the mechanism that makes CELF pretreatment so effective and guide the tuning of appropriate reaction conditions and process parameters in future work to enable economically viable cellulosic biofuel production. Thus far, experimental evidence indicates that structural transformation of biomass, including delignification, and chemical conversion (i.e., the hydrolysis of covalent bonds) both take place and lead to solubilization of lignin and cellulose. However, the exact mechanism and sequence of these events and the effect of temperature on this pretreatment method, all remain unknown.

Molecular dynamics (MD) simulation is an excellent tool to examine at atomic detail the structural basis of biomass recalcitrance and its morphological changes during pretreatment. MD studies have provided detailed insight into hydrogen bonding and stacking interactions in cellulose fibers[38−41] and revealed how solvents, such as ammonia and ionic liquids, can modify these interactions.[41−43] Recently, we reported a MD study of the hydrophobic polymer lignin in THF−water, which showed that the lignin is preferentially solvated by THF and adopts a random coil conformation, indicating that it is dissolved in the pretreatment reactor liquid and can thus be easily removed.[44,45] Hence, in contrast to the aqueous pretreatment methods, in which lignin molecules collapse and aggregate on the cellulose surface,[46,47] the addition of THF in CELF may contribute to the delignification of biomass by preferential dissolution. Details of cellulose solvation by THF and water, however, remain unknown.

Here, we present mechanistic evidence of enhanced cellulose deconstruction in THF−water mixtures by combining simulation and experimental results to demonstrate this unique co-solvent's ability to significantly improve the solubilization of cellulose compared to aqueous-only environments. We perform a series of all-atom MD simulations to analyze the structure of cellulose and its interactions with the co-solvents for both a whole multiple-chain cellulose fiber and a single-chain cellulose polymer. In the monophasic co-solvent mixture, THF and water spontaneously separate into two phases on the fiber surfaces while a single cellulose chain is preferentially solvated by water molecules. We supplement these simulation findings by demonstrating the enhanced hydrolysis of purified cellulose and the increased solubilization of cellulose and lignin from maple wood in equivolume mixtures of THF and water at conditions suitable for economic processing of biomass. The results indicate that the differential cellulose solvation by THF and water promotes its deconstruction and facilitates the delignification of biomass at very reasonable co-solvent concentrations and reaction temperatures, possibly allowing the solubilization of lignin and cellulose to be decoupled by tuning the reaction conditions. This study highlights that the versatility of the THF−water co-solvent system in the solvation of cellulose is crucial for efficient biomass pretreatment and fractionation.

■ METHODS

Simulations. For the fiber simulations, a 36-chain Iβ-cellulose crystal with 20 glucose units per chain (DP20) was solvated in an equilibrated rectangular box of THF and water molecules. The THF:H$_2$O molar ratio of the solution was 2:9, which roughly corresponds to the experimental 1:1 v/v concentration. The total number of atoms in this simulation system was ∼98 000. For comparison, the same cellulose structure was solvated in pure water. Due to greater rotation of the entire fiber in pure water, it was solvated in a larger cubic box of water, resulting in a system of ∼330 000 atoms. Also, single strand simulations were performed, in which a single cellulose 20mer was solvated in cubic simulation boxes of pure water or THF−water, at the same concentration as the cellulose fiber. The co-solvent single-strand system consists of ∼260 000 atoms, and the pure water system has ∼330 000 atoms.

CHARMM36 parameters[48,49] were used for the cellulose along with the latest CHARMM-additive ether parametrization for the THF molecules,[50] and the TIP3P model for the water.[51] All simulations were performed and analyzed with the GROMACS software version 5.0.1.[52] Long-range electrostatic interactions were calculated using the Particle-Mesh Ewald summation.[53,54] Short-range Coulomb and van der Waals (vdW) interactions were cutoff at 10 Å.

The simulations were run at three temperatures, 303, 378, and 445 K, which are identical to those used in previous simulations of lignin in THF−water,[44] allowing us to examine comparatively the effect of temperature on the cellulose solvation. The temperature was maintained by velocity rescaling with a stochastic term to ensure proper sampling.[55] After equilibrating the density with an NPT simulation of 1 ns applying the Berendsen barostat,[56] the systems were further relaxed for up to 20 ns in the NVT ensemble. The production runs were performed for over 150 ns with a time step of 2 fs in the
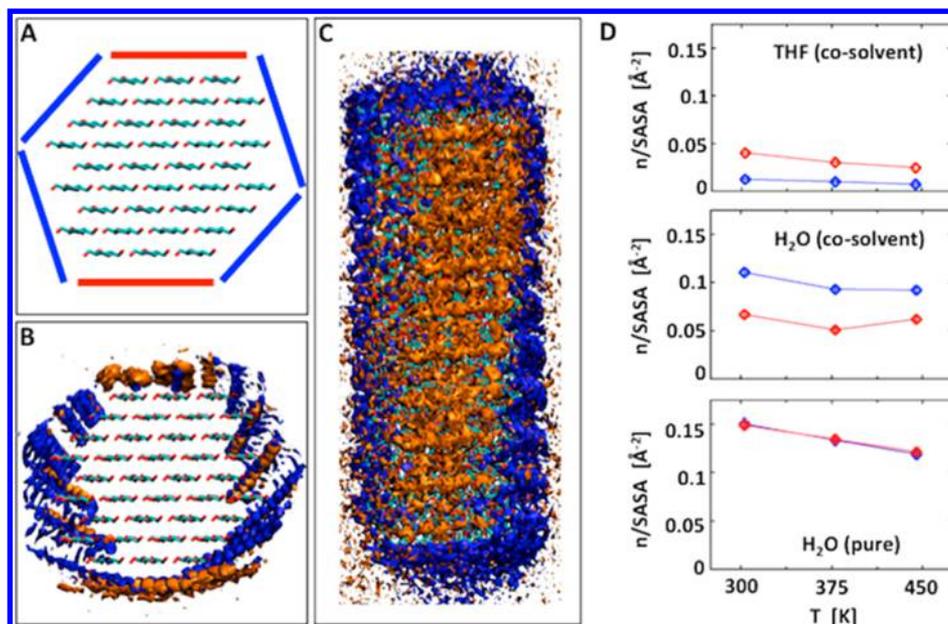
Figure 1. (a) View along the polymerization axis of the simulation starting structure, a crystalline cellulose Iβ fiber. The colored lines define the solvent-accessible hydrophilic (blue) and hydrophobic (red) fiber surfaces. (b) Spatial distributions of THF (orange contours) and water molecules (blue contours) around the cellulose fiber based on the simulation in the co-solvent system at 303 K. The contours define regions in space with ∼4 times the concentration of THF (orange) and ∼3 times the concentration of water (blue), respectively, in a bulk THF−water solution at the same conditions. The fiber is shown in its starting structure along the polymerization axis. (c) Spatial distributions as in (b). The fiber is shown in its starting structure with a view onto a hydrophobic face. (d) Average coordination numbers of solvent molecules around single glucose monomers divided by the solvent-accessible surface area, n/SASA, of the hydrophilic (blue) and hydrophobic (red) cellulose surfaces as a function of simulation temperature. n is derived from the center-of-mass RDFs as described in Figure S3. The corresponding values of the SASA are shown in Figure S4.

