ACS Sustainable Chemistry & Engineering

Research Article

Cite This: ACS Sustainable Chem. Eng. 2019, 7, 18313−18322

pubs.acs.org/journal/ascecg

Lignin-KMC: A Toolkit for Simulating Lignin Biosynthesis

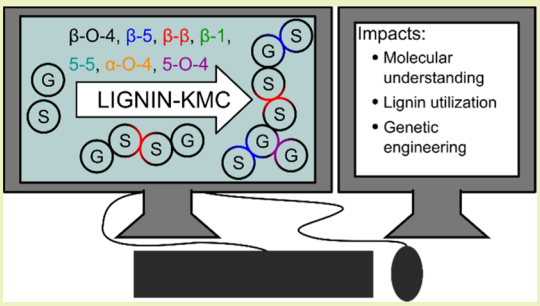
Michael J. Orella,[†] Terry Z. H. Gani,[†] Josh V. Vermaas,[‡] Michael L. Stone,[†] Eric M. Anderson,[†] Gregg T. Beckham,[§] Fikile R. Brushett,*[,†] and Yuriy Román-Leshkov*[,†]

[†]Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

[‡]Biosciences Center and [§]National Bioenergy Center, National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, Colorado 80401, United States

Supporting Information

ABSTRACT: Lignin is an abundant biopolymer of phenylpropanoid monomers that is critical for plant structure and function. Based on the abundance of lignin in the biosphere and interest in lignin valorization, a more comprehensive understanding of lignin biosynthesis is imperative. Here, we present an open-source software toolkit, Lignin-KMC, that combines kinetic Monte Carlo and first-principles calculations of radical coupling events to model lignin biosynthesis in silico. Lignification is simulated using the Gillespie algorithm with rates derived from density functional theory calculations of individual fragment couplings. Using this approach, we confirm experimental findings regarding the impact of lignification conditions on the polymer structure such as (1) the positive correlation between sinapyl alcohol fraction and depolymerization yield and (2) the primarily benzodioxane linked structure of C-lignin. Additionally, we identify the *in planta* monolignol supply rate as a possible control mechanism for lignin biosynthesis based on evolutionary stresses. These examples not only highlight the robustness of our modeling framework but also motivate future studies of new lignin types, unexplored monolignol chemistries, and lignin structure predictions, all with an overarching aim of developing a more comprehensive molecular understanding of native lignin, which, in turn, can advance the biological and chemistry communities interested in this important biopolymer.



KEYWORDS: *lignin biosynthesis, lignin structure, polymerization, radical coupling, lignin polymerization, kinetic Monte Carlo*

1. INTRODUCTION

Lignin is an alkyl-aromatic polymer that gives structural integrity to plant matter and protects it from microbial attack.[1−5] As a highly abundant natural biopolymer, lignin has substantial promise as a renewable feedstock for chemical synthesis of aromatics.[6−10] To economically convert lignocellulosic biopolymers, effective utilization of both polysaccharides and lignins is imperative as wasted feedstocks rapidly decrease the profitability of any biorefinery venture.[11] In current generation biorefineries, we are well equipped to upgrade the polysaccharide fraction through a number of catalytic[12−17] and biological[18−20] strategies. However, our understanding of lignin biosynthesis remains incomplete, limiting our ability to fully valorize lignin. A better quantitative understanding of the complex native lignin structure can enable the advancement, optimization, and, ultimately, deployment of recently proposed depolymerization techniques,[8,21−23] improve the suite of natural lignin-based polymer products,[24−29] provide a starting point for analysis of higher molecular weight lignin fragments,[30−34] and guide the efforts of plant biologists interested in genetically modifying plants to produce designed lignins.[35−39]

Lignin biosynthesis, or lignification, begins in the cytoplasm where a series of metabolic pathways convert phenylalanine and tyrosine into primarily coniferyl alcohol (G-unit), sinapyl alcohol (S-unit), and *p*-coumaryl alcohol (H-unit), which are collectively known as the monolignols.[4,5,9,40−46] These molecules are subsequently transported to the cell wall through a number of mechanisms,[47,48] where peroxidase and laccase enzymes mediate their oxidation into phenolic radicals. The radicals delocalize over the aromatic structure, resulting in coupling reactions that generate a complex phenolic polymer network containing a variety of intermolecular C−O or C−C bonds.[1,2] Figure 1 illustrates the most common chemical linkages found in native lignin fragments and their associated nomenclature that will be used throughout the remainder of this manuscript.

Previously, multiple research groups have attempted to model the lignification process to obtain fundamental understanding via a variety of stochastic methods parameterized by experimental data on bond distributions. For example, SIMREL, one of the earliest models of lignification, added monomers to growing chains of lignin so long as the chains conformed to experimentally measured lignins.[49] Similarly, to

Received: June 21, 2019
Revised: September 23, 2019
Published: September 24, 2019

DOI: 10.1021/acssuschemeng.9b03534
ACS Sustainable Chem. Eng. 2019, 7, 18313−18322

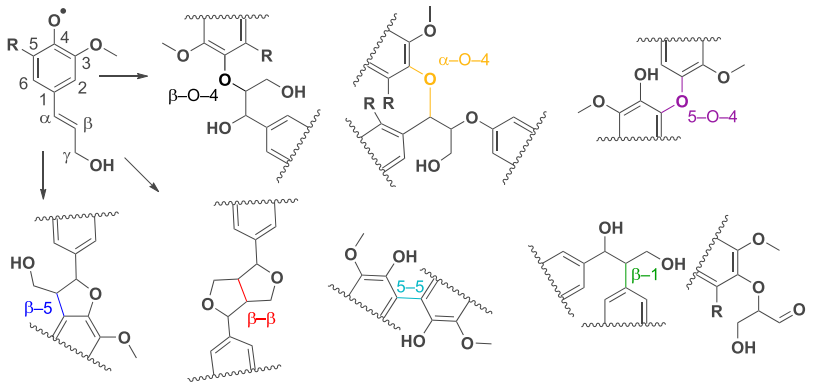


Figure 1. Lignification reactions that can occur in H, S, and G lignin between oxidized fragments. These connections represent the linkages that are most commonly found in native lignins and the linkages represented in the Lignin-KMC code. The characteristic bond for each linkage is color coded, and this color code is used throughout the main text. The β-1 bond is drawn in the form found under acidic conditions, with the aldehyde separated from the remainder of the bond (bottom right).

better understand the geometry of lignin within the middle lamella, Jurasek's model of lignification in the cell wall relied on bond formation probabilities obtained from experimental data.[50,51] Most recently, Yanez et al. used Metropolis Monte Carlo methods to generate lignin topologies that satisfy measured molecular weight, monomer, and bond distributions.[52,53] Leveraging these topologies, Broadbelt and coworkers have gained an improved understanding of lignin pyrolysis through the use of kinetic Monte Carlo methods.[54] Further building upon this work, Vermaas et al. developed methods to convert the topological libraries produced by Yanez et al. to structural libraries resembling those of Jurasek, which simulate the three-dimensional structure of lignin within cell-wall-like environments.[55] Parijs et al. were among the first to utilize kinetic Monte Carlo methods for predictive lignin simulations, but again their model was parameterized by information obtained from synthetic dehydrogenation polymers (DHPs).[56] While the abovementioned methods are parameterized primarily by experimental data representing a population of macromolecules that may have changed during the extraction process, an alternative approach is to parameterize models with data from ab initio predictions. Several prior density functional theory (DFT) calculations investigating monolignol coupling have focused on the thermodynamics of bond formation.[57−61] If kinetics are invoked at all, it is only in the control of β-O-4 reactions, as thermodynamics alone cannot explain their qualitative frequency.[62−64] Recently, we demonstrated that the full suite of radical coupling activation energies qualitatively reproduce many experimentally observed lignin findings,[65] motivating the parameterization of stochastic models with these energies for deeper, more quantitative, understanding of the lignin structure. To date, no lignification model currently exists that: (1) can be readily extended to newly-discovered monolignols, (2) is compatible with parameters derived from ab initio calculations that can be refined as the accuracy of such calculations improve, and (3) is easily accessible to the diverse community of chemists, engineers, and biologists focused on lignin utilization.

