JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Simulation Analysis of the Temperature Dependence of Lignin Structure and Dynamics

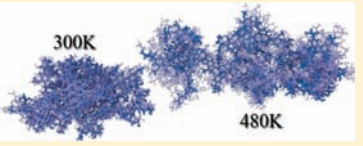
Loukas Petridis,[†,§] Roland Schulz,[†,‡,§] and Jeremy C. Smith[*,†]

[†]UT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, P.O.Box 2008, Oak Ridge, Tennessee 37831-6309, United States

[‡]Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996, United States

Supporting Information

ABSTRACT: Lignins are hydrophobic, branched polymers that regulate water conduction and provide protection against chemical and biological degradation in plant cell walls. Lignins also form a residual barrier to effective hydrolysis of plant biomass pretreated at elevated temperatures in cellulosic ethanol production. Here, the temperature-dependent structure and dynamics of individual softwood lignin polymers in aqueous solution are examined using extensive (17 μs) molecular dynamics simulations. With decreasing temperature the lignins are found to transition from mobile, extended to glassy, compact states. The polymers are composed of blobs, inside which the radius of gyration of a polymer segment is a power-law function of the number of monomers comprising it. In the low temperature states the blobs are interpermeable, the polymer does not conform to Zimm/Stockmayer theory, and branching does not lead to reduction of the polymer size, the radius of gyration being instead determined by shape anisotropy. At high temperatures the blobs become spatially separated leading to a fractal crumpled globule form. The low-temperature collapse is thermodynamically driven by the increase of the translational entropy and density fluctuations of water molecules removed from the hydration shell, thus distinguishing lignin collapse from enthalpically driven coil−globule polymer transitions and providing a thermodynamic role of hydration water density fluctuations in driving hydrophobic polymer collapse. Although hydrophobic, lignin is wetted, leading to locally enhanced chain dynamics of solvent-exposed monomers. The detailed characterization obtained here provides insight at atomic detail into processes relevant to biomass pretreatment for cellulosic ethanol production and general polymer coil−globule transition phenomena.



INTRODUCTION

Lignins are branched hydrophobic heteropolymers that provide mechanical strength to plant cell walls, regulate water conduction, and play an important role in plant defense against enzymatic or microbial degradation.[1,2] Lignins are also of central interest in biofuels research. Although plant biomass has the potential to be a renewable feedstock for industrial biofuel production, its recalcitrance to hydrolysis, which arises in part from lignin, necessitates expensive pretreatment prior to fermentation that significantly increases cost.[3] Pretreatment technologies commonly employed increase biofuel yield by disrupting the lignin−carbohydrate network that physically prevents enzymes from reaching and hydrolyzing cellulose.[4] A variety of pretreatment methods exist, the most common of which, such as dilute-acid, ammonia-based, and hydrothermal, involve temperatures $\gtrsim$100 °C.[4] Understanding the temperature dependence of lignin structure and dynamics is thus of particular importance to next-generation biofuel production.

Lignin is known to be "hard" and glassy at room temperature and to "soften" above its glass transition temperature, which ranges between 80 and 100 °C for softwoods.[5,6] Temperature influences the quality of water as a solvent for polymers: whereas at high temperatures water can be a "good" solvent, leading to the tendency of polymers to assume extended chain conformations, as the temperature decreases, water can become a "poor" solvent, and polymers tend to collapse to dense globules. This polymer coil−globule transition has been studied extensively experimentally,[7−10] theoretically,[11−13] and with simulation.[14−17]

The collapse at low temperatures is attributed to the hydrophobic effect, an interaction that drives diverse biological self-assembly phenomena, such as protein folding and membrane formation.[18−23] Hydrophobicity of idealized solutes, i.e. that do not exert attractive interactions with water, manifests itself differently on small and large length scales.[18,24,25] The free energy cost of solvating small idealized hydrophobic particles is entropic, driven by the formation of small water cavities that accommodate the solute without destroying water hydrogen bonds.[26,27] In contrast, large hydrophobic solutes break water−water hydrogen bonds, leading to the creation of a liquid−vapor interface and a large positive enthalpy of solvation.[25,28]

Water in the hydration layer of biomolecules has been extensively studied. Near proteins water has been shown to be denser than the bulk[29,30] and to exhibit slower dynamics (for a review see ref 31). Simulations have also found the subnanosecond translational and rotational entropy of hydration water

Received: July 21, 2011
Published: October 28, 2011

molecules near polyamidoamine dendrimers,[31] lipid bilayers,[32] and DNA[33] to be lower than the bulk. However, the structural and thermodynamic properties of lignin and its hydration water have so far not been investigated.

To our knowledge, there have been to date only two previous MD simulation studies of lignin. The first examined the organization of lignin oligomers on a cellulose surface. Due to computational limitations these simulations were limited to short trajectories performed in vacuum.[34] More recently, a combination of small angle neutron scattering and MD simulation examined the surface morphology of softwood lignin aggregates, which were found to exhibit self-similar surface properties constant over 3 orders of magnitude in length.[35]

Here, the effect of temperature and branching on the structure and dynamics of individual softwood lignin molecules in aqueous solution is investigated with the use of extensive (17.5 μs) atomistic molecular dynamics simulations. Due to their highly aggregating nature, individual lignin polymers can be found only in very dilute aqueous solutions, thus presenting steep challenges for their experimental characterization.[36] However, thermochemical pretreatment dissociates plant-cell wall lignin, and due to their hydrophobic character, the residual lignin polymers then collapse and coalesce to form clumps.[35,37−41] The computational investigation of the behavior of lignin polymers as a function of temperatures can provide insight into collapse and coalescence processes relevant to pretreatment of plant biomass.

The paper presents a comprehensive analysis of the temperature dependence of the size, shape, scaling properties, hydration, and dynamics of softwood lignin with varying degrees of branching. The lignin polymers are found to transition from high-temperature mobile and extended states to glassy and compact states at low temperatures. Hydration shell entropy is found to be at the thermodynamic origin of the lignin collapse at low temperatures.

METHODS

Model Systems. Structural models of individual lignin molecules were generated by using available experimental information on the average chemical composition of softwood lignins.[35] Softwood lignins are composed primarily of guaiacyl (G) monomers connected by various linkages, leading to the formation of branched and unbranched biopolymers.[42] Here, nine different lignin molecules were simulated.

Each molecule comprised 61 G units, with a molecular weight of ~13 kDa, within the experimentally determined range.[43] The average interunit-linkage composition was that of softwoods: β-O-4′ 50%, 5−5′ 30%, α-O-4′ 10%, and β-5′ 10%.[42] The number of branch points and their location along the chain were assigned randomly using a computer algorithm: two lignins, termed $L0_a$ and $L0_b$, have zero branch points, lignins $L1_a$ and $L1_b$ have one, *L2* two, *L3* three, *L4* four, *L5* five, and *L6* six. This distribution is consistent with the experimentally determined average linkage density for spruce wood, which is 0.052, or 3.2 branch points per 61 monomers.[44]

Hence, the primary structures of the nine lignins simulated here are different from each other but consistent with the average chemical composition of softwood lignin. For example, although for all molecules 50% of the linkages are of the β-O-4′ kind, the position of these linkages along the chains varies between molecules, as does the position of the branch points and the lengths of the branches. The sequence of linkages of each lignin can be found in Tables S2−S10 of the Supporting Information.

The resulting nine lignin molecules were subsequently individually solvated and subject to molecular dynamics simulations. Each lignin polymer was solvated in a rhombic dodecahedron with an inscribed sphere radius of 52.5 Å containing in total ~81 600 atoms.

The predicted properties of the model lignin molecules depend in general on the types of monomers and linkages used. For example, grasses contain a significant proportion of syringyl monomers that cannot form 5−5 linkages. Since monomers participating in 5−5 linkages contain one more hydroxyl group than those participating in β-O-4 and α-O-4 linkages, grass lignin models may be slightly less hydrophobic than the models presented here.

Molecular Dynamics Simulation Details. The CHARMM force field for lignin[45] and the TIP3P water model[46] were employed. Periodic boundary conditions were employed, and the PME algorithm[47,48] was used for electrostatic interactions. A reciprocal grid of 80 × 80 × 80 cells (84 × 84 × 84 for 480 K) was used with fourth order B-spline interpolation. A cutoff of 12 Å was used for the neighbor searching and real-space electrostatics. Charge groups were used for water but not for lignin. For the van der Waals interactions the switch function was used for distances 9−10 Å.

The simulations were performed with the program GROMACS 4.5.1[49−52] using a time step of 2 fs. Atoms involving hydrogens were constrained using the LINCS[53] algorithm (fourth order with one iteration), and for water the Settle algorithm was used.[54] Neighbor searching was performed every 10 steps. Temperature coupling was performed with the V-rescale algorithm[55] (τ = 0.1 fs) and pressure coupling with the Parrinello−Rahman algorithm[56] (τ = 1 fs).