NVT ensemble. Bonds to hydrogen atoms were constrained using the LINCS algorithm.[57]

Convergence of the simulations was verified by obtaining the population of the preferred rotamers of the torsion angle ω and the root-mean-squared fluctuations (RMSFs) of the torsion angles ϕ and ψ, which describe the conformations of the cellulose side-chains and main-chains, respectively, and the solvent coordination number of cellulose surface monomers for three independent parts of the simulations. The reported values (Tables S1−S4) are comparable for any of the three time blocks, confirming the convergence of local cellulose structural properties and of solute−solvent interactions. In all graphs of this work, the mean values were obtained from block-averaging and error bars are the standard errors of the mean values.

Experiments. Avicel cellulose reactions were carried out at 5 wt% solids in a 1 L Hastelloy Parr reactor with twin blade impellers rotated at 250 rpms. Maple wood reactions were carried out using air-dried (<7 wt% moisture content) maple wood chips obtained in upper New York State by Mascoma Corporation (Lebanon, NH). The chips were milled and sieved to 1 mm particle size. The co-solvent mixture was prepared volumetrically with THF (>99% purity, Fisher Scientific, Pittsburgh, PA) and deionized (DI) water to a 1:1 volume ratio. Concentrated sulfuric acid (72 wt%, Ricca Chemical Co.) was diluted in the co-solvent solution to obtain the prescribed acid loadings for each reaction. The highest system pressure was 225 psig at 445 K. The reactions were performed in a sealed reactor and supercritical conditions were not reached at these temperatures. The vapor composition of THF and water were not tracked, but it is expected to follow regular vapor−liquid equilibrium values at these pressures. The reaction vessel and all other experimental procedures are mentioned previously.[33]

The solids were collected from the reaction liquor by vacuum filtration through glass fiber filter paper (Fisher Scientific, Pittsburgh, PA), dried in an oven, and weighed. Maple wood and pretreated material composition was measured according to the established National Renewable Energy Laboratory (NREL) procedure (version 8-03-2012) in triplicates. Liquid samples were analyzed by HPLC

(Agilent 1200 system equipped with a Bio-Rad Aminex HPX-87H column and RI detector).

■ RESULTS AND DISCUSSION

THF and Water Locally Phase Separate on the Surface of Cellulose Fibers. As cellulose fibers constitute the largest fraction of a plant cell wall,[58] the analysis of their interaction with the THF−water solvent molecules is critical to understanding the molecular mechanism of CELF pretreatment. We performed MD simulations of a cellulose fiber in THF−water and pure water at three different temperatures (303, 378, and 445 K) to assess the solute−solvent interaction patterns upon the addition of THF and the changes in equilibrium cellulose conformations over the broad operating temperature range of this co-solvent system. The simulated fiber has six surfaces as visualized in Figure 1a. Four of these surfaces are referred to as hydrophilic because the polar hydroxyl groups of their glucose units are exposed to the solvent, while the other two surfaces are more hydrophobic because they expose to the solvent the less polar aliphatic hydrogen atoms of the glucose rings and the glycosidic bonds between the monomers. These solvent-exposed surface hydroxyls and glucose rings are the first sites of interaction between the cellulose fiber and its surrounding solvent.

In Figure 1b,c, the average spatial distributions of THF and water about the cellulose fiber are shown at 303 K. The orange contours correspond to regions of higher THF density and the blue contours to water-dense regions around the cellulose fiber. A striking observation is that, contrary to the miscible bulk environment, the two types of co-solvent molecules phase separate on the cellulose surface: THF molecules accumulate on the hydrophobic regions of the surface, while the hydrophilic fiber faces are mainly solvated by water molecules.

The average solvent coordination number of cellulose surface monomers weighted by their solvent-accessible surface area, $n/$SASA, allows for a direct comparison between the hydrophobic and hydrophilic cellulose surfaces (Figure 1d). These normalized coordination numbers clearly show that, at all three simulation temperatures, the hydration of cellulose, and in particular of its hydrophobic faces, is reduced upon the addition of THF to a water-only solvent environment, suggesting a perturbing effect of THF on water−cellulose interactions.

To obtain a better understanding of the molecular interactions between the solvent molecules and the cellulose fiber surface, we examined the orientation of THF and water molecules with respect to the surface glucose monomers, defined by the angle α between the normal vectors of glucose rings and those of the nearest THF molecules or the nearest water dipole direction, respectively (Figure 2a). This analysis
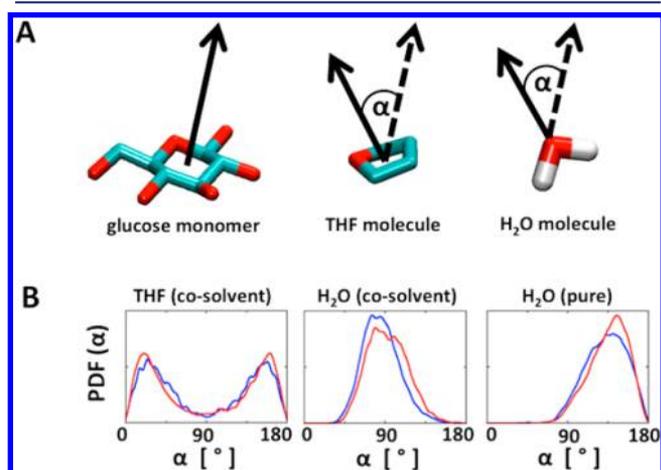

Figure 2. Orientational properties of solute−solvent interactions. (a) α is defined as the angle between the glucose ring normal (left) and the THF ring normal (middle) or the water dipole moment (right). Solvent molecules with a center-of-mass separation from a surface glucose monomer of 4 Å (for THF) or 3 Å (for water), respectively, were taken into consideration, which results in no more than two solvent molecules per glucose monomer being evaluated at any time. (b) Probability density function of α for solvent molecules with respect to hydrophilic (blue) and hydrophobic (red) cellulose surface monomers at 303 K.

quantifies which specific molecular orientations, and thus interactions, are preferred, leading to the observed phase separation. The angle α varies between 0° and 180° and, if no preferred interaction between the molecules exists, a solute-glucose orientation with $\alpha = 90°$ would be observed most frequently. The corresponding probability distributions at 303 K reveal that THF molecules have maximum values of $\alpha \approx 25°$ and 155° (Figure 2b), indicating offset stacking interactions between THF and glucose rings. In contrast, the distribution for water molecules shows a single peak at ∼135° in the pure water solution, which is shifted to ∼90° in the co-solvent solution, indicating that the water structure on the surface of cellulose changes when THF is added to an aqueous cellulose solution.

