



CHARMM force field parameters for 2'-hydroxybiphenyl-2-sulfinate, 2-hydroxybiphenyl, and related analogs

Yue Yu^a, Ishan A. Fursule^a, Landon C. Mills^a, Derek L. Englert^b, Brad J. Berron^a, Christina M. Payne^{a,*}

^a Department of Chemical and Materials Engineering, University of Kentucky, 177 F. Paul Anderson Tower, Lexington, KY 40506, USA

^b Department of Chemical and Materials Engineering, University of Kentucky, 4810 Alben Barkley Drive, Paducah, KY 42002, USA

ARTICLE INFO

Article history:

Received 27 October 2016

Received in revised form 6 December 2016

Accepted 7 December 2016

Available online 15 December 2016

Keywords:

Biodesulfurization

Dibenzothiophene

Parameterization

Petroleum refining

Organosulfur compound

ABSTRACT

2'-Hydroxybiphenyl-2-sulfinate (HBPS) desulfinase (DszB) catalyzes the cleavage of the carbon-sulfur bond from HBPS in the final step of microbial 4S pathway desulfurization reactions. DszB is notable for its substrate specificity and exhibits product inhibition, both of which hinder the overall 4S pathway turnover rate. To understand the molecular-level contributions to substrate and inhibitor binding to DszB, we plan to perform molecular dynamic simulations bound to an array of naphthenic molecules and biphenyl analogues of HBPS. However, many of the small molecules we are interested in are not included in standard force field packages, and thus, we must first produce accurate molecular mechanics force fields. Here, we develop and validate CHARMM-compatible force field parameters for the HBPS substrate, the 2-hydroxybiphenyl product, and potential inhibitors including: 2,2'-biphenol, 2-biphenyl carboxylic acid, 1,8-naphthosultam, and 1,8-naphthosultone. The selected molecules represent biphenyl compounds having both a single and double functional group and the planar naphthenic molecule class, all likely present in the oil-rich environment surrounding DszB-producing microorganisms. The Force Field Toolkit (ffTK) in VMD was used to optimize charge, bond distance, angle, and dihedral parameters. Optimized geometries were determined from quantum mechanical calculations. Molecular simulations of the molecules in explicit and implicit water solutions were conducted to assess the abilities of optimized parameters to recapitulate optimized geometries. Calculated infrared (IR) spectra were obtained and compared with experimental IR spectra for validation of the optimized MM parameters.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Currently, a majority of the world's transportation fuels are derived from sulfur-containing crude oil. The concentration of sulfur largely depends on the region of origin. Sulfur remaining in finished fuels, post-refining, is converted to sulfur oxides (SO_x) upon combustion, which are then emitted as a toxic pollutant. Accordingly, sulfur concentration in refined petroleum products is strictly regulated at extremely low levels to minimize SO_x emissions [1]. Hydrosulfurization (HDS) is the traditional industrial approach to liberating the majority of sulfur containing compounds from crude oil; this process occurs in the presence of metallic catalysts and hydrogen gas at high pressures and temperatures. Unfortunately, recalcitrant sulfur heterocycles, such as dibenzothiophene (DBT), often remain intact after HDS treatment [2–7].

Biodesulfurization is an enzymatic process by which organosulfur compounds recalcitrant to HDS can be freed of sulfur at ambient conditions. Naturally, this is a promising complementary process for implementation downstream of traditional HDS units [2,3,8]. However, improving the overall rate of enzymes that desulfurize thiophenic molecules remains a primary technical challenge preventing economical implementation of biodesulfurization [9,10].

There are two primary microbial DBT desulfurization pathways [11]: (1) one in which initial attack is directed at one of the ring carbon atoms (the Kodama pathway) [12], and (2) a sulfur-specific pathway resulting in carbon-sulfur bond cleavage without ring opening, which maintains fuel heating value (4S pathway, Fig. 1) [13,14]. In the 4S pathway, which is most suitable for industrial application, two monooxygenases (DszA and DszC), a desulfinase (DszB), and an oxidoreductase (DszD) convert DBT to 2-hydroxybiphenyl (HBP). In the first and second steps, DszC with DszD oxidize DBT to DBT-sulfone (DBTO₂) via a DBT-sulfoxide (DBTO) intermediate. DszA and DszD catalyze the conversion of DBT sulfone to 2'-hydroxybiphenyl-2-sulfinate (HBPS). In the final

* Corresponding author.

E-mail address: christy.payne@uky.edu (C.M. Payne).

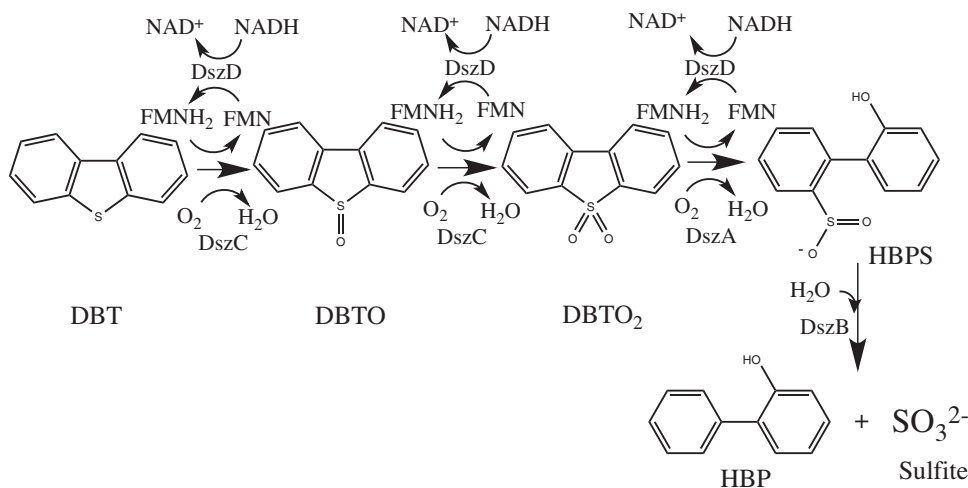


Fig. 1. Schematic of 4S pathway for the biodesulfurization of dibenzothiophene (DBT) [18]. The first three steps of the pathway are catalyzed by flavin-dependent monooxygenases, DszA and DszC. The flavin reductase, DszD, uses nicotinamide adenine dinucleotide (NADH to NAD⁺) to reduce flavin mononucleotide (FMN to FMNH₂) [13,19,20]. Dibenzothiophene (DBT) is converted to DBT-sulfoxide (DBTO), DBT-sulfone (DBTO₂), and then to 2'-hydroxybiphenyl-2-sulfinate (HBPS). In the final step, DszB cleaves the carbon-sulfur bond of the HBPS intermediate to produce 2-hydroxybiphenyl (HBP) and sulfite (SO₃²⁻). This last step is the rate-limiting step of the entire 4S pathway [18,15].

