Atomistic Simulations of Polydisperse Lignin Melts using Simple Polydisperse Residue Input Generator

Vaidyanathan Sethuraman[1], Josh V. Vermaas[2], Luna Liang[3], Arthur J. Ragauskas[3,4,5], Jeremy C. Smith[1,6], Loukas Petridis[1]

[1]Center for Molecular Biophysics, Oak Ridge National Laboratory, 1-Bethel Valley Road, Oak Ridge, TN-37830, USA

[2]MSU-DOE Plant Research Laboratory and Department of Biochemistry and Molecular Biology, Michigan State University, 612 Wilson Road, East Lansing, MI-48824, USA

[3]Department of Chemical & Biomolecular Engineering, University of Tennessee Knoxville, Knoxville, TN 37996, USA

[4]Department of Forestry, Wildlife, and Fisheries, Center for Renewable Carbon, University of Tennessee Institute of Agriculture, Knoxville, TN 37996, USA

[5]UTK–ORNL Joint Institute for Biological Science, Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

[6]Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee, TN 37996, USA

Abstract

Understanding the physics of lignin will help rationalize its function in plant cell walls as well as aiding practical applications such as deriving biofuels and bioproducts. Here, we present SPRIG (Simple Polydisperse Residue Input Generator), a program for generating atomic-detail models of random polydisperse lignin copolymer melts i.e., the state most commonly found in nature. Using these models, we use all-atom molecular dynamics (MD) simulations to investigate the conformational and dynamic properties of polydisperse melts representative of switchgrass (*Panicum virgatum* L.) lignin. Polydispersity, branching and monolignol sequence are found to not affect the calculated glass transition temperature, T_g. The Flory-Huggins scaling parameter for the segmental radius of gyration is 0.42 ± 0.02, indicating that the chains exhibit statistics that lie between a globular chain and an ideal Gaussian chain. Below T_g the atomic mean squared displacements are independent of molecular weight. In contrast, above T_g, they decrease with increasing molecular weight. Therefore, a monodisperse lignin melt is a good approximation to this polydisperse lignin when only static properties are probed, whereas the molecular weight distribution needs to be considered while analyzing lignin dynamics.

1. Introduction

The secondary cell walls of plants are primarily composed of three different polymers: cellulose, hemicellulose and lignin.[1–5] Lignin accounts for ~5-30% of plant biomass by weight and lends mechanical strength to the cell wall while also providing resistance to microbial attack. From an industrial standpoint, lignins are potential renewable sources of aromatic compounds used in the

resins and adhesives industry and can be used as a feedstock for novel materials during biofuel production.[3–11]

To use lignin as raw material for the above applications requires the lignin and polysaccharide cell wall components to be separated from each other.[7,10] However, strong intermolecular interactions between the cellulose and lignin molecules make them recalcitrant to traditional deconstruction treatments such as acid hydrolysis.[5,12–14] Thus, a thorough understanding of the structure and interactions between different components of the plant cell walls is required to design potential separation methods.[15–17]

Lignin in plant cell walls exists as amorphous, polydisperse random copolymers composed predominantly of *p*-hydroxyphenyl (H), guiacyl (G) and syringyl (S) lignin units.[1,5,17–19] These are formed by the radical polymerization of hydroxycinnamyl alcohol monolignol precursors.[18–20] In addition to the above monolignol units, some monocots such as switchgrass also contain *p*-hydroxycinnamate units such as *p*-coumarate (pCA), ferulate (FA) and tricin (T).[21–23] The units are bonded by a variety of linkages including β-O-4, β-β, β-5, γ-O-γ, 5-5 and α-O-4 with β-O-4 being the most common linker.[23,24] Due to their stochastic synthesis by radical coupling chemistry, lignin structures lack specific stereochemistry, and the ratios between monomeric units vary significantly between different plants. The G:S ratio correlates with biomass recalcitrance for some plant species.[17,25,26] In addition, the hydroxycinnamates (pCA and FA) in switchgrasses are hypothesized to form strong covalent bonds with the carbohydrates (celluloses and hemicelluloses) thus increasing their recalcitrance.[17]

Lignin in plant cell walls is highly polydisperse.[23,26] Polydispersity measures the distribution of molecular mass of the chains in a polymeric system and is characterized by the

polydispersity index[27–29] which is defined as the ratio between the weight-averaged and number-averaged molecular weights. For monodisperse polymers the polydispersity value equals 1.0. Lignins in plant cell walls have polydispersity values ranging between 3.0 and 5.0.[23,26] To achieve polydispersity values this large, some polymers must be much larger than others, suggesting substantial variation in the molecular weight distribution for individual lignin chains extracted from plant cell walls.

While there exist several experimental studies characterizing the structural and dynamical properties of lignin, computer simulations and theoretical treatments pertaining to understanding lignin melts (without solvents) are limited.[28,30–34] Further, the lignin structures that were probed in those studies that exist were limited to certain specific sequences[34] of monodisperse linear lignin molecules. Furthermore, openly available software tools to generate an ensemble of branched lignin molecules with an average composition and polydispersity reflecting experimental studies are lacking.

To address this limitation, a small number of open-source software packages[35–40] have been developed which incorporate significant heterogeneity in the structure of lignin. The LigninBuilder[35] tool developed by Vermaas *et* al.,[35] converts existing topology libraries for monolignol units and linkages published in the literature to build atomistic models lignin models. The Lignin-KMC[36] tool developed by Orella *et al*.,[36] combines density functional theory calculations for radical coupling with kinetic Monte Carlo simulations to model lignin biosynthesis. The LigninGraphs[37] tool developed by Wang *et al*., uses graph theory-based methods in conjunction with a rejection-free Metropolis Monte Carlo algorithm to build lignin polymer structures. Eswaran *et. al*.,[38] developed a molecular tool which can generate random lignin topologies using a graph-based generative model based on a combinatorial approach based on

experimental inputs such as the average molecular weight and S:G ratio. The Broadbelt group developed stochastic models[39,40] for generating lignin libraries, which can be validated against experimental parameters such as monomer, bond, molecular weight distributions and branching coefficients.

While some of the aforementioned tools[37,40] have the capability of generating branched lignin chains, other tools model lignin as a linear polymer. Except for the models generated by Broadbelt's group,[39,40] , other common tools for lignin topology generation do not explicitly set a given polydispersity by default. While specific lignins can be selected from a pool of generated lignins to achieve a specific polydispersity,[41] this post-hoc processing can alter other desired lignin properties, such as the S:G ratio or the linkage distribution. The resulting differences between the plant-based experimental and theoretical literature are an impediment to direct comparison with experimental measurements on cell wall lignins. To model more realistic plant cell wall lignins software that can build branched and polydisperse systems that respect other topological constraints is required.

To model polydisperse and branched lignin melts we have developed a versatile program, *viz*., Simple Polydisperse Residue Input Generator (SPRIG), wherein users can input experimental data such as the polydispersity and lignin composition of the systems. Further, SPRIG can be used to generate branched lignin system for any given target experimental composition. In addition, experimental data from gas permeation chromatography (GPC) can be directly input to the software to generate a lignin melt with a molecular weight distribution that is similar to the experimental data. In this fashion, researchers can leverage the software to compare the properties of lignin domains across different plant species. SPRIG also generates topology and forcefield

parameters of the lignin molecules at the atomic level which can be used directly with standard molecular dynamics packages to derive their conformational and dynamic properties.

Motivated by the need to understand the conformational and dynamic properties of lignin in various plant cell walls, in this work we present SPRIG then demonstrate its usefulness with all-atom molecular dynamics simulations of lignin domains. For this we start from recent solid-state nuclear magnetic resonance (ss-NMR) spectroscopy experiments by Kang *et al.,*[42] who proposed that lignin in intact maize stems self-aggregates into nanodomains with extensive surface contacts with hemicelluloses and minimal contact with cellulose. Combining this idea with the known polydispersity of lignin in plants[23,26] we model switchgrass (*Panicum virgatum* L.) lignin nanodomains in cell walls as polydisperse lignin melts. We describe the structure and use of the code and generate lignin melt structures with a composition consistent with that observed experimentally for switchgrass plants.[23] Subsequently, we compute various static and dynamic properties of the lignin melts using all-atom molecular dynamics simulations and compare these results with the models for ideal polymer melts.[43]

2. Methods

2.1 Simple Polydisperse Residue Input Generator (SPRIG)

To generate lignin melt configurations we developed a program called the Simple Polydisperse Residue Input Generator (SPRIG). SPRIG leverages the functionalities of the LigninBuilder,[35] Nanoscale Molecular Dynamics[44] (NAMD) and Visual Molecular Dynamics[45] (VMD) packages to generate linear or branched lignin melts of desired composition, polydispersity and number averaged molecular weight. SPRIG is written in Python 3 wrapped around a FORTRAN 90 code.

