# A Molecular Mechanics Force Field for Lignin

**LOUKAS PETRIDIS, JEREMY C. SMITH**
*Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831*

*Received 14 February 2008; Revised 8 May 2008; Accepted 12 June 2008*
*DOI 10.1002/jcc.21075*
*Published online 1 August 2008 in Wiley InterScience (www.interscience.wiley.com).*

**Abstract:** A CHARMM molecular mechanics force field for lignin is derived. Parameterization is based on reproducing quantum mechanical data of model compounds. Partial atomic charges are derived using the RESP electrostatic potential fitting method supplemented by the examination of methoxybenzene:water interactions. Dihedral parameters are optimized by fitting to critical rotational potentials and bonded parameters are obtained by optimizing vibrational frequencies and normal modes. Finally, the force field is validated by performing a molecular dynamics simulation of a crystal of a lignin fragment molecule and comparing simulation-derived structural features with experimental results. Together with the existing force field for polysaccharides, this lignin force field will enable full simulations of lignocellulose.

 J Comput Chem 30: 457–467, 2009

**Key words:** lignin; force field parameterization; biofuels; plant cell wall; CHARMM

## Introduction

The plant cell wall is made of cellulose microfibrils, which are embedded in a matrix of polysaccharides (hemicellulose and pectins) and lignin.[1] The main function of the cell wall is to provide structural rigidity and protection to the cell.

Plant cell wall structure has come under renewed interest recently in the context of the production of bioethanol from the enzymatic hydrolysis of lignocellulosic biomass (see[2–6] and references within). Cellulosic ethanol production is a two stage process involving first the hydrolysis of cellulose by cellulases to smaller oligosaccharides and then fermentation of the sugars to ethanol. The hydrolysis step is the bottleneck of the process because of the natural resistance, or "recalcitrance", of plant cell walls to degradation.[3]

There are two main physical factors contributing to biomass recalcitrance.[6] Firstly, cellulose is found in crystalline fibrils the compact structure of which impedes enzymatic access. In comparison, amorphous cellulose is readily digested by enzymes.[7] Secondly, matrix polysaccharides and lignin coat the cellulose fibril and act as a physical barrier preventing enzymes from reaching the cellulose. There is evidence that lignin also binds to the cellulose-binding-module of the enzymes, further inhibiting enzymatic action.[8] Confirmation of the contribution of lignin to biomass recalcitrance is provided by the finding that removing lignin from biomass increases the cellulose-hydrolysis yield from ∼20% to 98%.[9]

Computer simulation is a powerful tool for complementing experiment in obtaining an understanding of the molecular-level structure and dynamics of lignocellulose. Although there is a large volume of simulation work on cellulose,[10–15] there have been relatively few computational studies of lignin. Moreover, the computational studies of lignin[16–19] employed the CHARMM27 empirical force field, which was developed to model proteins rather than lignin. In this work we present the first essential step towards the accurate computer simulation of lignin: the derivation of an empirical molecular mechanics (MM) force field. Together with the existing force field for polysaccharides,[20] this force field will enable full simulations of lignocellulose.

Lignin is a heterogeneous aromatic biopolymer found in both the primary and secondary cell walls, for a review see Ref. 21. It is formed by radical polymerization of three alcohol monolignols: *p*-coumaryl, coniferyl, and sinapyl, shown in Figure 1a. The chemical composition and structure of lignin is highly heterogeneous, varying significantly between different plant species and even within different parts of the same plant cell wall. Although complex, lignin is composed primarily of three units *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S), derived by oxidation of *p*-coumaryl, coniferyl, and sinapyl, respectively. An illustration of one G and one S units is shown in Figure 1b.

There are five types of linkages that connect the units, leading to the formation of the long lignin biopolymer. The most common linkage (50−80% probability) is $\beta$-O-4′, connecting the oxygen of the hydroxyl on the phenyl ring in one unit with the second tertiary carbon of the other, see Figure 1b. Other common linkages are $\beta$-$\beta'$, $\alpha$-O-4′, and $\beta$-5′.

In this article we present a parameterization of a molecular mechanics force field of lignin, complementing the CHARMM

***Correspondence to:*** L. Petridis; e-mail: petridisl@ornl.gov

Contract/grant sponsor: U.S. Department of Energy (DOE) Bioenergy Science Center and DOE Field Work Proposal ERKP704.

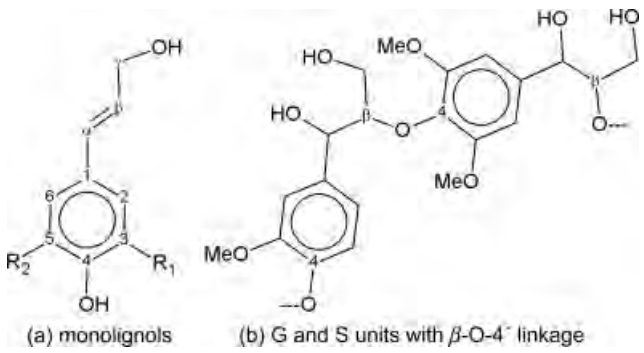


**Figure 1.** (a): The three monolignols: *p*-coumaryl ($R_1 = R_2 = H$), coniferyl ($R_1 = H, R_2 = OCH_3$), and sinapyl ($R_1 - R_2 = OCH_3$). (b) A guaiacyl unit connected with a $\beta$-*O*-4′ linkage to a syringyl unit. The dashed lines imply the continuation of the lignin chain.

empirical force field. The electrostatic interactions were optimized by assigning partial atomic charges so as to reproduce quantum chemical data. Care was taken to ensure the charges account for the electronic polarization present in condensed phase simulations. Dihedral force constants were determined by examining potential energy surfaces. The interactions between bonded atoms were subsequently optimized with respect to quantum chemical vibrational data using the Automated Frequency Matching Method.[22] As a final step, a molecular dynamics (MD) simulation of a relevant small-molecule crystal was performed. The satisfactory agreement between the time-averaged simulated and experimental structures suggests the present force field is suitable for use in MD simulation.

## Materials and Methods

### *Parameterization Strategy*

In this section we outline the general strategy employed to obtain the force field of lignin. The CHARMM potential energy function[23] of a molecule is approximated by eq. (1):

$$E = \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{U-B} K_{UB}(s - s_0)^2 + \sum_{\text{dihedrals}} K_\phi[1 + \cos(n\phi - \delta)] + \sum_{\text{improper}} K_\psi(\psi - \psi_0)^2 + \sum_{\text{nonbonded}} \left\{ \epsilon_{ij} \left[ \left( \frac{R_{\min,ij}}{r_{ij}} \right)^{12} - \left( \frac{R_{\min,ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi D r_{ij}} \right\}, \quad (1)$$

where $K_b$, $K_\theta$, $K_{UB}$, $K_\phi$, $K_\psi$ are respectively the bond, angle, Urey-Bradley, dihedral, and improper dihedral force constants; $b$, $\theta$, $s$, $\phi$, $\psi$ are the bond length, bond angle, Urey-Bradley 1, 3-distance, dihedral angle, and improper dihedral angle, with the subscript zero denoting the equilibrium value of the respective terms. The dihedral term also involves $n$, its periodicity, and $\delta$, its phase. Nonbonded terms between pairs of atoms denoted $i$ and $j$ at a distance $r_{ij}$ include the Lennard–Jones 6–12 potential for the van der Waals interactions and standard Coulomb interactions. At the Lennard–Jones minimum, $R_{\min}$ is the distance between atoms and $\epsilon$ the well depth. The effective dielectric constant is $D$ and is equal to one in this study and the partial atomic charges are given by $q$. The above empirical parameters are molecule dependent and must be optimized to model the specific molecule prior to performing simulations. This optimization step is generally referred to as parameterization of the force field and its application to the case of lignin is the main task of this work.

This parameterization of lignin follows the main procedure of parameterization of proteins[24] and linear and cyclic ethers[25] for the CHARMM force field. Of relevance for the present work, a recent ether force field[25] validated the use of existing tertiary carbon parameters for linear ethers and developed new non-bonded parameters for the ether oxygen based on reproducing both gas- and condensed-phase properties. Lignin also has a linear ether bond, but with the difference to those examined in reference 25 that the oxygen is bonded to a phenyl ring, thus creating an extensive $\pi$-bond that influences bond and angle vibrations, dihedral rotations and charge distribution. For this reason new parameters were needed for lignin.

