Molecular Simulation as a Tool for Studying Lignin

Amandeep K. Sangha,[a,b] **Loukas Petridis,**[a,b] **Jeremy C. Smith,**[a,b,c] **Angela Ziebell,**[d,e] **and Jerry M. Parks**[a,b]

[a] UT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6309; parksjm@ornl.gov (for correspondence)

[b] BioEnergy Science Center, Oak Ridge, TN

[c] Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, M407 Walters Life Sciences, 1414 Cumberland Avenue, Knoxville, TN 37996

[d] National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO 80401-3393; angela.ziebell@nrel.gov (for correspondence)

[e] BioEnergy Science Center, Golden, CO

Lignocellulosic biomass provides a sustainable source of sugars for biofuel and biomaterial production. However, biomass resistance to degradation imposes difficulties for economical conversion of plant carbohydrates to fermentable sugars. One of the key contributors to recalcitrance is lignin. Understanding the properties of lignin macromolecules in the cell wall matrix is useful for manipulating biomass structure to generate more easily degradable biomass. Along with experimental techniques such as 2D-NMR and mass spectrometry, computational techniques can be useful for characterizing the structural and energetic properties of the biomass assembly and its individual constituents. Here, we provide a brief introduction to lignin, review some of the recent, relevant scientific literature, and give our perspectives on the role of molecular simulation in understanding lignin structure. © 2011 American Institute of Chemical Engineers Environ Prog, 31: 47–54, 2012*

Keywords: biomass, recalcitrance, quantum chemistry, molecular dynamics

INTRODUCTION

Lignin, a major component of the plant cell wall matrix, has a remarkably heterogeneous structure. Second only to cellulose in its natural abundance, lignin penetrates into the space between cellulose, hemicellulose, and pectin, and provides rigidity and mechanical strength to the cell wall [1]. It also participates in efficient water conduction through the plant stem and provides defense against microbial, enzymatic, and chemical degradation of the wall.

Biofuel production from lignocellulosic biomass is one of several paths toward alternative transportation fuels. Lignocellulosic biomass is rich in carbohydrate content (cellulose and hemicellulose), which can be converted into fermentable sugars. Thus, it provides a renewable alternative to petroleum-based fuels. Biomass is resistant to deconstruction by enzymatic and chemical hydrolysis [2], and lignin is a major contributor [3]. Conversion of biomass to biofuels and bioproducts involves pretreatment, resulting in the depolymerization of lignin and removal of hemicellulose, followed by enzymatic hydrolysis and fermentation in subsequent steps. During pretreatment, hemicellulose is solubilized and converted into mono- and oligosaccharides, which can then undergo fermentation. Removal of hemicellulose renders cellulose accessible to hydrolytic enzymes for breakdown into sugars. Lignin, however, is removed only to a limited extent, and is redistributed over the surface of cellulose, making the conversion process less efficient by hindering access to cellulose by hydrolytic cellulases [4–6]. Furthermore, numerous studies have demonstrated the non-specific binding of cellulolytic enzymes to lignin [7–9], which creates a further obstacle to efficient cellulose hydrolysis. A recent study found that ~50% of added enzyme binds unproductively to lignin [9].

Determining the structural characteristics of the biomass assembly is essential for understanding the resistance of biomass toward degradation. Molecular simulation using quantum chemistry or classical molecular dynamics (MD) simulation can provide accurate structures, energies, and molecular properties that may or may not be accessible using conventional experimental techniques. However, relatively few molecular simulation studies have been applied to lignin and lignin derivatives. Some of the topics studied to date include electron delocalization in monolignol radicals [10, 11], the energetics of alkyl aryl ether (β-O-4) linkage formation of coniferyl alcohol (CA)-based dehydrodilignols [12], interactions of lignin with cellulose [13, 14] and absorptive resins [15], template polymerization of lignin guided by noncovalent interactions [16], conformations and intramolecular hydrogen bonding of monolignols [17], and the effects of pretreatment on biomass [18]. Computational studies of the pyrolysis of model lignin molecules [19–23], although related to lignin degradation rather than biosynthesis, have also provided important insight into chemical bonding in lignin. In this article, we review selected applications of quantum chemical calculations and MD to understanding lignin synthesis and structure, and provide our perspectives on areas that may benefit from molecular simulation in the near future.

LIGNIN STRUCTURE AND COMPOSITION

Lignin is a complex, heterogenous, optically inactive macromolecule composed mainly of three hydroxycinnamyl alcohols: *p*-coumaryl, coniferyl, and sinapyl alcohol (Figure 1). Other constituents are also observed in lesser amounts,

especially in grassy and herbaceous species, but we limit our discussion to the three most prevalent monolignols. These monomers, when incorporated into a lignin polymer, are referred to as p-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units, respectively. All three monomers share the same p-coumaryl alcohol core, but vary in their degree of methoxylation (Figure 1) [24].

The proportion of each subunit in lignin depends on the species, age, cell type, and tissue type [1, 25]. The presence of each additional methoxy group results in one fewer reactive site (S < G < H), and therefore fewer possible linkage
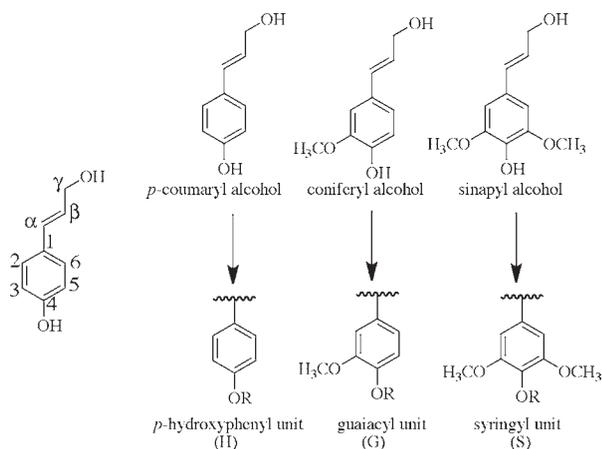

Figure 1. The three major monolignols, p-coumaryl, coniferyl, and sinapyl alcohol, differ in their extent of methoxylation. When incorporated into lignin, these species are referred to as p-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units.

combinations during polymerization. In addition, the bonds that form as a result of reaction at C3 or C5 tend to be C—C bonds, which are more difficult to break down than the alternative C—O bonds that dominate when methoxy content is high (i.e., when the S/G ratio is high). H lignin subunits are present in low levels in natural materials (4.9% of lignin in alfalfa [3] and 0.8% in Norway spruce [26]). Consequently, less is known about lignin high in H subunits and its impact on recalcitrance to processing.