A flowchart depicting the required inputs and algorithm for SPRIG is provided in Figure 1. The inputs comprise residue fractions, backbone linkages fractions, the polydispersity index of the melt and constraints (if any) pertaining to the residue-linker placement. First, SPRIG generates a polydisperse melt based on either the Schulz-Zimm distribution[46,47] or an experimental distribution provided by the user. The Schulz-Zimm model assumes an ideal step-growth polymerization reaction for the polymer synthesis with a unimodal distribution of molecular weights. Next, SPRIG assigns the unit and linkage types based on input fractions using an Inverse Transform Method (see discussion below). The 3D structure corresponding to the 2D topology is then generated using LigninBuilder.[35]

The initial structure generated may have chain overlaps and unphysical ring penetrations. To remove these an unbiased energy minimization is performed initially for 500 steps using NAMD software, following the protocol developed in LigninBuilder. If bonds are stretched beyond 1.65 Å, nearby residues are pushed away from the overlap point using a combination of volumetric grids in VMD and the `gridForce` feature in NAMD. In addition, the alchemical module in NAMD is turned on to decouple the atoms from the rest of the system to reduce the barrier to moving pierced atoms from the ring. This step is analogous to a thermodynamic integration calculation. For more details refer to Ref. 34. Outputs from SPRIG consist of the structure, force field and topology files compatible with both GROMACS and NAMD software, using TopoTools as a translation layer for GROMACS simulation.[48]

The open-source codes and explanations for the input keywords used are provided in https://github.com/vaidyanathanms/SPRInG_PolydispersePolymerBuilder. Here, we provide an overview of the salient features and the algorithm used in SPRIG. Details pertaining to each block in Figure 1 are described below.

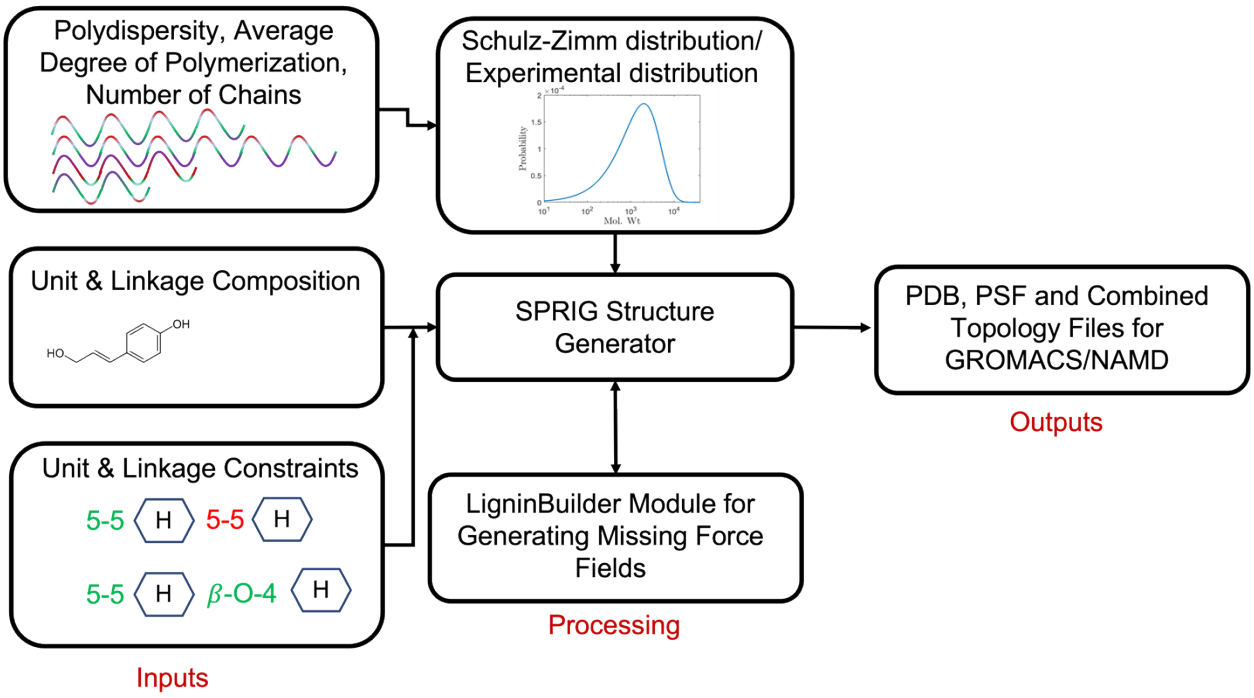


Figure 1*:* A simple flowchart for SPRIG with the required inputs, structure generation details and outputs produced.

2.1.1 Constraints

There are two specific constraints that ensure a specific 2D lignin sequence that is compatible with the user's inputs is also physically realizable.

i. Unit-Linkage constraint: This constraint ensures the compatibility of attaching a particular kind of linkage at a specific position in the unit. For instance, if a 5-5 linakge is made between two p-hydroxyphenyl (H) residues, the C5 position of the both residues becomes occupied. Therefore, the latter p-hydroxyphenyl (H) residue cannot be linked to a third residue with a 5-5 linker. However, in this case, a β-O-

4 linker is allowed to link this p-hydroxyphenyl residue with another residue since β-O-4 links only to the C_β position of the latter residue.

ii. Linkage-Linkage constraint: This constraint avoids incompatible bonds between residues. For instance, a γ-O-γ linker can be used to link the backbone and a branched unit. This modifies the C_β carbon of the backbone chain. Thus, a β-β linker cannot be present as the linker between this backbone unit and the next backbone unit.

2.1.2 Polydispersity Modeling

The Polydispersity index (Đ) is defined as

$$Đ = \frac{\overline{M_w}}{\overline{M_n}} = \frac{\sum_i (N_i M_i)^2}{\sum_i (N_i M_i)}, \quad (1)$$

where $\overline{M_w}$ and $\overline{M_n}$ are the weight-averaged and the number-averaged molecular weight, respectively, and N_i is the number of chains that have molecular weight M_i. To obtain a polydisperse lignin melt with a random distribution of residues and linkers with a target number average molecular weight, we use an iterative method based on either the theoretical distribution or a pre-determined experimental molecular weight distribution.

There are multiple theories used to model polydispersity[29,46,47,49–51] in polymeric systems. Here we use the Schulz-Zimm distribution, which is widely used to model polydispersity in synthetic polymers with polydispersities between 1.0 and 2.0. The probability density function of the molecular weights for a Schulz-Zimm[29,46,47] distribution is given by

$$p_{SZ}(\mathrm{y}) = \frac{\alpha^{\alpha} y^{\alpha-1} exp(-\alpha y)}{\Gamma(y)}, \tag{2}$$

where

$$y_i = \frac{M_i}{\overline{M_n}}, \tag{3}$$

is the ratio between the molecular weight of the i^{th} chain and the number averaged molecular weight. The constant α is related to the polydispersity index, Đ, through $\alpha = 1/(Đ - 1)$ and Γ represents the Gamma function.