Two model compounds were used to save computational time, since extensive quantum mechanical (QM) calculations on a full lignin dimer (for example that shown in Figure 1b) are computationally expensive. The two model systems are shown in Figure 2: The first compound methoxybenzene, also known as anisole, in Figure 2a, incorporates the basic features of a $\beta$-O-4′ link, an ether oxygen bonded to a tertiary and an aromatic carbon. Anisole was used to obtain all parameters involving the ether oxygen atom. The second compound, in Figure 2d, is *p*-hydroxyphenyl (PHP), the simplest lignin unit. PHP was used to obtain all lignin parameters not involving the ether oxygen.

Parameters were optimized by considering two factors. Firstly, the target data was reproduced as closely as possible. Effectively this ensures that the force field describes accurately specific properties of lignin. Secondly, compatibility with the existing CHARMM force field was ensured by restricting optimization to parameters that did not already exist in the force field. For example, the charges of the carbon atoms on the phenyl ring were maintained to the existing values of $-0.115q$, although electrostatic potential analysis suggested carbons not bonded to oxygen might have slightly lower charges. Similarly, the nonpolar hydrogen atoms of the methyl group were assigned the standard CHARMM force field charge of $0.09q$. It was, however, found necessary to create a new atom type, OET, to represent the ether oxygen bonded to the phenyl ring and a tertiary carbon.

The optimization strategy for the new parameters is summarized in the diagram in Figure 3. Internal parameters (equilibrium values for bond lengths, angles and dihedrals) were taken from MP2/6-31G* QM optimized geometries and were not further revised. The van der Waals parameters were taken unaltered from the existing protein[24] and ether[25] CHARMM force fields. Parameterization of the van der Waals parameters for the new atom type, OET, was not deemed necessary since all three ether oxygen types in reference 25 have the same Lennard–Jones parameters. This is a strong indication that the new atom-type in the present work, which is also an ether oxygen, will have the same parameters. Initial values for the partial atomic charges of $O_1$, $C_1$, and $C_\alpha$ were deduced from a restricted fit to the QM Electrostatic Potential (RESP)[26] that derives charges

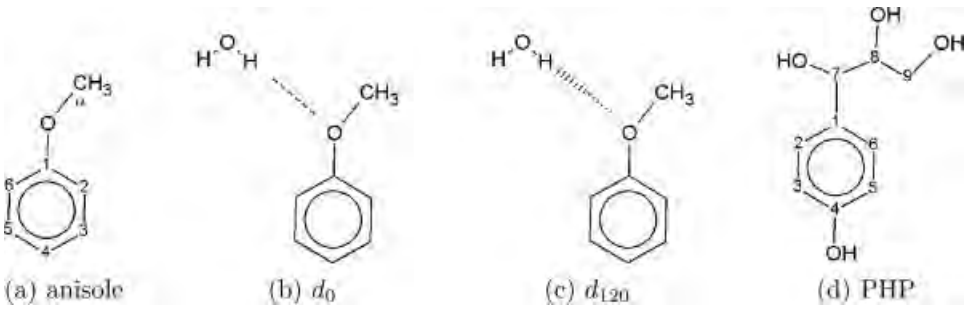


**Figure 2.** Sketches of model compounds used in parameterization: (a) methoxybenzene (or anisole) with atom names indicated, (b) and (c) the two supramolecular geometries used to obtain the partial charges, (d) *p*-hydroxyphenyl.

that fit the quantum mechanical electrostatic potential calculated on selected grid points. As mentioned earlier, all other charges were kept to their original CHARMM values. An iterative procedure, described in the next paragraph, was followed until convergence was achieved.

Charges were further optimized with respect to the QM interaction energies using a supramolecular approach with a model compound (anisole) interacting with one water molecule. The convergence criterion for charges was a $< 5\%$ difference with the QM energies. After completing the nonbonded interactions, parameters for dihedral rotations were deduced from QM potential energy surfaces. Fitting was performed by varying the dihedral force constants in increments, first of 0.1 kcal/mol and then 0.01 kcal/mol. In cases where QM surfaces were well reproduced (Figs. 4–6 and 8) the value of the force constant that gave the smallest error was chosen. In the cases where the shapes were not as well reproduced (Figs. 7 and 9) emphasis was placed on the low-energy region <3 kcal/mol. Finally, the remaining bond and angle parameters were derived by reproducing QM vibrational frequencies and normal modes with the Automated Frequency Matching Method (AFMM).[22] Good starting values of force constants were provided for AFMM by considering similar existing parameters in ethers, phenol and alcohols. Bond and angle optimization was repeated until the merit function [eq. (2)] reached a value of less than 60 $cm^{-1}$. The iterative procedure of Figure 3 was repeated until all convergence criteria were satisfied. In practice four iterations were found to be required.

### *Computational Details*

All QM calculations were performed with the NWChem 5.0 package.[27,28] Geometry optimizations were performed at the MP2[29] level of theory and to keep consistency with previous CHARMM parameterizations, the 6-31G* basis set was used.[24] To reduce computational time the frozen core approximation was invoked in all MP2 calculations, in which core orbitals (1*s* for O and C) are constrained to remain doubly-occupied. Optimization of the dihedral parameters was based on reproducing QM (MP2/6-31G*) adiabatic energy surfaces, in which the selected dihedral is held constant while the remaining degrees of freedom are allowed to relax to a constrained energy minimum.

Two methods were employed for partial charge determination: a restricted Electrostatic Potential (RESP) fit[26] and supramolecular calculations. The NWChem Electrostatic Potential module derives partial atomic charges that fit the quantum mechanical electrostatic potential calculated on selected grid points. Improved fitting was achieved by imposing additional constraints, where the molecule was grouped into subsets of atoms, which were each constrained to have zero total charge.[30] In anisole (Fig. 2a) the five hydrogens of the phenyl ring were constrained to have the opposite charges of the carbons they are bonded to. Additionally, the $C_1$-$O$-$C_\alpha H_3$ group of atoms was also constrained to net zero charge. The maximum distance between a grid point and any of the atomic centers was 3 Å and the grid spacing was 0.05 Å.

Supramolecular calculations of minimum interaction distances and energies between anisole and water were performed by keeping the molecules fixed at their MP2 optimized and TIP3P[31] geometries, respectively. The water molecule was also fixed at two different orientations: $d_0$ in which the water lies on the phenyl plane and $d_{120}$ in which water hydrogen points towards the ether oxygen lone pair,

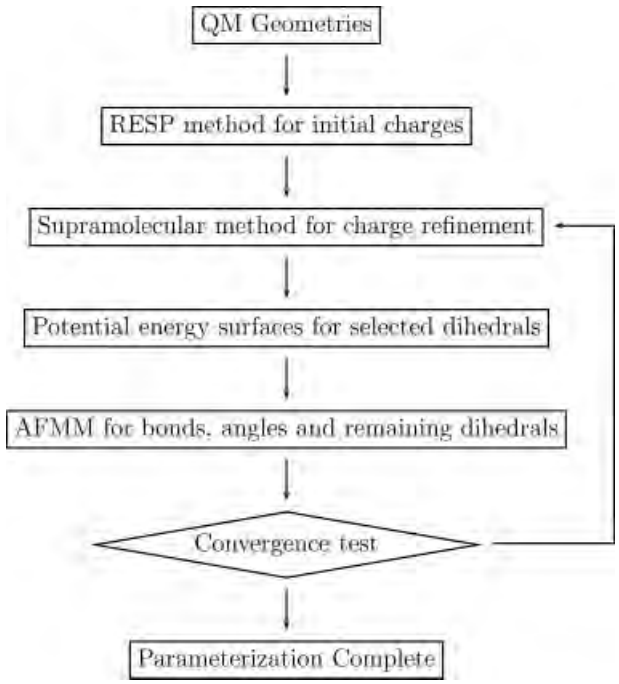


**Figure 3.** Schematic representation of the parameterization strategy.

**Table 1.** List of Anisole Atoms and Their Respective RESP Charges.

| Atoms | RESP charges |
|---|---|
| $C_1$ | 0.042 |
| $C_2$ | −0.128 |
| $C_3$ | −0.126 |
| $C_4$ | −0.070 |
| $C_5$ | −0.075 |
| $C_6$ | −0.093 |
| $O$ | −0.261 |
| $C_\alpha$ | 0.093 |
| $H_\alpha$ | 0.042 |

see Figure 2. The HF/6-31G* energy was computed varying only the intermolecular distance (in increments of 0.01 Å) so as to locate the minimum energy separation and interaction energy (the energy of complex minus the energy of the two isolated molecules). For comparison between the QM and MM data, the QM energies were scaled by a factor of 1.16 so as to correct for the absence of polarization and attractive Lennard–Jones contributions in the Hartree–Fock calculations, but no correction was made for basis set superposition error. This approach is widely used in CHARMM force field parameterizations, for example references 32 and 25. Similarly, MM minimum distances are expected to be about 0.2 Å shorter than the *ab initio* values.[32] The above procedure was performed for both orientations. To verify that the partial charges, derived from calculations on the model compound, are suitable for lignin a further calculation was performed. The minimum interaction energy and distance were calculated between whole lignin dimer (G and S units connected with a $\beta$-O-4′ linkage) shown in Figure 1b and water.