For the four lignins with either one or zero branch points, each system was first heated from 0 to 480 K in 1 ns and subsequently simulated at 480 K for a total of 100 ns. Ten structures were taken from this trajectory at 10 ns, 20 ns, ..., 100 ns. Starting from these 10 structures each of the four lignins was simulated at four different temperatures: 300, 360, 420, and 480 K. The cooling from 480 K to the target temperature was performed in 5 ns, and then the simulation was equilibrated for 55 ns, followed by 50 ns of production used for the analysis. The above protocol was chosen in order to obtain 10 very different starting structures, since preliminary simulations had shown that at 480 K large structural changes take place on the 100 ns time scale.

The five separate MD simulations for each of the *L2*, *L3*, *L4*, *L5*, and *L6* lignins were performed at 300 K. The simulations were equilibrated for 55 ns, followed by 50 ns of production used for the analysis.

The total simulation length was 17.5 μs. All simulations were performed on the Jaguar Cray XT5 supercomputer at Oak Ridge National Laboratory.

Analysis of Molecular Dynamics Simulation. *Structural Properties of Lignin.* The radius of gyration R_g and the gyration tensor were computed using GROMACS g_polystat using all the lignin atoms and without mass-weighting. The asymmetry of the chain conformations was described by the asphericity, Δ

$$\Delta = \left\langle \frac{(L_1 - L_2)^2 + (L_2 - L_3)^2 + (L_1 - L_3)^2}{2(L_1 + L_2 + L_3)^2} \right\rangle \quad (1)$$

where L_1, L_2, and L_3 are the eigenvalues of the radius of gyration tensor of the particular lignin molecule and $\langle ... \rangle$ represents an ensemble average, replaced here by a time average. The shape of the molecule was further characterized by calculating the prolateness, Σ

$$\Sigma = \left\langle \frac{(2L_1 - L_2 - L_3)(2L_2 - L_3 - L_1)(2L_3 - L_1 - L_2)}{2(L_1^2 + L_2^2 + L_3^2 - L_1L_2 - L_2L_3 - L_3L_1)^{3/2}} \right\rangle \quad (2)$$

For a solid ellipsoid with unit mass density and radius of gyration $r_{g,i}$ around axis i it follows[57]

$$r_{g,i}^2 = \frac{1}{5}\sum_{j\neq i} r_j^2 = \sum_{j\neq i} L_j \Rightarrow r_i = \sqrt{5L_i} \quad (3)$$

where r_i is the radius and L_i is the eigenvector of the gyration tensor introduced before. Thus, the volume of the ellipsoid is

$$V = \frac{4}{3}\pi r_1 r_2 r_3 = \frac{4}{3}\pi S^{3/2}\sqrt{L_1 L_2 L_3} \quad (4)$$

The solvent accessible surface area (SASA) was computed using GROMACS g_sas with a probe radius of 1.4 Å. The van der Waals radii for the lignin atom types were set to C, 1.5 Å; H, 1 Å; and O, 1.3 Å. A lignin atomic contact was defined as any pair of lignin atoms separated by <3 Å. Hydrogen bonds were defined by a < 3.5 Å donor−acceptor distance and a donor-H−acceptor angle between 150° and 210°.

Lignin Dynamics. Monomer mean square displacements

$$\langle \Delta r_n^2 \rangle = \langle [r_n(t) - r_n(0)]^2 \rangle \quad (5)$$

where $r_n(t)$ is the position of the nth monomer at time t were calculated after aligning each trajectory frame to the preceding frame to remove whole lignin molecule translation and rotation. Data for calculating $\langle \Delta r_n^2 \rangle$ were collected over the last 50 ns of 40 MD 100 ns trajectories (four molecules simulated 10 times each).

Structure of Hydration Water. The water molecules were classified into (a) "hydration shell" water, defined by a 4.9 Å distance cutoff between water oxygen atoms and any lignin non-hydrogen atom, where this cutoff distance was determined as the minimum in the proximal distribution function and (b) "bulk" water, defined by a distance of water oxygen atom to all lignin heavy atoms >4.9 Å.

The proximal distribution function $g_{prox}(r)$ is given by[58]

$$g_{prox} = \frac{\langle n \rangle}{A(r)\Delta r \rho_{bulk}} \quad (6)$$

where $\langle n \rangle$ is the average number of water oxygen atoms found at a distance $[r, r + \Delta r]$ from a non-hydrogen atom on the surface of the lignin (here Δr = 0.1 Å), $A(r)$ is the SASA of the lignin calculated with a probe radius r, and ρ_{bulk} = 0.0327, 0.0307, 0.0280, and 0.0242 Å^{-3} are the bulk number-densities of the TIP3P water model determined from four separate pure water simulations at T = 300, 360, 420, and 480 K, respectively. The product $A(r)\Delta r$ is the approximate volume of the shell of water molecules whose distance to the lignin surface is between r and $(r + \Delta r)$ Å. The number of hydration water molecules, W, is determined by counting the number of oxygen water atoms that are within a distance of 4.9 Å to any lignin non-hydrogen atom.

The isothermal compressibility, χ, is a measure of the density fluctuations of water in the grand canonical ensemble[59]

$$\chi = \frac{V}{k_B T}\frac{\langle W^2 \rangle - \langle W \rangle^2}{\langle W \rangle^2} \quad (7)$$

where W is the number of water molecules in (fixed) volume V. The compressibility of the lignin hydration water, χ_{hydr}, was calculated by counting the number of water molecules in the hydration shell as a function of time. To derive a value of χ_{bulk}, the compressibility of bulk water, that can be directly compared to χ_{hydr}, the lignin surface was superimposed onto the results of a control simulation of pure water,[30] and the number of water molecules inside the hydration shell volume was calculated as a function of time. Values of both χ_{hydr} and χ_{bulk} were derived from 60 independent 50 ps simulations at each temperature: in 30 of these lignin was in extended states and in the other 30 lignin was collapsed, see Figure S1 of the Supporting Information.

Entropy of Water. A two-phase thermodynamic model was employed that partitions the translational and rotational density of states of water molecules $g(\omega)$ into gas-like, $g^g(\omega)$, and solid-like, $g^s(\omega)$ components:[60,61]

$$g(\omega) = g^g(\omega) + g^s(\omega) \quad (8)$$

$$= \frac{2}{k_B T}\lim_{\tau \to \infty} \int_\tau^\tau C(t) e^{-i\omega t}\, dt \quad (9)$$

Here $g(\omega)$ is the Fourier transform of the velocity-autocorrelation function (VACF) $C(t)$. $C(t)$ is either the mass weighted VACF of the center of mass velocities

$$C_T(t) = \sum_{i=1}^{W} m\langle \vec{v}_i(t) \cdot \vec{v}_i(0) \rangle \quad (10)$$

or the moment of inertia weighted angular VACF

$$C_R(t) = \sum_{j=1}^{3}\sum_{i=1}^{W} \langle I_j \omega_{ij}(t)\omega_{ij}(0) \rangle \quad (11)$$

W is the number of water molecules in the system, m is the mass, $\vec{v}_i(t)$ is the velocity of the ith water molecule at time t, and I_j and ω_{ij} are the jth principal moment of inertia and angular velocity of water molecule, i. The translational entropy of water can be calculated by assigning the appropriate weight, λ, to the gas- and solid-like components:

$$S = \frac{1}{2\pi}\int_0^\infty g^g(\omega)\lambda^g(\omega) + \frac{1}{2\pi}\int_0^\infty g^s(\omega)\lambda^s(\omega) \quad (12)$$

Details on the decomposition of $g(\omega)$ and the derivation of the weighting functions W^g and W^s can be found in the Supporting Information. $C(t)$ (eq 10) for bulk water was determined from 10 pure water 20 ps simulations, where velocities and coordinates were saved every 4 fs. $C(t)$ for the hydration-shell water molecules was determined from 10 lignin-water 20 ps simulations, each starting from a different lignin configuration, from which data were saved every 4 fs. The sum in eqs 10 and 11 was taken over hydration-shell water molecules only.

A summary of the simulations used In the analysis of the thermodynamics of the collapse transition at 300 K is provided in the Supporting Information.

RESULTS

Structure. The radius of gyration, R_g, (Figure 1a) for all the lignin molecules exhibits a strong temperature dependence above $T \approx 420$ K with $R_g \approx 15$ Å for $T \leq 420$ K and $R_g \approx 17$ Å for T = 480 K. Figure 2 graphically represents the polymer chain at 300 and 480 K. The probability distribution of R_g at T = 480 K not only has a higher mean, but also is considerably broader than at T = 300 K (Figure S2a in the Supporting Information), thus allowing the chain to increase entropy by exploring more conformations at high temperatures. Conversely, the highest number of intermolecular lignin contacts is found in the low temperature collapsed state (Figure 1d). Surprisingly, lignin also makes the highest number of hydrogen bonds to water at low temperatures (Figure 1d, orange), although possessing the lowest solvent accessible surface area (SASA), Figure S2b. As discussed later, this is due to a higher hydration-water density at low temperatures.