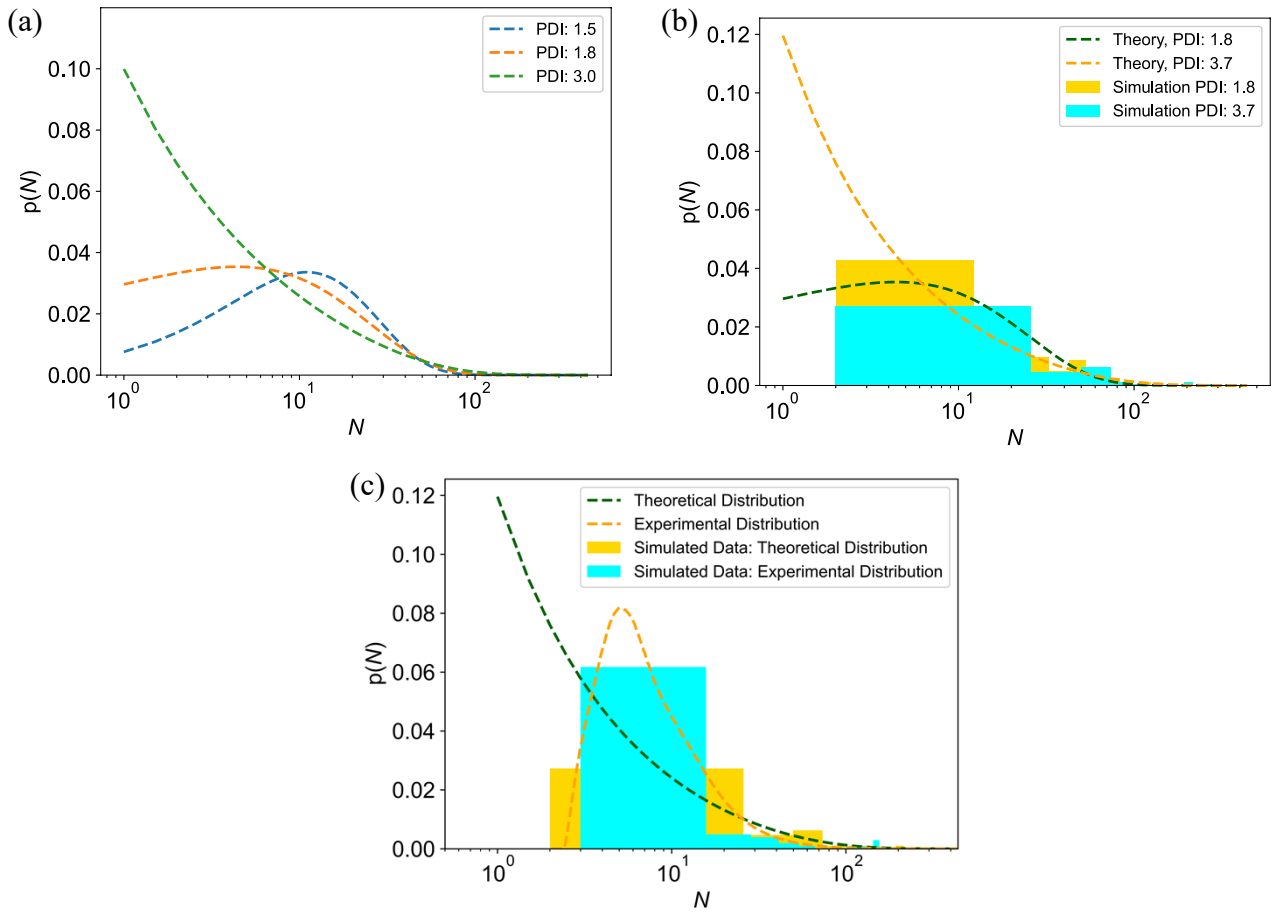


***Figure 2**:* (a) Schulz-Zimm distributions for three dispersity, Đ, values (Eq. 6); (b) Comparison between Schulz-Zimm distribution theory with the simulated molecular weights for lignin using SPRIG. The molecular weights from four different starting configurations for each dispersity value were used to compute the histogram; (c) Comparison between molecular weight distributions from theory, SPRIG simulation and experiments for Đ = 3.7.

Figure 2(a) shows the theoretical Schulz-Zimm distribution for three different polydispersity values. When the polydispersity is greater than 2.0, the Schulz-Zimm distribution decays exponentially at low molecular weights.

In practice, initially the molecular weights of the chains are generated by sampling the Schulz-Zimm distribution with uniform random numbers using the Inverse Transform technique. To this end, the Schulz-Zimm probability density function (Equation (2)) is integrated using the trapezoidal rule and is normalized to compute the cumulative distribution function. Further, uniform random numbers from the domain [0,1] are generated and mapped on to the ordinate of the normalized cumulative distribution function. The abscissa value corresponding to this ordinate value in the cumulative distribution function is multiplied by the number average molecular weight to obtain the molecular weight of this chain. Since the number of units in each chain is determined using Schulz-Zimm distribution before assigning their identities, the degree of polymerization is used as a surrogate for the molecular weight of the unit.

In addition, the minimum number of residues allowed for each chain is fixed at two and the maximum allowed molecular weight of the chain is fixed at ten times the number-averaged molecular weight. The former constraint is to avoid generating monolignols in the system and the latter constraint is for computational efficiency, since the value of the tail of the Schulz-Zimm distribution decays to approximately zero within five times the number averaged molecular weight. This process is iterated until the relative error between the polydispersity for the simulated system generated using SPRIG ($Đ_{sim}$) and the target polydispersity ($Đ_{tar}$) is within a predefined tolerance (default = 5%).

In Fig. 2(b), we show the results obtained from the theoretical Schulz-Zimm distribution and that from simulating a polymer melt with 20 chains using SPRIG. The theoretical distribution

matches the distribution obtained from SPRIG for $Đ_{tar}$ = 1.8 (dotted green lines) and for $Đ_{tar}$ = 3.7 (dotted yellow lines). The discrepancy between the theory and simulations arises from the small number of chains in our system. Note that, owing to the long-tail nature of Schulz-Zimm distribution, the number of chains with large molecular weights tends to be small.

If an experimental molecular weight distribution obtained from gas permeation chromatography (GPC) is available, this can be used in lieu of the Schulz-Zimm distribution for a direct match with experiment. For completeness, we recap the connection between the GPC experiments and the molecular weight distributions here.[52] For the i^{th} chain with a molar mass M_i within a melt of total mass m, the differential mass distribution, $w(M_i)$ is defined as

$$w(M_i) = \frac{dm}{dM_i}. \tag{4}$$

The differential mass distribution outputs from the GPC experiments can be either in terms of $w\left(\log(M)\right)$ or $w(M)$. These quantities are related to each other as follows:

$$w(\log(M_i)) = \frac{dm}{d\left(\log(M_i)\right)} = M_i\frac{dm}{dM_i} = M_i w(M_i). \tag{5}$$

Finally, the probability density function $p(M_i)$ can be computed using

$$p(M_i) = \frac{w(M_i)}{M_i} = \frac{w\left(\log(M_i)\right)}{M_i^2}, \tag{6}$$

where the number averaged molecular weight ($\overline{M_n}$) is defined by

$$\overline{M_n} = \frac{\int M_i p(M_i) dM_i}{\int p(M_i) dM_i}, \tag{7}$$

and the weight averaged molecular weight is defined by

$$\overline{M_w} = \frac{\int M_i^2 p(M_i) dM_i}{\int p(M_i) dM_i}. \quad (8)$$

In practice, a base molecular weight is defined for a unit in the system (default = 200 g/mol) and the target average molecular weights are converted to average degrees of polymerization by dividing by the base molecular weight. Subsequently, the input $w(M)$ or $w(\log M)$ is converted to $p(M)$ using Eqs (4) - (6). Lastly, the Inverse Transform method, similar to that used for generating the Schulz-Zimm distribution, is used here to generate chains with the desired polydispersity.

Note that the Schulz-Zimm distribution is formally valid for systems with Đ between 1.0 and 2.0, while larger Đ indicates complex polymerization reactions that are not modelled accurately by this simple Schulz-Zimm kinetic scheme. In Figure 2(c), we compare the Schulz-Zimm distribution with that of the experimental distribution for $Đ_{tar}$ = 3.7. It can be observed that the Schulz-Zimm distribution deviates significantly from the experimental distribution at low molecular weights. However, if the average properties of the melt obtained from using Schulz-Zimm or experimental distributions are similar, then the Schulz-Zimm distribution can be used in lieu of an experimental distribution if experimental data for the molecular weight distribution are unavailable. To test this hypothesis, in the results we compare the molecular weight distribution (Figure 2(c)) from SPRIG simulations for both the theoretical Schulz-Zimm distribution and the experimental distribution.

2.1.3 Unit and Linkage Generation

Users can input the average fraction of the units and linkages in their lignin system. The topology, parameter, unit-linkage and/or linkage-linkage constraints for the most common lignin units such as p-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) monophenolic units, ferulate (FA) and p-coumarate (pCA) hydroxycinnamate units and linkages such as β-O-4, α-O-4, γ-O-γ, β-β, β-5 and

5-5 linkages are pre-built into the SPRIG repository. Users can add new units or linkages by adding their corresponding topology, forcefield parameters and the constraint details into the respective files.

The units and linkages are assigned molecular details in a similar fashion. First, the sum of the input unit fractions is re-normalized to unity. If the abscissa of the random number generated corresponds to a ferulate or a pCA residue, they are assigned as branches to the polymer. Tricin units are assigned only if both the random number match is obtained and if the residue number corresponds to the initiator of the chain. Further, a discrete cumulative probability distribution is defined according to the input weight fractions of the residues and linkers and the mapping procedure like that explained above for generating polydisperse residues is followed.

In each iteration, the Euclidean (L2) norm of the error between the simulated and target residue (or linker) compositions is computed. The final residues and linkers are chosen such that the relative error between the fraction of each moiety in the simulated system and the target fraction is within a pre-specified relative tolerance (default = 10%). If the maximum number of iterations is exceeded, then the system corresponding to the minimum L2 norm is generated as output.

2.2 Model

As a test case, we used SPRIG to model a lignin melt with a composition consistent with that observed in the secondary cell walls of wild-type (WT) switchgrass.[23] The backbone of these lignin molecules is comprised of H, G and S monophenolic units bonded with β-O-4, β-β, β-5 and 5-5 linkages. Ferulate (FA) and p-coumarate (pCA) hydroxycinnamate units were modeled as side chains bonded to the backbone using γ-O-γ linkages.[53,54] The sequences of the backbone units, side chain units and linkages were chosen at random such that the average composition is within 10% of the experimentally observed values.[23] We fixed the total number of chains for all the

systems at 20. The distributions of molecular weights are given in Figure 2 for two of the polydispersity values studied in this paper. The average composition of the lignin melts simulated in this work is provided in Table 1. The influence of finite size effects on the static and dynamic properties of the system is not explicitly examined in this work.

Table 1: Composition of wild-type switchgrass studied.