For the vibrational frequency analysis the DFT/B3LYP level of theory was used with 6-31G* basis set together with the SBKJC effective core potential[33] and the Hessian was computed numerically. This combination of basis set and effective core potential has been shown to be a successful QM procedure for obtaining vibrational frequencies, since the deviation of the theoretical frequencies from the experimental frequencies in test molecules is relatively small.[34] The frequencies were calculated numerically. A frequency scaling factor of 0.9614 was used to compensate for the use of the harmonic approximation to the potential energy surface.[34]

All MM calculations were performed using the CHARMM 32b software.[23] MD simulation of the small-molecule crystal examined here (see Results) was done with the NAMD software.[35] Periodic boundary conditions were used to mimic the actual crystal environment and the integration time step was 1 fs. Starting with the experimental coordinates obtained with X-ray diffraction,[36] a crystal of 64 (4 × 4 × 4) unit cells was generated. One hundred twenty-eight dimers were simulated during the run, a total of 6912 atoms. The system was first energy minimized and then heated up to the experimental temperature (173 K) with 1 K temperature steps. During the first 82.3 ps equilibration run velocity rescaling was allowed. During the second 100 ps equilibration run and the 1 ns production run, velocity rescaling was not allowed and the system was kept under constant temperature (173 K) and pressure (1 atm). Temperature and pressure controls used the in-build capabiities of NAMD. Largevin Dynamics were used to maintain constant temperature a capability based on that implemented in X-PLOR.[37] For constant pressure, the Nosé-Hoover Langevin piston algorithm was used

## Results

### *Partial Atomic Charges*

Initial partial charges for the ether oxygen and the two carbons bonded to it ($C_1$ and $C_\alpha$) were derived by applying the RESP method to anisole (see Fig. 2). The full list of RESP charges obtained is shown in Table 1.

The above RESP charges needed further refinement for the following reasons. The RESP charges are based on a gas phase wavefunction and may not necessarily be consistent with the condensed phase.[38] Furthermore, the absence of symmetry between carbons $C_2$ and $C_6$ is due to the frozen, minimum energy, orientation of the methoxy group. Asymmetry is undesirable in the MM force field since this group is expected to rotate. Finally, the phenyl and methyl RESP charges are not consistent with the remaining CHARMM force field and their transferability to lignin molecules is not guaranteed.

Given the above considerations, the partial charges were further adjusted so as to reproduce minimum distances and interaction energies between anisole and a water molecule (see for example ref. 39). Two geometries were considered in this supra-molecular approach, the first ($d_0$) with water lying on the phenyl plane and the second ($d_{120}$) with the water hydrogen pointing at the position of the lone pair of the ether oxygen. A list of all final atomic charges is shown in Table 2. Only three charges ($C_1$, $O$, $C_\alpha$) in Table 2 were optimized, the rest being kept to their previous CHARMM values. The anisole oxygen has smaller charge than the value of −0.34 obtained in the previous linear ether study.[25] This can be attributed to the resonance effect present in anisole, in which the electron cloud of the extended $\pi$ bond is shared between the aromatic ring and the more electronegative oxygen. Apparently, the resonance effect has a greater effect than the inductive effect in which the electronegative oxygen exerts a pull on the electron density in the benzene ring through the sigma bond.

To mimic the effect of electronic polarizability, which is not explicitly taken into account in additive force fields, atomic charges were purposely overestimated. This leads to an enhanced molecular dipole moment, with the QM gas-phase dipole moment being 1.42

**Table 2.** A List of the Anisole Atoms with Their Respective Charges.

| Atom name | Atom type | Charge |
|---|---|---|
| $C_\alpha$ | CT3 | −0.060 |
| $H_{\alpha1}$, $H_{\alpha2}$, $H_{\alpha3}$ | HA | 0.090 |
| $O$ | OET | −0.280 |
| $C_1$ | CA | 0.070 |
| $C_2$, $C_3$, $C_4$, $C_5$, $C_6$ | CA | −0.115 |
| $H_2$, $H_3$, $H_4$, $H_5$, $H_6$ | HP | 0.115 |

Atom names refer to Figure 2 and atoms types follow the CHARMM27 force field, with the new atom type labeled as OET.

**Table 3.** Minimum Interaction Energies (kcal/mol) and Distances (Å) Between Water:Anisole (Fig. 2) and Water:Lignin-Dimer (Fig. 1b).

| | Interaction energies | | Interaction distances | |
|---|---|---|---|---|
| Orientations | QM | MM | QM | MM |
| $d_0$ | −4.01 | −3.96 | 2.15 | 1.82 |
| $d_{120}$ | −3.18 | −3.09 | 2.16 | 1.87 |
| dimer | −3.93 | −4.02 | 2.10 | 1.81 |

QM interaction energies were scaled by 1.16 as described in the text. Orientation geometries considered have the dihedral between the water molecule and the phenyl ring being 0, 120, and 60 degrees, respectively.

**Table 4.** *p*-Hydroxymethyl (PHP) Atoms with Their Respective Charges.

| Atom name | Atom type | Charge |
|---|---|---|
| $C_1$ | CA | 0.000 |
| $C_2$, $C_3$, $C_5$, $C_6$ | CA | −0.115 |
| $H_2$, $H_3$, $H_5$, $H_6$ | HP | 0.115 |
| $C_4$ | OH1 | 0.110 |
| $O_4$, $O_7$, $O_8$, $O_9$ | OH1 | −0.540 |
| $H_{O4}$, $H_{O7}$, $H_{O8}$, $H_{O9}$ | H | 0.430 |
| $C_7$, $C_8$ | CT1 | 0.140 |
| $C_9$ | CT2 | 0.050 |
| $H_7$, $H_8$, $H_{91}$, $H_{92}$ | HA | 0.090 |

Atom names refer to Figure 2 and atom types are according to the CHARMM27 force field.

Debyes whereas the MM value is 1.66 Debyes. Table 3 shows the comparison of the MM and QM interaction energies and distances, which were used to optimize the anisole charges. The empirical calculations reproduce well the scaled QM interaction energies, with the error being less than 3%. The empirical model gives distances about ∼0.3 Å shorter than the QM values, a result of intentionally overestimating the gas phase charges to obtain good condensed phase properties. In the previous general force field for ethers a similar behavior was observed with a ∼0.3 Å difference between QM and MM.[25] Finally, charge transfer was examined by Mulliken analysis (using the NWChem software) of the electronic charge density and was found to be not significant.

A further calculation was performed to ensure that the partial atomic charges of Table 2, which were derived using a model compound, can be transferred to lignin. The minimum interaction energies and distances between a whole lignin dimer (G and S units connected with a $\beta$-O-4′ linkage), shown in Figure 1b, and a TIP3P water molecule were obtained without further refinement of the parameters. This supramolecular complex is labelled as "dimer" in Table 3. As in the case of anisole and water, the agreement between the QM and MM interaction energies was excellent, further justifying the use of the charges in Table 2 for the $\beta$-O-4′ lignin linkage. Furtheremore, The lignin dimer had two methoxy groups at position 3 and 5 of the phenol ring, indicating that the presence of the methoxy group does not alter the charge of the ether oxygen.

For completeness the atom types and partial charges of PHP are included in Table 4, although these were taken unchanged from the existing CHARMM force field. To build a guaiacyl or a syringyl unit, methoxy groups must be added to the phenol rings and their respective atom types and charges can be obtained from the anisole in Table 2.