Here, we introduce Lignin-KMC,[66] an open-source toolkit for the simulation of lignification that combines the benefits of stochastic chemical kinetic methodologies and fundamental chemical behavior obtained from ab initio calculations. Rather than relying on traditional empirical parameterizations, Lignin-KMC uses relative rates estimated from the Eyring equation and transition-state energetics calculated at the M06-2X/def2-TZVP level of theory in implicit water.[65] By using this hybrid strategy, we achieve monomer resolution within the polymer while incurring a lower computational cost as compared to ab initio simulations of the lignification process in a solvent. Close agreement between the simulated outputs and experimentally measured structural properties of lignins with typical hardwood and softwood compositions illustrates the predictive capability of the model. We then apply the Lignin-KMC method to investigate the impact of the S/G ratio on depolymerization monomer yields and to explore in detail the importance of monolignol addition rate as a parameter in controlling the resulting lignin chemistry *in planta*. With minimal additions, we extend the model to simulate the formation of homopolymers of caffeyl alcohol (C-unit) and accurately predict their bond formation tendencies. With the source code publicly available on GitHub (https://github.com/michaelorella/lignin-kmc),[66] this toolkit can be readily used and built upon by researchers in areas focusing on lignin, ranging from plant biology and genetic engineering to catalytic depolymerization and renewable polymers derived from lignin, to gain deeper understanding of the lignification process and the implications of the lignin structure on valorization strategies.

2. RESULTS AND DISCUSSION

2.1. Model Development. Lignin-KMC relies on four main concepts: (1) a sparse matrix-based data structure to represent lignin in silico, (2) generation of possible reactions between lignin fragments, (3) computation of transition probabilities between two states; and (4) repetition of kinetic Monte Carlo simulations to obtain statistical distributions from which average lignin topologies can be calculated.

2.1.1. Data Structure. A graph structure representing the chemical connectivity of lignin allows straightforward mapping of lignification into a computational framework. This graph contains N nodes representing the simulated monomers, where each node is an object from the created "monomer" class in Python, connected by edges whose weights signify the types of bonds connecting them. The Supporting Information contains additional details on this data structure. Figure 2a shows the numbering and types of monolignols that constitute lignin.[1,2]

The β-O-4, β-5, β–β, β-1, α-O-4, 5-5, and 5-O-4 linkages that occur with measurable frequency in extracted lignins are

DOI: 10.1021/acssuschemeng.9b03534
ACS Sustainable Chem. Eng. 2019, 7, 18313−18322

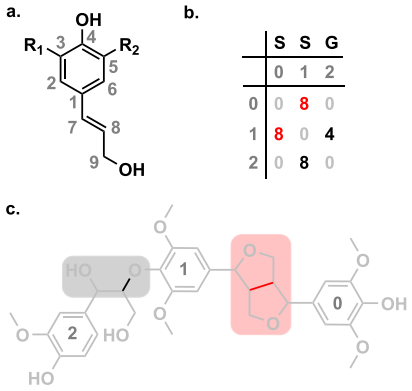


Figure 2. Data structure used for lignin topology throughout the kinetic Monte Carlo simulations. The generic monolignols shown in (a) are sinapyl alcohol (S-unit, R_1 = OMe, R_2 = OMe), coniferyl alcohol (G-unit, R_1 = OMe, R_2 = H), *p*-coumaryl alcohol (H-unit, R_1 = H, R_2 = H), and caffeyl alcohol (C-unit, R_1 = H, R_2 = OH). The adjacency matrix shown in (b) contains two bonds, one β–β (highlighted in red) between monomers 0 and 1, and one β-O-4 (highlighted in black) between monomers 1 and 2. The chemical structure of a lignin fragment with this adjacency matrix is shown in (c) with the same color highlighting, where the colored bonds represent the named bonds that are formed and the highlighted boxes show bonds around the formed bond that have changed as a consequence.

easily stored in the adjacency matrix, shown by example in Figure 2b. This matrix, A, corresponds to the fragment shown in Figure 2c and contains two bonds: red β–β between S-units 0 and 1, $A_{01} = A_{10} = 8$, and black β-O-4 between S-unit 1 and G-unit 2, $A_{12} = 4$, $A_{21} = 8$. The effects of rapid rearrangements accompanying the aforementioned bonds are highlighted by colored boxes in Figure 2c. As compared to the string-based representations of lignin,[56] the sparse adjacency matrix-based graph structure provides a method for storing lignin graphs that easily allows branching and monolignol indexing.

2.1.2. Possible Simulation Events. Oxidized fragments are known to react from the β, 1, 5, and 4 positions with any other nearby oxidized fragments at these same positions to produce the observed β-O-4, β-5, β–β, β-1, α-O-4, 5-5, and 5-O-4 linkages.[3,67] However, by allowing any monomer to react with any other monomer, the number of possible events grows rapidly with the number of monomers in the simulation. To avoid this, previous models have restricted the number of possible reactions by forcing reactions to occur only between monomers and a growing chain[49] or by limiting the number of reactive sites within the monomers.[56] Rather than following this approach, in Lignin-KMC, we allow all possible active sites on the monolignols to react with growing chains of any size to replicate the lignification process as it occurs within the cell walls of a plant. As such, we stipulate that only oxidized species may react with other oxidized species. We choose to neglect possible radical propagation reactions between a single oxidized species and a nonoxidized counterpart as these are predicted to incur large energetic penalties as compared to radical coupling,[62] and are not expected to be overcome by concentration effects. After the radical coupling, tautomerizations of all bonds other than β-O-4 were assumed to be instantaneous and irreversible, as none of these intermediate products are observed experimentally.[1,2] For the β-O-4 bond, however, we explicitly considered the difference in barriers leading from the unstable quinone methide intermediate to (1) the β-O-4 linkage by hydration, and (2) the α-O-4 bond by chain propagation.

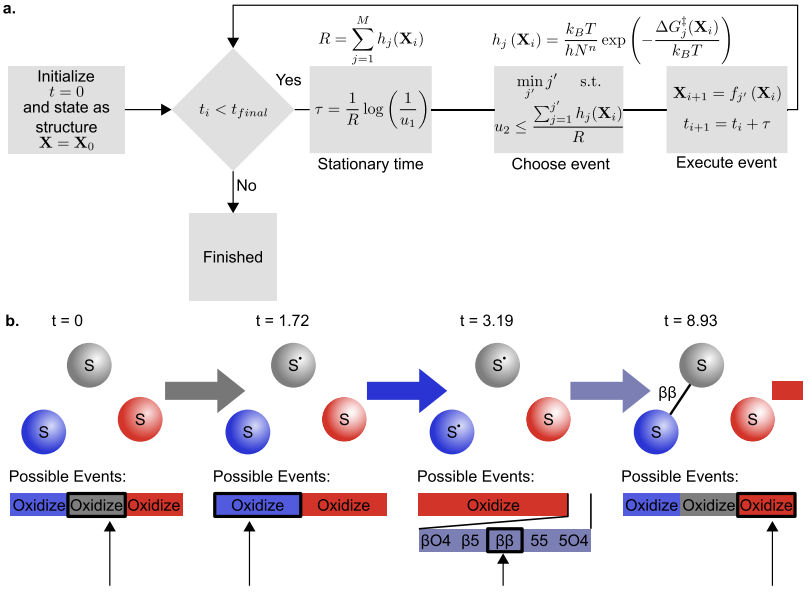


Figure 3. (a) Logic flow of the Gillespie algorithm used to perform the stochastic simulation of lignification. The state **X** is represented by the graph previously described. The stationary time equation translates the number u_1 sampled from a uniform distribution to a random number sampled on the exponential distribution. The event choice equation minimizes the event index j' such that the partial sum up to and including j' surpasses the value u_2. Finally, updating the state takes on a specific function $f_{j'}$ depending on the event chosen. (b) Schematic evolution of a lignin state using the Gillespie algorithm. In this schematic, three S-units in their natural state are initialized in the simulation at $t = 0$. Each frame of the schematic shows the possible events stacked onto a line normalized between 0 and 1, where the length of the individual bar is proportional to the rate of the event listed. The arrow below this number line represents a random number selection that determines which event of the listed possibilities will be performed.