Several different C—C and C—O bond linkages are possible between monomer units, with the most abundant being the β-O-4 alkyl aryl ether linkage. In addition, β-β, β-5, β-1, and 5-5 bonds (Figure 2) occur in natural lignin. The β-O-4 and β-β linkages generally lead to linear chains of lignin polymer, whereas others such as the α-O-4, 5-O-4 and 8-membered ring dibenzodioxocin linkages [27, 28] participate in branching of lignin chains.

Other less commonly occurring monomeric units are also found in natural lignin. Hydroxycinnamaldehydes can substitute for monolignols, especially in plants deficient in the enzyme cinnamyl alcohol dehydrogenase (CAD), and polymerize with the growing chain through a β-O-4 linkage. Dihydroconiferyl alcohol (DHCA), a minor component in gymnosperms, can undergo coupling only through O4 and C5, thereby always being a chain-terminating monomer unit. 5-Hydroxyconiferyl alcohol, prevalent in caffeic acid O-methyltransferase (COMT) down-regulated plants, can undergo linkage between Cβ and O4 of sinapyl or CA and vice versa to form dibenzodioxane (β-O-4, α-O-4) structures (Figure 2). Ferulates and diferulates are also present in grassy species [29].

Given the significant challenge posed by the presence of lignin during pretreatment, lignin down-regulation using modern molecular biology techniques has become a popular target for biofuels feedstock improvement. Down-regulation of all steps in the lignin biosynthetic pathway has been
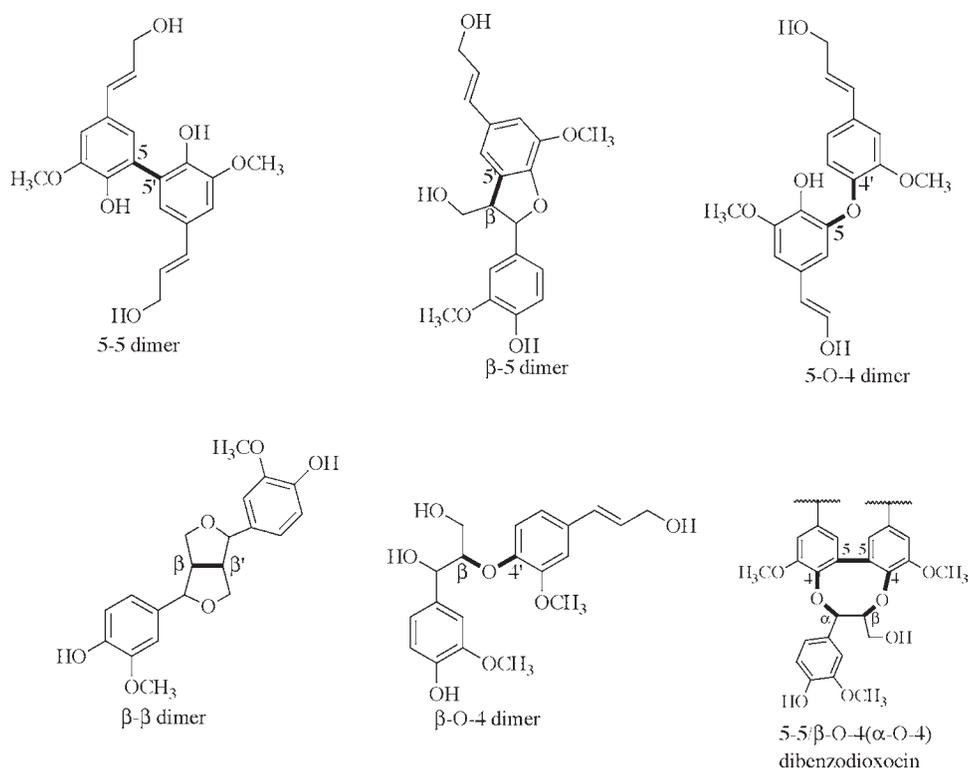
Figure 2. Various types of representative linkages in lignin.

achieved as summarized in a recent review on the genetics of lignin biosynthesis [30]. In addition to targeting reduction of total lignin content, lignin genetic modification can be used to increase the S/G ratio to take advantage of the lower recalcitrance of lignin with higher S content. A number of notable examples have produced lines that yield an increase in the amount of sugar released during pretreatment and saccharification [3, 31]. Although some lines suffer from severe growth problems, this is not always the case and more subtle down-regulation events can create normal phenotypes [32]. Some lines even boast improvements in pest resistance and tolerance to drought [33]. As the monomer ratio and total lignin concentration vary more radically with lignin genetic modification, transgenic lines will provide interesting data for input into lignin simulation studies.