Units	**Composition Fraction**
p-hydroxyphenol (H)	0.027
Gauaicyl (G)	0.434
Syringyl (S)	0.266
Ferulate (F)	0.039
p-coumarate (pCA)	0.234
Linkages	**Composition Fraction**
β-O-4	0.888
β-β	0.011
β-5	0.078
5-5	0.023

2.3 Simulation Methodology

Using the SPRIG-derived structures we performed molecular dynamics simulations using GROMACS version 2020.6[55,56] and CHARMM-36 forcefields.[57–59] We performed the simulations at temperatures between 300 K and 500 K with an increment of 10 K. Periodic boundary conditions were employed in all three directions. The leap-frog integrator algorithm with a time-step of 1 or 2 fs was used to integrate the equations of motion.[60,61] The particle mesh Ewald (PME) method was used to compute the electrostatic interactions.[62,63] The relative strength of the electrostatic interaction at the cut-off was set at 10^{-5}. The cutoffs for the Van der Waals and electrostatic interactions were set at 1 nm. All bonds with the hydrogen atoms were constrained using LINCS constraint algorithm.[64] The molecular dynamics simulations consisted of the following steps:

- *Step 1*: Energy minimization using the conjugate-gradient method.
- *Step 2*: NVT simulations at 600 K using a velocity rescaling temperature coupling algorithm with a coupling constant of 0.2 ps.[65]
- *Step 3*: NPT simulations at 1 bar and 600 K using barostat with a Berendsen pressure coupling[66,67] algorithm and a velocity rescaling temperature coupling algorithm. The temperature and pressure coupling constants were 0.1 ps and 0.5 ps, respectively.
- *Step 4*: Output configurations from *Step 3* were used as input conditions for target temperatures. *Steps 1* and *2* were performed again followed by NPT simulations using a barostat with Berendsen coupling at 1 bar and at target temperature.

- *Step 5*: NPT simulations at target temperature and 1 bar pressure using the Parinello-Rahman barostat[68,69] and velocity rescaling temperature coupling algorithm. The temperature and pressure coupling constants were 0.2 ps and 5 ps, respectively.

A total of at least 400 ns, including equilibration and production cycles, were run for each system. The last 280 ns were used to analyze the properties. The reported results were averaged from four different simulations starting with independent initial configurations and velocities.

2.4 Quantification Measures

2.4.1 Glass Transition Temperature

The glass transition temperature is defined as the point where the polymer melts change from a glassy state to a liquid-like state and is characterized by a change in slope in the density-temperature curves. The glass transition temperature T_g was calculated using a broken-stick regression,

$$T_g = T_i, \text{ such that } \frac{d(\delta_i + \delta_{M-i})}{di} = 0. \quad (9)$$

where T_i is the temperature at the i^{th} temperature. We plotted the specific volume of the melt, defined by the inverse of the density, as a function of temperature for a given system. Further, we divided the entire data into two groups and implemented linear fits separately to the two groups of points (n_i and n_{M-i}) where M represents the total number of points in the density-temperature curve. Weighted linear regression was performed on both sets of points with weights for each specific volume equal to the inverse of the standard mean of the four independent simulations (normalized to the sum of the weights). The weighted errors between the linear fits and the points for the two lines (δ_i and δ_{M-i}) were calculated for all possible combinations of i and $M - i$. The

glass transition point was then determined as the temperature T_i corresponding to the point at which a change in slope observed from the linear fits has the least sum of the errors to the linear fits (derivative is zero).

2.4.2 Segmental Radius of Gyration

The segmental radius of gyration (R_g) of a chain segment of N units was calculated using

$$R_g = \left\langle \sqrt{\frac{\sum_{i=1}^{N} m_i |\boldsymbol{r}_i|^2}{\sum_{i=1}^{N} m_i}} \right\rangle, \tag{10}$$

where $\langle ... \rangle$ represents the ensemble average, m_i represents the mass of the atom in the chain segment and the $\boldsymbol{r}_i$ represents the position of the atom in the Cartesian coordinates. We considered only the backbone of the lignin chains (H, G and S) units while computing the segmental R_g. R_g is calculated using `gmx gyrate`.

2.4.3 Shape Anisotropy Factor

To understand the conformations of the chains in space, we computed the shape anisotropy factor of the i^{th} chain, κ_i defined by

$$\kappa_i = \frac{3}{2}\left(\frac{\sum_{i=1}^{3}\left(\lambda_i - \overline{\lambda}\right)^2}{\sum_{i=1}^{3}\lambda_i}\right)$$

where λ_i represents the eigenvalues of the not mass weighted radius of gyration tensor. The eigenvalues λ_i are computed using the module `gmx polystat`. κ_i = 0 corresponds to a spherically symmetric system (globular chain) and κ_i = 1 corresponds to a line (stretched linear chain).

2.4.4 Mean Squared Displacement

The mean squared displacement of the atoms were computed by `gmx msd` using

$$\boldsymbol{r}^2(t) = \left\langle \frac{1}{N_c} \sum_{i}^{N_c} (\boldsymbol{r}_i(t) - \boldsymbol{r}_i(0))^2 \right\rangle, \tag{11}$$

where $\boldsymbol{r}_i(t)$ represents the position of the i^{th} atom at time t. For this all C_1 atoms were included. N_c represents the total number of C_1 atoms in the melt. Further, to compare the displacements between chains of different molecular weights, we computed the mean squared displacement of the chains at $t = 100$ ns denoted by d_{100} using $d_{100} = \boldsymbol{r}^2(t = 100 \text{ ns})$.

3. Results and Discussion

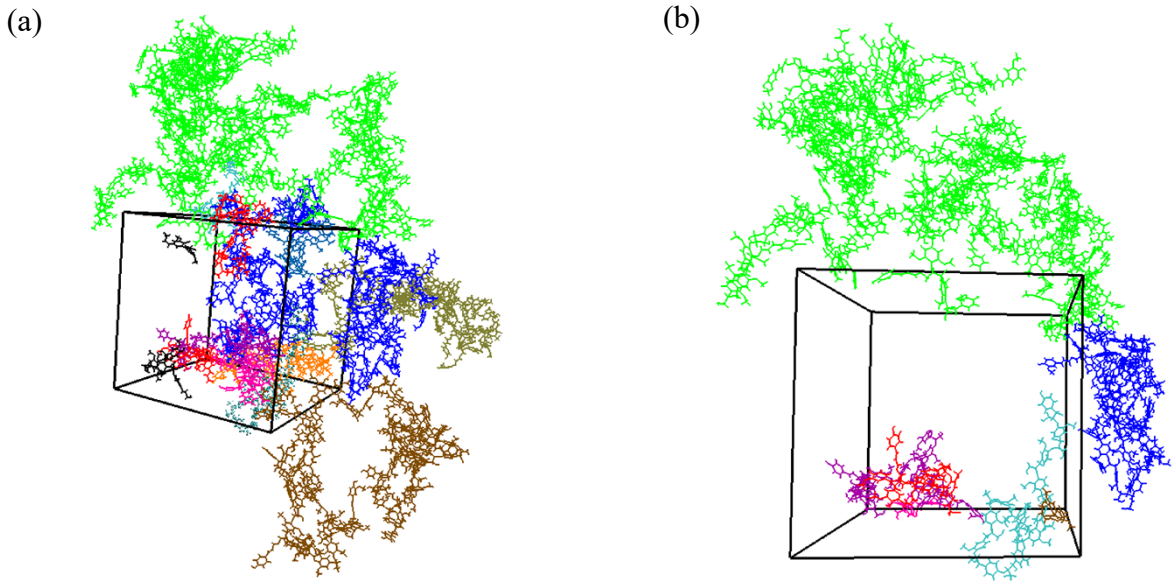


***Figure 3**:* (a) Equilibrated configuration of polydisperse lignin melt. (b) Zoomed-in picture from (a) showing a few selected chains.

Figure 3(a) shows the equilibrated configuration of a lignin melt after 400 ns for Đ = 3.7. Different chains are labelled in different colors to highlight chain variations. Polydispersity can be clearly seen in Figure 3(b) wherein only a few selected chains from Figure 3(a) are displayed.

Herein, the chain colored in green has a much higher molecular weight compared to other chains and the brown chain a very low molecular weight.

3.1 Properties of Lignin Melts

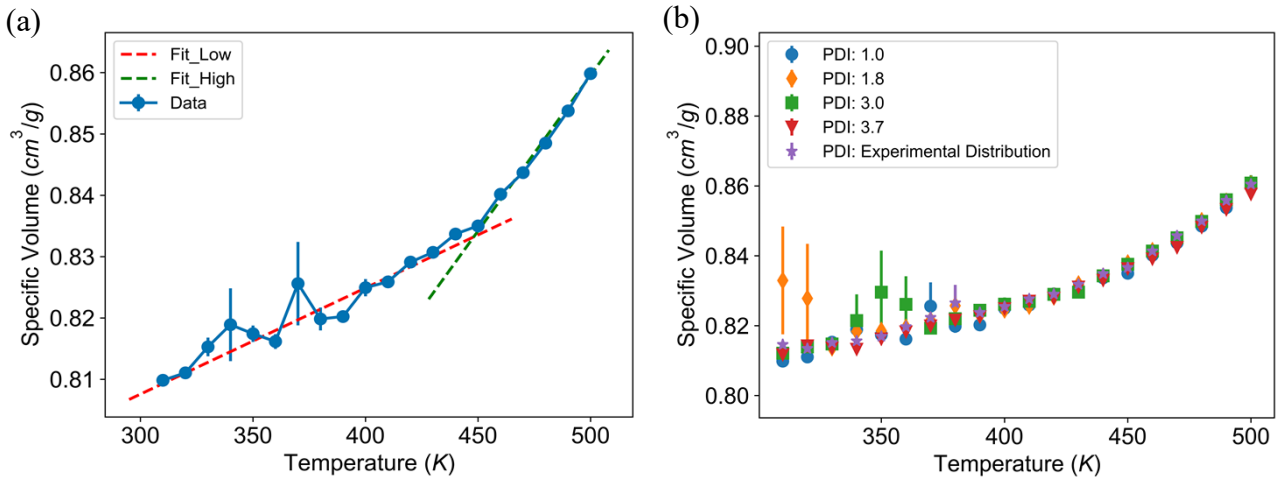


Figure 4*:* (a) Specific volumes as a function of temperature for monodisperse melts. The red and green dotted lines are obtained by minimizing the total error according to Equation (4). (b) Specific volumes as a function of temperature for all the systems studied. Error bars correspond to standard error of the mean.