DOI: 10.1021/acssuschemeng.9b03534
ACS Sustainable Chem. Eng. 2019, 7, 18313−18322

In addition to chemical reaction rates, the rate of monolignol transport to the cell wall has been previously predicted to strongly influence lignin polymerization by altering instantaneous monolignol concentrations.[56] To allow Lignin-KMC to dynamically add monolignols to the simulation, the adjacency matrix can grow in time. By combining the flexibility of the data structure with the extensibility of the event definitions, it is straightforward to incorporate new lignin understanding into this model structure.

2.1.3. Transition Probabilities. To accurately represent monomer coupling chemistry, we incorporated intrinsic chemical reaction rates associated with oxidation and coupling into the model. The coupling rates were approximated by applying the Eyring equation to energy barriers obtained from DFT calculations (Table S1)[65] for the reactions of monomer and dimer couplings to form β-O-4, β-5, $\beta-\beta$, β-1, α-O-4, 5-5, and 5-O-4 bonds. Similarly, oxidation rates were approximated by calculating a radical transfer activation barrier with an H-unit. Although DFT calculations are known to be error prone for absolute rate predictions, the relative rates, which benefit from error cancellation,[61,68] dictate the probabilities of reaction. Equation 1 shows the rate calculation, where $h_j(\mathbf{X})$ is the rate of event j in state $\mathbf{X}$, $\Delta G_j^{\ddagger}(\mathbf{X})$ is the difference between reactant and transition state energy obtained from DFT calculations, k_B is the Boltzmann constant, T the absolute temperature of the reaction (T = 298.15 K for all results presented here), N is the number of monomers in the simulation, n is the molecularity of the reaction, and h is Planck's constant. Because each monomer is treated as an individual species, the lignin species concentration dependence of these rates is handled implicitly, while any solvent and catalyst concentration effects are not included in the model.

$$h_j(\mathbf{X}) = \frac{k_B T}{h N^n} \exp\left(-\frac{\Delta G_j^{\ddagger}(\mathbf{X})}{k_B T}\right) \quad (1)$$

We treat the rate of monolignol transport as a parameter in lignification control. Rather than focus on particular transport rates, we parametrically sweep through the possible rates, which can be estimated at about 10^8 monomers per second from the entire cell.[69] However, with a cell area of about 100 μm^2 and growing chains requiring 1 nm^2 for growth, each chain could get approximately 1 monomer per second added to the chain. Previous models' reliance on experimental data to obtain transition probabilities only allows for regressive uses. However, based on new data available through DFT calculations,[65] we can obtain transition probabilities that enable predictive calculations. Although significant improvements have been realized in these DFT calculations, further insights could be offered by obtaining reaction rates from ab initio molecular dynamics-based simulations that can explicitly account for solvation or confinement effects during polymerization. We use the Gillespie algorithm,[70,71] a well-known methodology for simulating polymerization chain reactions,[72−74] to perform our stochastic simulation of the lignification process. Each iteration of the algorithm updates the state by drawing the time between events (stationary time) from an exponential distribution and selecting an event based on the weights determined through eq 1. In the representation of the Gillespie algorithm shown in Figure 3a, $\mathbf{X}_i$ is the simulation state and t_i is the simulation time at iteration i, R is the net rate of all possible events in the current state, M is the number of events at the current state, u_1 and u_2 are pseudo-random numbers on the uniform distribution, τ is the time that passes between the occurrence of individual events, j' is the index of the specific event that will occur, and $f_{j'}$ is the specific function that updates the state to execute event j'. For the mechanics of the implementation, see the source code.[66]

Figure 3b shows a schematic of the time evolution of the algorithm for a simple case of 3 S-units. The simulation begins by considering neutral, disconnected monomers. The only possible events are the oxidations of the three monomers. The rates of these events are normalized by the sum-total rate, stacked onto a number-line between 0 and 1, and compared to u_2 to perform the event selection. Meanwhile, the elapsed time for an individual step can be calculated by sampling an exponential distribution in the total rate, which can be obtained by transforming u_1. In the first iteration, the generated random numbers are $u = [0.1791, 0.6138]^T$, which advances the simulation time to 1.72 and elects to oxidize the gray monomer. At the next iteration, only the blue and the red monomers can be oxidized, and the choice of $u = [0.3789, 0.1012]^T$ oxidizes the blue monomer while advancing the time to 3.19. Now that there are two oxidized monomers, bonds could form between them, or the red monomer could be oxidized. For the sake of this demonstration, $u = [0.1341, 0.9772]^T$, causing bond formation between the two oxidized monomers. The algorithm continues selecting events and updating simulation state until either there are no further events or a specified time threshold is passed. Based on the specific implementation, we avoid regeneration of events that have already been accounted for, limiting the scaling to $O(N^{2.54})$, as shown in Figure S1. To ease the adaptation of these simulation results, we have also developed a method of converting the computational graph structure into a molfile format for visual output and analysis, which is described in detail in the Supporting Information. In addition, conversion tools have also been implemented to interface with other lignin analysis and modeling tools, such as LigninBuilder[55] and NetworkX[75] (Figure S3).

2.1.4. Sampling Process. As individual simulations are inherently stochastic by design, we rely on a method of repeated simulations and aggregated properties, as shown in Figure 4, where n indicates the simulation number and n_{max} the total number of simulations. The most important properties

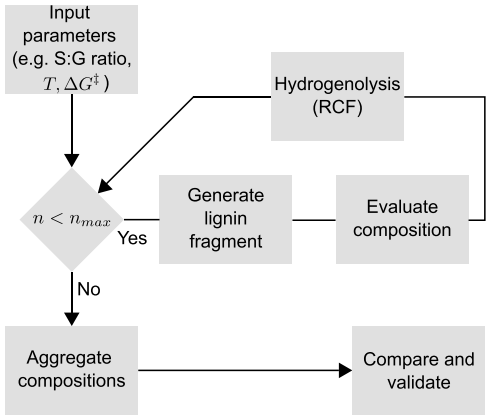


Figure 4. Schematic visualization of sampling methodology used for generation of in silico lignin compositions. Bonds are counted directly from the adjacency matrix, while C−O bonds formed during synthesis are broken instantaneously in this simulation to attempt to observe the theoretical monomer yields obtained from a RCF process.

DOI: 10.1021/acssuschemeng.9b03534
ACS Sustainable Chem. Eng. 2019, 7, 18313−18322