### *Bonded Parameters*

#### *Dihedral Rotations*

Dihedral rotations around the $\beta$-O-4′ linkage play a significant role in determining the configuration of the lignin macromolecule and therefore special care was taken to obtain good parameters for the equivalent dihedrals of the model system. These are $\omega_1$: $C_2$-$C_1$-$O$-$C_\alpha$ (or $C_6$-$C_1$-$O$-$C_\alpha$ equivalently) and $\omega_2$: $C_1$-$O$-$C_\alpha$-$H$. In the lignin dimer shown in Figure 1b, molecule rotation around these dihedrals quickly leads to severe steric hindrance between the two aromatic rings, thus rendering separate determination of the intrinsic dihedral potential impractical. Therefore the use of a smaller model compound is necessary.

The optimization was based on reproducing quantum-chemically obtained adiabatic energy surfaces, where the selected dihedral ($\omega_1$ or $\omega_2$) is held constant while the remaining degrees of freedom are allowed to relax to a constrained energy minimum. As seen in Figures 4 and 5 the MM surfaces closely follow the target QM data. Three terms with periodicity $n = 1, 2, 3$ were used to ensure that the minima of $\omega_2$ at 120 and 240 degrees are not zero and accurately describe the low energy regions that the system frequently samples during a MD run (Fig. 5). Similarly, two dihedrals with periodicity $n = 2, 4$ were required to describe the flat high energy regions of $\omega_1$ which are a result of the resonance effect (Fig. 4).

The remaining dihedral parameters of lignin that do not involve the ether oxygen were deduced from the more complex rotational potential energy profiles of the second model compound, PHP. The following four dihedrals were examined: $\omega_3 = C_2$-$C_1$-$C_7$-$X$, $\omega_4 = C_1$-$C_7$-$O_7$-$H_{O7}$, $\omega_5 = C_1$-$C_7$-$C_8$-$X$, and $\omega_6 = X$-$C_8$-$C_9$-$X$; where $X$ refers to wild atom types. All other dihedral parameters were obtained from the existing CHARMM force field. In initial calculations only the dihedral under examination was kept constant, while the rest of the molecule was allowed to relax. This method, however, produces very complex energy profiles as variation of the dihedrals examined induces large distortions of the molecule. Therefore, to accurately model a specific dihedral potential (e.g. $\omega_3$), it was found to be necessary to constrain the remaining three ($\omega_4$, $\omega_5$ and $\omega_6$) to their global equilibrium values.

To maintain consistency with the existing CHARMM force field, parameters for atom types that already existed in the force field were not optimized. Therefore, our initial approach was not to reparameterize the hydroxyl dihedrals $C_7$-$C_8$-$O_8$-$H_{O8}$ and $C_8$-$C_9$-$O_9$-$H_{O9}$. In practice, however, the MM optimized value of dihedral $C_8$-$C_9$-$O_9$-$H_{O9}$ was $12^o$ different than the QM optimized value. Furthermore, if there is large deviation between the QM and MM geometries, the AFMM automated method used to parametrize bonds and dihedrals will not be able to properly match the normal mode eigenvectors. This discrepancy indicated that optimization is necessary to adjust the $C_8$-$C_9$-$O_9$-$H_{O9}$ dihedral parameters. In the case of the $C_7$-$C_8$-$O_8$-$H_{O8}$ dihedral, good agreement was found between the QM and MM optimized geometries and the parameters were taken unaltered from the existing CHARMM force field for alcohols.

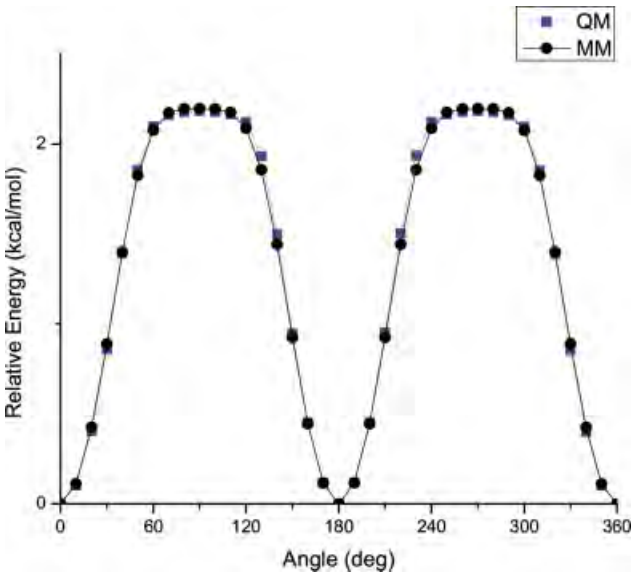


**Figure 4.** Potential energy for rotation around the $\omega_1 = C_2\text{-}C_1\text{-}O\text{-}C_\alpha$ dihedral of anisole. MM and QM (MP2/6-31G*) data are shown. The $x$-axis represents deviations from the equilibrium value of the dihedral, which was 180.0°.

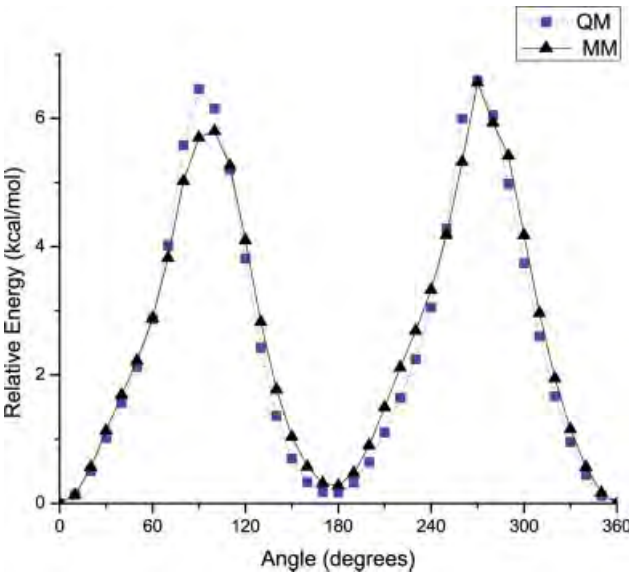


**Figure 6.** Potential energy for rotations around the $\omega_3 = C_2\text{-}C_1\text{-}C_7\text{-}X$ dihedral of PHP. MM and QM (MP2/6-31G*) data are shown. The $x$-axis represents deviations from the equilibrium value of the dihedral, which for $X = C_8$ was −64.5°.

The resulting energy surfaces are shown in Figures 6–9. Although the agreement between the QM and MM data is not perfect, the rather complex shapes are reproduced satisfactorily. During parameterization, special care was taken to reproduce as accurately as possible the low energy (≤3 kcal/mol) regions since it is these regions that will be thermally most-frequently sampled. For dihedrals $\omega_4$ (Fig. 7) and $\omega_5$ (Fig. 8) the reproduction of the low-energy regions was achieved at the expense of reproducing perfectly the high-energy regions. In the case of $\omega_6$, the rotational barriers are overestimated. This was deemed necessary since smaller dihedral constants which lower the barriers also shift the position of the MM minimum away from zero degrees. A complete list of the dihedral parameters of lignin can be found in Table 5.

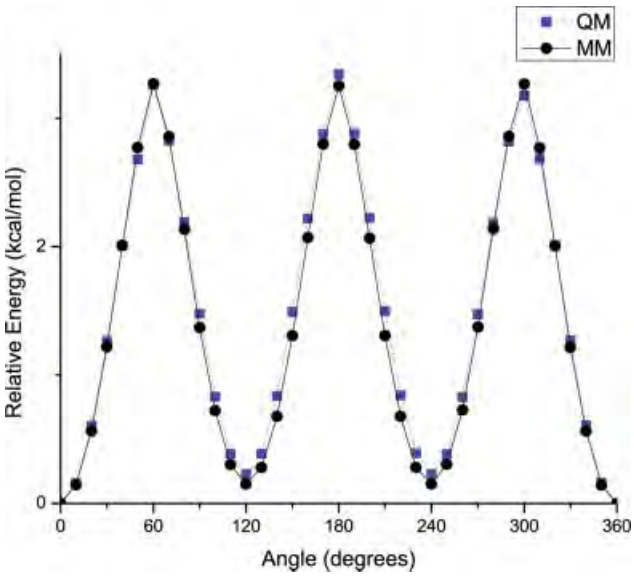


**Figure 5.** Potential energy for rotations around the $\omega_2 = C_1\text{-}O\text{-}C_\alpha\text{-}H$ dihedral of anisole. MM and QM (MP2/6-31G*) data are shown. The $x$-axis represents deviations from the equilibrium value of the dihedral, which was 180.0°.