In this section we examine the conformational and dynamical properties of lignin melts. First, we estimate the glass transition temperature of simulated lignin melts as a function of the polydispersity. Figure 4(a) shows the specific volume as a function of temperature for monodisperse melts. The red and green dotted lines show the low and high temperature best fit lines to the specific volumes which satisfy Equation (9). Larger error bars are seen for temperatures below the glass transition temperature. This arises from the presence of local minima in the free-energy landscapes especially at temperatures below the glass-transition temperatures. Unlike synthetic polymers, lignin displays only a weak first-order transition in the specific volume change as a function of temperature. This is consistent with experimental findings which suggest that lignin does not have a sharp glass transition temperature. The glass transition temperature for monodisperse lignin melts, obtained from the intersection of the two fitted lines, is approximately

448 K. Vural et al., showed that the glass transition temperature for 25% hydrated lignin melt is 360 K and that for 5% hydrated lignin melt is 425 K.[31] A linear extrapolation of these results to 0% hydration yields a T_g of 441 K which is in close agreement with the present work (448 K).

Table 2: Glass transition temperatures for all the systems studied.

Polydispersity index ($Đ_{ideal}$)	**Glass transition Temperature (T_g)**
1.0	448 K
1.8	433 K
3.0	433 K
3.7 (Schulz-Zimm distribution)	438 K
3.7 (experimental distribution)	444 K

Figure 4(b) displays the specific volumes as a function of temperature for different polydispersity values and Table 1 shows the associated computed glass transition temperatures. T_g varies between 430K and 445K, consistent with experimental measurements indicating that T_g varies between 50°C and 160°C (323 – 433 K).[70–72] The difference between the experimental and simulation results can be attributed to two factors. One is the presence of water in experimental conditions. Water acts as a plasticizer thereby reducing the effective glass transition temperature. The lignin melt simulations here are solvent-free. The second factor is that the cooling rate in MD is several orders of magnitude faster than in experiment, leading to a higher apparent T_g.

The glass transition temperatures for both monodisperse and polydisperse melts are within $440 \pm 10\ K$, suggesting that polydispersity does not affect T_g. This is also evidenced by the fact that above $350\ K$ the average specific volume values are insensitive to the polydispersity. The variation in T_g values at low temperatures ($T < 320\ K$) arises from kinetic trapping. Further,

although the backbone sequences and the branching positions are chosen at random for each simulation, T_g computed for each case (Table 2) is similar. Further, our results ($T_g \approx 440\ K$) are similar to the glass transition temperature for Vanilla lignin ($T_g = 448\ K$).[31] Together, these results suggest that the lignin glass transition temperatures, as calculated from rapidly-cooled solvent-free melts, are independent of backbone sequence or branching positions.

To better understand the conformation of the lignin melts at the chain level, we computed the segmental radius of gyration (R_g) as function of the segmental degree of polymerization for all the systems studied. For ideal linear chains, R_g is related to the degree of polymerization N via

$$R_g = C_b N^{\nu}, \tag{12}$$

where C_b and ν represent the characteristic length (Kuhn segment length) and the Flory scaling

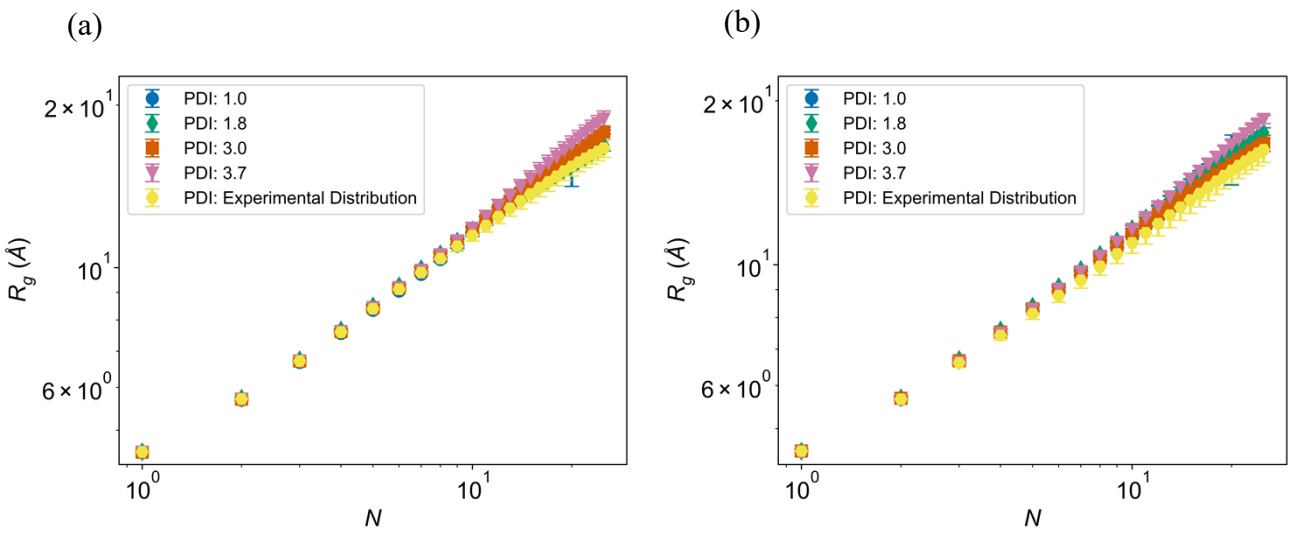


***Figure 5**:* Segmental radius of gyration as a function of the degree of polymerization at (a) 500 K and (b) 350 K. The data is truncated at N=30 since data beyond this value has poor statistics.

exponent (inverse fractal dimension), respectively. The fractal dimension, ν is 0.5 for ideal polymer melts (Gaussian melts) or polymers in a θ-solvent and $1/3$ for polymers in collapsed globular state. **Figure** 5(a) and (b) show the segmental radius of gyration for 500 K and 350 K,

respectively. The data at 500 K (**Figure** 5(a)) correspond to the R_g values for a temperature above T_g . The scaling is independent of the polydispersity. The differences at high N arise from low statistics. Similar trends are observed for the data at 350 K (**Figure** 5(b)), which is below T_g, suggesting that the radius of gyration scaling is independent of both polydispersity and temperature.

The fitting parameters to the segmental radius of gyration using Equation (12) were found to be independent of the temperature and polydispersity and the values are $\nu = 0.42 \pm 0.02$ and $C_b = 4.5$ Å, respectively. The values for ν suggest that lignin melts behave somewhere between an ideal melt and a globular melt. This is consistent with the observations for branched copolymers

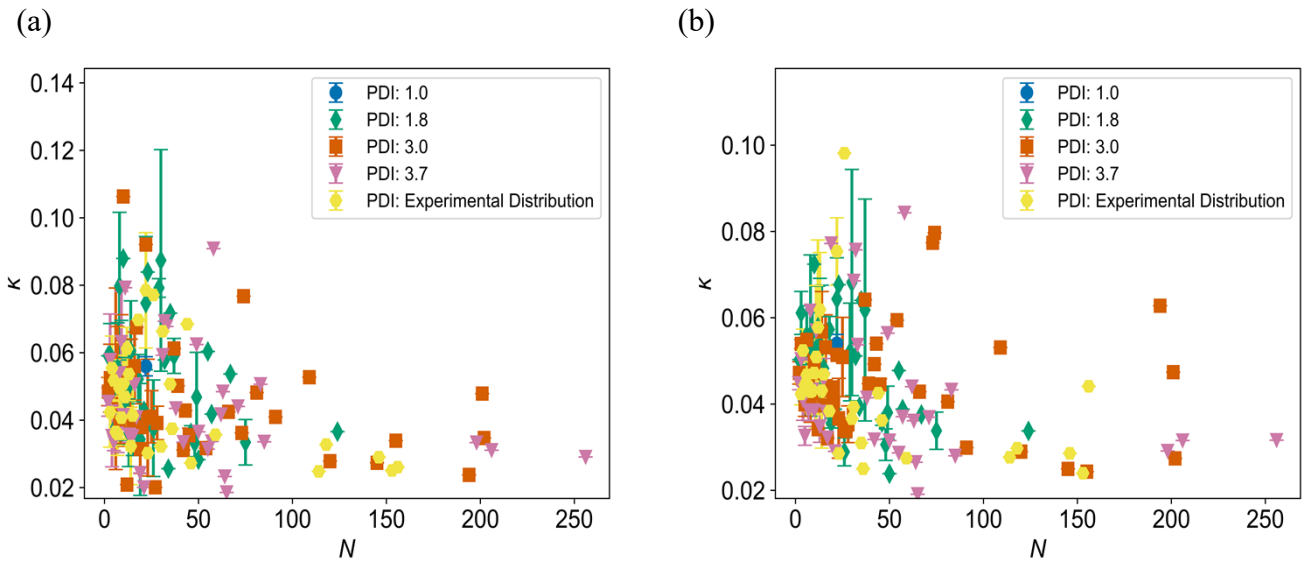


Figure 6*:* Shape anisotropy factors as a function of the number of backbone monomers in the melt at (a) 350 K; (b) 500 K.