LIGNIN BIOSYNTHESIS

Lignin has been studied for several decades, but a unified mechanism for its biosynthesis is still lacking. Two major hypotheses for lignin biosynthesis are frequently debated in the literature. Combinatorial random coupling [34, 35] of monolignol radicals is the most widely accepted mechanism for lignin polymerization, but protein-directed synthesis involving dirigent proteins has also been hypothesized [36, 37]. Before polymerization can begin, monolignols must be transferred to the lignifying zone of the cell wall where they are oxidized by peroxidases or laccases [38–40]. The transport mechanism, however, is unclear [41], although adenosine triphosphate (ATP)-dependent transport may be involved [42]. The first step in lignin polymerization is the oxidation of lignin monomers at the O4 position to form phenoxy radicals. The ease of oxidation of a given monolignol depends on its extent of methoxylation [43]. Sinapyl alcohol undergoes oxidation more readily than coniferyl or *p*-coumaryl alcohols because of the two methoxy groups at the *meta* positions relative to the allyl alcohol group. Enzymatic oxidation of monolignols and monolignol analogs by horseradish peroxidase (HRP), a model peroxidase enzyme, was found to depend on three factors: steric hindrance originating from the size of the substrate, steric hindrance from protein residues near the redox-active heme, and hydrophobicity of the substrate [40].

In the endwise polymerization mechanism [34, 44] of lignin chain growth, a growing lignin chain is formed through cross-coupling between a monolignol radical and a terminal phenoxy radical on a growing lignin chain. According to this hypothesis, the end unit of a growing chain must first undergo oxidation. Méchin *et al*. have shown that a low ratio of HRP to CA (HRP/CA) results in the occurrence of trimers, tetramers, and pentamers containing at least one β-O-4 bond [45]. Long-range electron transfer has also been suggested to take place to oxidize long lignin chains by enzymes that are not required to make physical contact with lignin substrates [46].

Another question that remains open is whether the monomers are transferred individually to the lignifying zone, where they subsequently undergo oxidation and coupling, or if all the monomers are transferred at the same time and they all undergo oxidation concurrently just before coupling. In the latter case, it is expected that the radicals would couple to form dimers instead of participating in endwise polymerization.

MOLECULAR SIMULATION: UNRAVELING THE STRUCTURE OF LIGNIN

Understanding lignin structure and composition is important for deciphering biosynthetic pathways and developing plants with more easily degradable cell walls. Experimental techniques for deconstructing biomass and analyzing its components can be augmented by complementary molecular simulation approaches. The choice of an appropriate computational method always depends on the specific molecular system being studied and the questions being asked. For condensed-phase phenomena that occur over longer time scales and do not involve bond breaking or forming, classical MD simulations can be performed on mesoscopic systems using standard software. To study chemical reactions such as oxidative coupling, however, quantum chemical, or quantum mechanical (QM), calculations are required, and in this case the balance between the desired level of accuracy and the computational cost of the calculations is a particularly important consideration. A comprehensive introduction to computational chemistry, covering both classical and quantum chemical approaches, is available in a recent book [47].

Quantum Chemical Calculations

One of the main challenges in quantum chemistry involves choosing a calculation method that provides accurate results in a reasonable amount of time. It is possible to calculate molecular energies that agree to within 1 kcal/mol or less of accurate experimental data, but these methods often require substantial computational resources for molecules containing more than a few atoms. Furthermore, such high accuracy may not be necessary to answer specific questions.

Ab initio quantum chemical methods are, in general, the most accurate, but also the most computationally demanding approaches, requiring substantial parallel computing resources for all but the smallest molecular systems. Density functional theory (DFT) [48, 49] is an extremely successful quantum chemical approach that has been applied in numerous studies of chemical structures, properties, and reactions [50]. Compared with more computationally intensive, correlated *ab initio* methods, much larger molecular systems can be studied using DFT at a lower computational cost on more modest computer systems. It is difficult to make general statements concerning the accuracy of DFT methods, partly because there are so many of them, but careful assessments have shown that the best DFT methods can rival the accuracy of much more computationally demanding *ab initio* methods for many molecular properties [50, 51].

In addition to *ab initio* and DFT approaches, semi-empirical quantum chemical methods provide an extremely efficient, albeit more approximate, approach for calculating molecular structures and their properties. Semi-empirical methods, which are on the order of 3–4 orders of magnitude faster than typical DFT methods, achieve their efficiency by introducing two additional approximations. First, many of the electron–electron interaction terms are simply neglected, and second, adjustable parameters are introduced to account, at least partially, for electron correlation and other effects. Although these methods, which have acronyms such as AM1 [52] and PM3 [53], were originally developed more than two decades ago, they are still actively developed and used because of their computational efficiency. Recently developed semi-empirical methods include PDDG/PM3 [54], RM1 [55], and PM6 [56].

Lignin-like molecules, in particular higher-order multimers of monolignol subunits, have numerous rotatable C—O and C—C bonds. It is therefore quite important to use conformational search techniques to locate the lowest-energy conformations when studying lignin structure and energetics [17, 57]. One useful strategy involves using a computationally affordable semi-empirical or molecular mechanics method (see below) coupled with a systematic or random conformer generation algorithm to find low-energy conformers. A more accurate quantum chemical method, such as DFT, is then used to refine the structures and energies of a limited number of these low-energy conformers, with the goal of obtaining the most important conformer(s).

Historically, the B3LYP [58, 59] DFT method has been the most widely used across all areas of computational chemistry, but a growing number of next-generation DFT methods with improved accuracy are now available [51]. Two examples are the Minnesota density functionals developed by Truhlar and coworkers [60–62] and the ωB97X-D density functional developed by Head-Gordon and coworker [63]. Although numerous other DFT methods exist and are commonly used, these two are singled out here because they yield accurate molecular structures and energies, they provide an adequate representation of nonbonded interactions such as hydrogen bonding and dispersion, which are likely to be important in molecular systems containing phenolic moieties such as those found in lignin, and they have already been applied successfully in computational studies of lignin-like systems [16, 17, 21, 22].

Quantum chemical calculations have been used to study the energies and electronic properties of lignin model compounds. Two key hypotheses related to lignin biosynthesis, the thermodynamic control and spin distribution hypotheses, have been investigated using quantum chemical calculations. According to the thermodynamic control hypothesis, the ratios of various linkages in lignin should correlate with the relative energies or thermodynamic stabilities of compounds bearing those linkages. In the spin distribution hypothesis, the sites bearing the most unpaired electron spin are expected to be the most reactive.