Figure 1b shows the volume within the SASA, V_{mol}, and the volume of the gyration ellipsoid, V_{elp}. No difference between branched or unbranched lignins is seen for either of these quantities. At the two lower temperatures (300 and 360 K) the volume of the gyration ellipsoid is a good approximation to V_{mol}, overestimating the volume by approximately the error bar. However, for the two higher temperatures (420 and 480 K) the gyration ellipsoid overestimates V_{mol} significantly. The fact that V_{elp} does not vary significantly with temperature excludes the presence of cavities at high temperatures. Therefore, the increase in R_g and V_{mol} can be understood by the polymer adopting more extended conformations.

Scaling Properties. The scaling concept of polymers[62] is now employed to examine the size of a linear segment of lignin

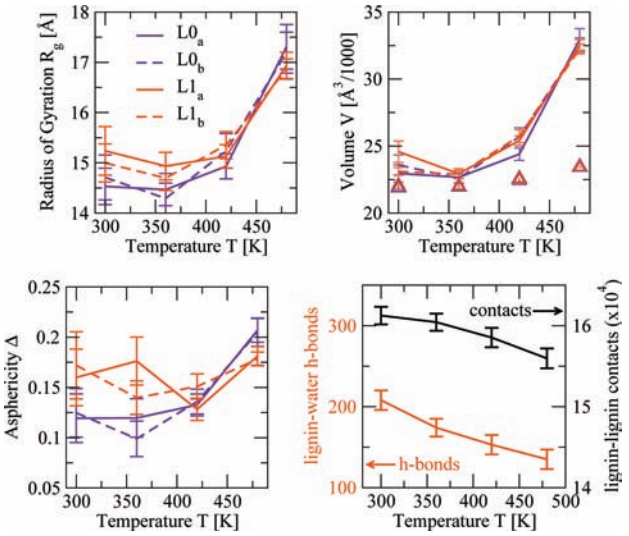


Figure 1. Temperature dependence of structural properties of unbranched ($L0$) and branched ($L1$) lignins: (a) radius of gyration R_g, (b) SASA volume (V_{sas}, triangles) and gyration ellipsoid volume (V_{elp}, lines), (c) asphericity Δ defined by eq 1, and (d) the lignin intramolecular contacts (black; squares) and lignin−water hydrogen bonds (orange; diamonds). Each data point is derived from 10 trajectories. The average values of each trajectory were first computed, and the means of these 10 average values are listed, with the error estimated as $\sigma\sqrt{9}$, where σ is the standard deviation of the mean. For V_{mol} the values of the four molecules mostly overlap, and the error estimate is not shown because it is similar to the line thickness.

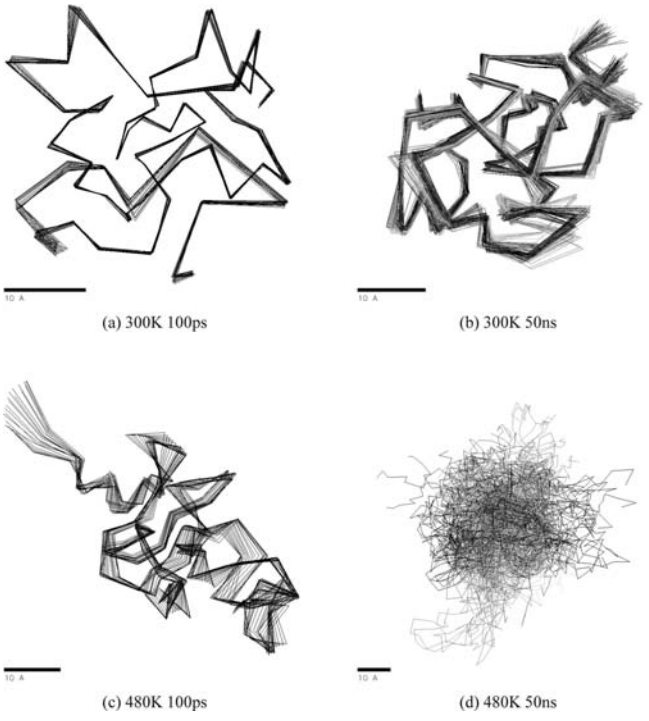


Figure 2. Structure of $L0_a$ at 300 and 480 K. Shown are lines connecting the center of mass of each residue. (a, c) The structures obtained from a 100 ps trajectory saved every 5 ps (smoothed over 5 frames). (b, d) Structures obtained from a 50 ns trajectory saved every 500 ps (smoothed over 100 frames). Time is mapped using a black (beginning of trajectory) to white (end of trajectory) scale.

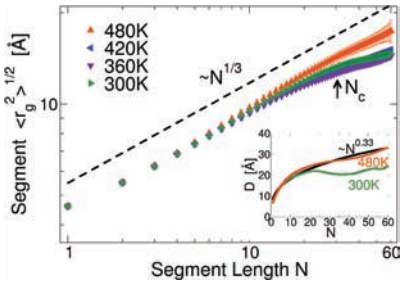


Figure 3. Scaling properties of the unbranched lignins at different temperatures. Root mean square of the radius of gyration $r_g(N)$ of a polymer segment comprising $(N + 1)$ monomers. Inset: Distance D between monomers i and j as a function of $N = |i - j|$. In all plots, the dashed black line is a $\sim N^{0.33}$ power-law function, and all plots are time averages over the last 50 ns of the ensemble of 20 MD trajectories. The error bars are the standard deviations of the ensemble distribution.

polymers comprising $(N + 1)$ monomers. For the ensemble of the two unbranched lignins at $T = 300$ K, the radius of gyration of the corresponding segment (denoted as $r_g(N)$ so as to distinguish it from $R_g = r_g(N_{tot})$ the radius of gyration of a whole polymer) is found to follow a power-law behavior for $N \lesssim N_c = 30$ (Figure 3)

$$r_g \propto N^{\nu} \tag{13}$$

where the scaling exponent $\nu = 0.34 \pm 0.01$. The power-law indicates self-similar packing density for short segments ($N \lesssim N_c$). For larger segment lengths, $N \gtrsim N_c$, the size of the segments increases more slowly with length and r_g approaches a plateau $r_g \sim N^0$.

The flattening of $r_g(N)$ for long segments ($N > N_c \simeq 30$) is a feature of linear homopolymers comprised of "blobs".[14] Blobs are polymer segments where monomers proximal along the chain are likely to be proximal in space. Inside a blob, the distance between polymers is a power-law function of their separation along the chain, and therefore, eq 13 is valid. Blobs can penetrate each other; i.e., monomers belonging to different blobs can be in close spatial proximity, giving rise to the plateau $r_g \sim N^0$ for low temperatures. At 300 K the blobs are in close spatial proximity, and the average distance, D, between monomers belonging to the same blob, is comparable to D between monomers on different blobs (inset of Figure 3, compare $N \simeq 11$ with $N \simeq 41$). The crossover length, equal to the characteristic segment length of a blob $\sim N_c$, can be determined by the condition that the size of the blob is of the order of the polymer size, i.e., $N_c \approx (R_g/a)^3 = 31$, where $a = 4.6$ Å is the radius of gyration of a single monomer.[14]

While $r_g(N)$ at $T = 300$, 360, and 420 K displays the crossover at N_c, at the highest temperature ($T = 480$ K) the power-law behavior of eq 13 is observed over the entire range of N (Figure 3). This self-similar fractal spatial chain structure, called a "crumpled globule", arises when the "blobs" are spatially segregated from each other.[63,64] The key difference between the crumpled and low-T globules is that, in the former, monomers distant along the chain are distant in space, whereas for the latter monomers distant along the chain have a relatively high probability of being proximal in space. Recently, chromatin was also found to exist in crumpled globules.[65]

The ensemble of the linear segments of the two branched lignins (called $L1$) also displays the power-law dependence of eq 13 for $N \lesssim N_c$, with an exponent $\nu = \nu_{L1} = 0.033 \pm 0.01$, and a

crossover behavior for $N \gtrsim N_c$ at low temperatures (Figure S3 of the Supporting Information). $r_g(N)$ was also computed for the ensemble (called $L1_{br}$) of short ($N \leq 6$) segments that contain the branch point, either on the ends of the segment or in its interior. $r_g(N)$ values of $L1$ and $L1_{br}$ are statistically indistinguishable, and therefore, the presence of the branch point does not alter the scaling behavior of collapsed lignins.