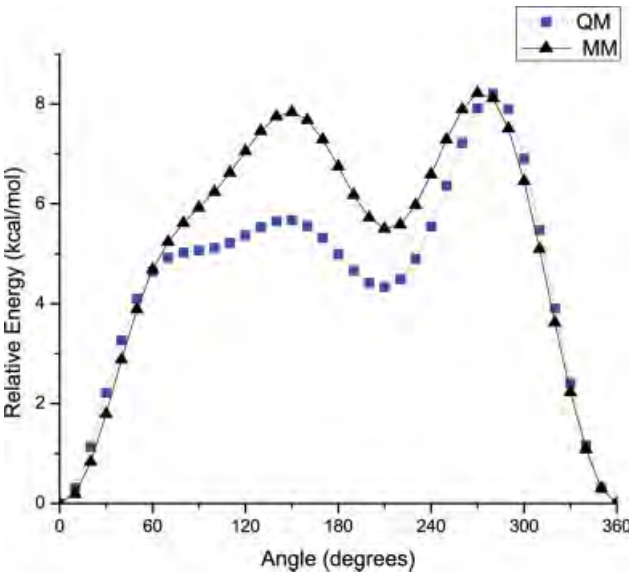


**Figure 7.** Potential energy for rotations around the $\omega_4 = C_1\text{-}C_7\text{-}O_7\text{-}H_{O7}$ dihedral of PHP. MM and QM (MP2/6-31G*) data are shown. The $x$-axis represents deviations from the equilibrium value of the dihedral, which was 82.9°.

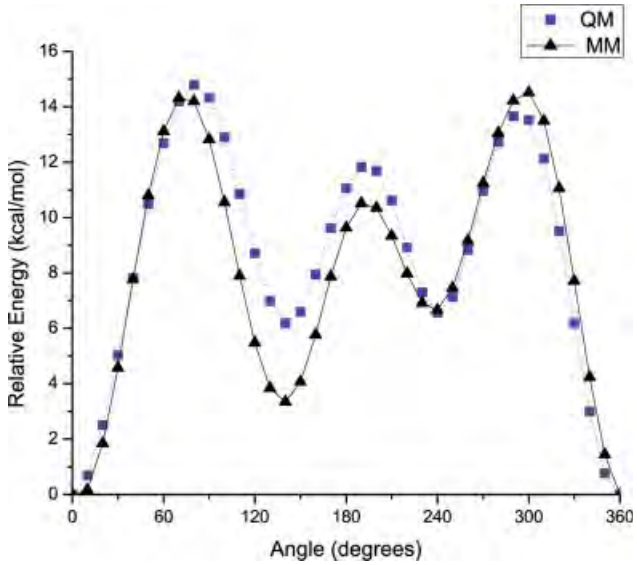


**Figure 8.** Potential energy for rotations around the $\omega_5 = C_1$-$C_7$-$C_8$-$X$ dihedral of PHP. MM and QM(MP2/6-31G*) data are shown. The $x$-axis represents deviations from the equilibrium value of the dihedral, which for $X = C_9$ was 176.2°.

### *Bond and Angle Vibrations*

The remaining bonded parameters (bonds and angles) were optimized to reproduce vibrational frequencies and eigenvector projections derived from QM calculations. For this the automated frequency matching method[22] (AFMM) was employed, which optimizes the MM parameter set until the best fit with the QM reference set is obtained. AFMM requires both the eigenfrequencies and eigenvectors of the MM set to match with the QM data. This is an important aspect of the method, since it avoids incorrect mode matching and thus misleading reproduction of vibrational frequencies. The merit function is given by:

$$\sigma^2 = \frac{\sum_i^{3N-6} \left(\omega_i^{\mathrm{qm}} - \omega_i^{\mathrm{mm}}\right)^2}{3N-6}, \qquad (2)$$

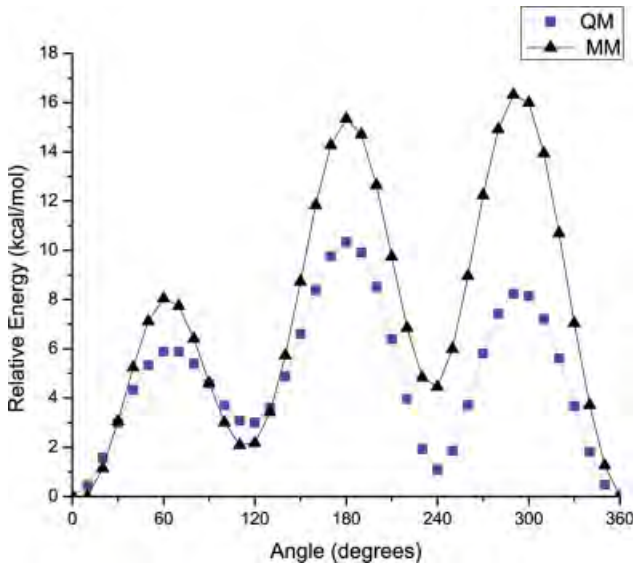


**Figure 9.** Potential energy for rotations around the $\omega_6 = X$-$C_8$-$C_9$-$X$ dihedral of PHP. MM and QM (MP2/6-31G*) data are shown. The $x$-axis represents deviations from the equilibrium value of the dihedral, which for $C_7$-$C_8$-$C_9$-$O_9$ was −175.1°.

**Table 5.** Dihedral Parameters for Lignin.

| Dihedrals | $K_\phi$ (kcal/mol) | $n$ | $\delta$ (deg) | Source |
|---|---|---|---|---|
| CA-CA-CA-CA | 3.10 | 2 | 180 | CHARMM |
| CA-CA-CA-CT3 | 3.10 | 2 | 180 | CHARMM |
| HP-CA-CA-CT3 | 4.20 | 2 | 180 | CHARMM |
| HP-CA-CA-CA | 3.70 | 2 | 180 | PHP |
| HP-CA-CA-HP | 2.40 | 2 | 180 | CHARMM |
| OH1-CA-CA-CA | 3.10 | 2 | 180 | CHARMM |
| OH1-CA-CA-HP | 4.20 | 2 | 180 | CHARMM |
| H-OH1-CA-CA | 0.99 | 2 | 180 | CHARMM |
| CT1-CA-CA-CA | 3.10 | 2 | 180 | CHARMM |
| CT1-CA-CA-HP | 4.20 | 2 | 180 | CHARMM |
| OET-CA-CA-OH1 | 3.10 | 2 | 180 | CHARMM |
| OET-CA-CA-OET | 3.10 | 2 | 180 | CHARMM |
| CA-CA-CT1-CT1 | 0.30 | 2 | 180 | PHP |
| CA-CA-CT1-HA | 0.15 | 2 | 180 | PHP |
| CA-CA-CT1-OH1 | 0.91 | 2 | 180 | PHP |
| CA-CT1-OH1-H | 0.40 | 1 | 0 | PHP |
| CA-CT1-OH1-H | 0.80 | 2 | 0 | PHP |
| CA-CT1-OH1-H | 0.65 | 3 | 0 | PHP |
| CA-CT1-CT1-OH1 | 0.50 | 3 | 0 | PHP |
| CA-CT1-CT1-CT2 | 0.50 | 3 | 0 | PHP |
| CA-CT1-CT1-HA | 0.50 | 3 | 0 | PHP |
| X-CT1-CT1-X | 0.20 | 3 | 0 | CHARMM |
| OH1-CT2-CT1-OH1 | 0.22 | 3 | 0 | PHP |
| OH1-CT2-CT1-CT1 | 0.22 | 3 | 0 | PHP |
| OH1-CT2-CT1-HA | 0.22 | 3 | 0 | PHP |
| HA-CT2-CT1-OH1 | 0.22 | 3 | 0 | PHP |
| HA-CT2-CT1-CT1 | 0.22 | 3 | 0 | PHP |
| HA-CT2-CT1-HA | 0.22 | 3 | 0 | PHP |
| X-CT2-OH1-X | 0.14 | 3 | 0 | CHARMM |
| X-CT1-OH1-X | 0.14 | 3 | 0 | CHARMM |
| CA-CA-OET-CT3 | 1.11 | 2 | 180 | anisole |
| CA-CA-OET-CT3 | 0.17 | 4 | 180 | anisole |
| CA-CA-OET-CT3 | 1.11 | 2 | 180 | anisole |
| CA-CA-OET-CT3 | 0.17 | 4 | 180 | anisole |
| CA-OET-CT3-HA | 4.00 | 1 | 0 | anisole |
| CA-OET-CT3-HA | 4.00 | 2 | 0 | anisole |
| CA-OET-CT3-HA | 0.11 | 3 | 0 | anisole |
| X-OET-CT1-X | 4.00 | 1 | 0 | anisole |
| X-OET-CT1-X | 4.00 | 3 | 0 | anisole |
| X-OET-CT1-X | 0.11 | 3 | 0 | anisole |
| OET-CA-CA-CA | 3.10 | 2 | 180 | anisole |
| OET-CA-CA-HP | 3.10 | 2 | 180 | anisole |

Also noted are the model compounds the parameters were derived from, CHARMM indicating that the parameters were not changed from the existing force field.