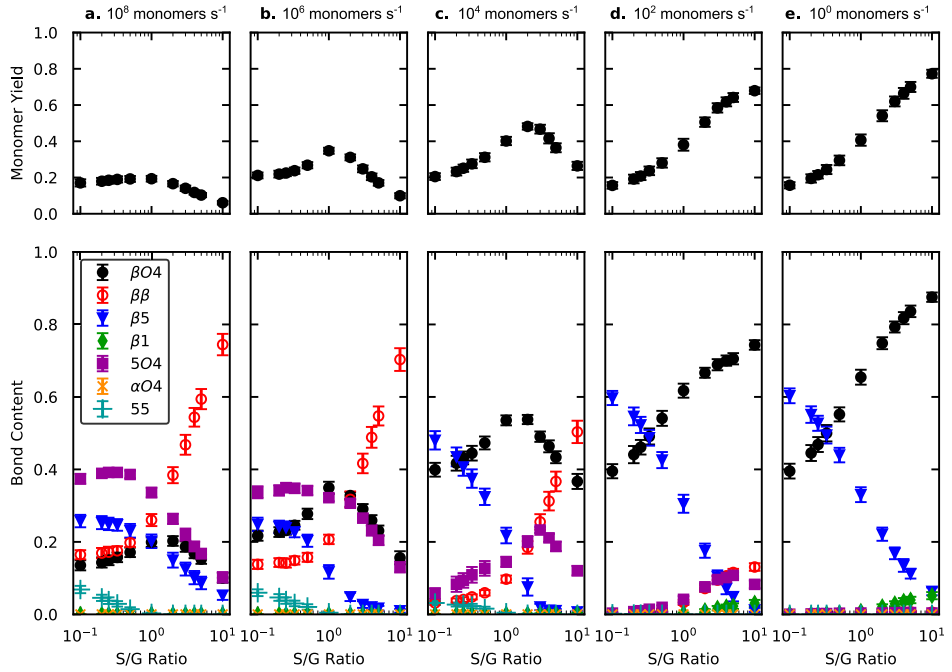


Figure 5. Bond and monomer compositions measured from lignin generated in silico under different polymerization conditions where (top) shows the yields of monomers obtained after degradation by ether cleavage treatments and (bottom) shows the fraction of bonds identified in the lignin fragments before hydrogenolysis. Each horizontal panel shows the effect of S/G ratio on these properties at a distinct monomer addition rate, measured in monomers per second, displayed at the top of each panel. Panel (a) mimics conditions of batch polymerization, where monolignol concentration is initially high, while panel (e) mimics conditions of transport-limited lignin polymerization. In the intermediate region (b−d), monolignol transport rates near the kinetic rates are used.

that we examine are the relative bond types and theoretical depolymerization monomer yields, as could be obtained through reductive catalytic fractionation (RCF). The bond contents can be obtained directly from the adjacency matrix, and the monomer yields are obtained by breaking all C−O bonds formed during lignification, as RCF depolymerization processes do,[76] while leaving C−C bonds intact. Once the C−O bonds have been removed from the adjacency matrix, we use depth-first search techniques described in the Supporting Information to identify and size the resulting lignin fragments.

2.2. Model Validation. In Section 2.1, we described the Lignin-KMC model of lignification processes; here, we validate the model and study important open questions in the literature. More specifically, we first investigate the polymerization of lignin composed of varying amounts of S- and G-monomers, with these monomers added to the simulation at varying rates. Through this study, we hope to understand the localization of bonds that are present in the final polymer and the impact that kinetics and transport have on final structure. Second, we study the homo-polymerization of caffeyl alcohol monomers, as these types of polymers have previously been difficult to rationalize computationally.

2.2.1. Analysis of S/G Lignins. Recent studies on the genetic engineering of lignins as feedstocks[35−38] for biorefineries have focused on increasing the ratio of S-units to G-units within the polymer as this is hypothesized to increase the fraction of β-O-4 bonds in the final structure, which should subsequently increase the monomer yield from RCF.[77] To date, studies on the RCF of different biomass species have identified a clear correlation between S/G ratio and depolymerization yield.[21] Genetically modified poplars have been found to be similarly correlated, where high S-content results in high monomer yields.[38,78−80] However, native variants of the same plant species were not similarly correlated, calling into question previous assumptions about how lignin forms from its constituent monomers.[81] Anderson et al. further investigated this effect by studying the distribution of dimers obtained from RCF of different poplars, showing that the S-units were more likely to form C−C dimers.[81] Accordingly, a deeper understanding of the lignification process would enable rational tuning of the control mechanisms that plants use to defend against biological degradation and would help differentiate between competing hypotheses formulated to explain the observed yields in RCF processing.

Based on its flexibility, our model can fully evaluate the two-dimensional space of monomer composition and transport rate to investigate these hypotheses. The discrepancies amongst the prior literature on the impact of S/G ratio suggest the existence of tunable mechanisms for controlling bond frequency and maximum monomer yields. To test the influence of the S/G ratio within the model framework, we systematically varied the S to G ratio between 0.1 and 10 at constant monolignol transport rates, as shown in Figure 5. The monolignol transport rates that are investigated here span 1 monomer per second to 10^8 monomers per second added to the simulation. This range of transport rates was selected to include the typical bond formation rates, which for the computed β-O-4 energetics is 1.5×10^8 reactions per second, such that both mass transport and kinetically limited regimes could be explored. In this study, the transport rate relative to the rates of bond formation events is more relevant than the absolute transport rates. The lowest selected transport rate of 1 monomer per second for individual chains is likely in the correct order of magnitude range for passive monolignol transport. The rate of monolignol transport from the entire cell can be estimated at about 10^8 monomers per second,[69] which

DOI: 10.1021/acssuschemeng.9b03534
ACS Sustainable Chem. Eng. 2019, 7, 18313−18322

would be reduced significantly by spatial restrictions on lignin coupling. At the high end, the transport rates correspond to in vitro processes or the rapid release of lignin stored within the vacuole.[82] Given the inherent uncertainty around the native addition rate, all of the chosen transport rates may be justified under certain conditions. Thus, rather than focus on any particular case, we parametrically sweep the transport rate and focus on the general trends that emerge in the bond distribution.

In Figure 5 there are clear trends in the theoretical monomer yields (top panels) and bond frequencies (bottom panels) as both the S/G ratio and monolignol transport rates are varied. Broadly speaking, as the monolignol transport rate increases, there is a larger maximum $\beta-\beta$ bond content and lower maximum β-O-4 and β-5 bond content. These changes are directly reflected by lower maximum monomer yields because of the decreased prevalence for C−O bonds.

The monolignol transport rate has been explored experimentally in the literature through the synthesis of DHP lignins via batch (Zulauf-type) and slow-feed (Zutropf-type) processes. The data observed under conditions mimicking Zulauf DHP experiments, where monomers are added quickly, are shown in Figure 5a. The trend of decreasing monomer yield with increasing S-content appears contrary to most observations and intuition. However, in this case, Lignin-KMC predicts a higher frequency of $\beta-\beta$ linkages created between S-units that terminate the polymer growth, whereas competing linkages at the 5- position are common in high-G-lignin, matching what has been simulated previously.[56] Most reports of Zulauf polymerization experiments suggest the presence of many low-molecular-weight fragments, with the occurrence of primarily $\beta-\beta$ bonds between S-units,[83] and β-5 bonds between G-units.[84] Interestingly, we also see a large fraction of 5-O-4 connections between high G-content lignins, but these tend to form between larger fragments after the formation of β-5 bonds (Figures S4−S6). To more closely match observed yields, we attempted simulating lignification under conditions that may more closely resemble those *in planta* (e.g., Zutropf experiments in vivo). Under these slow monomer addition conditions, shown in Figure 5e, increasing S-content drastically increases β-O-4 content, which subsequently increases the depolymerization monomer yield (Figure S7), as previously predicted.[77,85] This appears to match what has been observed with transgenic lignins,[38,78] and those generated during Zutropf DHP experiments,[86−88] as well as the lignins simulated by Merks and co-workers.[56] In addition, under these conditions, the bond distributions that we simulate match closely what has been reported previously.[9,67] At an S/G ratio of 1 and transport rate of 1 monomer per second, we observe a β-O-4 bond frequency near 65%, while typical reported values are between 50 and 62%.[9,67] Similarly, we observe $\beta-\beta$ bond contents of near 1% while reported values are usually between 2 and 12%.[9,67] The largest deviations that we see exist in β-5 and β-1 bonds, where we expect 3−11% but observe 33% and expect 1−7% but observe 1% respectively.[9,67] The relatively high frequency of β-5 bonds is more similar to what has been observed in DHP lignins,[84] but the differences between the in vitro polymers and native lignins remain unclear. Based on the stark contrast between these two limiting cases, we found it important to investigate some of the intermediate transport rate behavior.