Effect of Branching. Lignins are randomly branched, and it is therefore of interest to examine how branching affects the size and shape of the polymers. The ratio, $g = \langle R_g^2 \rangle_{L1} / \langle R_g^2 \rangle_{L0}$, of the mean square radius of gyration of the branched polymer ($L1$) to that of a linear polymer having the same molecular weight ($L0$) quantifies the effect of branching on the polymer size.

For polymers in a good solvent, branched chains assume a smaller R_g than their linear counterparts of the same molecular weight.[66,67] The theory by Zimm and Stockmayer (ZS)[68] predicts $g < 1$ for various branching configurations of polymers in ideal solvents with the assumption of isotropy, and has been employed successfully to interpret experimental data, e.g.[67,69] In the Supporting Information we modify the ZS theory to render it applicable to collapsed polymers in bad solvents (eq 0.8) and demonstrate that, while the decrease in polymer size due to branching is predicted to be smaller for bad solvents than for ideal solvents, g remains <1. For example, with the modified ZS for a star polymer of three equal-length arms, $g = 0.86$ for a bad solvent compared to $g_{ZS} = 0.77$ for an ideal solvent.

Interestingly, the present simulations show the unbranched lignins to mostly have smaller R_g than the branched lignins. Hence, for lignin at 300 K $g > 1$. This is because, in contrast to the assumption of anisotropy made in the analytical theory, the simulated lignins are not spherical. The anisotropy of the chain conformations can be described by the asphericity Δ (eq 1), that takes values between $\Delta = 0$, corresponding to a spherical shape, and $\Delta = 1$, for a rod-like shape. Table 1 offers an explanation of the trend in the calculated radii of gyration: the more aspherical the molecule, the higher its R_g (see also Figure S4 in the Supporting Information).

To compare the current simulations to the ZS theory, which assumes $\Delta = 0$, nearly spherical ($\Delta < 0.05$) polymer configurations were further examined. At 300 K (Figure S4c) the branched lignins have larger R_g than the unbranched, and thus $g > 1$ and the collapsed polymers do not conform to the ZS theory. In contrast, at 480 K (Figure S4d) $g < 1$ and ZS prediction is verified as expected for near-spherical extended polymers.

The shape of a polymer is further characterized by the prolateness Σ (eq 2). Most of the present lignins have a prolate "melon-like" configuration with $\Sigma > 0$, the exception being molecule $L0_a$, which adopts both prolate ($\Sigma > 0$) and oblate "disk-like" ($\Sigma < 0$) configurations.

The temperature dependence of R_g (Figure 1a) is different for the branched and unbranched lignins, such that at low T the R_g is higher for the branched molecules ($g > 1$) whereas for high T it is higher for the unbranched lignins ($g < 1$). This difference in the temperature trend between the $L0$ and $L1$ lignins correlates with the asphericity, Δ, shown in Figure 1c. Hence, at low T the branch points lead to increased asphericity and thus a larger R_g. The asphericity for the branched lignins has no significant temperature dependence. For both the branched and unbranched lignins the temperature dependence is smaller than what would be expected for a constant-mass ellipsoid the R_g of which displays the temperature dependence of Figure 1a.

Structure of Hydration Water. Structural properties of water close to the surface of the lignin are quantified by the proximal distribution function, $g_{prox}(r)$, given by eq 6. At low temperatures $g_{prox}(r)$ is significantly structured displaying one peak at 2.8 Å and another at 3.5 Å (Figure 4a). Increasing temperature leads to a gradual loss of local hydration layering, as shown by the decrease of the hydration peaks in Figure 4a. However, the

Table 1. Structural Properties of the Lignin Molecules with Various Number of Branch Points (Indicated by the Number in the Molecule Name)[a]

name	R_g	$\langle\Delta\rangle$	$\langle\Sigma\rangle$
$L0_a$	14.0 ± 0.1	0.08 ± 0.01	−0.25 ± 0.23
$L0_b$	15.5 ± 0.3	0.15 ± 0.03	0.49 ± 0.16
$L1_a$	14.4 ± 0.4	0.08 ± 0.03	0.51 ± 0.16
$L1_b$	15.2 ± 0.8	0.12 ± 0.03	0.02 ± 0.30
$L2$	16.0 ± 0.6	0.20 ± 0.04	0.65 ± 0.06
$L3$	15.1 ± 0.5	0.12 ± 0.04	0.25 ± 0.35
$L4$	14.4 ± 0.2	0.10 ± 0.02	0.30 ± 0.11
$L5$	17.6 ± 0.3	0.20 ± 0.02	0.11 ± 0.21
$L6$	14.0 ± 0.2	0.06 ± 0.01	0.49 ± 0.28

[a] R_g is the radius of gyration, Δ the asphericity (eq 1), and Σ the prolateness (eq 2) at $T = 300$ K. Quantities are averaged over the last 50 ns of the simulations. For molecules with zero and one branch points, the average value of R_g, Δ, and Σ of each trajectory was first computed and the mean of these 10 average values is listed, with the error estimated as $\sigma\sqrt{9}$, where σ is the standard deviation of the mean. For molecules with more than two branch points the error is estimated from the standard deviation of the averaging of the single trajectory.

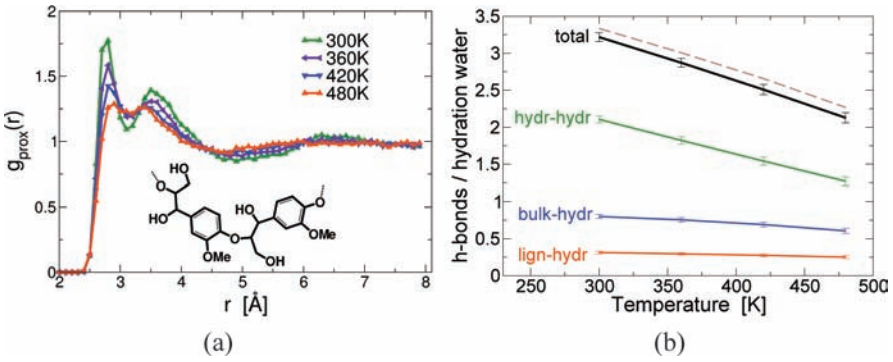


Figure 4. (a) Proximal distribution function of water oxygen atoms at a distance r from the surface of the lignin. Data are averaged over the last 50 ns of the 10 MD trajectories of the $L0_a$ lignin with no branch points. Inset: Chemical structure of two bonded lignin monomers. (b) Average number of hydrogen bonds a hydration water molecule makes with other hydration-shell waters, bulk water, and lignin. The dotted brown line is the number of H-bonds per bulk water molecule.

average fraction of the SASA which is hydrophilic is $\langle\phi\rangle_{pol}$ = 0.43 ± 0.01 for all temperatures, for a "hydrophilic" atom crudely defined as having partial charge $|q| > 0.2e$. Therefore, lignin collapse at low T is not associated with decreased exposure of hydrophobic moieties, in contrast to what is seen for proteins.

The density of the lignin hydration shell, ρ_{sh}, was derived using eqs 0.9 and 0.10 in the Supporting Information. Figure S5 in the Supporting Information shows the hydration shell density of lignin to be smaller than the bulk by 2%, 4%, 7%, and 10% at T = 300, 360, 420, and 480 K, respectively. This is consistent with the surface of lignin being "wet" at 300 K and lignin making fewer hydrogen bonds to water at high temperatures (Figure 1d). A hydration shell water molecule participates on average in only 3% fewer hydrogen bonds than does a bulk water molecule (Figure 4b), similar to the hydration of small hydrophobic solutes that also do not perturb significantly the H-bond network of the surrounding water.[18]

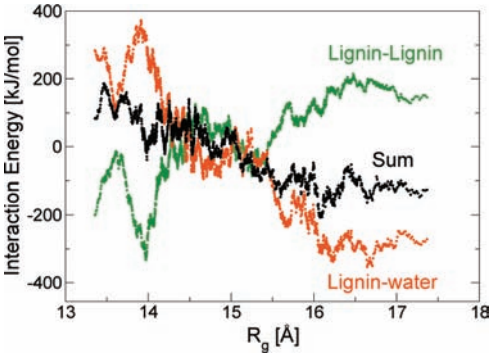


Figure 5. Lignin−lignin and lignin−water interactions energies as a function of the lignin R_g at 300 K. Data represent ensemble average of the unbranched and one-branch lignins.

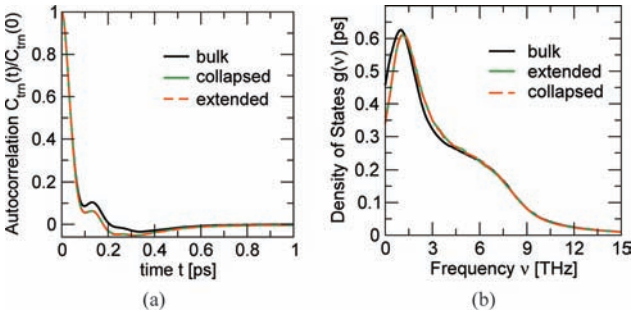


Figure 6. (a) Translational velocity autocorrelation functions and the respective (b) density of states of water.