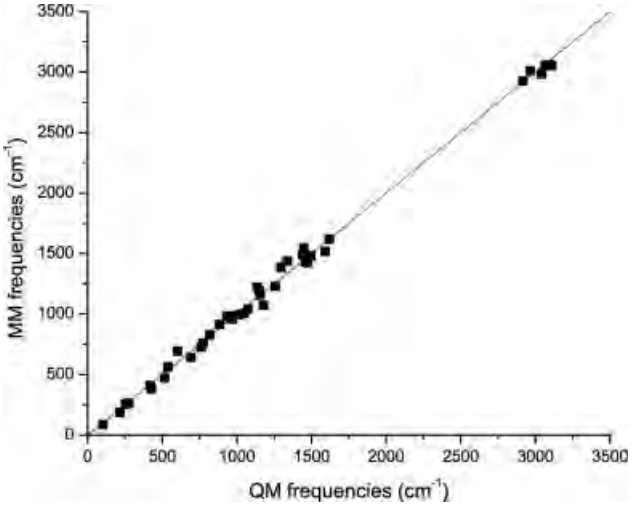


**Figure 10.** Vibrational frequencies of anisole. The plotted line shows the ideal fit between QM and MM data.

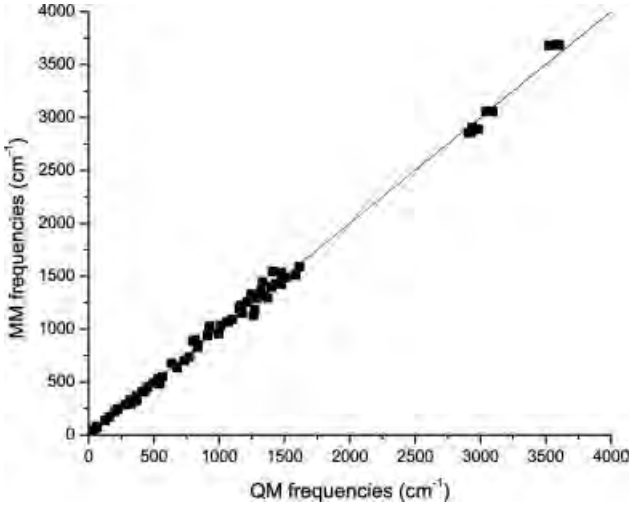


**Figure 11.** Vibrational frequencies of *p*-hydroxyphenyl (PHP). The plotted line shows the ideal fit between QM and MM data.

where $\omega^{qm}$ and $\omega^{mm}$ refer to the vibrational frequencies obtained with QM and the MM methods respectively and $N$ is the total number of atoms of the molecule. Initial values for bond and angle parameters were taken from chemically similar compounds that have been already parameterized: benzene and linear ethers for anisole and phenol and alcohols for PHP.

The MM and QM eigenvectors were inspected using VMD[40] and it was ensured that modes were correctly matched. In the case of anisole (the smaller of the two model compounds) there was a one-to-one correspondence. In the case of PHP, some QM modes were assigned to more than one MM mode. In these cases a visual inspection was used to correctly match the modes.

The resulting plots of $\omega^{qm}$ and $\omega^{mm}$ for anisole and PHP are shown in Figures 10 and 11, respectively. The closer to the diagonal the points are, the better the fit. In both model compounds there were modes that could not be matched properly by only adjusting the bond and angle terms. The MM frequencies of these modes (901 and 940 $cm^{-1}$ in PHP and 985 $cm^{-1}$ in anisole) were $\sim$100 $cm^{-1}$ higher than these from QM. Visualization of the motions showed that these modes involve dihedral rotations of the phenyl hydrogens. By decreasing the *HP-CA-CA-CA* dihedral force constant from 4.2 to 3.7 kcal/mol a better match was obtained for these modes. A full list of the bond and angle parameters of lignin can be found in Tables 6 and 7.

After optimization the root mean square deviation from the reference set was found to be $\sigma = 51.6$ $cm^{-1}$ for anisole and $\sigma = 55.6$ $cm^{-1}$ for PHP, similar in range to previous parameterizations involving AFMM, which quote values of $\sigma = 40$ $cm^{-1}$ [41] and $\sigma = 47 - 94$ $cm^{-1}$.[42]

### *Force Field Validation*

In the final part of this work, the parameter set was tested without further adjustment against a condensed phase experimental property of lignin that was not used during the parameterization. Because of the highly heterogeneous structure of lignin, the most appropriate experimental data to use is the crystal structure of a lignin-subunit dimer, erythro-2-(2,6-Dimethoxy-4-methylphenoxy)-1-(4-hydroxy-3,5-dimethoxyphenyl)propane-1,3-diol (EPD),[36] see Figure 12. The chosen compound is very similar to two syringyl units connected with a $\beta$-$O$-4′ linkage, but with the hydroxy group of one of the phenol rings substituted by a methyl group. The single crystal X-ray diffraction study revealed a triclinic $P\bar{1}$ structure whose unit cell dimensions are listed in Table 8. $P\bar{1}$ symmetry implies that for each atom in position $(x, y, z)$ there is an equivalent atom in position $(-x, -y, -z)$.

**Table 6.** Bond Parameters for Lignin.

| Bonds | $K_b$ (kcal/mol Å$^2$) | $b_0$ (Å) | Source |
|---|---|---|---|
| OH1-CA | 364 | 1.371 | PHP |
| CT1-CT1 | 213 | 1.500 | PHP |
| CT2-CT1 | 194 | 1.520 | PHP |
| CT1-CA | 313 | 1.490 | PHP |
| OET-CT3 | 347 | 1.424 | anisole |
| OET-CT3 | 347 | 1.424 | anisole |
| OET-CA | 230 | 1.372 | anisole |
| CT3-CA | 230 | 1.372 | anisole |
| HA-CT3 | 341 | 1.111 | anisole |
| OH1-CT1 | 428 | 1.420 | CHARMM |
| OH1-CT2 | 428 | 1.420 | CHARMM |
| HA-CT1 | 309 | 1.110 | CHARMM |
| HA-CT2 | 309 | 1.110 | CHARMM |
| OH1-H | 545 | 0.960 | CHARMM |
| CA-CA | 305 | 1.375 | CHARMM |
| CA-HA | 340 | 1.083 | CHARMM |
| CA-HP | 340 | 1.080 | CHARMM |

Also noted are the model compounds the parameters were derived from, CHARMM indicating that the parameters were not changed from the existing force field.

**Table 7.** Angle Parameters for Lignin.

| Angles | $K_\theta$ [kcal/(mol rad$^2$)] | $\theta_0$ (deg) | $K_{ub}$ [kcal/ (mol Å$^2$)] | (Å) $S_0$ | Source |
|---|---|---|---|---|---|
| CT1-CA-CA | 35.5 | 120.7 | | | PHP |
| HA-CT1-CA | 37.7 | 109.6 | | | PHP |
| CT1-CT1-CA | 80.0 | 109.8 | | | PHP |
| OH1-CT1-CA | 82.8 | 112.5 | | | PHP |
| OET-CA-CA | 38.8 | 120.0 | | | anisole |
| CT3-OET-CA | 91.2 | 116.8 | | | anisole |
| OET-CT3-HA | 59.3 | 109.9 | | | anisole |
| HA-CT3-HA | 35.5 | 109.4 | 4.7 | 1.790 | anisole |
| CA-CA-CA | 40.0 | 120.0 | 35.0 | 2.416 | CHARMM |
| CA-CA-HP | 30.0 | 120.0 | 22.0 | 2.153 | CHARMM |
| HA-CT1-CT1 | 34.5 | 110.2 | 20.2 | 2.179 | CHARMM |
| HA-CT1-CT2 | 34.5 | 110.1 | 22.8 | 2.179 | CHARMM |
| HA-CT2-CT1 | 33.4 | 110.1 | 22.4 | 2.179 | CHARMM |
| HA-CT2-HA | 33.5 | 108.8 | 5.4 | 1.802 | CHARMM |
| CT1-CT1-CT2 | 68.4 | 113.5 | 11.2 | 2.561 | CHARMM |
| OH1-CT1-CT2 | 75.7 | 104.9 | | | CHARMM |
| OH1-CT2-CT1 | 75.7 | 110.5 | | | CHARMM |
| OH1-CT1-CT1 | 75.7 | 108.4 | | | CHARMM |
| OH1-CA-CA | 45.2 | 120.0 | | | CHARMM |
| H-OH1-CA | 65.0 | 108.5 | | | CHARMM |
| OH1-CT1-HA | 45.9 | 108.2 | | | CHARMM |
| OH1-CT2-HA | 45.9 | 109.4 | | | CHARMM |
| H-OH1-CT1 | 57.5 | 105.7 | | | CHARMM |
| H-OH1-CT2 | 57.5 | 105.0 | | | CHARMM |
| OET-CT1-CT1 | 75.7 | 110.1 | | | CHARMM |
| OET-CT1-CT2 | 75.7 | 110.1 | | | CHARMM |
| OET-CT1-HA | 59.3 | 109.9 | | | CHARMM |
| CT1-OET-CA | 91.2 | 116.8 | | | CHARMM |