step, DszB cleaves the carbon-sulfur bond of HBPS, resulting in HBP and sulfite (SO₃²⁻). The HBP product is reincorporated in the oil fraction, and sulfite is removed in the aqueous fraction. This final catalytic step has been widely identified as rate-limiting, making DszB the most logical target for activity improvements [8,15]. Product inhibition also affects DszB conversion rates, wherein increasing concentrations of product, and possibly HBP derivatives, inhibit turnover [16,17]. Despite this knowledge, we have relatively little mechanistic understanding of DszB activity, substrate binding, or product inhibition, making rational protein design efforts challenging.

We anticipate molecular simulation is capable of addressing many questions underlying poor substrate turnover and product inhibition. Prior to that, however, we must first develop force field parameters for a subset of chemically relevant aromatic molecules with which to complete these studies. The molecules we parameterize here (Fig. 2) have been selected based upon an experimental study by Watkins et al. [15], where the authors established DszB rate and inhibition constants for several functionalized biphenyl compounds and planar naphthenic molecules that exist in petroleum. In addition to the HBPS substrate and the HBP product, we examine 2,2'-biphenol (BIPH), 2-biphenyl carboxylic acid (BCA), 1,8-naphthosultam (NTAM), and 1,8-naphthosultone (NAPO). These molecules represent a variety of functional groups that are anticipated to mediate different molecular-level interactions. Analogs with a single functional group attached, such as 2-biphenyl carboxylic acid, were consistently found to be non-inhibitory at concentrations up to 100 μM. Within the two functional group class, the most inhibitory of the studied analogs was 2,2'-biphenol ($K_I = 17 \mu\text{M}$). In the planar naphthenic class, the most strongly inhibiting compound was 1,8-naphthosultam ($K_I = 1.8 \mu\text{M}$). On the other hand, 1,8-naphthosultone was not inhibitory [15]. With accurate force-field parameters for these molecules, we will be capable of developing molecular models to describe substrate binding and inhibition phenomena in DszB.

The CHARMM all-atom additive force field is a widely accepted parameter set for biomolecular simulation [21–23]; it is also easily expanded upon given the availability of the compatible CHARMM General Force Field (CGenFF) parameter set for organic, 'drug-like' molecules and force field fitting tools [24]. Accordingly, we build upon the CHARMM parameter set, ideal for modeling DszB, by using CGenFF nomenclature and parameter analogies as a starting point

for biphenyl and planar naphthenic analog parameter generation. As a result, all the conventions and recommendations for usage of the biomolecular CHARMM force field and CGenFF apply to the newly generated parameter set [24].

Both CHARMM and CGenFF use the Class I potential energy function, which is the sum of bonded and nonbonded contributions (Eqs. (1) and (2)) [22–24]. The bonded term of the potential energy function, U_{bonded} , includes contributions from bonds (b), angles (θ), dihedrals (ϕ), improper dihedrals (ϕ), and a Urey–Bradley 1,3-term ($r_{1,3}$). Equilibrium values for these components are denoted with a subscript of 0 (i.e., b_0 , θ_0 , ϕ_0 and $r_{(1,3)_0}$). In the dihedral term, n and δ are the dihedral multiplicity and phase, respectively. Each term has a corresponding force constant denoted by K . The non-bonded potential, $U_{\text{nonbonded}}$, includes electrostatic and van der Waals (vdW) terms. In the electrostatics term, q_i and q_j are the partial atomic charges of atom i and j , respectively, r_{ij} is the distance between atoms i and j , and D is the dielectric constant. The van der Waals contribution is described as a Lennard-Jones 6–12 function, where ϵ_{ij} is the potential energy well depth and $R_{\text{min},ij}$ is the arithmetic mean of the Lennard-Jones radii of atoms i and j [22].

Intramolecular (bonded terms):

$$U_{\text{bonded}} = \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_\phi(1 + \cos(n\phi - \delta)) + \sum_{\text{improper dihedral}} K_\phi(\phi - \phi_0)^2 + \sum_{\text{Urey-Bradley}} K_{UB}(r_{1,3} - r_{(1,3)_0})^2 \quad (1)$$

Intermolecular (nonbonded terms):

$$U_{\text{nonbonded}} = \sum_{\text{nonbonded}} \frac{q_i q_j}{4\pi\epsilon D r_{ij}} + \epsilon_{ij} \left[\left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^6 \right] \quad (2)$$

Single functional biphenyl Double functional biphenyl Planar naphthenic

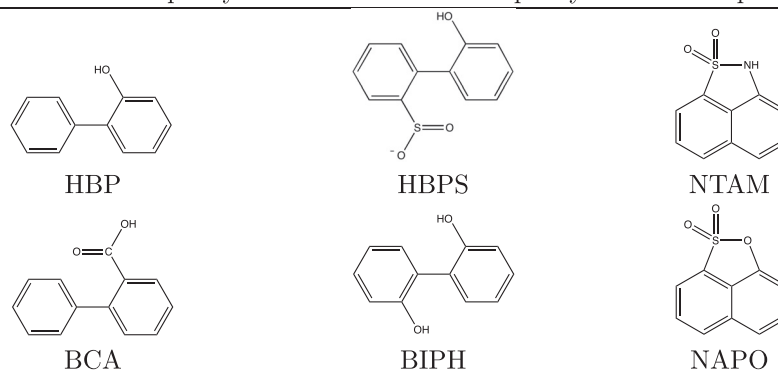


Fig. 2. Functionalized biphenyls and naphthenic molecules parameterized in this study. 2'-hydroxybiphenyl-2-sulfinate (HBPS) is the native DszB substrate, and 2-hydroxybiphenyl (HBP) is the product. The remaining four molecules represent potential inhibitors of DszB. Two molecules from each class have been selected. From the single functional group biphenyls, we parameterize HBP and 2-biphenyl carboxylic acid (BCA). HBPS and 2,2'-biphenol (BIPH) represent the double functional group biphenyls, and 1,8-naphthosultam (NTAM) and 1,8-naphthosultone (NAPO) represent planar naphthenic substrate/inhibitors.