The data corresponding to these intermediate transport regimes, which are yet to be computationally investigated, are shown in Figure 5b−d. These data show the gradual changes in predicted bond content and depolymerization yields as a function of the transport rate. Interestingly, it is possible to observe extrema in the monomer yields of Figure 5b,c at S/G ratios between 1 and 5, which can be explained by a tradeoff between high $\beta-\beta$ bond content between S-units and high β-5 bond content between G-units. Perhaps unsurprisingly given the behavior in the limiting cases, at intermediate monolignol transport rates, regions of relative invariance appear in the monomer yield as the S/G ratio changes. The existence of such plateaus under specific monolignol transport rates could help explain why monomer yields do not appear to change as the S/G ratio changes in wild-type lignins.[81]

2.2.2. Analysis of Nontraditional Lignins. Recent discoveries of lignins composed of building blocks beyond traditional S- and G-monomers have inspired a number of different applications.[39,44,89−91] These monomers are still typically phenylalanine-derived species but are undermethylated as compared to their S-, G-, and H-counterparts. Homopolymers of caffeyl alcohol (C-lignin) hold particular promise as a feedstock for fuel and chemical production because of the near-exclusive presence of interunit benzodioxane linkages, a specific type of β-O-4 linkage that can easily be cleaved.[89,90,92,93] Based on this discovery, researchers have sought to understand the thermodynamic origin of the selectivity of benzodioxane linkages compared to a hydrated β-O-4 linkage.[90] In addition, recent efforts have been made to take advantage of the well-defined ether linkage structure by using C-lignins in RCF treatments, where, despite the high theoretical yields, catalyst deactivation limits the actual yields.[92,93] Therefore, in the present model, it is more informative to compare the results of bond contents than those of monomer yields as would be possible for S/G-lignin.

The tunability of the model enables simulation of these novel lignins by defining new monomer types within the model and updating the set of transition state energies obtained from kinetic DFT calculations (Table S2). Notably, with those two changes, the same methodology employed in the previous section and knowledge gained for modeling S/G lignins can be applied to guide the exploration of this polymer. Specifically, C-lignin was synthesized in silico using a transport rate of 1 C-units per second, as at this transport rate, long chains were preferred over dimer products in S/G lignin. The distribution for the simulations run is shown in Figure 6. Despite the lower predicted activation barrier of β-5 bond formation for the coupling of two monomers, our calculations show the resulting lignin is highly enriched in β-O-4 content, matching the native lignin found in seed coats as determined by experimentation.[89] This counter-intuitive result can again be understood by comparing the energetics of monomer−monomer additions, with monomer−chain additions (Table S2). Thus, with minimal adaptation, we can accurately describe the structure of this lignin type. We anticipate that, as new lignin forms are discovered, or additional parameters of interest are identified, this model can be easily adjusted to incorporate this new behavior.

3. CONCLUSIONS

In this work, we demonstrated the general framework and broad applicability of the Lignin-KMC toolkit for simulating lignification using both traditional S/G-lignin and the recently discovered C-lignin. Through these applications, we have verified the conclusions drawn from previous models, garnered

DOI: 10.1021/acssuschemeng.9b03534
ACS Sustainable Chem. Eng. 2019, 7, 18313−18322

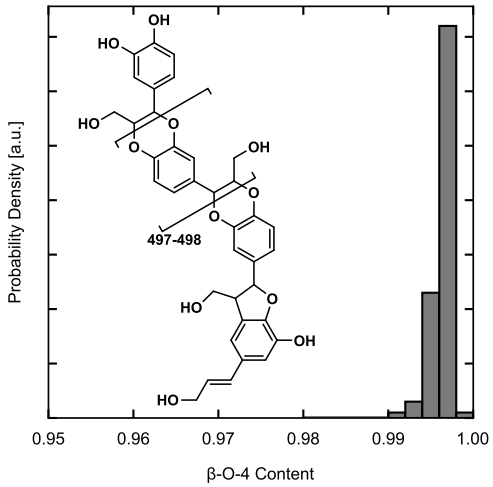


Figure 6. β-O-4 bond content distribution for the prediction of C-lignin generation. For these simulations, monolignols were added at a rate of 1 unit/s, starting with 2 monomers, with a maximum of 500 monomers included.

better understanding of the chemistry that occurs during lignification, and have illustrated the impact of intermediate monolignol supply rates on the lignin structure. Specifically, we have shown, for the first time, the behavior of lignification chemistry where transport processes occur on timescales similar to the bond formation kinetics. Through this study, regions of relative invariance of depolymerization yields with changing S/G ratio were identified that could have significant implications for plant biologists attempting to create "designer" lignins. Further, our studies allow us to understand the structure of these S/G lignins that will eventually be used by both the catalytic depolymerization scientists and polymer scientists, providing an improved starting point for their optimization and engineering. We have also shown that, with minimal adaptations, this model can accurately simulate the formation of the C-lignin polymer, which has previously not been attempted. From this point, the Lignin-KMC model can be further used to explore different biosynthetic parameters of interest to help provide a better starting point for more realistic models of designer lignins for use in detailed computational studies. We note that this model is currently unable to explore the effects of spatial orientation or confinement of lignin fragments within the cell wall, which we hypothesize could also have significant impact on the final lignin structure. To that end, Lignin-KMC serves as a foundation for future studies and extensions that incorporate these types of effects, including time-dependent monomer delivery rates, leading to a more comprehensive, molecular understanding of the lignification process.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssuschemeng.9b03534.

Full computational details; model parameters; and illustrations of supporting results (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: brushett@mit.edu (F.R.B.).
*E-mail: yroman@mit.edu (Y.R.-L.).

ORCID

Michael J. Orella: 0000-0003-1207-4704
Terry Z. H. Gani: 0000-0003-0357-6390
Josh V. Vermaas: 0000-0003-3139-6469
Gregg T. Beckham: 0000-0002-3480-212X
Fikile R. Brushett: 0000-0002-7361-6637
Yuriy Román-Leshkov: 0000-0002-0025-4233

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work performed was supported by the National Science Foundation, CBET Award no 1454299, and the Center for Bioenergy Innovation, a U.S. Department of Energy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. This work was authored in part by Alliance for Sustainable Energy, LLC, the manager and operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work used the Extreme Science and Engineering Discovery Environment (XSEDE)[94] Comet at the San Diego Super Computer Cluster through allocations TG-CTS180027, which is supported by National Science Foundation grant number ACI-1548562; and TG-MCB090159 which is supported by National Science Foundation grant number ACI-1053575. M.J.O. gratefully acknowledges the National Science Foundation Graduate Research Fellowship Program for financial support. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under grant no. 1122374. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors thank Connor Coley and McLain Leonard for insightful discussion and constructive feedback.

■ REFERENCES

(1) Boerjan, W.; Ralph, J.; Baucher, M. Lignin Biosynthesis. *Annu. Rev. Plant Biol.* **2003**, *54*, 519−546.
(2) Ralph, J.; Lundquist, K.; Brunow, G.; Lu, F.; Kim, H.; Schatz, P. F.; Marita, J. M.; Hatfield, R. D.; Ralph, S. A.; Christensen, J. H.; et al. Lignins: Natural Polymers from Oxidative Coupling of 4-Hydroxyphenyl-Propanoids. *Phytochem. Rev.* **2004**, *3*, 29−60.
(3) Vanholme, R.; Demedts, B.; Morreel, K.; Ralph, J.; Boerjan, W. Lignin Biosynthesis and Structure. *Plant Physiol.* **2010**, *153*, 895−905.
(4) Zhao, Q.; Dixon, R. A. Transcriptional Networks for Lignin Biosynthesis: More Complex than We Thought? *Trends Plant Sci.* **2011**, *16*, 227−233.
(5) Weng, J.-K.; Chapple, C. The Origin and Evolution of Lignin Biosynthesis. *New Phytol.* **2010**, *187*, 273−285.
(6) Huber, G. W.; Iborra, S.; Corma, A. Synthesis of Transportation Fuels from Biomass: Chemistry, Catalysts, and Engineering. *Chem. Rev.* **2006**, *106*, 4044−4098.
(7) Ragauskas, A. J.; Williams, C. K.; Davison, B. H.; Britovsek, G.; Cairney, J.; Eckert, C. A.; Frederick, W. J.; Hallett, J. P.; Leak, D. J.; Liotta, C. L.; et al. The Path Forward for Biofuels and Biomaterials. *Science* **2006**, *311*, 484−489.