DFT calculations have been used to investigate the structure, electronic properties, and reactivity of monolignols, dehydrodilignols, and other lignin-like compounds. Durbeej et al. [11, 12, 64] used B3LYP calculations with a polarizable continuum representation of the solvent to study the energetics and electronic structure of oxidative coupling of coniferyl radicals. Their major findings, along with other lignin simulation studies, are summarized in a thorough and informative review on the challenges of using computation to study lignin biosynthesis and degradation [65]. Briefly, Durbeej and Eriksson calculated electron spin densities of monolignol radicals and dehydrogenated dilignol radicals. They found that the majority of unpaired spin is localized on the phenoxy radical (O4) in both mono- and dilignol radicals [11], which may help to explain why β-O-4 linkages are so prevalent in natural lignin. These authors also compared the relative energies of the β-O-4 linked dimer with β-β, β-5, 5-5, and 5-O-4 lignans [64]. A lack of correlation was observed between the calculated relative energies of the dimers and the experimentally observed ratios of intermonomeric unit linkages. The β-β linkage was found to be the most favorable (i.e., lowest energy) in the calculations, followed by β-5 (2.6 kcal/mol higher) and 5-5 (6.6 kcal/mol higher), whereas the most abundant linkage in natural lignin, β-O-4, was found to be significantly higher in energy (11.6 kcal/mol) than the β-β dilignol.

Although the lack of agreement between computation and experiment may seem discouraging, it is important to remember that many factors are likely to be involved in determining the linkages observed in vitro and in vivo. Nevertheless, these results provide an important baseline study of the intrinsic reactivities of monolignol radicals in oxidative coupling reactions in vitro. They also provide the necessary framework for future simulations using more accurate quantum chemical methods and more sophisticated multiscale simulation techniques.

Using DFT calculations, Beste et al. [19–22] have studied pyrolysis or thermal decomposition reactions of lignin-like model compounds. They investigated reaction pathways for hydrogen abstraction from phenethyl phenyl ether (PPE) by benzyl and phenoxy radicals [22]. Hydrogen can be abstracted from the α or β position during pyrolysis, and the selectivity of α/β hydrogen abstraction as well as the rates of these reactions vary with the substituents on PPE [19, 20]. The bond dissociation enthalpies of C—O and C—C bonds have also been studied [21] to shed light on the mechanism of breakdown of lignin into smaller chains during pyrolysis.

Recently, the CBS-QB3 method was applied to pyrolysis reactions involving PPE [23]. Complete basis set (CBS) methods [66] are complex, ab initio-based, composite energy calculations that provide highly accurate structures, energies, and thermo-chemical quantities. Computed energies were used along with transition state theory (TST) to obtain rate constants for various reaction pathways, with the computational results in near-quantitative agreement (within 1–2 kcal/mol) with experimental measurements.

Chen and Sarkanen [16] used DFT calculations to investigate how a double-stranded lignin chain could be replicated to form a complementary daughter chain. In their study, noncovalent interactions were found to be sufficiently strong to direct the orientation of monolignol radicals and a double-stranded lignin template for oxidative coupling.

Very recently, single-conformation spectroscopy and DFT calculations were used to probe monolignol conformations in the gas phase. Rodrigo et al. showed that the allyl alcohol side chains of monolignols are "locked" into rigid configurations in which the hydroxyl group is oriented toward the vinyl π cloud [17]. For coniferyl and sinapyl alcohol, the phenolic hydroxyl group was found to form an intramolecular hydrogen bond with an adjacent methoxy group. The agreement between theory and experiment further validated the use of DFT calculations in characterizing structural properties of lignin-like species.

Molecular Dynamics Simulation

MD simulation involves stepwise integration of Newton's second law ($F = ma$) for a molecular system. MD is a powerful tool to complement experiments in obtaining atomic-detail understanding of the structure and dynamics of biomolecules. It is now common to simulate systems of 10^5 or more atoms for hundreds of nanoseconds. An empirical potential function, also called a molecular mechanics force field, is employed to describe the energy of the system as a function of interatomic distances. Electronic degrees of freedom are not considered as in computationally more intensive quantum chemical calculations.

One of the most commonly used empirical potential functions for MD simulations is the CHARMM (CHemistry At Harvard Molecular Mechanics) force field [67]. The form of the CHARMM potential energy function is:

$$E = \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angles}} K_\Theta(\Theta - \Theta_0)^2 + \sum_{U-B} K_{\text{ub}}(s - s_0)^2$$
$$+ \sum_{\text{dihedrals}} K_\varphi[1 + \cos(n\varphi - \delta)] + \sum_{\text{impropers}} K_\psi(\psi - \psi_0)^2$$
$$+ \sum_{\text{nonbonded}} \left\{ \varepsilon_{ij} \left[\left(\frac{R_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{ij}}{r_{ij}}\right)^6 \right] + \frac{q_i q_j}{4\varepsilon_0 \pi r_{ij}} \right\}, \quad (1)$$

in which contributions to the energy include bonded (bond, angle, Urey–Bradley (U-B), dihedral, and improper dihedral) and non-bonded (Lennard-Jones and Coulombic) terms. Various parameters such as the equilibrium bond distances, angles, dihedrals, their corresponding force constants K, Lennard-Jones (i.e., van der Waals) parameters, and partial atomic charges q_i are molecule-dependent and must be optimized for specific types of systems prior to performing the simulation.

MD simulation of a β-O-4 linked guaiacyl dimer in aqueous solution showed that intramolecular hydrogen bonds do not persist significantly and are preferentially exchanged with intermolecular hydrogen bonds to the surrounding solvent [68]. Findings were confirmed by nuclear magnetic resonance (NMR) spectroscopy experiments, which revealed significant

conformational flexibility and weak, transient intermolecular hydrogen bonds [69].