Here, we report parameters for the six molecules shown in Fig. 2 that are compatible with CHARMM and CGenFF potential energy functions. Initial parameters were obtained by analogy with molecules available in CGenFF version 1.0.0 with force field version 3.0.1. Using Visual Molecular Dynamics (VMD) version 1.9.2 and the Force Field Tool Kit (ffTK) plugin [25,26], we optimized the missing CGenFF parameters, iterating until molecular mechanics (MM) simulations were capable of reproducing quantum mechanical (QM)-derived target data. We have optimized bond, angle, and dihedral parameters for each of the molecules. ffTK does not currently provide support for optimization of improper dihedrals or the Urey-Bradley 1,3-term [26], but for aromatic compounds, these terms are not likely to significantly contribute to the potential energy given their relatively stiff torsion potentials and cyclic constraints. The optimized parameters were validated by comparing calculated MM geometries to the optimized molecular geometries and through recapitulation of experimental infrared (IR) spectra.

2. Methods

2.1. Parameterization

The parameterization process used in this study follows the general ffTK procedure outlined by Mayne et al. [26]. We discuss each of the steps, including the file types generated and use of external software and webservers, in the sections that follow. The inherently iterative force field optimization procedure is illustrated as a flow chart in Fig. 3. Briefly, the procedure entails (1) developing an initial parameter set based on analogy with existing CGenFF parameters, (2) optimizing the geometry of the molecule, (3) optimizing partial atomic charges against QM-derived water interaction profiles, (4) optimizing bonds and angles against the QM-derived Hessian matrix, and (5) optimizing the dihedrals against QM-derived potential energy surfaces (PES). This procedure is repeated, returning to the charge optimization step, as many times as necessary to obtain convergence. Here, we defined our convergence tolerance as partial atomic charge differences of less than 0.01 between the previous parameter set and the current set. As mentioned above, we did not optimize the Urey-Bradley term or the improper dihedral term, as ffTK does not currently support optimization of these terms. Improper dihedral terms were not employed in the selected molecules with the exception of BCA (C13-C1-O2-O1), the parameters of which were well represented by initial estimates. Additionally, chemical intuition suggests the cyclic and planar nature of our molecules will result in only minor contributions from these terms to the bonded potential energy term. Nonbonded

parameters were obtained from the CGenFF parameter set and were not optimized, per standard protocol.

2.1.1. Initialize parameters by analogy with CGenFF

The initial molecular structure was generated using the Avogadro molecular editor (v.1.1.1) [27], resulting in .mol2 and .pdb formatted files containing coordinates and bond orders, in the case of the former. The .mol2 formatted file was then used as input to the ParamChem server (<https://cgenff.paramchem.org>), which assigned atom types, charges, and parameters by analogy to existing molecules in the CGenFF parameter set [24,28,29]. From ParamChem, we obtained a stream file (.str) containing the CHARMM/CGenFF-compatible topology and bonded parameters for the input molecule. We have provided the initial .str files obtained from ParamChem in .str format files. ParamChem assigned a 'penalty value' to each generated parameter, where parameters with penalties should be validated.

Starting from the .pdb from Avogadro, VMD Molefactory was used to assign CGenFF atom types in an updated .pdb file. The initial .str file from ParamChem was used, in Molefactory, to initialize charges and generate the .psf file that required for ffTK optimization. Finally, nonbonded parameters were assigned by analogy with CGenFF using ffTK and were combined with the initial bonded parameters from the .str file in a .par file. The .par file, with both nonbonded and bonded parameters, was used in bond, angle, and dihedral optimization and was updated with new parameters following each step.

2.1.2. Geometry optimization

The optimized geometry for each molecule in Fig. 2 was determined from quantum mechanical calculation using Gaussian09 [30]. The geometry optimization started with a moderate-size basis set, MP2/6-31G*, with polarization functions applied to only heavy atoms; this initial optimization served as an input for a more extensive optimization using the MP2/6-31++G** basis set, adding polarization functions and diffuse functions to all atoms. This dual minimization approach was necessary to reach a global minimum in the case of sulfur-containing molecules. For consistency, the approach was applied in all geometry optimizations.

2.1.3. Charge optimization

Partial atomic charges were optimized in ffTK on the basis of calculated water-interaction profiles, consistent with existing CHARMM force field parameters [24,22]. For each hydrogen bond donor or acceptor atom, a water molecule was inserted within hydrogen bonding distance. The ffTK algorithm optimized the