Cellulose is hydrolyzed during the CELF pretreatment, which involves the cleavage of glycosidic bonds. This can take place in dilute acid solutions mediated by a single water-splitting event,[59,60] and thus the proximity of water molecules

to the glycosidic linkages is critical for efficient hydrolysis. As mentioned above, glycosidic bond oxygens are more solvent-exposed on the hydrophobic than on the hydrophilic surfaces of a cellulose fiber. Hence, the occupancy or probability of hydrogen bonds, P(H-bond), formed between water hydrogens and these oxygens is generally higher for cellulose chains on the hydrophobic (12−20%) than for those on the hydrophilic surfaces (8−10%) of a fiber in pure water, as shown in Figure 3a. In the co-solvent system, this H-bond probability for the
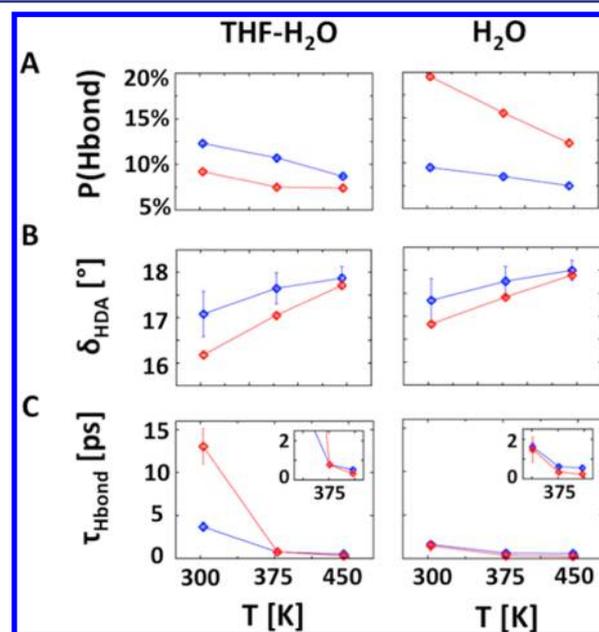

Figure 3. (a) Probabilities of hydrogen bonds formed between the hydrophilic (blue) or hydrophobic (red) cellulose surface glycosidic bond oxygens accepting water hydrogens in the co-solvent (left) and in the pure water solution (right) as a function of simulation temperature. A hydrogen bond is defined by the standard geometrical criterion of donor−acceptor distance smaller than 3.5 Å and the hydrogen-donor−acceptor angle smaller than 60°. (b) Average hydrogen-donor−acceptor angle, δ_{HDA}, of the same hydrogen bonds as in (a) as a function of simulation temperature. (c) Average hydrogen bond lifetimes, τ_{H-bond}, of the same hydrogen bonds as in (a). The analysis is performed for 5 independent parts of the simulation of length 1 ns. The average lifetime is obtained as the average over all autocorrelation functions of the binary hydrogen bond existence function during that part.[68]

hydrophobic surfaces is reduced to roughly the same value as for the hydrophilic surfaces (see Figure 3a), as a result of the THF−water phase separation on the cellulose surface. Interestingly, the H-bonds formed between water hydrogens and the glycosidic oxygens in the cellulose−co-solvent system have a smaller hydrogen-donor−acceptor angle, δ_{HDA}, than those in the cellulose−water system, indicating a more linear and thus stronger H-bond,[61] particularly on the hydrophobic surfaces and at 303 K (Figure 3b). The calculation of H-bond lifetimes, τ_{H-bond}, reveals that these H-bonds are also longer-lived (Figure 3c), further emphasizing that water molecules form stronger H-bonds with the glycosidic linkages of cellulose when THF is added. This shows that the presence of THF removes excess water molecules from the hydrophobic cellulose surfaces, but enhances the binding of remaining water molecules to the cellulose glycosidic bonds, consistent with the greater hydrolytic cleavage.
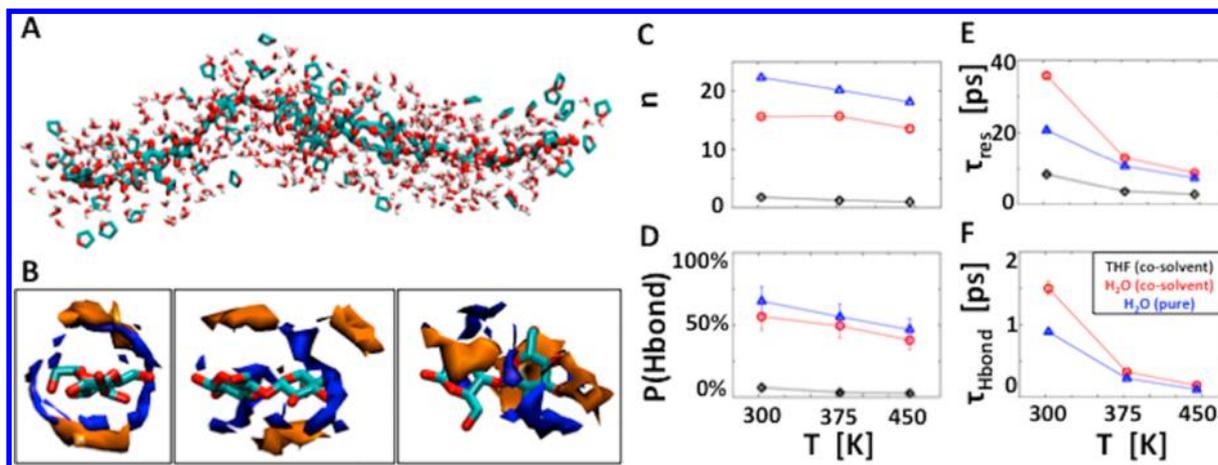
Figure 4. (a) Snapshot of the single cellulose chain with THF and water molecules within 5 Å after ∼100 ns of MD simulation time at 303 K. (b) Spatial distributions of THF (orange contours) and water molecules (blue contours) around cellobiose units in the simulation of the co-solvent system at 303 K. The contours define regions in space with ∼5 times the concentration of THF (orange) and ∼3 times the concentration of water (blue), respectively, in a bulk THF−water solution at the same conditions. The cellobiose is shown along the polymerization axis (left), in the glucose ring plane (middle), and along the ring normal (right). (c) Average coordination numbers, n, of solvent molecules around single glucose monomers of the cellulose strand as a function of simulation temperature. n is derived from the center-of-mass RDFs shown in Figure S13 through integration as in Figure S3. The corresponding RDF first minima are 6.9 Å for THF and 6.4 Å for water in both the THF−water and the pure water simulations. (d) Probabilities of the solvent oxygen atoms accepting a hydrogen bond from the cellulose hydroxyl groups as a function of simulation temperature. A hydrogen bond is defined as in Figure 3a. (e) Average cellulose strand solvation shell residence times, τ_{res}, of solvent molecules as a function of simulation temperature. The solvation shell is defined as the space within 5 Å of a cellulose monomer, which corresponds to the maximum in water−cellulose RDFs (see Figure S13). The residence time for a solvent molecule is the time it stays in the cellulose solvation shell. (f) Average hydrogen bond lifetimes, $\tau_{H\text{-bond}}$, of the water oxygen atoms in the co-solvent (red) and in pure water (blue) accepting a hydrogen bond from the cellulose strand hydroxyls as a function of simulation temperature. The values are obtained as described in Figure 3c.