Subsequent calculations by the same group probed the interactions between lignin and cellulose. Estimated enthalpies of absorption of β-O-4 linked guaiacyl dimer on hydrophilic and hydrophobic cellulose surfaces were found to be similar [13]. However, the structure of cellulose was found to influence the preferred geometries of the lignin phenol rings. For the hydrophilic surfaces, where electrostatic interactions dominate, only one of the two aromatic rings was adsorbed. In contrast, both aromatic rings were adsorbed on the hydrophobic surface, and van der Waals interactions accounted for half the adsorption energy.

In vacuo simulations of two β-O-4 guaiacyl oligomers with degree of polymerization 10 and 20 in the presence of a 36-chain cellulose whisker also found lignin to adsorb in a flat orientation on the surface of the cellulose [14]. A significant increase of the van der Waals contribution was found, now dominating the adsorption interaction for hydrophobic and hydrophilic surfaces. However, the studies in Refs. 13, 14, 68, 69 involve relatively short trajectories (<0.5 ns) and may therefore suffer from inadequate sampling.

Molecular mechanics parameters optimized specifically for lignin were recently developed [70] following the standard procedure used to develop the CHARMM force field for proteins and nucleic acids [67]. Two molecules, methoxybenzene (anisole) and *p*-hydroxyphenol, were used as simple models of lignin. Here, we provide only a brief summary of the strategy used to optimize the new parameters. Equilibrium values for bonds, angles, and dihedrals were taken from quantum chemical MP2/6-31G(d) optimized geometries. Lennard-Jones parameters were assigned based on analogy to existing atom types in the CHARMM force field. Atomic partial charges were optimized with respect to quantum chemical interaction energies using a supramolecular approach with a model compound (anisole) interacting with a single water molecule. Charges were adjusted to reproduce equilibrium intermolecular distances and interaction energies between anisole and a TIP3P water molecule [71]. After determining the non-bonded terms, dihedral parameters were derived from quantum chemical potential energy surfaces by fitting to torsional profiles. The remaining bond and angle parameters were then optimized to reproduce vibrational frequencies and eigenvector projections derived from QM calculations.

Extensive MD simulation was recently combined with small-angle neutron scattering (SANS) to characterize the structure of softwood lignin aggregates in aqueous solution [18]. The surface of the aggregates was found to be highly folded and characterized by surface fractal properties that are invariant under a change of scale from ∼1−1000 Å, demonstrating excellent agreement between SANS and MD. Furthermore, MD revealed that deeply buried water molecules preferentially hydrate oxygen atoms of the lignin, and the aggregates were found to be dynamically heterogeneous with enhanced monomer mobility at the surface, a feature attributed to local depression of the glass transition temperature. The contribution of lignin to biomass recalcitrance is closely associated with temperature-dependent changes in the polymer structure upon heating and cooling. MD simulation can be employed to examine the temperature dependence of the structure and dynamics of individual lignin polymers. Multi-million-atom systems of cellulose and lignin in aqueous solution were recently used as benchmarks to test the parallel efficiency of MD simulations on massively parallel supercomputers [72]. The authors were able to perform parallel simulations on up to ∼30,000 processors.

MD simulation also offers valuable insight into the behavior of water near biomolecules, often called hydration water. The enormous number and very fast dynamics of water molecules pose technical challenges in studies of hydration water of biomolecules in aqueous solution. On the other hand, studying the structure and dynamics of hydration water molecules is tractable using MD simulation, and has already been carried out for cellulose [73] and lignin [18].

FUTURE DIRECTIONS

MD simulation is currently being used in our group to simulate the re-precipitation of lignin onto crystalline and amorphous cellulose after pretreatment. The aim of that study is to obtain atomic-level understanding of dilute-acid and flow-through pretreatment, a process in which lignin is depolymerized but then re-precipitates back onto the cellulose. Although these simulations are intended to study the effects of a specific pretreatment process, other conditions could, in principle, also be simulated. High-temperature MD simulations could also be used in combination with experimental data to characterize the temperature-dependent dynamics of cell wall polymers such as lignin, cellulose, and hemicellulose. Combining molecular simulation techniques with more conventional, experimental studies is a powerful approach for answering many of the open questions related to lignin structure, composition, and biosynthesis at the molecular level.

Coarse-grained MD simulation enables large systems to be studied on significantly larger time scales than those typically achieved in standard all-atom simulations [74, 75]. In this approach, groups of several atoms are modeled as single particles to reduce the complexity of the system. Coarse grain simulations have recently been performed in our lab for cellulose [76], and a similar approach could be used to study the structure and dynamic properties of lignin on microsecond time scales.

NMR chemical shifts of lignin model compounds have been calculated using quantum chemical methods [77], and this approach should be reinvestigated further as an interpretive tool to aid in spectral assignments. Combined experimental and computational NMR studies of various aromatic compounds similar to lignin have been reviewed elsewhere [57].

Semi-empirical quantum chemical methods will likely continue to play a significant role in computational studies of lignin structure and reactivity, particularly as new methods with improved accuracy are developed.

Another promising approach that should be applied in lignin studies is the combined quantum mechanical/molecular mechanical (QM/MM) method [78, 79]. In QM/MM calculations, the majority of a molecular system is described using a molecular mechanics force field, but a relatively small portion of the system (∼50−200 atoms) is described using quantum mechanics. The advantage of the QM/MM approach is that it enables chemical reactions to be simulated within a realistic molecular environment. This approach has been used extensively to study numerous enzyme-catalyzed reactions [80], and is particularly well suited to studying enzymes involved in lignin synthesis such as laccases and peroxidases, or non-enzymatic, solution-phase reactions involving monolignols, lignans, and the like.

Successful application of QM/MM methods is far from trivial [79]. For example, in addition to choosing an accurate and efficient quantum chemical approach, special attention must be paid to the interface between the QM and MM subsystems, particularly when the linkage occurs across a covalent bond. Also, atoms included in the QM subsystem must be carefully chosen. In general, only groups that participate directly in the reaction process are described using quantum mechanics. A proper description of electrostatics and dynamic effects is necessary to obtain accurate free energies, which can then be compared directly with experimental data.