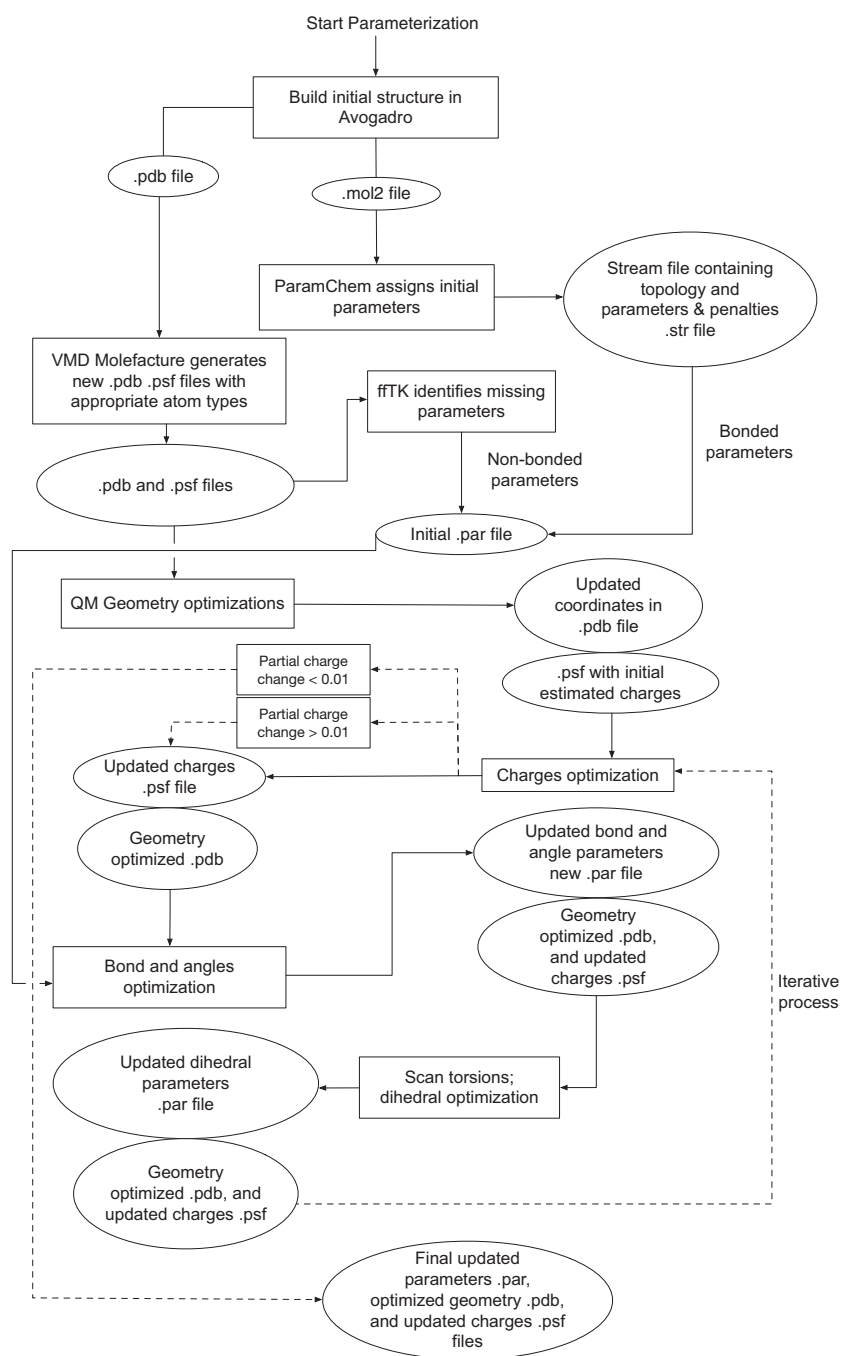


Fig. 3. Flow chart illustrating the ffTK parameterization process used to obtain MM force fields for each molecule in Fig. 2. Rectangular boxes represent a step in the workflow. Ovals are used to describe the set of files generated from the preceding step. Text below or next to the lines between steps provides additional contextual information. Dashed lines represent the iterative decision point in the workflow, where convergence of the partial charges results in a final set of MM parameters.

placement of water molecules such that they were automatically oriented for hydrogen bonding and steric repulsion between water molecules was minimized [26]. The generated water-interaction Gaussian input files were visually inspected prior to QM calculation to prevent over constraining orientation of the water molecules. In cases where molecules contained aromatic rings (i.e., where aromatic carbons are sp^2 hybridized), aromatic carbons were assigned as both hydrogen bond donors and acceptors. The distance between water molecules and the target atoms and the rotation angle of water molecules were optimized quantum mechanically at the HF/6-31G* level of theory. HF/6-31G* was used in this step to maintain consistency with the CHARMM force field [26].

The output from QM calculations was imported into ffTK for MM force field fitting. For every water–atom interaction, a MM interaction energy as a function of distance was calculated using the provided parameters. Partial atomic charges were varied as the deviation between the QM and MM minimum distances between the atoms, interaction energies, and dipole moments were minimized [26]. Constraints for the upper and lower charge boundaries were manually adjusted according to the assigned ParamChem penalty, where smaller penalties were more strictly constrained. In each charge optimization step, the default ffTK advanced setting parameters were applied for all molecules except for HBPS, which is a negatively charged molecule. For HBPS, the interaction

energy scaling factor was manually changed from 1.16 to 1 to better approximate the bulk phase [24]. The convergence tolerance of each calculation was set to 0.005. An updated .psf file was generated at this point.

2.1.4. Bond and angle optimization

When available, experimental vibrational spectra should be used for bond and angle parameter optimization. However, this information was not available for any of the compounds in Fig. 2; thus, bond and angle parameters were optimized against the QM Hessian matrix calculated at the MP2/6-31G^{*} level of theory. From this matrix of second derivatives of energy with respect to coordinates, a QM PES was determined for comparison with the MM-derived PES, using the input parameter set. The equilibrium bond and angle values and the force constants, K_b and K_θ , changed as the deviations in the QM and MM energies were minimized. A new .par file, with optimized bond and angle parameters, was generated following optimization.

2.1.5. Dihedral optimization

Dihedral parameters were the last to be optimized, as they critically depend on charges and bond and angle parameters. Only dihedrals with ParamChem penalties were optimized (Table 1). Planar naphthenic ring structures were not well represented among existing CGenFF molecules; therefore, many more dihedrals were optimized in these two molecules than the other four. A MM PES was calculated using the updated parameter set, and dihedral parameters were optimized against a QM calculated PES, allowing the force constant, K_ϕ , of each dihedral to change. The periodicity, n , and phase shift angle, δ , were set according to initial parameters from ParamChem and were not varied in the optimization [26].

The QM PES were generated in Gaussian09. Each dihedral angle was scanned bidirectionally from equilibrium value in 5° increments at the MP2/6-31G^{*} level of theory. The resulting QM log files, containing data from the equilibrium value to the final scanned dihedral, were used as input to fTK for the optimization.