In the Supporting Information, we present more detailed results, such as H-bonds between the solvent-exposed cellulose hydroxyls and the oxygen atoms of the co-solvent molecules and total interaction energies between the co-solvents and the cellulose fiber, providing a quantitative description for the observed local THF−water phase separation on cellulose surfaces. The results indicate that, at any simulation temperature, THF molecules preferentially bind to the hydrophobic cellulose surfaces, thus interfering with their hydration, and water molecules preferably form H-bonds with the hydrophilic surfaces. The addition of THF to aqueous cellulose solutions generally reduces the binding of water to cellulose and weakens their interactions (Figure S5).

Altogether, the results presented show that THF and water phase separate on the cellulose surface and that this phenomenon involves specific interactions between the solvent molecules and the hydrophobic and hydrophilic fiber surfaces. Co-solvent demixing has been observed before in bulk THF−water mixtures at temperatures between ∼335 and ∼415 K.[62,63] However, the local phase separation on the cellulose surface occurs at all simulated temperatures, including 303 and 445 K, which lie in the miscibility regime of bulk THF−water. Hence, we conclude that this differential solvation of cellulose is induced by the chemical heterogeneity of its surface with both hydrophobic and hydrophilic regions. As a consequence of this, the structure of water on the fiber surface is altered in the presence of THF, mainly resulting in reduced interactions except for the stronger H-bonds with the glycosidic bond oxygens, which may facilitate the hydrolysis of cellulose.

The demixing of THF and water on the surface of cellulose also provides a plausible explanation for the efficient lignocellulose fractionation during the CELF pretreatment. It has been shown before that lignin molecules preferentially aggregate onto hydrophobic faces of crystalline cellulose fibers.[64,65] Therefore, the accumulation of THF on the same faces could shield them from lignin binding, thus possibly enhancing the delignification process.

Individual Cellulose Strands Are Preferentially Bound by Water. A single fully solvated cellulose strand (Figure 4a) assumes different conformations than those packed in a large cellulose fiber because of the lack of H-bonds and stacking interactions with neighboring chains. Therefore, with respect to understanding the mechanism of the co-solvent pretreatment, the simulation of a single cellulose chain provides complementary information to the presented results on a whole fiber in THF−water.

Spatial distributions of THF and water molecules around cellobiose units (Figure 4b) show that both co-solvent molecules bind to specific sites of glucose monomers: as with the fiber, water binds to the cellulose chain's hydroxyl groups and THF to its ring structures. The average densities show that water is less dense around the hydroxymethyl groups than around the ring hydroxyls, which are in close proximity to each other and possibly even allow for multiple H-bonds to water molecules. Interestingly, high THF densities are observed on only one side of any glucose ring and the THF molecules are preferentially located on opposing sides of adjacent monomers, which reduces THF−THF interactions in proximity to cellulose and allows for greater solvation by water. The coordination number of glucose units (Figure 4c) reveals that up to two THF molecules coordinate one cellulose monomer on average, reducing the number of possible water molecules surrounding cellulose from $n \approx 20$ in pure water to $n \approx 15$ in the co-solvent. Still, the larger number of water molecules coordinating cellulose indicates that it is the primary solvent of individual cellulose chains in THF−water.

Probabilities of H-bonds, formed between the cellulose hydroxyls and the solvent oxygens, confirm that water binds to

the single strand more than THF does (Figure 4d). This number is between 40 and 60% for water and <10% for THF molecules, the sum of which is lower than for solvent-exposed hydroxyl groups on a cellulose fiber (see Figure S5a). This may be explained by a number of hydroxyls not accessible to solvent, as indicated by the spatial densities presented in Figure 4b, possibly due to the formation of intramolecular H-bonds within a flexible cellulose strand that are not observed for the fiber. Interestingly, this analysis reveals that the total number of cellulose H-bonds with solvent oxygen atoms is similar in both solvents, as the sum of P(H-bond) for THF and water in the co-solvent is roughly the same as P(H-bond) in the pure water. This shows that the presence of the larger THF molecule, when H-bonded to cellulose, does not block nearby hydroxyl groups and still allows cellulose to form as many H-bonds with solvent molecules as in pure water.

We also computed the residence time, τ_{res}, of solvent molecules on cellulose and the lifetime of hydrogen bonds, $\tau_{H\text{-}bond}$, between solvent and cellulose. Figure 4e shows the average time a THF or water molecule remains within a sphere of radius 5 Å around individual monomers. Although water diffuses faster in bulk than THF does ($D_{H_2O} \approx 20 \times 10^{-5}$ cm^2/s and $D_{THF} \approx 10 \times 10^{-5}$ cm^2/s at $T = 445$ K[63]), its residence time is 3 times longer than that of THF; i.e., cellulose-bound water does not move to the bulk solution as often as THF. At the lowest simulation temperature, the water residence time in the co-solvent is 2 times longer than in the pure water solution, suggesting that the presence of THF results in longer cellulose–water interactions, which may also increase the likelihood of hydrolysis reactions to take place. The same trend is also observed when comparing the cellulose H-bond lifetimes of water in the two solvents (Figure 4f). At 303 K, $\tau_{H\text{-}bond} \approx 1.5$ ps in the presence of THF, but only ~1 ps in its absence. At higher temperatures, however, the solvation residence and the H-bond lifetimes of water molecules both are independent of the presence of THF, which is consistent with the lower configurational flexibility of the cellulose chain in THF–water than in water observed only at 303 K (see Figure S10b).