Thermodynamics of the Collapse Transition. The thermodynamics of the transition of lignin from an extended conformation ($R_g = R_{ext}$, SASA= A_{ext}, $W = W_{ext}$) to a collapsed conformation ($R_g = R_{col} < R_{ext}$, SASA = $A_{col} < A_{ext}$, $W = W_{col} < W_{ext}$) at 300 K are now examined.

Enthalpy Change. For the compact lignin structures, $R_g \lesssim$ 14.2 Å, that are most frequently sampled in the 300 K simulation, the intralignin interaction energy is negative (Figure 5, see Figure S6 for a similar plot at 480 K), but the lignin−water interaction energy is positive and of larger magnitude. In contrast, for extended lignin structures ($R_g \gtrsim$ 15.3 Å) that are rarely sampled, the lignin−water interaction energy is now negative and outweighs the positive intralignin interaction, thus leading a net negative interaction energy. Thus, the collapse transition of lignin from R_g = 16.3 to 14.1 Å (Figure S1) is enthalpically strongly disfavored, by $\Delta H \sim$ 200 kJ/mol (see Figure 5).

Entropy of Hydration Water. Differences in water dynamics between the bulk and the hydration shell water molecules in the collapsed and extended lignin simulations can be analyzed by computing translational and rotational velocity autocorrelation functions (VACF) and the associated density of states, Figure 6. The translational VACFs of the hydration shell of extended and collapsed lignins are both similar to and different from the bulk (Figure 6a). Negative values of the VACF are due to water molecules rebounding from collisions with their neighbors. The deeper negative minimum in the hydration shell arises from water confinement by the lignin surface, and has been also observed for hydration water of proteins,[70,71] DNA,[33] and lipid bilayers.[32]

A peak in the translational density of states represents the population of a mode of a given frequency (Figure 6b). Therefore, the slight shift of the main peak of the hydration water toward higher frequencies arises from water molecules on the surface of the lignin that librate at a higher frequencies than the bulk.[32,71] Also, the translational diffusion coefficient

$$D_T = \frac{k_B T g(\omega = 0)}{12Wm} \tag{14}$$

of hydration water is smaller than the bulk, supporting the idea that water molecules on the lignin surface are translationally restricted (see Table 2). The rotational VACF and $g(\omega)$ spectra show almost no variation between the bulk and hydration water (Figure S7 in the Supporting Information), although the rotational diffusion coefficient is also higher for the bulk (see Table 2).

Consistent with the above, the translational and orientational entropies per water molecule, calculated using eq 12, increase when going from the hydration water to the bulk (Table 2).

Table 2. Comparison of Entropy and Fluidity of Bulk and Hydration Water of Collapsed and Extended Lignin Structures at 300 K

		bulk	hydration collapsed	hydration extended
trans. diffusion const. D_T	[Å^2/ps]	0.52 ± 0.03	0.41 ± 0.04	0.39 ± 0.03
rotational diffusion const. D_R	[1/ps]	0.48 ± 0.02	0.43 ± 0.02	0.41 ± 0.01
trans. fluidicity factor f_T		0.34 ± 0.01	0.30 ± 0.01	0.29 ± 0.01
rotational fluidicity factor f_R		0.078 ± 0.001	0.074 ± 0.002	0.071 ± 0.001
translational entropy TS_T	[kJ/mol]	16.32 ± 0.06	15.79 ± 0.03	15.76 ± 0.08
rotational entropy TS_R	[kJ/mol]	3.952 ± 0.008	3.951 ± 0.009	3.925 ± 0.006
total entropy TS	[kJ/mol]	20.28 ± 0.06	19.74 ± 0.04	19.69 ± 0.09
entropy diff WT($S_{bulk} - S$)	[kJ/mol]		333 ± 22	424 ± 34

dx.doi.org/10.1021/ja206839u |J. Am. Chem. Soc. 2011, 133, 20277–20287

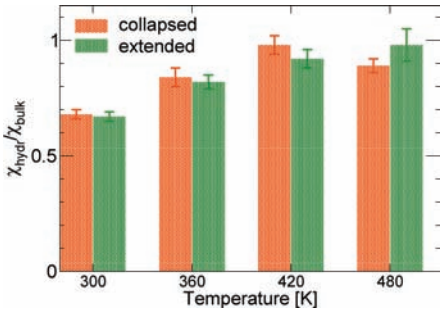


Figure 7. Relative compressibility of hydration water. For each temperature, data are derived from 60 50 ps MD trajectories: 30 with lignin in a collapsed state and 30 in an extended state. The uncertainty bars correspond to the error of the mean over the 30 trajectories.

Therefore, the hydration water molecules are entropically unfavorable compared to the bulk, by an associated free energy of ~0.57 kJ/mol per molecule. However, no difference is found between the water hydrating extended and collapsed lignins. Consequently, when lignin collapses from the extended state, with W_{ext} water molecules in its solvation shell, to the collapsed state with $W_{col} < W_{ext}$ water molecules, then a number $(W_{ext} - W_{col})$ of hydration waters are released to bulk, and the associated free energy change is thus given by

$$\Delta G_{trns} = (W_{ext} - W_{col})T(S_{hydr} - S_{bulk}) \quad (15)$$

Thus, ΔG_{trns} favors lignin collapse by $\Delta G_{trns} = -91$ kJ/mol at $T = 300$ K.

Hydration Water Density and Compressibility. The proximal distribution functions and, therefore, the density of the hydration water of extended and collapsed lignins are identical at 300 K (Figure S5e in the Supporting Information). This demonstrates that the lignin collapse is not accompanied by dewetting. Similar behavior is found for "small" hydrophobic solutes, as defined in the length-scale dependent theory of hydrophobicity,[18,24] where hydrogen bonds between water molecules solvating a hydrophobic solute remain intact. Therefore, the solvation free energy has an entropic origin and is given by the excess chemical potential, $\Delta\mu$, required to create a cavity of volume V in water of compressibility χ[26,27]

$$\Delta\mu = \frac{V}{2\chi} + \frac{k_B T}{2}\log(2\pi\sigma_W^2) \quad (16)$$

where $\sigma_W = \langle W^2\rangle - \langle W\rangle^2$. $\Delta\mu$ in eq 16 is roughly proportional to V.[18] Therefore, since the volumes of the extended and collapsed states are the same, the entropic penalty for creating a cavity to accommodate the collapsed conformation is equal to that for the extended state.

However, at 300 K the compressibility of the lignin hydration water, χ_{hydr}, is lower than that of bulk water, χ_{bulk} (Figure 7). The entropic cost of surrounding lignin with the less compressible hydration water equals the excess chemical potential of creating a cavity in bulk water (the volume of which equals that of the solvation shell) and filling it with hydration water

$$G_{fluc} = \frac{W}{2\rho}\left(\frac{1}{\chi_{hydr}} - \frac{1}{\chi_{bulk}}\right) + \frac{k_B T}{2}\log\left(\frac{\chi_{hydr}}{\chi_{bulk}}\right) \quad (17)$$

where the hydration shell has volume $V = W\rho$, density ρ, and contains W water molecules. Thus, the associated free energy change when lignin collapses from the extended state, the solvation shell of which has W_{ext} water molecules, to the collapsed state with $W_{col} < W_{ext}$ is given by

$$\Delta G_{fluc} = \frac{W_{col} - W_{ext}}{2\rho}\left(\frac{1}{\chi_{hydr}} - \frac{1}{\chi_{bulk}}\right) \quad (18)$$

The ratio χ_{hydr}/χ_{bulk} is 0.67 ± 0.01, 0.83 ± 0.02, 0.95 ± 0.03, and 0.93 ± 0.04 at 300, 360, 420, and 480 K, respectively. Thus, collective solvent density fluctuations favor the lignin collapse ($\Delta G_{fluc} < 0$) at $T \leq 360$, but do not influence the transition at high temperatures where $\chi_{hydr} \simeq \chi_{bulk}$. Substituting the parameters of Figures 7 and S1 in eq 18 gives a $\Delta G_{fluc} = -565 \pm 70$ kJ/mol at 300 K.

Water compressibility has been recently identified as a measure of the hydrophobicity of molecular surfaces:[72−74] with the more hydrophobic the surface the larger the compressibility. Figure 7 would then imply that at 480 K lignin is slightly more hydrophobic than at 300 K. The above temperature trend is consistent with the hydration water density being only 2% lower than bulk at 300 K, but 10% lower at 480 K. Interestingly, the fraction of hydrophilic SASA, $\langle\phi\rangle_{pol} = 0.42 \pm 0.01$, is the same for both the collapsed and extended conformations at all temperatures. This suggests that although the polarity of the surface groups is the main factor determining hydrophobicity of a lignin polymer, the total SASA of the polymer, which is larger at high temperatures (Figure 3), may play a secondary role.