The QM and MM PES results were visually inspected using a VMD plotting utility to determine if further refinement was necessary based on the Root Mean Square Error (RMSE) value and general fit of the curves. All optimizations started with the simulated annealing protocol [31]. Once the general shape of the MM

Table 1
Dihedrals of each molecule included in the optimization process. The dihedrals were scanned 90° in either direction about the equilibrium value in 5° increments with the exception of HBPS; for HBPS, the dihedral angles marked with an asterisk were scanned 40° about the equilibrium value. Atom numbers correspond to the convention illustrated in Fig. 4

Scan torsions			
HBP	HBPS	BCA	BIPH
O C1 C2 C3	O1 C8 C7 C12	O2 C13 C1 C2	O1 C8 C7 C12
O C1 C2 C7	O1 C8 C7 C2	O1 C13 C1 C2	O1 C8 C7 C2
C2 C1 O H	S C3 C2 C1*	C13 C1 C2 C3	C7 C8 O1 H
	S C3 C2 C7*	C13 C1 C2 C7	
	C7 C8 O1 H*		
	C2 C3 S O2*		
NTAM	NAPO		
C1 C2 C3 S	H C2 C3 S	C3 C2 C1 S	H C2 C1 S
C2 C3 C4 C5	C9 C4 C3 S	C2 C1 C10 C9	C5 C10 C1 S
C3 C4 C5 N	S C4 C3 S	C1 C10 C9 O3	C9 C10 C1 S
C6 C5 N S	C4 C5 N S	C8 C9 O3 S	C10 C9 O3 S
C2 C3 S N	C2 C3 S O2	C2 C1 S O2	C2 C1 S O3
C4 C3 S N	C4 C3 S O2	C10 C1 S O2	C10 C1 S O3
C5 N S C3	C5 N S O2	C9 O3 S C1	C9 O3 S O2
H N S C3	H N S O2		

PES matched that of the QM PES, the optimization method was switched to the downhill simplex method and the PES cutoff was decreased from 10 kcal/mol to 5 kcal/mol, further enhancing convergence [32,33]. The optimization of the dihedral parameters was considered converged when the RMSE was less than 0.5 and the minimum energy difference of the QM and MM PES basins was less than 0.2 kcal/mol. After the optimization converged, the .par parameter file was again updated.

All the parameters including partial charges, bonded, and non-bonded parameters influence MM behavior. If the optimization is incomplete, changing any bonded parameter will influence partial charges and vice versa. Thus, the parameterization process is inherently iterative. Following dihedral optimization, we returned to the charge optimization step (Fig. 3), using the new .par file to optimize the partial charges. When the partial charges did not change by more than 0.01, the parameterization process was considered complete. At completion, the parameterization process yielded a .par file with the optimized parameters, a .pdb file with the optimized geometry, and a .psf file with the molecular topology and charges.

2.2. Validation of optimized parameters

2.2.1. Explicit solvent molecular dynamics simulations

Molecular dynamics simulation of the explicitly-solvated molecules were conducted using the optimized parameter sets for comparison to the optimized geometries. Each molecule was solvated in a cubic periodic cell of TIP3P water using the Multiscale Modeling Tools for Structural Biology (MMTSB) Tool Set [34]; the buffer between the molecule and the edge of the periodic cell was 13 Å. For HBPS, a charged molecule, one sodium ion was included to neutralize the overall system charges for application of long-range electrostatic methods. In CHARMM, the system was minimized for 2000 steps using the steepest descent algorithm followed by 5000 steps using the Adopted-basis Newton–Raphson algorithm. The minimized system was then heated from 90 K to 300 K in 25 K increments. This was followed by isothermal-isobaric (NPT) simulation for 100 ps at 300 K and 1.01325 bar to equilibrate the system density. The density-equilibrated systems were then simulated for 100 ns at 300 K in the NVT ensemble in NAMD [35].

2.2.2. Preparation of HBPS and experimental IR spectra

Except for HBPS, IR spectra for all molecules in Fig. 2 are available in the literature. The experimental HBP, BIPH, and BCA spectra were obtained from the NIST Chemistry WebBook [36]. NTAM and NAPO spectra were from BIORAD (BIORAD: Bio-Rad/Sadtler IR Data Collection obtained from Bio-Rad Laboratories, Philadelphia, PA (US)). The HBPS IR spectrum was experimentally determined. HBPS was synthesized by suspending 20 mM biphenosultine in 44 mM NaOH at room temperature overnight on a Fisher Scientific Reliable Tube Rotator. The biphenosultine was a gift from the Texas A&M University LINCHPIN Laboratory. The resulting aqueous HBPS was lyophilized at −35 °C and 60 mT in a Benchtop Freezedryer (SP Scientific VirTis Advantage Plus XL 85). Lyophilized HBPS was analyzed via Attenuated Total Reflectance (ATR) FTIR using a Varian Inc. 7000e Spectrometer. The lyophilized sample was placed on the diamond ATR crystal, and the spectrum was collected for wavelengths between 700 and 4000 cm^{−1} and averaged over 32 scans.