by Douglas and coworkers.[73,74]

A further question is whether the conformations for individual lignin chains varies with size. For instance, could the shorter chains in a polydisperse melt adopt stretched conformations if the persistence length of the chain is greater than the end-to-end distance and the longer chains

adopt a globular conformation to create a uniform shape? In Figure 6(a), we display the shape anisotropy factor (κ) of all the chains in the melt at two different temperatures. Shape anisotropy factors vary between 0 for a spherical and 1 for rod-like geometry. The shape anisotropy factor is observed to be always below 0.1, irrespective of the molecular weight of the chains. However, larger chains have lower shape anisotropy factors, and are closer to spherical than smaller chains. We suspect that in part, this bias is due to packing constraints, as small collapsed globules formed from one polymer are more irregular in shape than larger with more rotatable bonds to minimize their surface area to water. The small shape anisotropy factors suggest that the chains are globular but not perfectly spherical which is in agreement with the results from the radius of gyration analysis in Figure 5. Further, by comparing Figures 6(a) and 6(b), we see no evidence for a temperature dependence.

Finally, probing the motion of the lignin chains, we look at the mean squared displacements at $t = 100$ ns as a function of their degree of polymerization for all the systems studied. At the low temperature (350K), the MSD is independent of, or weakly dependent on, the molecular weight (Figure 7(a)). Further, the maximum MSD at this temperature is ~2.5 Å^2, suggesting that the

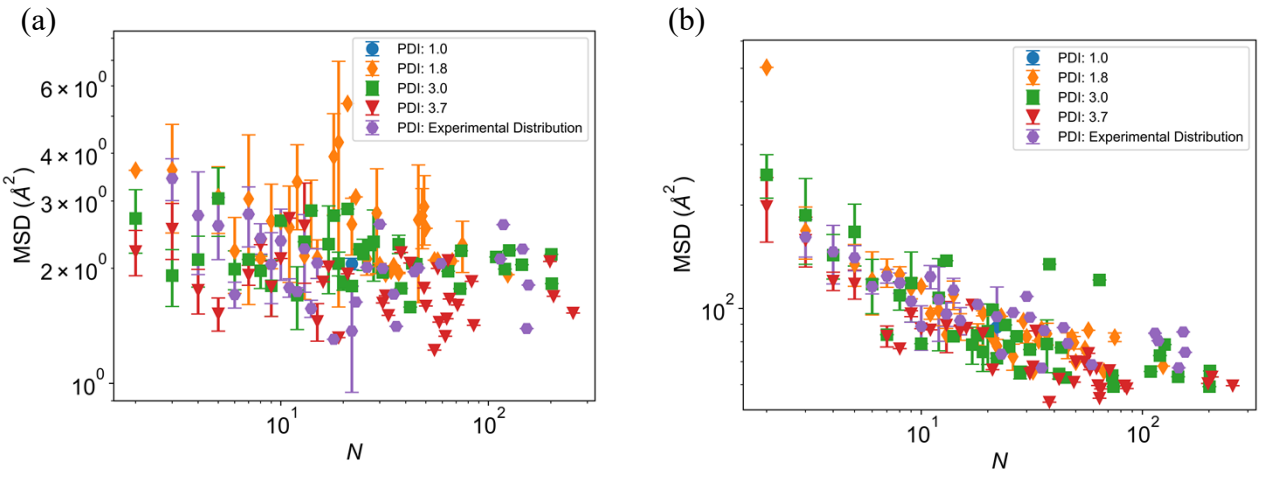


Figure 7: Mean squared displacements as a function of the number of backbone monomers for (a) 350 K; (b) 500 K

polymeric monomers move distances smaller than the Kuhn segment length ($C_b = 4.5$ Å) during the 100 ns sampled. This confinement arises from the glassy nature of the chains at this temperature. In contrast, at 500K a strong dependence on the molecular weight is observed with the low molecular weight chains moving faster than the high molecular weight chains. This suggests that the dynamics of low molecular weight chains become decoupled from the dynamics of high molecular weight chains above T_g.

4. Conclusions

We developed a generic software, SPRIG, for generating initial configurations of polydisperse lignin melt systems. The melts can be generated corresponding to a target polydispersity value from either a theoretical Schulz-Zimm distribution or directly from the experimental probability distributions obtained from a GPC experiment. The outputs generated by SPRIG can be directly used in software such as GROMACS or NAMD to perform molecular dynamics simulations. Compared to other[35–40][35,36] tools for generating lignin molecules, the flexibility in generating lignin with realistic chemical structure and inhomogeneity to reflect experimental conditions makes SPRIG a novel tool for understanding properties of lignin at the atomistic level. In addition, users can directly input the data obtained from GPC experiments into the SPRIG software to generate lignin melts with a molecular weight distribution which is similar to the ones obtained from GPC experiments. Currently, SPRIG can generate lignin molecules comprising p-hydroxyphenyl, guaiacyl, syringyl and tricin monophenolic units along with β-O-4, α-O-4 β-β, β-5 and 5-5 linkages. Ferulate and p-coumarate hydroxycinnamate units can also be added as branches to the lignin backbone. Also, SPRIG can be readily extended to other units or linkages upon addition of topology and forcefield parameters associated with them to the library.

To demonstrate the utility of SPRIG we computed static and dynamic properties for a

model switchgrass polydisperse lignin melt using all-atom molecular dynamics simulations. We showed that the polydispersity, or the distribution of chain molecular weights, does not affect the calculated glass transition temperature. The scaling of the radius of gyration showed that both monodisperse and polydisperse melts both behave somewhere between a globular melt and an ideal Gaussian melt. Further, the shape of the polymer chains is insensitive to the molecular weight distribution and the temperature. However, the mean-square displacement is strongly dependent on the molecular weight distribution of the chains and the temperature at which the dynamic properties are probed. Together, we show that a monodisperse melt is a good approximation to polydisperse cases if only static properties are probed whereas the molecular weight distribution needs to be considered while analyzing dynamic properties of the system above T_g.

Notes

The open-source codes for the SPRIG software are available in https://github.com/vaidyanathanms/SPRInG_PolydispersePolymerBuilder.git. Codes for generating initial configurations, molecular dynamics simulations and trajectory analysis are available in https://github.com/vaidyanathanms/PureLigninMelt.git.

Acknowledgements

This research was supported by the Genomic Science Program, Office of Biological and Environmental Research, U.S. Department of Energy (DOE), under Contract FWP ERKP752. Computations were performed using the resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

References

[1] J. Yan, Z. Hu, Y. Pu, E. Charles Brummer, and A.J. Ragauskas, "Chemical compositions of four switchgrass populations," Biomass Bioenergy **34**(1), 48–53 (2010).

[2] J.D. Kubicki, H. Yang, D. Sawada, H. O'Neill, D. Oehme, and D. Cosgrove, "The Shape of Native Plant Cellulose Microfibrils," Sci. Rep. **8**(1), 13983, (2018).

[3] J. Ralph, C. Lapierre, and W. Boerjan, "Lignin structure and its engineering," Curr. Opin. Biotechnol. **56**, 240–249 (2019).

[4] A.J. Ragauskas, C.K. Williams, B.H. Davison, G. Britovsek, J. Cairney, C.A. Eckert, W.J. Frederick, J.P. Hallett, D.J. Leak, C.L. Liotta, J.R. Mielenz, R. Murphy, R. Templer, and T. Tschaplinski, "The path forward for biofuels and biomaterials," Science (1979) **311**(5760), 484–489 (2006).

[5] A. Zoghlami, and G. Paës, "Lignocellulosic Biomass: Understanding Recalcitrance and Predicting Hydrolysis," Front. Chem., **7:**874 (December), (2019).

[6] M.J. de la Torre, A. Moral, M.D. Hernández, E. Cabeza, and A. Tijero, "Organosolv lignin for biofuel," **45**, 58-73, Ind. Crops Prod., (2013).

[7] S. Laurichesse, and L. Avérous, "Chemical modification of lignins: Towards biobased polymers," Prog. Polym. Sci., **39**(7), 1266–1290 (2014).