CONCLUDING REMARKS

The presence of lignin presents a major challenge for obtaining biofuels and bioproducts from biomass. Part of that challenge involves understanding fundamental aspects of lignin structure and synthesis so that rational approaches to its efficient depolymerization can be developed. Molecular simulation can provide atomistic detail and accurate energetics of reactions and processes involving lignin, the exact details of which are experimentally unobservable at this time. Quantum chemical calculations and MD will continue to provide fundamental insight that will hopefully result in improved technologies for biomass conversion.

LITERATURE CITED

1. Whetten, R.W., MacKay, J.J., & Sederoff, R.R. (1998). Recent advances in understanding lignin biosynthesis, Annual Review of Plant Physiology and Plant Molecular Biology, 49, 585–609.
2. Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W., & Foust, T.D. (2007). Biomass recalcitrance: Engineering plants and enzymes for biofuels production, Science, 315, 804.
3. Chen, F., & Dixon, R.A. (2007). Lignin modification improves fermentable sugar yields for biofuel production, Nature Biotechnology, 25, 759–761.
4. Jørgensen, H., Kristensen, J.B., & Felby, C. (2007). Enzymatic conversion of lignocellulose into fermentable sugars: Challenges and opportunities, Biofuels Bioproducts and Biorefining, 1, 119–134.
5. Selig, M.J., Viamajala, S., Decker, S.R., Tucker, M.P., Himmel, M.E., & Vinzant, T.B. (2007). Deposition of lignin droplets produced during dilute acid pretreatment of maize stems retards enzymatic hydrolysis of cellulose, Biotechnology Progress, 23, 1333–1339.
6. Donohoe, B.S., Decker, S.R., Tucker, M.P., Himmel, M.E., & Vinzant, T.B. (2008). Visualizing lignin coalescence and migration through maize cell walls following thermochemical pretreatment, Biotechnology and Bioengineering, 101, 913–925.
7. Eriksson, T., Borjesson, J., & Tjerneld, F. (2002). Mechanism of surfactant effect in enzymatic hydrolysis of lignocellulose, Enzyme and Microbial Technology, 31, 353–364.
8. Palonen, H., Tjerneld, F., Zacchi, G., & Tenkanen, M. (2004). Adsorption of Trichoderma reesei CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin, Journal of Biotechnology, 107, 65–72.
9. Berlin, A., Gilkes, N., Kurabi, A., Bura, R., Tu, M., Kilburn, D., & Saddler, J. (2005). Weak lignin-binding enzymes, Applied Biochemistry and Biotechnology, 121, 163–170.
10. Russell, W.R., Forrester, A.R., Chesson, A., & Burkitt, M.J. (1996). Oxidative coupling during lignin polymerization is determined by unpaired electron delocalization within parent phenypropanoid radicals, Archives of Biochemistry and Biophysics, 332, 357–366.
11. Durbeej, B., & Eriksson, L.A. (2003). Spin distribution in dehydrogenated coniferyl alcohol and associated dilignol radicals, Holzforschung, 57, 59–61.
12. Durbeej, B., & Eriksson, L.A. (2003). A density functional theory study of coniferyl alcohol intermonomeric cross linkages in lignin: Three-dimensional structures, stabilities and the thermodynamic control hypothesis, Holzforschung, 57, 150–164.
13. Besombes, S., & Mazeau, K. (2005). The cellulose/lignin assembly assessed by molecular modeling. Part 1: Adsorption of a threo guaiacyl β-O-4 dimer onto an Iβ cellulose whisker, Plant Physiology and Biochemistry, 43, 299–308.
14. Besombes, S., & Mazeau, K. (2005). The cellulose/lignin assembly assessed by molecular modeling. Part 2: Seeking for evidence of organization of lignin molecules at the interface with cellulose, Plant Physiology and Biochemistry, 43, 277–286.
15. Mardis, K.L., Glemza, A.J., Brune, B.J., Payne, G.F., & Gilson, M.K. (1999). Differential adsorption of phenol derivatives onto a polymeric sorbent: A combined molecular modeling and experimental study, Journal of Physical Chemistry B, 103, 9879–9887.
16. Chen, Y., & Sarkanen, S. (2010). Macromolecular replication during lignin biosynthesis, Phytochemistry, 71, 453–462.
17. Rodrigo, C.P., James, W.H., III, & Zwier, T.S. (2011). Single-conformation ultraviolet and infrared spectra of jet-cooled monolignols: p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol, Journal of the American Chemical Society, 133, 2632–2641.
18. Petridis, L., Pingali, S.V., Urban, V., Heller, W.T., O'Neil, H.M., Foston, M., Ragauskas, A.J., & Smith, J.C. (2011). Self-similar multiscale structure of lignin revealed by neutron scattering and molecular dynamics simulation, Physical Review E, 83, 061911-4.
19. Beste, A., Buchanan, A.C., III, Britt, P.F., Hathorn, B.C., & Harrison, R.J. (2007). Kinetic analysis of the pyrolysis of phenethyl phenyl ether: Computational prediction of α/β-selectivities, Journal of Physical Chemistry A, 111, 12118–12126.
20. Beste, A., Buchanan, A.C., III, & Harrison, R.J. (2008). Computational prediction of α/β selectivities in the pyrolysis of oxygen-substituted phenethyl phenyl ethers, Journal of Physical Chemistry A, 112, 4982–4988.
21. Beste, A., & Buchanan, A.C., III. (2009). Computational study of bond dissociation enthalpies for lignin model compounds. Substituent effects in phenethyl phenyl ethers, Journal of Organic Chemistry, 74, 2837–2841.
22. Beste, A., & Buchanan, A.C., III. (2010). Substituent effects on the reaction rates of hydrogen abstraction in the pyrolysis of phenethyl phenyl ethers, Energy and Fuels, 24, 2857–2867.
23. Jarvis, M.W., Daily, J.W., Carstensen, H.-H., Dean, A.M., Sharma, S., Dayton, D.C., Robichaud, D.J., & Nimlos, M.R. (2011). Direct detection of products from the pyrolysis of 2-phenethyl phenyl ether, Journal of Physical Chemistry A, 115, 428–438.
24. Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P.F., Marita, J.M., Hatfield, R.D., Ralph, S.A., Holst Christensen, J., & Boerjan, W. (2004). Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids, Phytochemistry Reviews, 3, 29–60.
25. Monties, B., & Fukushima, K. (2010). Occurrence, function and biosynthesis of lignins. In M. Hofrichter and A. Steinbuchel (Eds.), Biopolymers: Lignin, humic substances and coal (pp. 107–115), Weinheim: Wiley-VCH.
26. Wadenbäck, J., Clapham, D., Gellerstedt, G., & von Arnold, S. (2004). Variation in content and composition of lignin in young wood of Norway spruce, Holzforschung, 58, 107–115.
27. Karhunen, P., Rummakko, P., Sipilä, J., Brunow, G., & Kilpeläinen, I. (1995). Dibenzodioxocins: A novel type of linkage in softwood lignins, Tetrahedron Letters, 36, 169–170.
28. Karhunen, P., Rummakko, P., Sipilä, J., Brunow, G., & Kilpeläinen, I. (1995). The formation of dibenzodioxocin structures by oxidative coupling. A model reaction for lignin biosynthesis, Tetrahedron Letters, 36, 4501–4504.