2.2.3. Calculated IR spectra

The optimized parameters were validated by comparing IR spectra calculated from simulation with experimental spectra. The IR Spectra Density calculator in VMD was used to compute spectra from simulation. Implicit solvent MD simulations were used as input to the IR calculations, as to reduce statistical error that may arise as a result of incomplete sampling in the presence of explicit

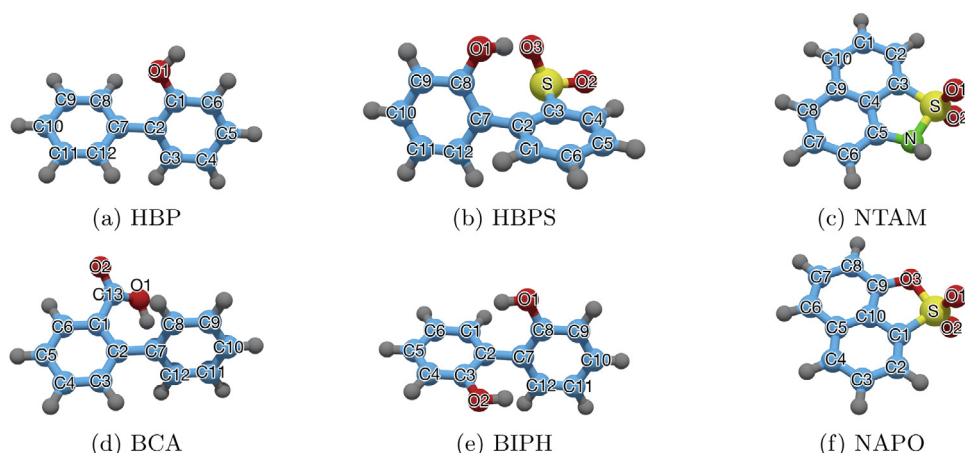


Fig. 4. Optimized geometry for each molecule obtained using Gaussian09 with the MP2/6-31++ G** level of theory. The atom labeling convention used in the parameterization process for each molecule is shown in black lettering over the atoms. Note, the atom names/numbers do not necessarily correspond to IUPAC naming conventions.

solvent over relatively short simulations [37]. As with explicit solvent MD, the implicit solvent MD simulations were conducted in NAMD for each molecule. The Generalized Born implicit solvent (GBIS) model was used to obtain 20 ps of data following a 2000 steps minimization. The hydrogen bonds were not constrained to obtain vibrational frequency data. The simulation timestep was 1 fs and the Born radius cutoff was set at 13 Å.

The VMD IR spectra density calculator computes the spectra from simulation trajectories by determining the dipole moment of all atoms in the molecule and calculating the auto-correlated Fourier transform. The default parameters in the IR Spectra Density Calculator were used with the exception of the time step and the maximum frequency, which were set to 1 fs and 4000 cm⁻¹, respectively.

3. Results and discussion

3.1. Geometry optimization

The optimized geometries of each of the six molecules is shown in Fig. 4. The molecules can be divided into two groups: biphenyl derivatives (HBPS, HBP, BCA, and BIPH) and naphthalene

derivatives (NTAM and NAPO). The aromatic rings of unsubstituted biphenyl in the solution and vapor phases are not coplanar and exhibit a characteristic twist about the central dihedral (e.g., dihedral C3-C2-C7-C8 of Fig. 4a), which experiment suggests is 44.4(1.2)° [38]. Substitution, particularly at the *ortho*-position, introduces steric effects that lead to an increase in this angle. As expected, QM geometry optimization gives a dihedral angle of 53.5°, 59.2°, and 59.4° for HBP, HBPS, and BCA, respectively. For HBP, experimental X-ray data suggests this values should be 54.3(8)° [39]. Unfortunately, experimental structures of HBPS and BCA are not available for comparison, but the proximity to the dihedral angle of *ortho*-substituted 2-hydroxybiphenyl suggests the calculated values are reasonable. The larger functional groups of HBPS and BCA introduce more steric effects on the biphenyl rings, slightly increasing the dihedral angle; this is also observed in the difunctionalized BIPH molecule. BIPH is a symmetric molecule, with hydroxyl groups in the *trans* position in the hydrated form. The X-ray structure of the 2,2'-biphenol monohydrate reveals the dihedral angle to be 67.6(1)° [40]. QM geometry optimization at the MP2/6-31++G** level of theory resulted in a dihedral angle of 70.7°, in relative agreement with experiment.

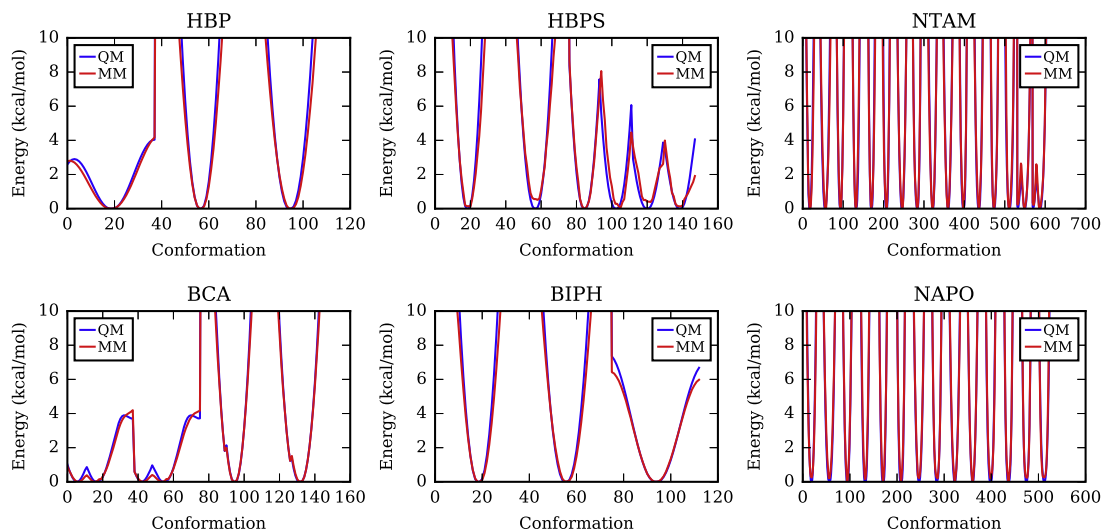


Fig. 5. Torsional profiles of HBP, HBPS, NTAM, BCA, BIPH, and NAPO, as labeled at the top of each plot. The QM PES (blue) and the MM PES (red) are shown for each of the scanned dihedrals (Table 1). As aromatic rings are relatively well described by existing molecules in CGenFF, the dihedrals scanned largely represent those of the functional groups, the connections of the functional group with the aromatic ring, and bridging aromatic carbons. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

As with several of the biphenyl derivatives, structural information for the two naphthalene derivatives is not available in the literature. However, the base naphthalene structure has been shown to be planar, and thus, we can expect that calculations will yield relatively planar geometry overall [41]. Some strain in the aromatic rings, as a result of the *peri* interactions, is expected to manifest in slight variations in bond lengths compared to unsubstituted naphthalene [42]. We did, in fact, observe this phenomenon in our geometry optimization, where the distance between C3 and C5 atoms was slightly shortened by 0.17 Å compared to the distance between C8 and C10 in NTAM; similarly the distance between C1 and C9 was slightly shortened by 0.19 Å compared to C4 and C6 in NAPO. Overall, the naphthenic molecules maintained planar geometries, with slight protrusion of the *peri* substitutions from the plane. The optimized geometry files (in .pdb format) have been provided in .pdb formatted files.