We report in the Supporting Information that the polymer's overall dimensionality and the local conformations of its functional groups are not affected by the addition of THF to a pure water solution. Moreover, we provide further evidence for the preferential binding of water to a single cellulose chain fully solvated in THF–water, supporting the solvent distribution and H-bond results presented in Figure 4.

The preferential hydration of cellulose chains may be crucial for driving the cellulose dissolution process, which involves the detachment and separation of oligomeric cellulose chains from a fiber into solution. Thus, it is desirable to know the degree of solvation of a cellulose chain by THF and by water as it detaches from the mainly THF-bound hydrophobic surface of a fiber (see Figure 1) and becomes preferentially hydrated when fully solvated. To this end, we simulated at 303 K a stack of three cellulose sheets with the middle chain on the top surface detached at varying degrees by separating its reducing end from the surface by $d = 5$ to 42.5 Å (Figure 5a). At the maximum separation distance, the detached chain is a cello-octamer. Clearly, the solvent-accessible surface area (SASA) of this octamer increases with d and thus the SASA-weighted solvent coordination number, n, of glucose monomers describes the degree of solvation independent of d. Even at small separations from the surface ($d = 5$ Å), water solvates the slightly detached
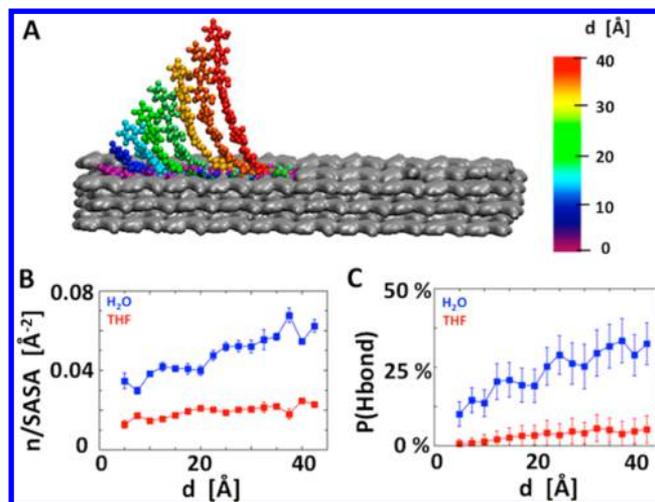

Figure 5. (a) A three-layered cellulose stack, in which the middle chain on the top hydrophobic surface is pulled perpendicular to that surface in steps of 2.5 Å from 5 to 42.5 Å. At maximum separation the detached oligomer is a cello-octamer. (b) Average coordination numbers of water (blue) and THF (red) molecules around single glucose monomers divided by the solvent-accessible surface area, n/SASA, of the cello-octamer. (c) Probabilities of the solvent oxygen atoms accepting a hydrogen bond from the cello-octamer hydroxyl groups as a function of simulation temperature. A hydrogen bond is defined as in Figure 3a.

chain more than THF does (Figure 5b). The THF solvation remains constantly low (n/SASA ≈ 0.02 Å^{-2}) compared to that by water (n/SASA $\approx 0.03–0.07$ Å^{-2}), which increases with d. The same trends are observed for the H-bond probabilities, P(H-bond), between hydroxyls of the octamer and solvent oxygen atoms (Figure 5c). This analysis shows that a single cellulose chain is preferentially bound by water as soon as it starts to detach from the fiber and that this effect grows as it becomes fully solvated.

The preferential binding of a cellulose strand by water over THF is notable considering the THF concentration (ca. 1:1, v/v) in our simulations. The longer hydration times in the co-solvent than in the pure water solution reveal that local water molecules are restrained and that the hydrolysis of a single cellulose chain may be enhanced in the presence of THF. Also, the alternating binding of THF to both sides of the glucose rings, in combination with the H-bonds of water to the hydroxyl groups, reveals how the properties of both co-solvents are exploited for fully solvating a cellulose chain and possibly preventing its association with other cellulose aggregates. This differential solvation of a single cellulose chain by THF and water may be important for the effectiveness of the CELF pretreatment.

The Differential Cellulose Solvation by THF and Water Promotes Its Solubilization. The primary action of traditional co-solvent reactions on biomass, such as ethanol–water organosolv pretreatment and pulping, is to promote delignification of biomass to allow unfettered access to cellulose. After removing the lignin shield, hydrolysis and solubilization of the remaining cellulose-rich material would closely match that of a pure cellulose substrate in both enzymatic and catalytic downstream applications. The addition of a strong acid, e.g. sulfuric acid, in dilute concentrations (usually 0.05 to 0.2 M) as a pretreatment catalyst is beneficial in practice as it reduces the activation energy of the hydrolytic cleavage of glycosidic bonds

and enhances the bulk deconstruction of lignocellulose. Until now, the mechanism of enhanced cellulose hydrolysis with certain co-solvents has been unknown and many organosolv co-solvent processes have been typically operated at relatively high acid loadings to achieve sufficient cellulose deconstruction.[66] However, selecting a high acid concentration may be undesirable for certain downstream applications because it also promotes the destruction of solubilized sugar and lignin intermediates.

To demonstrate the effects of local THF−water phase separation and preferential hydration of cellulose observed in our simulations on a real model substrate, we performed heated reactions at 445 K of microcrystalline cellulose (Avicel) in aqueous dilute sulfuric acid mixtures both with THF as a co-solvent and without THF at 0.1 and 0.2 M acid concentrations and measured the percent solubilization of the Avicel cellulose. The results reveal that the presence of THF as a water-miscible co-solvent greatly enhanced the solubilization of cellulose in comparison to the water-only solvent system at both acid concentrations (Figure 6). As expected, higher sulfuric acid
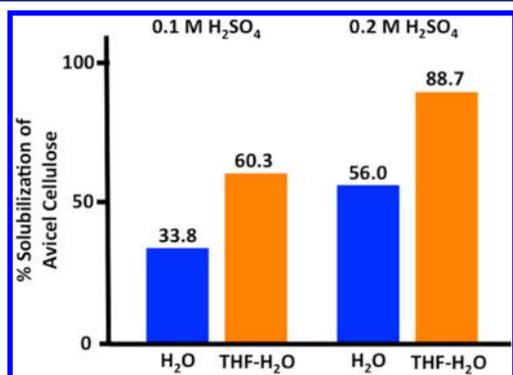

Figure 6. Comparison of the % solubilization of Avicel cellulose in heated batch reactions solvated in water-only or in THF−water co-solvent mixtures at two dilute sulfuric acid concentrations. The reaction temperature was 445 K and the duration was 30 min for all reactions.

concentration also increased the total solubilization of cellulose in both water-only and co-solvent reactions. However, a greater amount of cellulose was solubilized in the co-solvent system at the lower acid loading (60.3% at 0.1 M) than in the water-only system at twice the acid concentration (56.0% at 0.2 M), which demonstrates the enhancement potential of the THF co-solvent. In fact, at 0.1 M acid loading, the rate of solubilization of Avicel cellulose was almost doubled in the THF−water co-solvent reaction compared to the water-only reaction. Thus, only about half of the sulfuric acid loading would be needed to achieve a comparable cellulose solubilization in the THF−water system as in the water-only case. These results demonstrate the advantage of THF−water co-solvent pretreatment.