(8) Ragauskas, A. J.; Beckham, G. T.; Biddy, M. J.; Chandra, R.; Chen, F.; Davis, M. F.; Davison, B. H.; Dixon, R. A.; Gilna, P.; Keller, M.; et al. Lignin Valorization: Improving Lignin Processing in the Biorefinery. *Science* **2014**, *344*, 1246843.
(9) Rinaldi, R.; Jastrzebski, R.; Clough, M. T.; Ralph, J.; Kennema, M.; Bruijnincx, P. C. A.; Weckhuysen, B. M. Paving the Way for Lignin Valorisation: Recent Advances in Bioengineering, Biorefining and Catalysis. *Angew. Chem., Int. Ed.* **2016**, *55*, 8164−8215.
(10) Schutyser, W.; Renders, T.; Van den Bosch, S.; Koelewijn, S.-F.; Beckham, G. T.; Sels, B. F. Chemicals from Lignin: An Interplay of Lignocellulose Fractionation, Depolymerisation, and Upgrading. *Chem. Soc. Rev.* **2018**, *47*, 852−908.
(11) Davis, R.; Tao, L.; Tan, E. C. D.; Biddy, M. J.; Beckham, G. T.; Scarlata, C.; Jacobson, J.; Cafferty, K.; Ross, J.; Lukas, J.; et al. *Process Design and Economics for the Conversion of Lignocellulosic Biomass to Hydrocarbons: Dilute-Acid and Enzymatic Deconstruction of Biomass to Sugars and Biological Conversion of Sugars to Hydrocarbons; NREL/TP-5100-60223*; National Renewable Energy Lab. (NREL): Golden, CO (United States), 2013.
(12) Su, Y.; Brown, H. M.; Huang, X.; Zhou, X.-d.; Amonette, J. E.; Zhang, Z. C. Single-Step Conversion of Cellulose to 5-Hydroxymethylfurfural (HMF), a Versatile Platform Chemical. *Appl. Catal., A* **2009**, *361*, 117−122.
(13) Luterbacher, J. S.; Alonso, D. M.; Rand, J. M.; Questell-Santiago, Y. M.; Yeap, J. H.; Pfleger, B. F.; Dumesic, J. A. Solvent-Enabled Nonenyzmatic Sugar Production from Biomass for Chemical and Biological Upgrading. *ChemSusChem* **2015**, *8*, 1317−1322.
(14) Luterbacher, J. S.; Rand, J. M.; Alonso, D. M.; Han, J.; Youngquist, J. T.; Maravelias, C. T.; Pfleger, B. F.; Dumesic, J. A. Nonenzymatic Sugar Production from Biomass Using Biomass-Derived γ-Valerolactone. *Science* **2014**, *343*, 277−280.
(15) Fatih Demirbas, M. Biorefineries for Biofuel Upgrading: A Critical Review. *Appl. Energy* **2009**, *86*, S151−S161.
(16) Alonso, D. M.; Wettstein, S. G.; Dumesic, J. A. Bimetallic Catalysts for Upgrading of Biomass to Fuels and Chemicals. *Chem. Soc. Rev.* **2012**, *41*, 8075−8098.
(17) Van de Vyver, S.; Geboers, J.; Jacobs, P. A.; Sels, B. F. Recent Advances in the Catalytic Conversion of Cellulose. *ChemCatChem* **2011**, *3*, 82−94.
(18) Zhang, Y.-H. P.; Lynd, L. R. Toward an Aggregated Understanding of Enzymatic Hydrolysis of Cellulose: Noncomplexed Cellulase Systems. *Biotechnol. Bioeng.* **2004**, *88*, 797−824.
(19) Alvira, P.; Tomás-Pejó, E.; Ballesteros, M.; Negro, M. J. Pretreatment Technologies for an Efficient Bioethanol Production Process Based on Enzymatic Hydrolysis: A Review. *Bioresour. Technol.* **2010**, *101*, 4851−4861.
(20) Langston, J. A.; Shaghasi, T.; Abbate, E.; Xu, F.; Vlasenko, E.; Sweeney, M. D. Oxidoreductive Cellulose Depolymerization by the Enzymes Cellobiose Dehydrogenase and Glycoside Hydrolase 61. *Appl. Environ. Microbiol.* **2011**, *77*, 7007−7015.
(21) Van den Bosch, S.; Schutyser, W.; Vanholme, R.; Driessen, T.; Koelewijn, S.-F.; Renders, T.; Meester, B. D.; Huijgen, W. J. J.; Dehaen, W.; Courtin, C. M.; et al. Reductive Lignocellulose Fractionation into Soluble Lignin-Derived Phenolic Monomers and Dimers and Processable Carbohydrate Pulps. *Energy Environ. Sci.* **2015**, *8*, 1748−1763.
(22) Van den Bosch, S.; Schutyser, W.; Koelewijn, S.-F.; Renders, T.; Courtin, C. M.; Sels, B. F. Tuning the Lignin Oil OH-Content with Ru and Pd Catalysts during Lignin Hydrogenolysis on Birch Wood. *Chem. Commun.* **2015**, *51*, 13158−13161.
(23) Anderson, E. M.; Katahira, R.; Reed, M.; Resch, M. G.; Karp, E. M.; Beckham, G. T.; Román-Leshkov, Y. Reductive Catalytic Fractionation of Corn Stover Lignin. *ACS Sustainable Chem. Eng.* **2016**, *4*, 6940−6950.
(24) Ten, E.; Vermerris, W. Recent Developments in Polymers Derived from Industrial Lignin. *J. Appl. Polym. Sci.* **2015**, *132*, 42069.
(25) Holmberg, A. L.; Nguyen, N. A.; Karavolias, M. G.; Reno, K. H.; Wool, R. P.; Epps, T. H. Softwood Lignin-Based Methacrylate Polymers with Tunable Thermal and Viscoelastic Properties. *Macromolecules* **2016**, *49*, 1286−1295.
(26) Naseem, A.; Tabasum, S.; Zia, K. M.; Zuber, M.; Ali, M.; Noreen, A. Lignin-Derivatives Based Polymers, Blends and Composites: A Review. *Int. J. Biol. Macromol.* **2016**, *93*, 296−313.
(27) Brzonova, I.; Kozliak, E. I.; Andrianova, A. A.; LaVallie, A.; Kubátová, A.; Ji, Y. Production of Lignin Based Insoluble Polymers (Anionic Hydrogels) by C. Versicolor. *Sci. Rep.* **2017**, *7*, 17507.
(28) Wang, S.; Shuai, L.; Saha, B.; Vlachos, D. G.; Epps, T. H. From Tree to Tape: Direct Synthesis of Pressure Sensitive Adhesives from Depolymerized Raw Lignocellulosic Biomass. *ACS Cent. Sci.* **2018**, *4*, 701−708.
(29) Grossman, A.; Vermerris, W. Lignin-Based Polymers and Nanomaterials. *Curr. Opin. Biotechnol.* **2019**, *56*, 112−120.
(30) Liitiä, T. M.; Maunu, S. L.; Hortling, B.; Toikka, M.; Kilpeläinen, I. Analysis of Technical Lignins by Two- and Three-Dimensional NMR Spectroscopy. *J. Agric. Food Chem.* **2003**, *51*, 2136−2143.
(31) Mansfield, S. D.; Kim, H.; Lu, F.; Ralph, J. Whole Plant Cell Wall Characterization Using 2D-NMR. *Nat. Protoc.* **2012**, *7*, 1579−1589.
(32) Sette, M.; Lange, H.; Crestini, C. Quantitative HSQC Analyses of Lignin: A Practical Comparison. *Comput. Struct. Biotechnol. J.* **2013**, *6*, No. e201303016.
(33) Wen, J.-L.; Sun, S.-L.; Xue, B.-L.; Sun, R.-C. Recent Advances in Characterization of Lignin Polymer by Solution-State Nuclear Magnetic Resonance (NMR) Methodology. *Materials* **2013**, *6*, 359−391.
(34) Ralph, J.; Landucci, L. L. NMR of Lignins. *Lignin and Lignans: Advances in Chemistry*; CRC Press: Boca Raton, Florida, 2016; pp 137−244.
(35) Fu, C.; Mielenz, J. R.; Xiao, X.; Ge, Y.; Hamilton, C. Y.; Rodriguez, M.; Chen, F.; Foston, M.; Ragauskas, A.; Bouton, J.; et al. Genetic Manipulation of Lignin Reduces Recalcitrance and Improves Ethanol Production from Switchgrass. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 3803−3808.
(36) Hisano, H.; Nandakumar, R.; Wang, Z.-Y. Genetic Modification of Lignin Biosynthesis for Improved Biofuel Production. In *Biofuels: Global Impact on Renewable Energy, Production Agriculture, and Technological Advancements*; Tomes, D., Lakshmanan, P., Songstad, D., Eds.; Springer New York: New York, NY, 2011; pp 223−235.
(37) Li, X.; Weng, J.-K.; Chapple, C. Improvement of Biomass through Lignin Modification. *Plant J.* **2008**, *54*, 569−581.
(38) Stewart, J. J.; Akiyama, T.; Chapple, C.; Ralph, J.; Mansfield, S. D. The Effects on Lignin Structure of Overexpression of Ferulate 5-Hydroxylase in Hybrid Poplar1. *Plant Physiol.* **2009**, *150*, 621−635.
(39) Oyarce, P.; De Meester, B.; Fonseca, F.; de Vries, L.; Goeminne, G.; Pallidis, A.; De Rycke, R.; Tsuji, Y.; Li, Y.; Van den Bosch, S.; et al. Introducing Curcumin Biosynthesis in Arabidopsis Enhances Lignocellulosic Biomass Processing. *Nat. Plants* **2019**, *5*, 225−237.
(40) Bonawitz, N. D.; Chapple, C. The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype. *Annu. Rev. Genet.* **2010**, *44*, 337−363.
(41) Liu, C.-J. Deciphering the Enigma of Lignification: Precursor Transport, Oxidation, and the Topochemistry of Lignin Assembly. *Mol. Plant* **2012**, *5*, 304−317.
(42) Shi, R.; Sun, Y.-H.; Li, Q.; Heber, S.; Sederoff, R.; Chiang, V. L. Towards a Systems Approach for Lignin Biosynthesis in Populus Trichocarpa: Transcript Abundance and Specificity of the Monolignol Biosynthetic Genes. *Plant Cell Physiol.* **2010**, *51*, 144−163.
(43) Moura, J. C. M. S.; Bonine, C. A. V.; de Oliveira Fernandes Viana, J.; Dornelas, M. C.; Mazzafera, P. Abiotic and Biotic Stresses and Changes in the Lignin Content and Composition in Plants. *J. Integr. Plant Biol.* **2010**, *52*, 360−376.
(44) Ralph, J. Hydroxycinnamates in Lignification. *Phytochem. Rev.* **2010**, *9*, 65−83.