Conformational Entropy of Lignin. The extended lignin structures have larger conformational entropy, and the entropy change going from an extended (R_{ext}) to a collapsed (R_{col}) state can be estimated from the entropy penalty of confining a Gaussian polymer of R_{ext} to a volume $L \sim D^3$ of characteristic dimension $L = R_{col}$, $\Delta G_{conf} \simeq k_B T(\pi^2/3)(D/R_{ext})^2 = k_B T(\pi^2/3)(R_{col}/R_{ext})^2$.[75] Using this method, here $\Delta G_{conf} = 11$ kJ/mol.

Free Energy Change of Lignin Collapse. Summing all the contributions discussed in detail above, the total free energy change of lignin collapse at 300 K is approximately

$$\Delta G \simeq \Delta H + \Delta G_{trns} + \Delta G_{fluc} + \Delta G_{conf} \quad (19)$$

The above calculations give $\Delta H \simeq 200$ kJ/mol, $\Delta G_{trns} \simeq -90$ kJ/mol, $\Delta G_{fluc} \simeq -565$ kJ/mol, and $\Delta G_{conf} \simeq 10$ kJ/mol, and therefore, the overall free energy of collapse is $\Delta G \simeq -450$ kJ/mol. Hence the release of entropically unfavorable hydration shell water molecules leading to ΔG_{trns} and ΔG_{fluc} drives the lignin collapse. This mechanism of collapse is different from that usually considered for coil to globule transitions, in which the favorable enthalpy gain arising from increased monomer−monomer contacts in the collapsed state compensates for the decrease in chain configurational entropy.[76]

Lignin Chain Dynamics. Time-dependent monomer mean-square displacements (MSDs, eq 5) from the ensemble of unbranched polymers are shown at four temperatures in Figure 8a. At low temperatures ($T = 300$ and 360 K) Figure 8a exhibits three regimes, typical of glass-forming polymers:[77,78] (i) ballistic ($t \lesssim 30$ ps); (ii) a region (100 ps $\lesssim t \lesssim$10 ns), in which $\langle\Delta r_n^2\rangle \sim t^\beta$ with $\beta_{300K} = 0.29$ and $\beta_{360K} = 0.42$, reflecting the temporary confinement of the monomers by their nearest neighbors in a "caging" effect, (iii) subdiffusive $t \gtrsim 10$ ns where $\beta \simeq 0.5$, consistent with the Rouse theory of unbranched polymer melts.[79] With increasing temperature the caging effect decreases, and at high temperatures ($T = 420$ and 480 K), the caging plateau disappears.[77] The MSD exponent of $\beta_{420K} = 0.58$

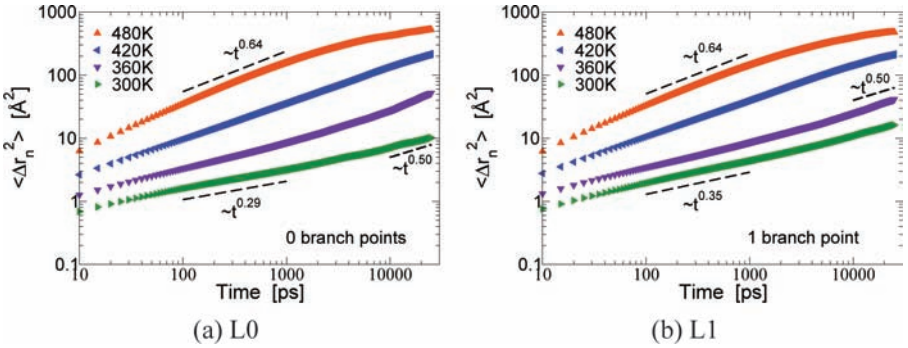


Figure 8. (a) Mean square displacements (MSDs) of the ensemble of lignins with no branch points, *L*0, at four temperatures, with translation and rotation of the entire molecule removed. MSD for the ensemble with one branch point, *L*1, is shown in part b.

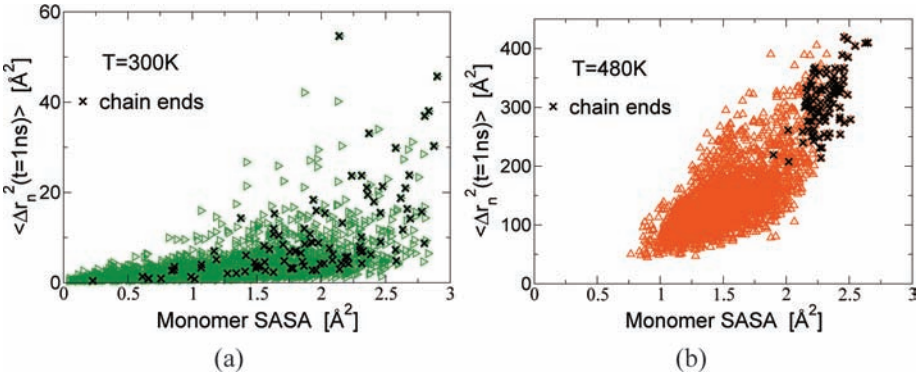


Figure 9. Monomer MSD at t = 1 ns of the ensemble of polymers with zero and one branch points versus the monomer SASA.

and $\beta_{480K} = 0.64$ is similar to that found for unbranched polymer melts above their glass transition temperature ($\beta = 0.61$).[80] The increased chain mobility is also apparent in the graphical representations of Figure 2.

The MSD of the one-branch-point ensemble, Figure 8b, displays characteristics similar to those for the zero branch polymers in Figure 8a. The power-law behavior exponents $\langle \Delta r_n^2 \rangle \sim t^{\beta}$ are now $\beta_{300K} = 0.35$, $\beta_{360K} = 0.40$, $\beta_{420K} = 0.58$, and $\beta_{480K} = 0.64$. Overall, the MSD observed here at T = 300 K is similar to that found in a previous MD study of lignin aggregates.[35]

Representative MSDs of individual monomers belonging to the same chain from a single MD simulation at T = 300 K for an unbranched and branched lignin molecules are shown in Figure S8 in the Supporting Information. Significant variations are seen. Not all monomers are equally exposed to the solvent, leading to a range of solvent accessible surface areas (SASA) per monomer. A general trend is observed, in which monomers with smaller SASAs tend to have the lower $\langle \Delta r_n^2 \rangle$. Interestingly, monomers at the ends of the chains do not display the fastest dynamics; nor do those at branch points exhibit the slowest dynamics. Similar behavior of $\langle \Delta r_n^2 \rangle$ is observed for lignins with branch points.

The dependence of the mobility of a monomer on its solvent exposure is stronger at high temperatures (Figures 9 and S9 in the Supporting Information). The slopes of approximate linear regressions at T = 300, 360, 420, and 480 K are 3, 6, 30, and 100, respectively. Furthermore, the strength of the correlation between monomer MSD and SASA also increases with temperature, with approximate linear regression χ^2 coefficients in Figures 9a, S9a, S9b, and 9b of 0.53, 0.55, 0.62, and 0.75, respectively.

DISCUSSION

Pretreatment is responsible for a significant fraction of the production cost of cellulosic ethanol from lignocellulosic biomass, due to the associated energy requirements as well as capital and operating costs. Available information on pretreated lignin includes how much remains after pretreatment,[81,82] its chemical composition,[83−86] and the size of the lignin aggregates that form after pretreatment.[35,37,39,41] Commonly employed experimental methods to examine lignin include analytical chemistry,[81,82] NMR,[83] neutron scattering,[35,37] and electron microscopy exploring the micrometer scale.[38−41] The atomistic simulations presented here complement these studies by providing a detailed description of the temperature-dependent change in structure and dynamics of individual lignin molecules on the nanometer scale.