29. Lu, F.-C., & Ralph, J. (2010). Lignin. In R.-C.Sun (Ed.), Cereal straw as a resource for sustainable biomaterials and biofuels: Chemistry, extractives, lignins, hemicelluloses and cellulose (pp 169–207), Oxford: Elsevier.

30. Bonawitz, N.D., & Chapple, C. (2010). The genetics of lignin biosynthesis: Connecting genotype to phenotype, Annual Review of Genetics, 44, 337–363.

31. Fu, C., Mielenz, J.R., Xiao, X., Ge, Y., Hamilton, C.Y., Rodriguez, M., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R.A., & Wang, Z.-Y. (2011). Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass, Proceedings of National Academy of Sciences USA, 108, 3803.

32. Reddy, M.S., Chen, F., Shadle, G., Jackson, L., Aljoe, H., & Dixon, R.A. (2005). Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (Medicago sativa L.), Proceedings of National Academy of Sciences USA, 102, 16573–16578.

33. Gallego-Giraldo, L., Jikumaru, Y., Kamiya, Y., Tang, Y., & Dixon, R.A. (2011). Selective lignin downregulation leads to constitutive defense response expression in alfalfa (Medicago sativa L.), New Phytologist, 190, 627–639.

34. Sarkanen, K.V., & Ludwig, C.H. (1971). Lignins: Occurrence, formation, structure and reactions. New York: Wiley-Interscience.

35. Freudenberg K., & Neish, A.C. (1968). Constitution and biosynthesis of lignin. Berlin: Springer-Verlag.

36. Davin, L.B., Wang, H.-B., Crowell, A.L., Bedgar, D.L., Martin, D.M., Sarkanen, S., & Lewis, N.G. (1997). Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center, Science, 275, 362–366.

37. Davin, L.B., & Lewis, N.G. (2005). Lignin primary structures and dirigent sites, Current Opinion in Biotechnology, 16, 407–415.

38. Whetten, R., & Sederoff, R. (1995). Lignin biosynthesis, The Plant Cell, 7, 1001.

39. Aoyama, W., Sasaki, S., Matsumura, S., Mitsunaga, T., Hirai, H., Tsutsumi, Y., & Nishida, T. (2002). Sinapyl alcohol-specific peroxidase isoenzyme catalyzes the formation of the dehydrogenative polymer from sinapyl alcohol, Journal of Wood Science, 48, 497–504.

40. Kobayashi, T., Taguchi, H., Shigematsu, M., & Tanahashi, M. (2005). Substituent effects of 3,5-disubstituted p-coumaryl alcohols on their oxidation using horseradish peroxidase–H_2O_2 as the oxidant, Journal of Wood Science, 51, 607–614.

41. Liu, C.J., Miao, Y.C., & Zhang, K.W. (2011). Sequestration and transport of lignin monomeric precursors, Molecules, 16, 710–727.

42. Miao, Y.C., & Liu, C.J. (2010). ATP-binding cassette-like transporters are involved in the transport of lignin precursors across plasma and vacuolar membranes, Proceedings of National Academy of Sciences USA, 107, 22728–22733.

43. Russell, W.R., Burkitt, M.J., Scobbie, L., & Chesson, A. (2006). EPR investigation into the effects of substrate structure on peroxidase-catalyzed phenylpropanoid oxidation, Biomacromolecules, 7, 268–273.

44. Brunow, G., Kilpeläinen, I., Sipilä, J., Syrjaenen, K., Karhunen, P., Setaelae, H., & Rummakko, P. (1998). Oxidative coupling of phenols and the biosynthesis of lignin, ACS Symp Ser, 697, 131–147.

45. Méchin, V., Baumberger, S., Pollet, B., & Lapierre, C. (2007). Peroxidase activity can dictate the in vitro lignin dehydrogenative polymer structure, Phytochemistry, 68, 571–579.

46. Johjima, T., Itoh, N., Kabuto, M., Tokimura, F., Nakagawa, T., Wariishi, H., & Tanaka, H. (1999). Direct interaction of lignin and lignin peroxidase from Phanerochaete chrysosporium. Proceedings of National Academy of Sciences USA, 96, 1989.

47. Cramer, C.J. (2004). Essentials of computational chemistry (2nd Edition: Theories and models), West Sussex, England: Wiley.

48. Parr, R.G., & Yang, W. (1994). Density-functional theory of atoms and molecules, USA: Oxford University Press.

49. Koch, W., Holthausen, M.C., & Holthausen, M.C. (2001). A chemist's guide to density functional theory (Volume 293). Weinheim: Wiley-VCH.