The residues from the 0.2 M acid loading reactions were then collected and imaged by SEM as shown in Figure 7. At the mild 1000× magnification shown, structural features between the THF−water and water-only reactions were already very distinct in the images. After the water-only reaction, the remaining cellulose residues still contained sharply defined edges and flat plate-like regions that resembled the same features present in unreacted Avicel cellulose suggesting that the solubilization of the cellulose fibers was uniform and mild. In contrast, residues from the THF co-solvent reaction appeared under SEM to be significantly more digested as evident by the smoother edges that are far less structurally defined and the lack of large sections containing flat surfaces as was found in the residues from the water-only reaction. The absence of sharp, distinct structural features in the co-solvent reactions indicates that greater solubilization had occurred around the edge and outside regions of the Avicel that typically contain smaller and more solvent-exposed fibers than the unsolubilized bulk region. The above experimental results correlate with the simulation finding that the preferential binding of water molecules to exposed cellulose structures in a THF co-solvent environment may be an important mechanism responsible for enhancing cellulose hydrolysis independent of acid concentrations.

The unique differential cellulose solvation by THF and water co-solvent observed in the simulations paired with the previous results on preferential lignin solvation by THF provides a mechanistic insight into how CELF pretreatment can be effective at achieving both enhanced cellulose hydrolysis as well as biomass delignification.[67] To investigate this further, we also performed reactions of raw maple wood chips with and without THF co-solvent at 0.05, 0.1, and 0.2 M concentrations of sulfuric acid. The mass and composition of raw maple wood and the remaining solids resulting from each reaction are shown
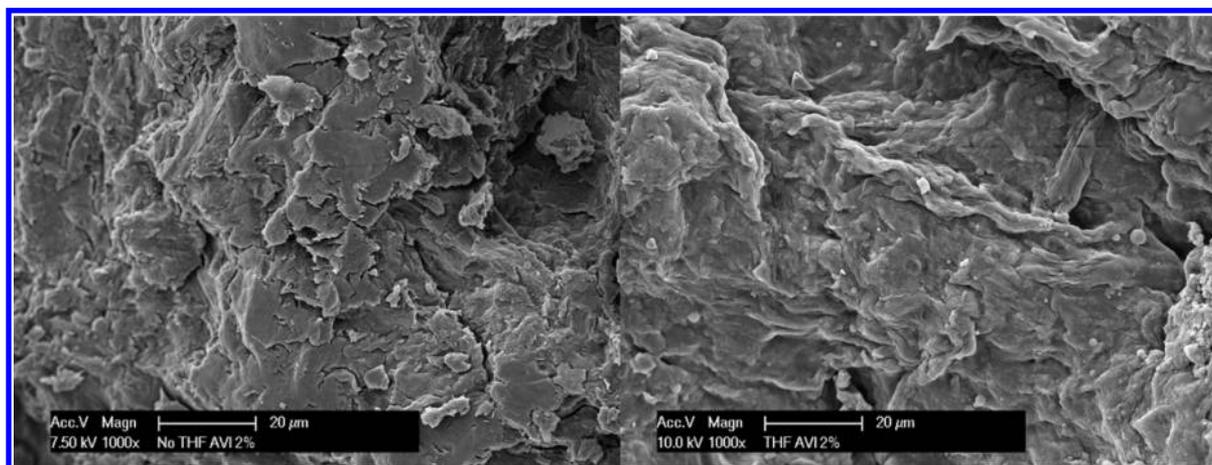

Figure 7. SEM images of Avicel cellulose solids after reaction in water (left) and in THF−water co-solvent (right). Reaction conditions: 445 K, 30 min, and 0.2 M sulfuric acid.

in Figure 8 on a 100 g basis of the raw material, demonstrating the extent of both cellulose and lignin solubilization. The
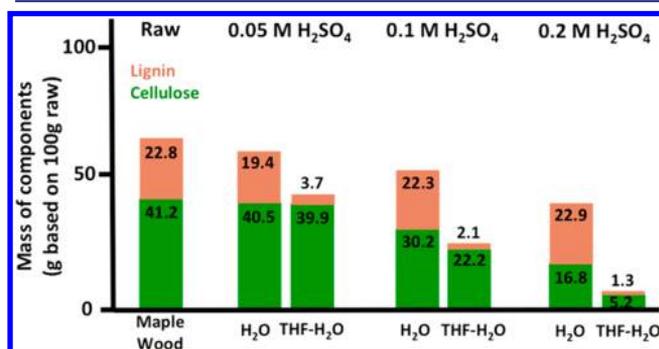

Figure 8. Comparison of the fate of cellulose and lignin components after experimental reactions of maple wood in both water-only and THF−water co-solvent mixtures with varying dissolved sulfuric acid concentrations. Tracked cellulose and lignin components are based on each of their mass compositions in 100 g of raw maple wood starting material. Other components such as hemicellulose sugars, extractives, organic acids, proteins, and ash etc. make up the remaining 36 g of the raw material, but are not shown as they get easily solubilized at mild conditions. Reaction temperatures (435−445 K) and durations (15−30 min) were held constant at each acid loading.

results show that the presence of THF in water greatly enhances the hydrolysis and solubilization of both cellulose and lignin from maple wood and confirms the observations from the Avicel reactions. Interestingly, at the very dilute 0.05 M acid concentrations, cellulose was only slightly more solubilized in THF−water than in water-only reactions, but lignin was nearly completely solubilized into the liquid phase in the presence of THF. This would allow for maximum recovery of cellulose as a solid from biomass without the interference of lignin. In contrast, at higher acid concentrations (0.1 and 0.2 M), both cellulose and lignin were significantly solubilized in the presence of THF, which is beneficial for greater total viscosity reduction of the biomass slurry to achieve efficient downstream catalytic upgrading of solubilized sugars.