At low temperatures ($T \lesssim 420$ K) lignins are found to be compact ellipsoidal objects, as indicated by the agreement between V_{mol} and V_{elp} (Figure 1), of low solvent accessible surface area. Contrary to what is found for polymers in good solvents, branching does not lead to reduction of the lignin R_g, and instead a strong correlation is found between the lignin R_g and its asphericity (Table 1), as one expects for compact ellipsoids. The polymers are composed of interpermeable ∼30-monomer blobs, with power-law chain statistics observed only inside the blobs (Figure 3). The polymer state described above is often termed "equilibrium globule".[63] The compact lignin structures at ≲420 K found here are consistent with small-angle neutron scattering studies showing lignin aggregation after dilute acid pretreatment at ∼120 K.[37]

Above 420 K the lignin molecules expand to nonellipsoidal, extended forms without cavities, leading to sharp increases in

R_g, V_{elp}, and A_{SAS} but no significant change in V_{mol} or the asphericity Δ (Figure 1). The blobs become spatially separated, leading to self-similar chain packing, i.e., the same for segments of all lengths. This self-similar polymer state is often termed "crumpled globule",[63] and the temperature-driven transition from equilibrium to crumpled globules has been reported previously in other polymer simulations.[14,64,87,88]

To characterize the thermodynamics of the lignin transition from extended to compact states at 300 K, the various contributions to the free energy of collapse were investigated. The enthalpy change is positive (unfavorable), because the lignin−water interaction, favoring the extended state, is stronger than the lignin−lignin interaction that favors the collapse (Figure 4). The lignin conformational entropy also does not favor the collapse, as extended states sample more configurational space. Hydration shell water molecules were found to have lower translational entropy than the bulk, while their rotational entropy was similar (Table 2). Additionally, hydration water has slightly lower density fluctuations than the bulk, making the latter entropically more favored. Protein unfolding, accompanied by exposure of core hydrophobic residues to water, has been recently associated with increase in hydration shell compressibility.[72] Here the hydration water of collapsed and extended lignins has a similar compressibility, which, critically, is lower than that of the bulk.

The collapse of lignin from extended (large SASA) to compact (small SASA) states is accompanied by the displacement of hydration water molecules to the bulk. Therefore, the simulations demonstrate that the release of entropically unfavorable hydration water molecules into the bulk is the driving force behind the collapse of lignin at 300 K.

Previous MD simulations of nonpolar hydrophobic polymers in water found the hydration contribution to the free energy of collapse, defined as the total free energy minus the polymer−water interaction energy and intrapolymer interaction energy and entropy, to favor the collapse and to be enthalpy dominated.[21] The present lignin simulations are of larger polymers than those of ref 21, and, unlike ref 21, include water−polymer attractive Coulombic interactions. The results also show the hydration contribution (here approximated by $\Delta G_{trns} + \Delta G_{fluc}$) to drive the collapse, but to be entropy driven. The present simulations also complement previous calculations on the kinetics of hydrophobic collapse that found collective water density fluctuations as the rate limiting step in the collapse.[22] Here, water density fluctuations are found to also contribute significantly to the thermodynamics of lignin collapse.

Although important biological functions of lignin, such as water conduction and cell wall defense against enzymatic digestion, stem directly from its hydrophobicity, the lignin surface is shown here to be wetted by water, as indicated by the lignin hydration water density being equal to the bulk (Figure S5a). Furthermore, the suppression of hydration water compressibility (Figure 7) is similar to that found near hydrophilic surfaces.

Temperature-induced changes in lignin dynamics are highlighted by the disappearance of the intermediate-time (100 ps to 10 ns) plateau in the monomer mean-square displacements for temperatures above 360 K. The low-temperature plateau, reflecting a temporary localization of the monomers by their nearest neighbors, is similar to that found for compact homopolymers that exhibit glassy behavior.[77] Interestingly, the onset of constrained dynamics below 360 K occurs at lower temperature than the collapse transition at 420 K. Monomers exposed to the solvent experience smaller friction than those buried in the core of the lignin, leading to a correlation between a monomer's SASA and its mobility (Figure 9). Therefore, the onset of the dynamic transition above 360 K coincides with an increase with the overall lignin solvent exposure (Figure S2b). For polymer melts above their Θ-temperature (the temperature at which the second viral coefficient disappears and the polymer acts as an ideal Gaussian coil) chain connectivity determines monomer mobility, with chain ends exhibiting faster dynamics,[80,89] something not always observed in Figures 9a and S9a.

MD simulation might conceivably be combined with a lignin polymerization model that predicts the reaction conditions controlling the primary structure of lignin polymers.[90] Depending on the primary sequence of the oligomer, only some of potentially bindable sites are considered free to participate in monomer−oligomer coupling that leads to extension of the lignin chain.[90] However, some of these free bindable sites may be sterically hindered by neighboring lignin monomers and therefore not available to form bonds with monomers. Information on such sterical hindrance could possibly be provided by MD simulation, such as the present, which provides information on the three-dimensional configurations of lignin polymers.

CONCLUSIONS

Extensive, 17.5 μs simulations of single lignin polymers in aqueous solution have probed the temperature-dependent structural and dynamic changes of this biomass component and its hydration water. With increasing temperature the lignin was found to transition from compact conformations with glassy dynamics to extended conformations with enhanced dynamics. The unfavorable translational entropy of lignin hydration water molecules and the low compressibility of the hydration shell were found to thermodynamically drive the transition from extended to collapsed states at 300 K. The present molecular level understanding of the structure, dynamics, and thermodynamics of lignin as a function of temperature may provide fundamental information needed to help understand biomass pretreatment and thus improve the efficiency of cellulosic ethanol production.

ASSOCIATED CONTENT

Supporting Information. Structural properties of collapsed and expanded lignins, details on the calculation of the entropy of water, analytic theory on the effect of branching on lignin size, correlation between R_g and asphericity, monomer mean square displacements at 300 K, and chain topologies of the nine lignin models. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
smithjc@ornl.gov

Author Contributions
[§]These authors contributed equally to this work.

ACKNOWLEDGMENT

We thank Shiang-Tai Lin for discussion on water entropy. This research was funded by the Genomic Science Program, Office of Biological and Environmental Research, Office of Science, U.S. Department of Energy under Contract FWP ERKP752.

Resources of the National Center for Computational Sciences at Oak Ridge National Laboratory were used under an INCITE award from the DOE Office of Science.