3.2. Optimized parameters

Charge, bond, angle, and dihedral optimization, conducted as described above, resulted in a unique set of parameters for each molecule. The optimized charges (Table 2) and bond and angle parameters (Table 3) were manually inspected to ensure optimization did not result in values or force constants enforcing physically unrealistic conformations. The dihedral parameters (Table 4) were visually inspected as described above. For the optimized dihedrals, i.e., those with unreasonable penalties assigned by ParamChem, the MM-derived PES was fit to the QM PES (Fig. 5). Good agreement with QM was achieved in all six cases. The MM-energy minima positionally aligned with the QM PES minima and were within 0.2 kcal/mol. In the sections that follow, we will describe how these parameters correspond to optimized geometries and the ability of the force constants to maintain these bond distances and angles.

3.3. Comparing QM and MM equilibrium bond and angle values

To determine whether the force field parameters were capable of sufficiently describing the optimized geometries, and thus the strength of the assigned force constants, we conducted 100-ns MD simulations of each molecule in explicit solvent; the newly optimized parameters were used as input to the simulations. From these simulations, we evaluated how the parameterized bonds and angles changed with respect to time. The data were binned into histograms for comparison with the value obtained from QM geometry optimization (Fig. 6). In general, we expect the MM-sampled distribution will be centered around the QM value, representing a good fit.

For each parameterized angle, the values sampled over the course of the 100-ns MD simulations exhibited a Gaussian distribution, which encompassed the QM-determined equilibrium value. For most, the distribution maximum aligned well with the QM equilibrium angle. In cases where the maximum was not identical to the QM-optimized geometry, for example the C13–C1–O2 angle of BCA (Fig. 6c), competing constraints on the surrounding geometry made it difficult to enforce the optimum angle without application of large force constants. Despite not ideally replicating the optimized angle in every case, the QM optimized angle values all lie within the 95% confidence interval of the Gaussian distribution, suggesting the optimized parameters sufficiently describe molecular geometries.

3.4. Comparison of calculated and experimental IR spectra

To validate the optimized parameters, we calculated IR spectra for each of the molecules from implicit solvent MD trajectories obtained using the parameters. The calculated spectra were directly compared with experimental spectra (Fig. 7). For all

Table 2

Optimized partial atomic charge parameters for each molecule. Atom names are as shown in Fig. 4. The atom type corresponds to the CGenFF naming convention. Atomic charges shown here include those with CGenFF-assigned penalties, which we have optimized according to the described procedure. For the complete set of atomic charges, refer to the supplemental material.

Atoms	Atom type	Charges
(a) HBP		
C1	CG2R61	0.130
C2	CG2R67	0.243
C3	CG2R61	−0.265
C7	CG2R67	−0.0223
O	OG311	−0.678
H6	HGP1	0.472
(b) HBPS		
C1	CG2R61	−0.627
C2	CG2R67	−0.0123
C3	CG2R61	0.321
C7	CG2R67	0.495
C8	CG2R61	−0.0813
C12	CG2R61	0.0227
O1	OG311	−0.599
H9	HGP1	0.298
S	SG302	0.353
O2	OG2P1	−0.701
O3	OG2P1	−0.701
(c) NTAM		
C1	CG2R61	−0.278
C2	CG2R61	−0.102
C3	CG2RC0	0.176
C4	CG2RC0	0.281
C5	CG2RC0	0.242
C6	CG2R61	−0.354
C9	CG2R61	−0.0748
H2	HGR61	0.244
S	SG302	−0.148
O1	OG2P1	−0.223
O2	OG2P1	−0.223
N	NG3C51	−0.338
H7	HGP1	0.378
(d) BCA		
C1	CG2R61	0.142
C2	CG2R67	−0.215
C3	CG2R61	−0.0363
C7	CG2R67	0.113
C12	CG2R61	−0.113
O1	OG311	−0.586
O2	OG2D1	−0.499
(e) BIPH		
C1	CG2R61	−0.155
C2	CG2R67	−0.00360
C3	CG2R61	0.212
C7	CG2R67	−0.00360
C8	CG2R61	0.212
C12	CG2R61	−0.155
O1	OG311	−0.547
H8	HGP1	0.374
O2	OG311	−0.547
H10	HGP1	0.374
(f) NAPO		
C1	CG2RC0	0.373
C2	CG2R61	−0.329
C3	CG2R61	−0.192
C5	CG2R61	−0.0514
C8	CG2R61	−0.449
C9	CG2RC0	0.399
C10	CG2RC0	0.103
H1	HGR61	0.305
S	SG302	−0.154
O1	OG2P1	−0.167
O2	OG2P1	−0.167
O3	OG3C51	−0.202