Also noted are the model compounds the parameters were derived from, CHARMM indicating that the parameters were not changed from the existing force field.

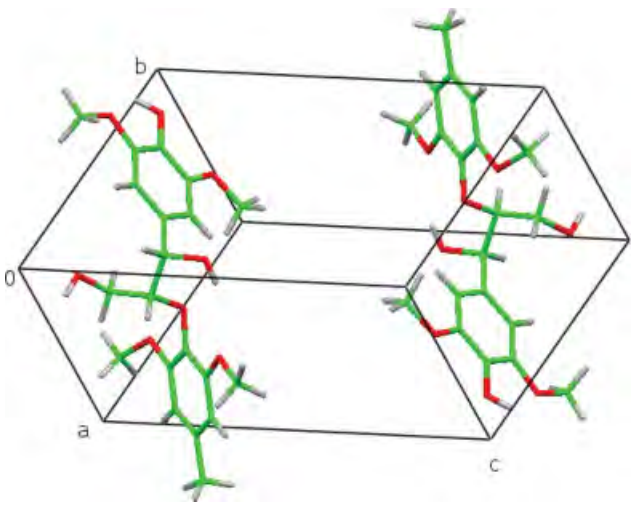


**Figure 12.** Crystallographic cell containing two EPD molecules, with the unit cell axes also shown. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

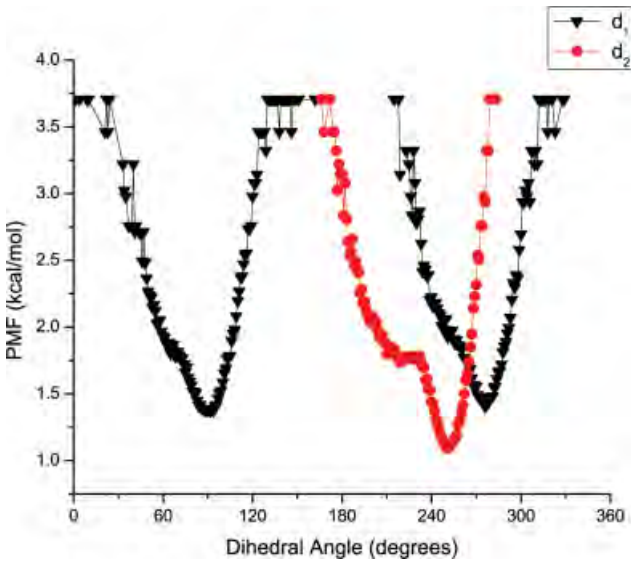


**Figure 13.** Potential of mean force [eq. (3)] calculation for dihedral angles $d_1 = C_5\text{-}C_4\text{-}O\text{-}C'_8$ and $d_2 = C_4\text{-}O\text{-}C'_8\text{-}C'_7$. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

To mimic as closely as possible the conditions under which the experiment was run, the MD simulation was performed for $4 \times 4 \times 4$ unit cells (128 dimers) using periodic boundary conditions while keeping the temperature and pressure fixed at their experimental values. The unit cell dimensions were allowed to vary during the simulation and their time averages are shown in Table 8.

The MD unit cell dimensions were found to be close to the experimental values and the system remained triclinic, see Table 8. The unit cell underwent a moderate expansion, with a 5% increase in volume. After aligning the MD coordinates with the experimental structure, the root means square deviation (RMSD) between the experimental and calculated structure was found to be $0.173 \pm 0.033$ Å.

It is also of particular importance that the current force field models well the $\beta$-$O$-$4'$ linkage that plays an important role in the conformation of the lignin macromolecule. For this reason the time average of the two dihedrals, $d_1$ and $d_2$, that define the $\beta$-$O$-$4'$ linkage was compared with the experimental crystal values, see Table 9. The two dihedrals are (numbering scheme in Fig. 2d): $d_1 = C_5\text{-}C_4\text{-}O\text{-}C'_8$ and $d_2 = C_4\text{-}O\text{-}C'_8\text{-}C'_7$, with the prime indicating an atom on

**Table 8.** Unit Cell Properties of Small-Molecule-Dimer for Experimental Crystal Structure and from Molecular Dynamics Simulation.

| Cell dimension | Experiment | MD |
|---|---|---|
| A (Å) | 8.69 | $8.73 \pm 0.02$ |
| B (Å) | 8.90 | $8.93 \pm 0.01$ |
| C (Å) | 13.11 | $13.68 \pm 0.03$ |
| $\alpha$ (deg) | 73.85 | $74.48 \pm 0.05$ |
| $\beta$ (deg) | 86.15 | $86.30 \pm 0.01$ |
| $\gamma$ (deg) | 83.06 | $83.06 \pm 0.02$ |
| Cell volume (Å$^3$) | 966 | 1020 |

**Table 9.** Dihedrals Defining the $\beta$-*O*-4 Linkage $d_1 = C_5$-$C_4$-*O*-$C'_8$ and $d_2 = C_4$-*O*-$C'_8$-$C'_7$, see Figure 2d.

| Dihedral | Experiment | MD |
|---|---|---|
| $d_1$ (deg) | 80.0 | $77.9 \pm 6.3$ |
| $d_2$ (deg) | $-152.8$ | $-148.5 \pm 5.5$ |

the second syringyl unit of the dimer. As with previous results, the simulation results are in accord with experiment.

To further probe the behavior of the two dihedrals ($d_1$ and $d_2$) a MD simulation was performed of a single EDP molecule in the gas phase. The Potential of Mean Force (PMF) was calculated as a function of the two dihedrals. The PMF is given by eq. (2):

$$w(\theta) = -k_B T \log P(\theta), \tag{3}$$

where $\theta$ is the dihedral angle in question and $P(\theta)$ the probability of finding the system at $\theta$. $P(\theta)$ was derived from the MD trajectories and $w(\theta)$ is shown in Figure 13.

Dihedral $d_1$ (related to $\omega_1$, Fig. 4) has two minima at 90 and 270 degrees with a barrier of the order of 2.3 kcal/mol. These data are consistent with Figure 4. Dihedral $d_2$ (related to $\omega_2$, Fig. 5) has one minimum at $-110$ degrees and a barrier greater than 2.6 kcal/mol. Because of the higher $d_2$ barrier the simulation sampled a limited range of dihedral angles, and its periodicity cannot be determined from these data. Dihedral $d_2$ has higher barrier than $d_1$ in accordance with Figures 4 and 5. The crystal environment seems to have a greater influence on $d_2$, where there is a 40 degrees difference in the equilibrium values between gas and crystal phases. The above points further validate the use of the dihedral constants derived from the model compound anisole.

## Discussion

This work presents a molecular mechanics force field for lignin, which is compatible with the CHARMM potential energy function. The parameterization was based on reproducing quantum-mechanically derived target data. Special care was taken to correctly describe the most common lignin linkage: the $\beta$-O-4′ bond. The partial atomic charge of the oxygen and carbon atoms participating in the linkage were derived by examining interactions between a lignin fragment model compound and a water molecule. Dihedral parameters were obtained by reproducing QM adiabatic rotational potential energy profiles, with emphasis placed on reproducing accurately the thermally sampled low energy regions. The remaining bond and angle parameters were derived using the AFMM method. To test the validity of the force field a simulation of a lignin-dimer crystal was performed. The overall good agreement between the structural properties of the MD run and the experiment provide confidence that the force field will be useful in simulation of lignocellulosic biomass.