[8] Y.N. Guragain, A.I. Herrera, P. V. Vadlani, and O. Prakash, "Lignins of bioenergy crops: A review," Nat. Prod. Commun., **10**(1), 201–208 (2015).

[9] A. Duval, and M. Lawoko, "A review on lignin-based polymeric, micro- and nano-structured materials," React. Funct. Polym., **85**, 78–96 (2014).

[10] A.J. Ragauskas, G.T. Beckham, M.J. Biddy, R. Chandra, F. Chen, M.F. Davis, B.H. Davison, R.A. Dixon, P. Gilna, M. Keller, P. Langan, A.K. Naskar, J.N. Saddler, T.J. Tschaplinski, G.A. Tuskan, and C.E. Wyman, "Lignin valorization: Improving lignin processing in the biorefinery," Science (1979) **344**(6185), (2014).

[11] S.L. Hilburg, A.N. Elder, H. Chung, R.L. Ferebee, M.R. Bockstaller, and N.R. Washburn, "A universal route towards thermoplastic lignin composites with improved mechanical properties," Polymer (Guildf), **55**(4), 995–1003 (2014).

[12] A.S. Gross, J.W. Chu, L. Petridis, and J.C. Smith, "On the molecular origins of biomass recalcitrance: The interaction network and solvation structures of cellulose microfibrils," Nat. Rev. Chem., **114**(42), 13333–13341 (2018).

[13] O.M. Terrett, and P. Dupree, "Covalent interactions between lignin and hemicelluloses in plant secondary cell walls," Curr. Opin. Biotechnol., **56**, 97–104 (2019).

[14] L. Petridis, and J.C. Smith, "Molecular-level driving forces in lignocellulosic biomass deconstruction for bioenergy," Nat. Rev. Chem., **2**(11), 382–389 (2018).

[15] P. Langan, L. Petridis, H.M. O'Neill, S.V. Pingali, M. Foston, Y. Nishiyama, R. Schulz, B. Lindner, B. Leif Hanson, S. Harton, W.T. Heller, V. Urban, B.R. Evans, S. Gnanakaran, A.J. Ragauskas, J.C. Smith, and B.H. Davison, "Common processes drive the thermochemical pretreatment of lignocellulosic biomass," Green Chem., **16**(1), 63–68 (2014).

[16] X. Zhao, K. Cheng, and D. Liu, "Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis," **82**, 815–827, Appl. Microbiol. Biotechnol., (2009).

[17] R. Samuel, M. Foston, N. Jiang, L. Allison, and A.J. Ragauskas, "Structural changes in switchgrass lignin and hemicelluloses during pretreatments by NMR analysis," Polym. Degrad. Stab., **96**(11), 2002–2009 (2011).

[18] J.H. Grabber, J. Ralph, R.D. Hatfield, and S. Quideau, "p-Hydroxyphenyl, Guaiacyl, and Syringyl Lignins Have Similar Inhibitory Effects on Wall Degradability," J. Agric. Food. Chem., **45**(7), 2530–2532 (1997).

[19] J.D. Crowe, N. Feringa, S. Pattathil, B. Merritt, C. Foster, D. Dines, R.G. Ong, and D.B. Hodge, "Identification of developmental stage and anatomical fraction contributions to cell wall recalcitrance in switchgrass," Biotechnol. Biofuels, **10**(1), 1–16 (2017).

[20] A.K. Sangha, B.H. Davison, R.F. Standaert, M.F. Davis, J.C. Smith, and J.M. Parks, "Chemical Factors that Control Lignin Polymerization," J. Phys. Chem. B, **118**(1), 164–170 (2014).

[21] G. Xu, Z. Shi, Y. Zhao, J. Deng, M. Dong, C. Liu, V. Murugadoss, X. Mai, and Z. Guo, "Structural characterization of lignin and its carbohydrate complexes isolated from bamboo (Dendrocalamus sinicus)," Int. J. Biol. Macromol., **126**, 376–384 (2019).

[22] M. Li, C.G. Yoo, Y. Pu, A.K. Biswal, A.K. Tolbert, D. Mohnen, and A.J. Ragauskas, "Downregulation of pectin biosynthesis gene GAUT4 leads to reduced ferulate and lignin-carbohydrate cross-linking in switchgrass," Nat. Commun. Biol., **2**(1), 1–11 (2019).

[23] L. Liang, Y. Wang, S. Bhagia, V. Sethuraman, Z. Yang, X. Meng, N. Bryant, L. Petridis, J.C. Smith, S.V. Pingali, N.C. Gallego, Y. Pu, B.R. Evans, H.M.O. Neill, B.H. Davison, and A.J. Ragauskas, "Chemical and Morphological Structure of Transgenic Switchgrass Organosolv Lignin Extracted by Ethanol , Tetrahydrofuran , and γ - Valerolactone Pretreatments," ACS Sustain. Chem. Eng., 2022, 10, 28, 9041–9052 (2022).

[24] N. Giummarella, Y. Pu, A.J. Ragauskas, and M. Lawoko, "A critical review on the analysis of lignin carbohydrate bonds," Green Chem., **21**(7), 1573–1595 (2019).

[25] R. Samuel, Y. Pu, B. Raman, and A.J. Ragauskas, "Structural characterization and comparison of switchgrass ball-milled lignin before and after dilute acid pretreatment," Appl. Biochem. Biotechnol., **162**(1), 62–74 (2010).

[26] R. Samuel, Y. Pu, N. Jiang, C. Fu, Z.Y. Wang, and A. Ragauskas, "Structural characterization of lignin in wild-type versus COMT down-regulated switchgrass," Front. Energy. Res., **1**(JAN), 1–9 (2014).

[27] V.A. Harmandaris, V.G. Mavrantzas, and D.N. Theodorou, "Atomistic molecular dynamics simulation of polydisperse linear polyethylene melts," Macromolecules, **31**(22), 7934–7943 (1998).

[28] D. Vural, J.C. Smith, and L. Petridis, "Polymer principles behind solubilizing lignin with organic cosolvents for bioenergy," Green Chem., **22**(13), 4331–4340 (2020).

[29] N.A. Lynd, M.A. Hillmyer, and M.W. Matsen, "Theory of polydisperse block copolymer melts: Beyond the Schulz - Zimm distribution," Macromolecules, **41**(12), 4531–4533 (2008).

[30] M.D. Smith, and D. Vural, "Dynamics of Lignin: Molecular Dynamics and Neutron Scattering," ACS Symposium Series **1338**, 49–67 (2019).

[31] D. Vural, J.C. Smith, and L. Petridis, "Dynamics of the lignin glass transition," Phys. Chem. Chem. Phys., **20**(31), 20504–20512 (2018).

[32] D. Vural, C. Gainaru, H. O'Neill, Y. Pu, M.D. Smith, J.M. Parks, S.V. Pingali, E. Mamontov, B.H. Davison, A.P. Sokolov, A.J. Ragauskas, J.C. Smith, and L. Petridis, "Impact of hydration and temperature history on the structure and dynamics of lignin," Green Chem., **20**(7), 1602–1611 (2018).

[33] P. Langan, L. Petridis, H.M. O'Neill, S.V. Pingali, M. Foston, Y. Nishiyama, R. Schulz, B. Lindner, B. Leif Hanson, S. Harton, W.T. Heller, V. Urban, B.R. Evans, S. Gnanakaran, A.J. Ragauskas, J.C. Smith, and B.H. Davison, "Common processes drive the thermochemical pretreatment of lignocellulosic biomass," Green Chem., **16**(1), 63–68 (2014).

[34] T.B. Rawal, M. Zahran, B. Dhital, O. Akbilgic, and L. Petridis, "The relation between lignin sequence and its 3D structure," Biochim. Biophys. Acta Gen. Subj., **1864**(5), 129547 (2020).

[35] J. V. Vermaas, L.D. Dellon, L.J. Broadbelt, G.T. Beckham, and M.F. Crowley, "Automated Transformation of Lignin Topologies into Atomic Structures with LigninBuilder," ACS Sustain. Chem. Eng., **7**(3), 3443–3453 (2019).

[36] M.J. Orella, T.Z.H. Gani, J. V. Vermaas, M.L. Stone, E.M. Anderson, G.T. Beckham, F.R. Brushett, and Y. Román-Leshkov, "Lignin-KMC: A Toolkit for Simulating Lignin Biosynthesis," ACS Sustain. Chem. Eng., **7**(22), 18313–18322 (2019).

[37] Y. Wang, J. Kalscheur, E. Ebikade, Q. Li, and D.G. Vlachos, "LigninGraphs: lignin structure determination with multiscale graph modeling," J. Cheminform., **14:**43(1), (2022).

[38] S. Cheranma devi Eswaran, S. Subramaniam, U. Sanyal, R. Rallo, and X. Zhang, "Molecular structural dataset of lignin macromolecule elucidating experimental structural compositions," Sci. Data, **9**(1), 647, (2022).

[39] L.D. Dellon, A.J. Yanez, W. Li, R. Mabon, and L.J. Broadbelt, "Computational Generation of Lignin Libraries from Diverse Biomass Sources," Energy and Fuels **31**(8), 8263–8274 (2017).