As lignin is rapidly removed, in particular in the presence of THF, the initial rate of cellulose solubilization from maple wood agreed well with the Avicel experiments at the same reaction conditions. The rate of cellulose hydrolysis to soluble glucose monomers was approximately doubled in THF−water with 0.1 M sulfuric acid, achieving over 46% cellulose removal vs 26% in water-only. At the higher sulfuric acid concentration of 0.2 M, cellulose solubilization was further increased to over 87.4% in THF−water vs about 59.2% in water-only. The co-solvent solubilization of 87.4% from maple wood very closely matched the 88.7% solubilization of Avicel at the same conditions. However, the water-only reactions did not match well with the Avicel solubilization, especially at 0.1 M sulfuric acid concentrations, likely due to the interference from lignin, which is not as thoroughly removed as in the presence of THF, and the existence of more acid-labile amorphous cellulose in maple wood.

The experimental results can be explained by the unique behavior of THF and water molecules in the proximity of cellulose fibers and strands observed in the simulations. Since the interactions of lignin with cellulose in hardwoods are not as strong as the covalent glycosidic bonds of cellulose, the enhanced biomass delignification and lignin disruption at mild

reaction conditions can be attributed to the local phase separation of THF and water and the preferential solvation of lignin by THF reported previously.[44] On the other hand, the preferential binding of water to cellulose chains, and in particular to solvent-exposed glycosidic bonds, could promote the hydrolysis of cellulose at higher acid concentrations. This dependence of lignin and cellulose solubilization on the reaction conditions (e.g., acid loading) could allow the two processes to be decoupled by tuning the conditions, which demonstrates the high versatility of the CELF pretreatment compared to other co-solvent pretreatment methods.

■ CONCLUSIONS

The pretreatment of lignocellulose in aqueous THF co-solvent is very efficient in separating the biomass components and catalyzing their hydrolysis to sugar monomers.[33] The present work sought to provide detailed molecular insight into the solvation of cellulose in THF−water. The simulated multiple-chain fiber and single-chain polymer are very different cellulose structures in terms of dimensionality and flexibility. However, since lignin has been shown to bind cellulose fibers in aqueous environment, and since a fully solvated cellulose chain is expected to be more amenable to hydrolysis than a fiber, the study of both cellulose structures was necessary to provide a more complete understanding of the underlying mechanism of THF−water-mediated biomass deconstruction.

Our results indicate that the co-solvent system exploits the different features of its two components and their ability to demix to solvate cellulose fibers and single chains in very special and different ways: THF and water phase separate on the surface of whole fibers while single chains are preferentially bound by water in the mixed co-solvent bath. This local co-solvent phase separation may be key to its efficient pretreatment observed in experiments. The binding of THF to the hydrophobic faces of cellulose fibers supports the hypothesis that THF may block lignin aggregation on those faces, which occurs in purely aqueous solutions,[64,65] thus promoting the delignification and fractionation of biomass. The strong hydration of single cellulose strands observed in our study indicates that cellulose can be more easily hydrolyzed in the presence of THF. Moreover, in combination with the preferential solvation of lignin by THF,[44] the preferential hydration at later stages of the pretreatment process, when cellulose strands are more accessible to solvent than they are in a fiber, may also advance the fractionation process.

The experimental results are consistent with those of the simulations, showing greater disruption of microcrystalline cellulose and increased solubilization of cellulose and lignin in the presence of the co-solvent. In particular, the greater biomass deconstruction at lower acid loadings and milder reaction conditions than with traditional aqueous pretreatment methods is consistent with the preferential hydration of the solvent-exposed glycosidic bonds observed in the simulations, providing a means to fine-tune the CELF pretreatment with downstream applications. The separation of lignin and cellulose solubilization by tuning reaction parameters would be a key advantage of the CELF method, allowing its integration as a platform pretreatment technology for producing renewable fuels from biomass by both biological and catalytic pathways

Overall, the results presented here show versatility of the THF−water co-solvent in the solvation of cellulose that may contribute to the effectiveness of the CELF pretreatment method. Further simulations of lignocellulose aggregates or

dissolved lignin and cellulose polymers in THF−water will provide more information on the effects of THF and water on more complex biomass materials, describing the underlying driving force for the delignification of cellulose and the solute−solvent interactions relevant to hydrolysis reactions during biomass pretreatment. Such fundamental knowledge, obtained from molecular simulations, can aid in improving reaction conditions for the effective conversion of biomass into biofuels.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b03285.

> Analyses of the cellulose fiber simulations, and the cellulose strand simulations, including Figures S1−S14 and Tables S1−S4 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*smithjc@ornl.gov