REFERENCES

(1) Grabber, J. H. *Crop Sci.* **2005**, *45*, 820–831.
(2) Jorgensen, H.; Kristensen, J. B.; Felby, C. *Biofuels, Bioprod. Biorefin.* **2007**, *1*, 119–134.
(3) 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.
(4) Yang, B.; Wyman, C. E. *Biofuels, Bioprod. Biorefin.* **2008**, *2*, 26–40.
(5) Irvine, G. M. *Wood Sci. Technol.* **1985**, *19*, 139–149.
(6) Bjurhager, I.; Olsson, A.-M.; Zhang, B.; Gerber, L.; Kumar, M.; Berglund, L. A.; Burgert, I.; Sundberg, B.; Salmen, L. *Biomacromolecules* **2010**, *11*, 2359–2365.
(7) Sun, S. T.; Nishio, I.; Swislow, G.; Tanaka, T. *J. Chem. Phys.* **1980**, *73*, 5971–5975.
(8) Stepanek, P.; Konak, C.; Sedlacek, B. *Macromolecules* **1982**, *15*, 1214–1216.
(9) Kubota, K.; Fujishige, S.; Ando, I. *J. Phys. Chem.* **1990**, *94*, 5154–5158.
(10) Wu, C.; Wang, X. H. *Phys. Rev. Lett.* **1998**, *80*, 4092–4094.
(11) Stockmayer, W. H. *Makromol. Chem.* **1960**, *35*, 54–74.
(12) Flory, P. *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, NY, 1966.
(13) DeGennes, P. *J. Phys., Lett.* **1975**, *36*, 55–57.
(14) Ma, J. P.; Straub, J. E.; Shakhnovich, E. I. *J. Chem. Phys.* **1995**, *103*, 2615–2624.
(15) Zhou, Y. Q.; Karplus, M.; Wichert, J. M.; Hall, C. K. *J. Chem. Phys.* **1997**, *107*, 10691–10708.
(16) Polson, J. M.; Zuckermann, M. J. *J. Chem. Phys.* **2002**, *116*, 7244–7254.
(17) Steinhauser, M. O. *J. Chem. Phys.* **2005**, *122*, 094901.
(18) Chandler, D. *Nature* **2005**, *437*, 640–647.
(19) ten Wolde, P. R.; Chandler, D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6539–6543.
(20) ten Wolde, P. R. *J. Phys.: Condens. Matter* **2002**, *14*, 9445–9460.
(21) Athawale, M. V.; Goel, G.; Ghosh, T.; Truskett, T. M.; Garde, S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 733–738.
(22) Miller, T. F.; Vanden-Eijnden, E.; Chandler, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14559–14564.
(23) Li, I. T. S.; Walker, G. C. *J. Am. Chem. Soc.* **2010**, *132*, 6530–6540.
(24) Lum, K.; Chandler, D.; Weeks, J. D. *J. Phys. Chem. B* **1999**, *103*, 4570–4577.
(25) Berne, B. J.; Weeks, J. D.; Zhou, R. H. *Annu. Rev. Phys. Chem.* **2009**, *60*, 85–103.
(26) Hummer, G.; Garde, S.; Garcia, A. E.; Pohorille, A.; Pratt, L. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8951–8955.
(27) Garde, S.; Hummer, G.; Garcia, A. E.; Paulaitis, M. E.; Pratt, L. R. *Phys. Rev. Lett.* **1996**, *77*, 4966–4968.
(28) Stillinger, F. H. *J. Solution Chem.* **1973**, *2*, 141–158.
(29) Svergun, D. I.; Richard, S.; Koch, M. H. J.; Sayers, Z.; Kuprin, S.; Zaccai, G. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2267–2272.
(30) Merzel, F.; Smith, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5378–5383.
(31) Lin, S. T.; Maiti, P. K.; Goddard, W. A. *J. Phys. Chem. B* **2005**, *109*, 8663–8672.
(32) Debnath, A.; Mukherjee, B.; Ayappa, K. G.; Maiti, P. K.; Lin, S.-T. *J. Chem. Phys.* **2010**, *133*, 174704.
(33) Jana, B.; Pal, S.; Maiti, P. K.; Lin, S. T.; Hynes, J. T.; Bagchi, B. *J. Phys. Chem. B* **2006**, *110*, 19611–19618.
(34) Besombes, S.; Mazeau, K. *Plant Physiol. Biochem.* **2005**, *43*, 277–286.
(35) Petridis, L.; Pingali, S. V.; Urban, V.; Heller, W.; O'Neill, H.; Foston, M.; Ragauskas, A.; Smith, J. C. *Phys. Rev. E* **2011**, *83*, 061911.
(36) Islam, M. F.; Jenkins, R. D.; Bassett, D. R.; Lau, W.; Ou-Yang, H. D. *Macromolecules* **2000**, *33*, 2480–2485.
(37) Pingali, S. V.; Urban, V. S.; Heller, W. T.; McGaughey, J.; O'Neill, H.; Foston, M.; Myles, D. A.; Ragauskas, A.; Evans, B. R. *Biomacromolecules* **2010**, *11*, 2329–2335.
(38) Selig, M. J.; Viamajala, S.; Decker, S. R.; Tucker, M. P.; Himmel, M. E.; Vinzant, T. B. *Biotechnol. Prog.* **2007**, *23*, 1333–1339.
(39) Donohoe, B. S.; Decker, S. R.; Tucker, M. P.; Himmel, M. E.; Vinzant, T. B. *Biotechnol. Bioeng.* **2008**, *101*, 913–925.
(40) Kristensen, J. B.; Thygesen, L. G.; Felby, C.; Jorgensen, H.; Elder, T. *Biotechnol. Biofuels* **2008**, *1*, 10.1186/1754-6834-1-5.
(41) Chundawat, S. P. S.; Donohoe, B. S.; da Costa Sousa, L.; Elder, T.; Agarwal, U. P.; Lu, F.; Ralph, J.; Himmel, M. E.; Balan, V.; Dale, B. E. *Energy Environ. Sci.* **2011**, *4*, 973–984.
(42) Pu, Y.; Zhang, D.; Singh, P.; Ragauskas, A. J. *Biofuels, Bioprod. Biorefin.* **2008**, *2*, 58–73.
(43) Brunow, G.; Kilpelainen, I.; Lapierre, C.; Lundquist, K.; Simola, L. K.; Lemmetyinen, J. *Phytochemistry* **1993**, *32*, 845–850.
(44) Yan, J. F.; Pla, F.; Kondo, R.; Dolk, M.; McCarthy, J. L. *Macromolecules* **1984**, *17*, 2137–2142.
(45) Petridis, L.; Smith, J. C. *J. Comput. Chem.* **2009**, *30*, 457–467.
(46) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Pastor, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(47) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
(48) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
(49) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comp.* **2008**, 435.
(50) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
(51) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Mod.* **2001**, *7*, 306–317.
(52) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
(53) Hess, B. *J. Chem. Theory Comput.* **2007**, *4*, 116–122.
(54) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952–962.
(55) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.
(56) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
(57) Lamb, H. *Higher Mechanics*; The University Press: Cambridge, UK, 1920.
(58) Ashbaugh, H. S.; Pratt, L. R.; Paulaitis, M. E.; Clohecy, J.; Beck, T. L. *J. Am. Chem. Soc.* **2005**, *127*, 2808–2809.
(59) Dadarlat, V. M.; Post, C. B. *J. Phys. Chem. B* **2001**, *105*, 715–724.
(60) Lin, S. T.; Blanco, M.; Goddard, W. A. *J. Chem. Phys.* **2003**, *119*, 11792–11805.
(61) Lin, S. T.; Maiti, P. K.; Goddard, W. A. *J. Phys. Chem. B* **2010**, *114*, 8191–8198.
(62) Gennes, P. G. d. *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca, NY, 1979.
(63) Grosberg, A. Y.; Nechaev, S. K.; Shakhnovich, E. I. *J. Phys.* **1988**, *49*, 2095–2100.
(64) Grosberg, A.; Rabin, Y.; Havlin, S.; Neer, A. *Europhys. Lett.* **1993**, *23*, 373–378.
(65) Lieberman-Aiden, E.; et al. *Science* **2009**, *326*, 289–293.
(66) Freire, J. J. *Branched Polymers II*; Advances in Polymer Science; Springer-Verlag Berlin: Berlin, 1999; Vol. 143, pp 35−112, Review.
(67) Wang, W. J.; Kharchenko, S.; Migler, K.; Zhu, S. P. *Polymer* **2004**, *45*, 6495–6505.
(68) Zimm, B. H.; Stockmayer, W. H. *J. Chem. Phys.* **1949**, *17*, 1301–1314.
(69) Yu, Y. L.; DesLauriers, P. J.; Rohlfing, D. C. *Polymer* **2005**, *46*, 5165–5182.
(70) Abseher, R.; Schreiber, H.; Steinhauser, O. *Proteins: Struct., Funct., Bioinf.* **1996**, *25*, 366–378.

(71) Rocchi, C.; Bizzarri, A. R.; Cannistraro, S. *Phys. Rev. E* **1998**, *57*, 3315–3325.
(72) Sarupria, S.; Garde, S. *Phys. Rev. Lett.* **2009**, *103*, 037803.
(73) Jamadagni, S. N.; Godawat, R.; Garde, S. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 147–171.
(74) Godawat, R.; Jamadagni, S. N.; Garde, S. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15119–15124.
(75) Edwards, S. F.; Freed, K. F. *J. Phys. A: Gen. Phys.* **1969**, *2*, 145–150.
(76) Grosberg, A.; Khokhlov, A. *Statistical Physics of Macromolecules*; AIP Series in Polymers and Complex Materials; AIP Press: College Park, MD, 1994.
(77) Dokholyan, N. V.; Pitard, E.; Buldyrev, S. V.; Stanley, H. E. *Phys. Rev. E* **2002**, *65*, 030801.
(78) Aichele, M.; Gebremichael, Y.; Starr, F. W.; Baschnagel, J.; Glotzer, S. C. *J. Chem. Phys.* **2003**, *119*, 5290–5304.
(79) Doi, M.; Edwards, S. F. *Theory of Polymer Dynamics*; OUP: Oxford, 1983.
(80) Paul, W.; Smith, G. D. *Rep. Prog. Phys.* **2004**, *67*, 1117–1185.
(81) Yang, B.; Wyman, C. E. *Biotechnol. Bioeng.* **2004**, *86*, 88–98.
(82) Wyman, C. E.; Dale, B. E.; Elander, R. T.; Holtzapple, M.; Ladisch, M. R.; Lee, Y. *Bioresour. Technol.* **2005**, *96*, 1959–1966.
(83) Samuel, R.; Pu, Y. Q.; Raman, B.; Ragauskas, A. J. *Appl. Biochem. Biotechnol.* **2010**, *162*, 62–74.
(84) Jung, S.; Foston, M.; Sullards, M. C.; Ragauskas, A. J. *Energy Fuels* **2010**, *24*, 1347–1357.
(85) Studer, M. H.; DeMartini, J. D.; Davis, M. F.; Sykes, R. W.; Davison, B.; Keller, M.; Tuskan, G. A.; Wyman, C. E. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6300–6305.
(86) Fu, C.; Mielenz, J. R.; Xiao, X.; Ge, Y.; Hamilton, C. Y.; Rodriguez, M.; Chen, F.; Foston, M.; Ragauskas, A.; Bouton, J.; Dixon, R. A.; Wang, Z.-Y. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 3803–3808.
(87) Abrams, C. F.; Lee, N. K.; Obukhov, S. P. *Europhys. Lett.* **2002**, *59*, 391–397.
(88) Rostiashvili, V. G.; Lee, N. K.; Vilgis, T. A. *J. Chem. Phys.* **2003**, *118*, 937–951.
(89) Binder, K.; Baschnagel, J.; Paul, W. *Prog. Polym. Sci.* **2003**, *28*, 115–172.
(90) van Parijs, F. R.; Morreel, K.; Ralph, J.; Boerjan, W.; Merks, R. M. *Plant Physiol.* **2010**, *153*, 1332–1344.