DOI: 10.1021/acssuschemeng.9b03534
ACS Sustainable Chem. Eng. 2019, 7, 18313−18322

(45) Mottiar, Y.; Vanholme, R.; Boerjan, W.; Ralph, J.; Mansfield, S. D. Designer Lignins: Harnessing the Plasticity of Lignification. *Curr. Opin. Biotechnol.* **2016**, *37*, 190−200.
(46) Barros, J.; Serrani-Yarce, J. C.; Chen, F.; Baxter, D.; Venables, B. J.; Dixon, R. A. Role of Bifunctional Ammonia-Lyase in Grass Cell Wall Biosynthesis. *Nat. Plants* **2016**, *2*, 16050.
(47) Liu, C.-J.; Miao, Y.-C.; Zhang, K.-W. Sequestration and Transport of Lignin Monomeric Precursors. *Molecules* **2011**, *16*, 710−727.
(48) Hatfield, R.; Vermerris, W. Lignin Formation in Plants. The Dilemma of Linkage Specificity. *Plant Physiol.* **2001**, *126*, 1351−1357.
(49) Glasser, W. G.; Glasser, H. R. Simulation of Reactions with Lignin by Computer (SIMREL). I. Polymerization of Coniferyl Alcohol Monomers. *Macromolecules* **1974**, *7*, 17−27.
(50) Jurasek, L. Toward a 3-Dimensional Model of Lignin Structure. *J. Pulp Pap. Sci.* **1995**, *21*, J274−J279.
(51) Jurasek, L. Molecular Modelling of Fibre Walls. *J. Pulp Pap. Sci.* **1998**, *24*, 209−212.
(52) Yanez, A. J.; Li, W.; Mabon, R.; Broadbelt, L. J. A Stochastic Method to Generate Libraries of Structural Representations of Lignin. *Energy Fuels* **2016**, *30*, 5835−5845.
(53) Dellon, L. D.; Yanez, A. J.; Li, W.; Mabon, R.; Broadbelt, L. J. Computational Generation of Lignin Libraries from Diverse Biomass Sources. *Energy Fuels* **2017**, *31*, 8263−8274.
(54) Yanez, A. J.; Natarajan, P.; Li, W.; Mabon, R.; Broadbelt, L. J. Coupled Structural and Kinetic Model of Lignin Fast Pyrolysis. *Energy Fuels* **2018**, *32*, 1822−1830.
(55) Vermaas, J. V.; Dellon, L. D.; Broadbelt, L. J.; Beckham, G. T.; Crowley, M. F. Automated Transformation of Lignin Topologies into Atomic Structures with LigninBuilder. *ACS Sustainable Chem. Eng.* **2019**, *7*, 3443−3453.
(56) van Parijs, F. R. D.; Morreel, K.; Ralph, J.; Boerjan, W.; Merks, R. M. H. Modeling Lignin Polymerization. I. Simulation Model of Dehydrogenation Polymers. *Plant Physiol.* **2010**, *153*, 1332−1344.
(57) Watts, H. D.; Mohamed, M. N. A.; Kubicki, J. D. Evaluation of Potential Reaction Mechanisms Leading to the Formation of Coniferyl Alcohol α-Linkages in Lignin: A Density Functional Theory Study. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20974−20985.
(58) Sangha, A. K.; Parks, J. M.; Standaert, R. F.; Ziebell, A.; Davis, M.; Smith, J. C. Radical Coupling Reactions in Lignin Synthesis: A Density Functional Theory Study. *J. Phys. Chem. B* **2012**, *116*, 4760−4768.
(59) Beste, A.; Buchanan, A. C. Computational Study of Bond Dissociation Enthalpies for Lignin Model Compounds. Substituent Effects in Phenethyl Phenyl Ethers. *J. Org. Chem.* **2009**, *74*, 2837−2841.
(60) Younker, J. M.; Beste, A.; Buchanan, A. C. Computational Study of Bond Dissociation Enthalpies for Substituted β-O-4 Lignin Model Compounds. *ChemPhysChem* **2011**, *12*, 3556−3565.
(61) Kim, S.; Chmely, S. C.; Nimlos, M. R.; Bomble, Y. J.; Foust, T. D.; Paton, R. S.; Beckham, G. T. Computational Study of Bond Dissociation Enthalpies for a Large Range of Native and Modified Lignins. *J. Phys. Chem. Lett.* **2011**, *2*, 2846−2852.
(62) Durbeej, B.; Eriksson, L. A. Formation of β-O-4 Lignin Models -A Theoretical Study. *Holzforschung* **2005**, *57*, 466−478.
(63) Shigematsu, M.; Kobayashi, T.; Taguchi, H.; Tanahashi, M. Transition State Leading to β-O′ Quinonemethide Intermediate of p-Coumaryl Alcohol Analyzed by Semi-Empirical Molecular Orbital Calculation. *J. Wood Sci.* **2006**, *52*, 128−133.
(64) Sangha, A. K.; Davison, B. H.; Standaert, R. F.; Davis, M. F.; Smith, J. C.; Parks, J. M. Chemical Factors That Control Lignin Polymerization. *J. Phys. Chem. B* **2014**, *118*, 164−170.
(65) Gani, T.; Orella, M.; Anderson, E.; Stone, M.; Brushett, F.; Beckham, G.; Román-Leshkov, Y. Computational Evidence for Kinetically Controlled Radical Coupling during Lignification. *ACS Sustainable Chem. Eng.* **2019**, *7*, 13270−13277.
(66) Orella, M.; Vermaas, J. V.; Gani, T. *Lignin-KMC*; MIT, 2019.
(67) Dimmel, D. Overview. *Lignins and Lignans: Advances in Chemistry*; CRC Press: Boca Raton, Florida, 2016; pp 1−10.
(68) Zhao, Y.; Truhlar, D. G. Density Functionals with Broad Applicability in Chemistry. *Acc. Chem. Res.* **2008**, *41*, 157−167.
(69) Vermaas, J. V.; Dixon, R. A.; Chen, F.; Mansfield, S. D.; Boerjan, W.; Ralph, J.; Crowley, M. F.; Beckham, G. T. Passive Membrane Transport of Lignin-Related Compounds. *Proc. Natl. Acad. Sci. U.S.A.* [Online early access] DOI: 10.1073/pnas.1904643116. Published Online: Oct 28, **2019**. https://www.pnas.org/content/early/2019/10/25/1904643116.
(70) Gillespie, D. T. A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. *J. Comput. Phys.* **1976**, *22*, 403−434.
(71) Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem.* **1977**, *81*, 2340−2361.