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Sun, J. *Polym. Degrad. Stab.* **2004**, *84*, 331−339.
(2) Fernandes, A. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, E1195−E1203.
(3) Lynd, L. R.; Cushman, J. H.; Nichols, R. J.; Wyman, C. E. *Science (Washington, DC, U. S.)* **1991**, *251*, 1318−1323.
(4) Lynd, L. R.; Wyman, C. E.; Gerngross, T. U. *Biotechnology progress* **1999**, *15*, 777−793.
(5) Ragauskas, A.; et al. *Science (Washington, DC, U. S.)* **2006**, *311*, 484−489.
(6) U.S. Department of Energy. 2016 Billion-Ton Report: Advancing Domestic Resources for a Thriving Bioeconomy, http://energy.gov/eere/bioenergy/downloads/2016-billion-ton-report-advancing-domestic-resources-thriving-bioeconomy.
(7) Dwivedi, P.; et al. *Environ. Sci. Technol.* **2015**, *49*, 2512−2522.
(8) Schubert, C. *Nat. Biotechnol.* **2006**, *24*, 777−784.
(9) Himmel, M.; et al. *Science* **2007**, *315*, 804−807.
(10) McCann, M.; Carpita, N. *J. Exp. Bot.* **2015**, *66*, 4109−4118.
(11) Zhao, X.; Zhang, L.; Liu, D. *Biofuels, Bioprod. Biorefin.* **2012**, *6*, 561−579.
(12) Zheng, Y.; Zhao, J.; Xu, F.; Li, Y. *Prog. Energy Combust. Sci.* **2014**, *42*, 35−53.
(13) Kratky, L.; Jirout, T. *Chem. Eng. Technol.* **2011**, *34*, 391−399.
(14) Cowling, E. B.; Kirk, T. K. *Biotechnol. Bioeng. Symp.* **1976**, 95−123.
(15) Ramos, L. P. *Quim. Nova* **2003**, *26*, 863−871.
(16) Marousek, J. *Revista Tecnica* **2012**, *35*, 1−9.
(17) Fengel, D.; Wegener, G. *Wood — Chemistry*, Walter de Gruyter: Berlin and New York, 1984.
(18) Agbor, V.; Cicek, N.; Sparling, R.; Berlin, A.; Levin, D. *Biotechnol. Adv.* **2011**, *29*, 675−685.
(19) Swatloski, R.; Spear, S.; Holbrey, J.; Rogers, R. *J. Am. Chem. Soc.* **2002**, *124*, 4974−4975.
(20) Lee, J. *J. Biotechnol.* **1997**, *56*, 1−24.
(21) Sun, Y.; Cheng, J. *Bioresour. Technol.* **2002**, *83*, 1−11.
(22) Behera, S.; Arora, R.; Nandhagopal, N.; Kumar, S. *Renewable Sustainable Energy Rev.* **2014**, *36*, 91−106.
(23) Tarkow, H.; Feist, W. *Adv. Chem. Ser.* **1969**, *95*, 197−218.
(24) Zheng, Y.; Pan, Z.; Zhang, R. *Int. J. Agric. Biol. Eng.* **2009**, *2*, 51−68.
(25) Fan, L.-t.; Gharpuray, M.; Lee, Y.-H. *Cellulose Hydrolysis*, Biotechnology Monographs; Springer: Berlin, 2011).
(26) Sannigrahi, P.; Ragauskas, A.; Miller, S. *BioEnergy Res.* **2008**, *1*, 205−214.
(27) Teramura, H.; et al. *PLoS One* **2015**, *10*, e0128417.
(28) Modenbach, A.; Nokes, S. *Biotechnol. Bioeng.* **2012**, *109*, 1430−1442.
(29) Zhu, J. Y.; Pan, X. J. *Bioresour. Technol.* **2010**, *101*, 4992−5002.
(30) Shuai, L.; Questell-Santiago, Y.; Luterbacher, J. *Green Chem.* **2016**, *18*, 937−943.
(31) Agnihotri, S.; Johnsen, I.; Bøe, M.; Øyaas, K.; Moe, S. *Wood Sci. Technol.* **2015**, *49*, 881−896.
(32) Cruz, A.; et al. *Biotechnol. Biofuels* **2013**, *6*, 52.
(33) Cai, C.; Zhang, T.; Kumar, R.; Wyman, C. *Green Chem.* **2013**, *15*, 3140−3145.
(34) Nguyen, T.; Cai, C.; Kumar, R.; Wyman, C. *ChemSusChem* **2015**, *8*, 1716−1725.
(35) Nguyen, T.; Cai, C.; Osman, O.; Kumar, R.; Wyman, C. *Green Chem.* **2016**, *18*, 1581.
(36) Banford, W. H.; Manes, M. M. Production of Tetrahydrofuran. U.S Patent 2846449, 1958.
(37) Cai, C.; Nagane, N.; Kumar, R.; Wyman, C. *Green Chem.* **2014**, *16*, 3819−3829.
(38) Gross, A.; Chu, J.-W. *J. Phys. Chem. B* **2010**, *114*, 13333−13341.
(39) Parthasarathi, R.; et al. *J. Phys. Chem. A* **2011**, *115*, 14191−14202.
(40) Petridis, L.; et al. *Biomacromolecules* **2014**, *15*, 4152−4159.
(41) Bellesia, G.; Chundawat, S.; Langan, P.; Dale, B.; Gnanakaran, S. *J. Phys. Chem. B* **2011**, *115*, 9782−9788.
(42) Mostofian, B.; Smith, J.; Cheng, X. *Interdiscip. Sci.: Comput. Life Sci.* **2011**, *3*, 308−320.
(43) Rabideau, B.; Agarwal, A.; Ismail, A. *J. Phys. Chem. B* **2013**, *117*, 3469−3479.
(44) Smith, M.; et al. *Green Chem.* **2016**, *18*, 1268.
(45) Smith, M.; Petridis, L.; Cheng, X.; Mostofian, B.; Smith, J. *Phys. Chem. Chem. Phys.* **2016**, *18*, 6394−6398.
(46) Petridis, L.; Schulz, R.; Smith, J. *J. Am. Chem. Soc.* **2011**, *133*, 20277−20287.
(47) Langan, P.; et al. *Green Chem.* **2014**, *16*, 63−68.
(48) Guvench, O.; Hatcher, E.; Venable, R.; Pastor, R.; MacKerell, A. *J. Chem. Theory Comput.* **2009**, *5*, 2353−2370.
(49) Guvench, O.; et al. *J. Chem. Theory Comput.* **2011**, *7*, 3162−3180.
(50) Vorobyov, I.; et al. *J. Chem. Theory Comput.* **2007**, *3*, 1120−1133.
(51) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926−935.
(52) Pronk, S.; et al. *Bioinformatics* **2013**, *29*, 845−854.
(53) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(54) Essmann, U.; et al. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(55) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.

(56) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(57) Hess, B.; Bekker, H.; Berendsen, H.; Fraaije, J. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(58) Heredia, A.; Jiménez, A.; Guillén, R. *Z. Lebensm.-Unters. Forsch.* **1995**, *200*, 24−31.

(59) Zhang, T. Glucose production from cellulose in subcritical and supercritical water. Ph.D. Thesis, The University of Iowa, 2008.

(60) Fleming, K.; Pfaendtner, J. *J. Phys. Chem. A* **2013**, *117*, 14200−14208.

(61) Arunan, E.; et al. *Pure Appl. Chem.* **2011**, *83*, 1637−1641.

(62) Matous, J.; Novak, J. P.; Sobr, J.; Pick, J. *Collect. Czech. Chem. Commun.* **1972**, *37*, 2653−2663.

(63) Smith, M.; Mostofian, B.; Petridis, L.; Cheng, X.; Smith, J. *J. Phys. Chem. B* **2016**, *120*, 740−747.

(64) Lindner, B.; Petridis, L.; Schulz, R.; Smith, J. *Biomacromolecules* **2013**, *14*, 3390−3398.

(65) Vermaas, J.; et al. *Biotechnol. Biofuels* **2015**, *8*, 217−16.

(66) Sannigrahi, P.; Ragauskas, A. *Aqueous Pretreatment of Plant Biomass for Biological and Chemical Conversion to Fuels and Chemicals* **2013**, 201−222.

(67) Cai, C.; Wyman, C.; Zhang, T.; Kumar, R. Co-solvent to produce reactive intermediates from biomass. PCT Patent No. WO2014176531, 2013.

(68) Luzar, A.; Chandler, D. *Nature* **1996**, *379*, 55−57.