50. Sousa, S.F., Fernandes, P.A., & Ramos, M.J. (2007). General performance of density functionals, Journal of Physical Chemistry A, 111, 10439–10452.

51. Goerigk, L., & Grimme, S. (2011). A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions, Physical Chemistry Chemical Physics, 13, 6670–6688.

52. Dewar, M.J.S., Zoebisch, E.G., Healy, E.F., & Stewart, J.J.P. (1985). Development and use of quantum mechanical molecular models. 76. AM1: A new general purpose quantum mechanical molecular model, Journal of the American Chemical Society, 107, 3902–3909.

53. Stewart, J.J.P. (1989). Optimization of parameters for semiempirical methods. I. Method, Journal of Computational Chemistry, 10, 209–220.

54. Repasky, M.P., Chandrasekhar, J., & Jorgensen, W.L. (2002). PDDG/PM3 and PDDG/MNDO: Improved semiempirical methods, Journal of Computational Chemistry, 23: 1601–1622.

55. Rocha, G.B., Freire, R.O., Simas, A.M., & Stewart, J.J.P. (2006) RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I, Journal of Computational Chemistry, 27, 1101–1111.

56. Stewart, J.J.P. (2007). Optimization of parameters for semiempirical methods. V. Modification of NDDO approximations and application to 70 elements, Journal of Molecular Model, 13, 1173–1213.

57. Elder, T., & Fort, R.C., Jr., Reactivity of lignin—Correlation with molecular orbital calculations. In C.Heitner, D.R.Dimmel, and J.A.Schmidt (Eds.), Lignin and lignans: Advances in chemistry (pp 321–347), Boca Raton, FL: CRC Press.

58. Becke, A.D. (1993). Density-functional thermochemistry. III. The role of exact exchange, Journal of Chemical Physics, 98, 5648–5652.

59. Lee, C., Yang, W., & Parr, R.G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, Physical Review B, 37, 785–789.

60. Zhao, Y., & Truhlar, D.G. (2008). The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals, Theoretical Chemistry Accounts, 120, 215–241.

61. Zhao, Y., & Truhlar, D.G. (2008). Density functionals with broad applicability in chemistry, Accounts of Chemical Research, 41, 157–167.

62. Zhao, Y., Schultz, N.E., & Truhlar, D.G. (2006). Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions, Journal of Chemical Theory and Computation, 2, 364–382.

63. Chai, J.D., & Head-Gordon, M. (2009). Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections, Physical Chemistry Chemical Physics, 10, 6615–6620.

64. Durbeej, B., & Eriksson, L.A. (2003). Formation of β-O-4 lignin models: A theoretical study, Holzforschung, 57, 466–478.

65. Durbeej, B., Wang, Y.N., & Eriksson, L. (2003). Lignin biosynthesis and degradation—A major challenge for computational chemistry, Lecture Notes in Computer Science, 2565, 267–278.

66. Petersson, G.A., Bennett, A., Tensfeldt, T.G., Al- Laham, M.A., Shirley, W.A., & Mantzaris, J. (1988). A complete basis set model chemistry. I. The total energies of closed-shell atoms and hydrides of the first-row elements, Journal of Chemical Physics, 89, 2193.

67. MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K, Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., & Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins, The Journal of Physical Chemistry B, 102, 3586–3616.

68. Besombes, S., & Mazeau, K. (2004). Molecular dynamics simulations of a guaiacyl β-O-4 lignin model compound: Examination of intramolecular hydrogen bonding and conformational flexibility, Biopolymers, 73, 301–315.

69. Besombes, S., Utille, J.P., Mazeau, K., Robert, D., & Taravel, F.R. (2004). Conformational study of a guaiacyl β-O-4 lignin model compound by NMR. Examination of intramolecular hydrogen bonding interactions and conformational flexibility in solution, Magnetic Resonance in Chemistry, 42, 337–347.

70. Petridis, L., & Smith, J.C. (2009). A molecular mechanics force field for lignin, Journal of Computational Chemistry, 30, 457–467.

71. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., & Pastor, M.L. (1983). Comparison of simple potential functions for simulating liquid water, Journal of Chemical Physics, 79, 926–935.

72. Schulz, R., Lindner, B., Petridis, L., & Smith, J.C. (2009). Scaling of multimillion-atom biological molecular dynamics simulation on a petascale supercomputer, Journal of Chemical Theory and Computation, 5, 2798–2808.

73. Matthews, J.F., Skopec, C.E., Mason, P.E., Zuccato, P., Torget, R.W., Sugiyama, J., Himmel, M.E., & Brady, J.W. (2006). Computer simulation studies of microcrystalline cellulose Iβ, Carbohydrate Research, 341, 138–152.

74. Nielsen, S.O., Lopez, C.F., Srinivas, G., & Klein, M.L. (2004). Coarse grain models and the computer simulation of soft materials, Journal of Physics: Condensed Matter, 16, R481.

75. Tozzini, V. (2005). Coarse-grained models for proteins, Current Opinion in Structural Biology, 15, 144–150.

76. Srinivas, G., Cheng, X., & Smith, J.C. (2011). A solvent-free coarse-grain model for crystalline and amorphous cellulose fibrils, Journal of Chemical Theory and Computation, 7, 2539–2548.

77. Cody, G.D., & Sághi-Szabó, G. (1999). Calculation of the ^{13}C NMR chemical shift of ether linkages in lignin derived geopolymers: Constraints on the preservation of lignin primary structure with diagenesis, Geochimica et Cosmochimica Acta, 63, 193–205.

78. Warshel, A., & Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, Journal of Molecular Biology, 103, 227–249.

79. Senn, H.M., & Thiel, W. (2009). QM/MM methods for biomolecular systems, Angewandte Chemie International Edition, 48, 1198–1229.

80. Senn, H.M., & Thiel, W. (2007). QM/MM studies of enzymes, Current Opinion in Chemical Biology, 11, 182–187.