(72) Luo, Z.-H.; Shi, D.-P.; Zhu, Y. Multiple Active Site Monte Carlo Model for Heterogeneous Ziegler-Natta Propylene Polymerization. *J. Appl. Polym. Sci.* **2010**, *115*, 2962−2970.
(73) Van Steenberge, P. H. M.; D'hooge, D. R.; Reyniers, M.-F.; Marin, G. B. Improved Kinetic Monte Carlo Simulation of Chemical Composition-Chain Length Distributions in Polymerization Processes. *Chem. Eng. Sci.* **2014**, *110*, 185−199.
(74) Mastan, E.; Li, X.; Zhu, S. Modeling and Theoretical Development in Controlled Radical Polymerization. *Prog. Polym. Sci.* **2015**, *45*, 71−101.
(75) Hagberg, A. A.; Schult, D. A.; Swart, P. J. *Exploring Network Structure, Dynamics, and Function Using NetworkX*, 2008; Vol. *5*.
(76) Ferrini, P.; Rinaldi, R. Catalytic Biorefining of Plant Biomass to Non-Pyrolytic Lignin Bio-Oil and Carbohydrates through Hydrogen Transfer Reactions. *Angew. Chem.* **2014**, *126*, 8778−8783.
(77) Phongpreecha, T.; Hool, N. C.; Stoklosa, R. J.; Klett, A. S.; Foster, C. E.; Bhalla, A.; Holmes, D.; Thies, M. C.; Hodge, D. B. Predicting Lignin Depolymerization Yields from Quantifiable Properties Using Fractionated Biorefinery Lignins. *Green Chem.* **2017**, *19*, 5131−5143.
(78) Shuai, L.; Amiri, M. T.; Questell-Santiago, Y. M.; Héroguel, F.; Li, Y.; Kim, H.; Meilan, R.; Chapple, C.; Ralph, J.; Luterbacher, J. S. Formaldehyde Stabilization Facilitates Lignin Monomer Production during Biomass Depolymerization. *Science* **2016**, *354*, 329−333.
(79) Parsell, T.; Yohe, S.; Degenstein, J.; Jarrell, T.; Klein, I.; Gencer, E.; Hewetson, B.; Hurt, M.; Kim, J. I.; Choudhari, H.; et al. A Synergistic Biorefinery Based on Catalytic Conversion of Lignin Prior to Cellulose Starting from Lignocellulosic Biomass. *Green Chem.* **2015**, *17*, 1492−1499.
(80) Luo, H.; Abu-Omar, M. M. Lignin Extraction and Catalytic Upgrading from Genetically Modified Poplar. *Green Chem.* **2018**, *20*, 745−753.
(81) Anderson, E. M.; Stone, M. L.; Katahira, R.; Reed, M.; Muchero, W.; Ramirez, K. J.; Beckham, G. T.; Román-Leshkov, Y. Differences in S/G Ratio in Natural Poplar Variants Do Not Predict Catalytic Depolymerization Monomer Yields. *Nat. Commun.* **2019**, *10*, 2033.
(82) Dima, O.; Morreel, K.; Vanholme, B.; Kim, H.; Ralph, J.; Boerjan, W. Small Glycosylated Lignin Oligomers Are Stored in Arabidopsis Leaf Vacuoles. *Plant Cell* **2015**, *27*, 695−710.
(83) Anahashi, M. T.; Higuchi, T. Dehydrogenative Polymerization of Monolignols by Peroxidase and H202 in a Dialysis Tube. I. Preparation of Highly Polymerized DHPs. *Wood Res. J.* **1981**, *67*, 15.
(84) Terashima, N.; Atalla, R. H.; Ralph, S. A.; Landucci, L. L.; Lapierre, C.; Monties, B. New Preparations of Lignin Polymer Models under Conditions That Approximate Cell Wall Lignification. I. Synthesis of Novel Lignin Polymer Models and Their Structural Characterization by ^{13}C NMR. *Holzforschung* **1995**, *49*, 521−527.
(85) Galkin, M. V.; Samec, J. S. M. Lignin Valorization through Catalytic Lignocellulose Fractionation: A Fundamental Platform for the Future Biorefinery. *ChemSusChem* **2016**, *9*, 1544−1558.
(86) Saake, B.; Argyropoulos, D. S.; Beinhoff, O.; Faix, O. A Comparison of Lignin Polymer Models (DHPs) and Lignins by 31P NMR Spectroscopy. *Phytochemistry* **1996**, *43*, 499−507.

DOI: 10.1021/acssuschemeng.9b03534
ACS Sustainable Chem. Eng. 2019, 7, 18313−18322

(87) Nimz, H. H.; Lüdemann, H.-D. Kohlenstoff-13-NMR-Spektren von Ligninen, 6. Lignin- Und DHP-Acetate. *Holzforschung* **1976**, *30*, 33−40.
(88) Nimz, H.; Mogharab, I.; Lüdemann, H.-D. 13C-kernresonanz-spektren von ligninen, 3. Vergleich von fichtenlignin mit künstlichem lignin nach freudenberg. *Makromol. Chem.* **1974**, *175*, 2563−2575.
(89) Chen, F.; Tobimatsu, Y.; Havkin-Frenkel, D.; Dixon, R. A.; Ralph, J. A Polymer of Caffeyl Alcohol in Plant Seeds. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 1772−1777.
(90) Berstis, L.; Elder, T.; Crowley, M.; Beckham, G. T. Radical Nature of C-Lignin. *ACS Sustainable Chem. Eng.* **2016**, *4*, 5327−5335.
(91) Carlos del Río, J.; Rencoret, J.; Gutiérrez, A.; Kim, H.; Ralph, J. Hydroxystilbenes Are Monomers in Palm Fruit Endocarp Lignins. *Plant Physiol.* **2017**, *174*, 2072−2082.
(92) Stone, M. L.; Anderson, E. M.; Meek, K. M.; Reed, M.; Katahira, R.; Chen, F.; Dixon, R. A.; Beckham, G. T.; Román-Leshkov, Y. Reductive Catalytic Fractionation of C-Lignin. *ACS Sustainable Chem. Eng.* **2018**, *6*, 11211−11218.
(93) Li, Y.; Shuai, L.; Kim, H.; Motagamwala, A. H.; Mobley, J. K.; Yue, F.; Tobimatsu, Y.; Havkin-Frenkel, D.; Chen, F.; Dixon, R. A.; et al. An "Ideal Lignin" Facilitates Full Biomass Utilization. *Sci. Adv.* **2018**, *4*, No. eaau2968.
(94) Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazlewood, V.; Lathrop, S.; Lifka, D.; Peterson, G. D.; et al. XSEDE: Accelerating Scientific Discovery. *Comput. Sci. Eng.* **2014**, *16*, 62−74.

DOI: 10.1021/acssuschemeng.9b03534
ACS Sustainable Chem. Eng. 2019, 7, 18313−18322