There are indications that lignin composition and structure influence the biodegradability of biomass stocks. For example, it has been found that hardwood, which contains more syringyl units, is less recalcitrant than softwood, which is made up mainly of guaiacyl.[43] Additionally, the less recalcitrant secondary cell wall is known to have more linear lignin with mostly $\beta$-O-4′ linkages, whereas the more recalcitrant primary wall has more *C*-*C* linkages and more highly branched lignin.[44] From the earlier discussions, it is plausible that examining different lignocellulose models with varying lignin composition and structure might shed light on the varying degrees of hydrolyzability of different biomass stocks.

However, the accurate computer simulation of lignin in lignocellulose will present significant challenges. Unlike many biological macromolecules that have been studied with molecular simulation, both the chemical and three-dimensional structures of lignin are relatively poorly researched. However, the present force field provides a basis for constructing molecular models of lignin systems, and, in combination with a range of biophysical measurements both existing and underway, significant progress in determining structures of lignocellulosic biomass systems can be expected in the near future.

## Acknowledgments

LP thank T. Splettstoesser for useful discussions on the usage of NAMD. NWChem calculations were run on the National Center for Computational Sciences SGI Altix supercomputer and NAMD calculations on the Oak Ridge National Laboratory Center for Molecular Biophysics 92-core AMD MolDyn cluster.

## References

1. Cosgrove, D. J. Nat Rev Mol Cell Biol 2005, 6, 850.
2. Lynd, L. R.; Laser, M. S.; Brandsby, D.; Dale, B. E.; Davison, B.; Hamilton, R.; Himmel, M.; Keller, M.; McMillan, J. D.; Sheehan, J.; Wyman, C. E. Nat Biotechnol 2008, 26, 169.
3. Himmel, M. E.; Ding, S. Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. Science 2007, 315, 804.
4. 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.; Mielenz, J. R.; Murphy, R.; Templer, R. Tschaplinski, T. Science 2006, 311, 484.
5. Gray, K. A.; Zhao, L. S.; Emptage, M. Curr Opin Chem Biol 2006, 10, 141.
6. Zhang, Y. H. P.; Ding, S. Y.; Mielenz, J. R.; Cui, J. B.; Elander, R. T.; Laser, M.; Himmel, M. E.; McMillan, J. R.; Lynd, L. R. Biotechnol Bioeng 2007, 97, 214.
7. Fan, L. T.; Lee, Y. H.; Beardmore, D. R. Biotech Bioeng 1981, 23, 419.
8. Berlin, A.; Gilkes, N.; Kurabi, A.; Bura, R.; Tu, M. B.; Kilburn, D.; Saddler, J. Appl Biochem Biotechnol 2005, 121, 163.
9. Mooney, C. A.; Mansfield, S. D.; Touhy, M. G., Saddler, J. N. Bioresour Technol 1998, 64, 113.
10. Matthews, J. F.; Skopec, C. E.; Mason, P. E.; Zuccato, P.; Torget, R. W.; Sugiyama, J.; Himmel, M. E.; Brady, J. W. Carbohydr Res 2006, 34, 138.
11. Nimlos, M. R.; Matthews, J. F.; Crowley, M. F.; Walker, R. C.; Chukkapalli, G.; Brady, J. V.; Adney, W. S.; Clearyl, J. M.; Zhong, L. H.; Himmel, M. E. Protein Eng Des Sel 2007, 20, 179.
12. Yui, T.; Hayashi, S. Biomacromolecules 2007, 8, 817.
13. Yui, T.; Nishimura, S.; Akiba, S.; Hayashi, S. Carbohydr Res 2006, 341, 2521.
14. Mazeau, K.; Heux, L. J Phys Chem B 2003, 107, 2394.

15. Vietor, R. J.; Mazeau, K.; Lakin, M.; Perez, S. Biopolymers 2000, 54, 342.
16. Besombes, S.; Robert, D.; Utille, J. P.; Taravel, F. R.; Mazeau, K. Holzforschung 2003, 57, 266.
17. Besombes, S.; Mazeau, K. Biopolymers 2004, 73, 301.
18. Besombes, S.; Mazeau, K. Plant Physiol Biochem 2005, 43, 299
19. Besombes, S.; Mazeau, K. Plant Physiol Biochem 2005, 43, 277.
20. Kuttel, M.; Brady, J. W.; Naidoo, K. J. J Comp Chem 2002, 23, 1236.
21. Davin, L. B.; Lewis, N. G Curr Opin Biotechnol 2005, 16, 407.
22. Vaiana, A. C.; Cournia, Z.; Costescu, I. B.; Smith, J. C. Comput Phys Commun 2005, 167, 34.
23. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J Comput Chem 1983, 4, 187.
24. 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. J Phys Chem B 1998, 102, 3586.
25. Vorobyov, I.; Anisimov, V. M.; Greene, S.; Venable, R. M.; Moser, A.; Pastor, R. W.; MacKerell, A. D. J Chem Theory Comput 2007, 3, 1120.
26. Breneman, C. N.; Wiberg, K. B. J Comput Chem 1990, 11, 361.
27. Harrison, R. J.; Nichols, J. A.; Straatsma, T. P.; Dupuis, M.; Bylaska, E. J.; Fann, G. I.; Windus, T. L.; Apra, E.; Anchell, J.; Bernholdt, D.; Borowski, P.; Clark, T.; Clerc, D.; Dachsel, H.; de Jong, B.; Deegan, M.; Dyall, K.; Elwood, D.; Fruchtl, H.; Glendenning, E.; Gutowski, M.; Hess, A.; Jaffe, J.; Johnson, B.; Ju, J.; Kendall, R.; Kobayashi, R.; Kutteh, R.; Lin, Z.; Littlefield, R.; Long, X.; Meng, B.; Nieplocha, J.; Niu, S.; Rosing, M.; Sandrone, G.; Stave, M.; Taylor, H.; Thomas, G.; van Lenthe, J.; Wolinski, K.; Wong, A.; Zhang, Z. A. Computational Chemistry Package for Parallel Computers, Version 4.0.1; Pacific Northwest National Laboratory: Richland, WA, 2001.
28. Kendall, R. A.; Apra, E.; Bernholdt, D. E.; Bylaska, E. J.; Dupuis, M.; Fann, G. I.; Harrison, R. J.; Ju, J. L.; Nichols, J. A.; Nieplocha, J.; Straatsma, T. P.; Windus, T. L.; Wong, A. T. Comput Phys Commun 2000, 128, 260.
29. Moller, C.; Plesset, M. S. Phys Rev 1934, 46, 618.
30. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. JACS 1996, 118, 2309.
31. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. J Chem Phys 1983, 79, 926.
32. Mackerell, A. D.; Karplus, M. J Chem Phys 1991, 95, 10559.
33. Stevens, W. J.; Basch, H.; Krauss, M. J Chem Phys 1984, 81, 6026.
34. Scott, A. P.; Radom, L. J Phys Chem 1996, 100, 16502.
35. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. J Comput Chem 2005, 26, 1781.
36. Langer, V.; Lundquist, K.; Miksche, G. E. Acta Crystallogr Sect E: Struct Rep Online 2005, 61, O1001.
37. Brunger, A. T. X-PLOR, The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, 1992
38. Mackerell, A. D. J Comput Chem 2004, 25, 1584.
39. Chen, I. J.; Yin, D. X.; MacKerell, A. D. J Comput Chem 2002, 23, 199.
40. Humphrey, W.; Dalke, A.; Schulten, K. J Mol Graphics 1996, 14, 33.
41. Cournia, Z.; Smith, J. C.; Ullmann, G. M J Comput Chem 2005, 26, 1383.
42. Vaiana, A. C.; Schulz, A.; Wolfrum, J.; Sauer, M.; Smith, J. C. J Comput Chem 2003, 24, 632.
43. Sesteghalian, A. R.; Srivastava, V.; Gilkes, N.; Gregg, D. J.; Saddler N. J. Glycosyl Hydrolases in Biomass Conversion, ACS; Symposium Series, No. 769, ACS; 2001, p. 100.
44. Jung, H. G.; Deetz D.A. Forage cell wall structure and digestibility, ASA-CSSA-SSSA; 1993, p. 315.