[40] A.J. Yanez, W. Li, R. Mabon, and L.J. Broadbelt, "A Stochastic Method to Generate Libraries of Structural Representations of Lignin," Energy and Fuels **30**(7), 5835–5845 (2016).

[41] R.C. Andeme Ela, S. Raza, P.A. Heiden, J. V Vermaas, and R.G. Ong, "Lignin Nanoparticle Morphology Depends on Polymer Properties and Solvent Composition: an Experimental and Computational Study," ACS Appl. Polym. Mater., **4**(10), 6925–6935 (2022).

[42] X. Kang, A. Kirui, M.C. Dickwella Widanage, F. Mentink-Vigier, D.J. Cosgrove, and T. Wang, "Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR," Nat. Commun., **10**(1), 1–9 (2019).

[43] M. Rubinstein, *Polymer Physics* (Oxford University Press, 2003).

[44] J.C. Phillips, D.J. Hardy, J.D.C. Maia, J.E. Stone, J. V Ribeiro, R.C. Bernardi, R. Buch, G. Fiorin, J. Hénin, W. Jiang, R. McGreevy, M.C.R. Melo, B.K. Radak, R.D. Skeel, A. Singharoy, Y. Wang, B. Roux, A. Aksimentiev, Z. Luthey-Schulten, L. V Kalé, K. Schulten, C. Chipot, and E. Tajkhorshid, "Scalable molecular dynamics on CPU and GPU architectures with NAMD," J. Chem. Phys., **153**(4), 44130 (2020).

[45] W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual molecular dynamics," J. Mol. Graph., **14**(1), 33–38 (1996).

[46] B.H. Zimm, "Apparatus and Methods for Measurement and Interpretation of the Angular Variation of Light Scattering; Preliminary Results on Polystyrene Solutions," J. Chem. Phys., **16**(12), 1099–1116 (1948).

[47] G. V Schulz, "Über die Kinetik der Kettenpolymerisationen. V.," Zeitschrift Für Physikalische Chemie **43B**(1), 25–46 (1939).

[48] J. V Vermaas, D.J. Hardy, J.E. Stone, E. Tajkhorshid, and A. Kohlmeyer, "TopoGromacs: Automated Topology Conversion from CHARMM to GROMACS within VMD," J. Chem. Inf. Model., **56**(6), 1112–1116 (2016).

[49] P.J. Flory, "Molecular Size Distribution in Linear Condensation Polymers1," J. Am. Chem. Soc., **58**(10), 1877–1885 (1936).

[50] Y. Kawagoe, D. Surblys, H. Matsubara, G. Kikugawa, and T. Ohara, "Construction of polydisperse polymer model and investigation of heat conduction: A molecular dynamics study of linear and branched polyethylenimine," Polymer (Guildf) **180**(August), 121721 (2019).

[51] H. Tobita, "Random Branching of Polymer Chains with Schulz–Zimm Distribution. 1. Bivariate Distribution and Related Formulae," Macromol. Theory Simul., **29**(2), 1900056 (2020).

[52] A.M. Striegel, W.W. Yau, J.J. Kirkland, and D.D. Bly, "Frontmatter, Modern Size-Exclusion Liquid Chromatography," i–xvi, John Wiley & Sons, Inc. (2009)

[53] C.G. Wilkerson, S.D. Mansfield, F. Lu, S. Withers, J.-Y. Park, S.D. Karlen, E. Gonzales-Vigil, D. Padmakshan, F. Unda, J. Rencoret, and J. Ralph, "Monolignol Ferulate Transferase Introduces Chemically Labile Linkages into the Lignin Backbone," Science (1979) **344**(6179), 90–93 (2014).

[54] J.H. Grabber, S. Quideau, and J. Ralph, *P-COUMAROYLATED SYRINGYL UNITS IN MAIZE LIGNIN: IMPLICATIONS FOR r-ETHER CLEAVAGE BY THIOACIDOLYSIS,* Phytochem.,

43 *(6), 1189-1194* (1996).

[55] H.J.C. Berendsen, D. van der Spoel, and R. van Drunen, "GROMACS: A message-passing parallel molecular dynamics implementation," Comput. Phys. Commun., **91**(1), 43–56 (1995).

[56] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, and H.J.C. Berendsen, "GROMACS: fast, flexible, and free.," J. Comput. Chem., **26**(16), 1701–1718 (2005).

[57] L. Petridis, and J.C. Smith, "A molecular mechanics force field for lignin," J. Comput. Chem., **30**(3), 457–467 (2009).

[58] O. Guvench, S.S. Mallajosyula, E.P. Raman, E. Hatcher, K. Vanommeslaeghe, T.J. Foster, F.W. Jamison, and A.D. MacKerell, "CHARMM Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and Carbohydrate–Protein Modeling," J. Chem. Theory Comput., **7**(10), 3162–3180 (2011).

[59] J. V. Vermaas, L. Petridis, J. Ralph, M.F. Crowley, and G.T. Beckham, "Systematic parameterization of lignin for the CHARMM force field," Green Chem., **21**(1), 109–122 (2019).

[60] M.P. Allen, and D.J. Tildesley, "Computer simulation of liquids: Second edition," Computer Simulation of Liquids: Second Edition, 1–626, Oxford Academic, (2017).

[61] D. Frenkel, Smit B "Understanding Molecular Simulation: From Algorithm to Applications" BBT-UMS. (Second E. Eds.; Academic Press, San Diego, 2002), pp. xix–xxii.

[62] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen, "A smooth particle mesh Ewald method," J. Chem. Phys., **103**(19), 8577–8593 (1995).

[63] T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems," J. Chem. Phys., **98**(12), 10089–10092 (1993).

[64] B. Hess, H. Bekker, H.J.C. Berendsen, and J.G.E.M. Fraaije, "LINCS: A linear constraint solver for molecular simulations," J. Comput. Chem., **18**(12), 1463–1472 (1997).

[65] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling.," J. Chem. Phys., **126**(1), 14101 (2007).

[66] H.J.C. Berendsen, "Transport properties computed by linear response through weak coupling to a bath" in edited by M. Meyer and V. Pontikis (Springer Netherlands, Dordrecht, 1991), pp. 139–155 in Computer simulation in material science.

[67] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, and J.R. Haak, "Molecular dynamics with coupling to an external bath," J. Chem. Phys., **81**(8), 3684–3690 (1984).

[68] S. Nosé, and M.L. Klein, "Constant pressure molecular dynamics for molecular systems," Mol. Phys., **50**(5), 1055–1076 (1983).

[69] M. Parrinello, and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," J. Appl. Phys., **52**(12), 7182–7190 (1981).

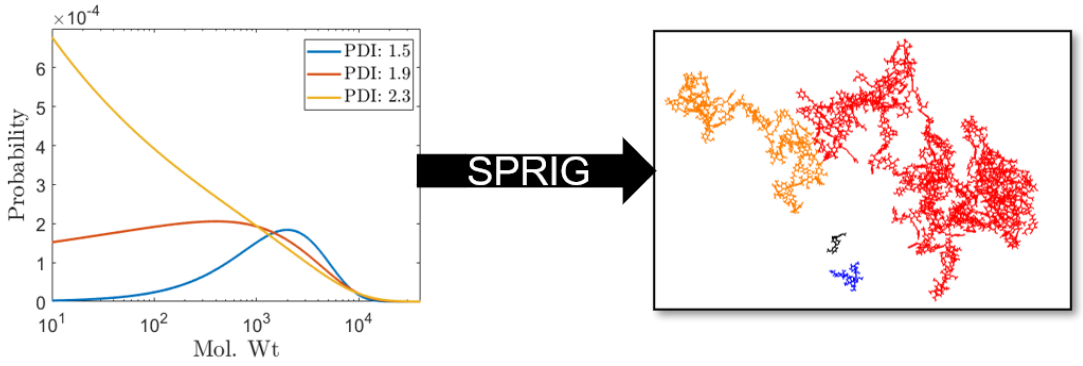
[70] Z. Börcsök, and Z. Pásztory, "The role of lignin in wood working processes using elevated temperatures: an abbreviated literature survey," Euro. J. Wood. Wood. Prod., **79**(3), 511–526 (2021).

[71] M. Poletto, "Assessment of the thermal behavior of lignins from softwood and hardwood species," Maderas: Cienc. Tecnol., **19**(1), 63–74 (2017).

[72] N. Kaliyan, and R.V. Morey, "Natural binders and solid bridge type binding mechanisms in briquettes and pellets made from corn stover and switchgrass," Bioresour. Technol., **101**(3), 1082–1090 (2010).

[73] A. Chremos, and J.F. Douglas, "Communication: When does a branched polymer become a particle?," J. Chem. Phys., **143**(11), 111104, (2015).

[74] J.F. Douglas, "Swelling and growth of polymers, membranes, and sponges," Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics, **54**(3), 2677–2689 (1996).



For Table of Contents Use Only.

"Atomistic Simulations of Polydisperse Lignin Melts using Simple Polydisperse Residue Input Generator" by Vaidyanathan Sethuraman, Josh V. Vermaas, Luna Liang, Arthur J. Ragauskas, Jeremy C. Smith and Loukas Petridis