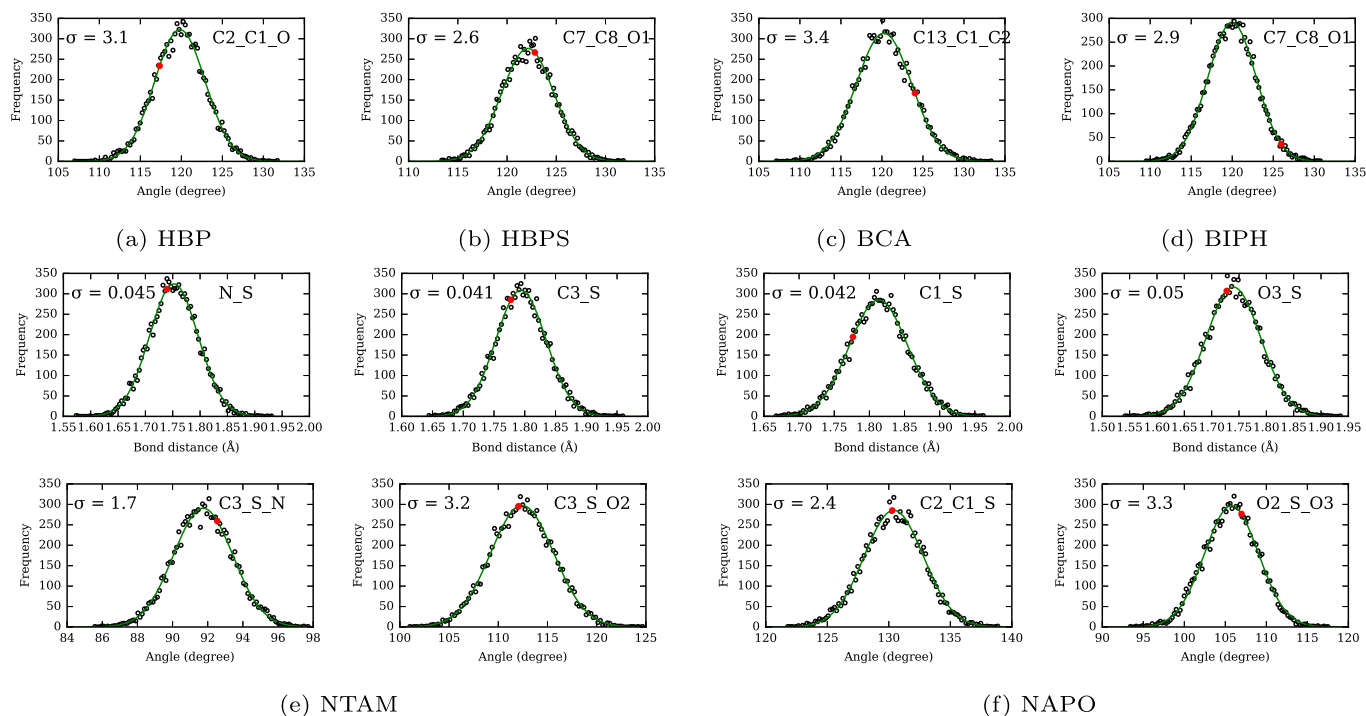


Fig. 6. Comparison of the QM-optimized bonds and angles to the bond distances and angles sampled over the course of a 100-ns MD simulation for (a) HBP, (b) HBPS, (c) BCA, (d) BIPH, (e) NTAM, and (f) NAPO. The parameterized bonds and angles were measured in each frame of the MD simulation; the measurements were binned into histograms (unfilled blue circles) and fit with a Gaussian distribution (green line). The σ given in the top left of each plot is the standard deviation of the histogram fit to a Gaussian distribution. The x-axes represent either the value of the bond distance (Å) or the angle (°) and the y-axes represent frequency of a measurement, as sampled over a 100-ns MD simulation. The red filled circles represent the QM optimum value for a given bond length or angle. For the complete set of optimized parameters comparison, please refer to the supplemental material. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

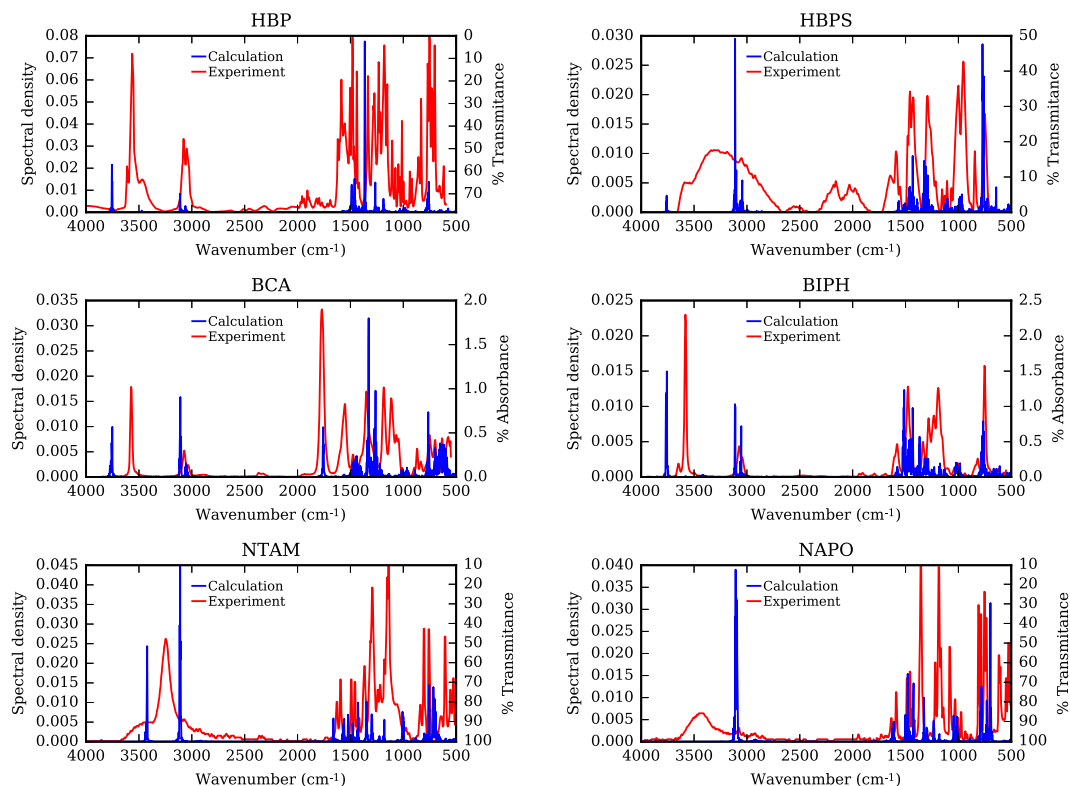


Fig. 7. Experimental and theoretical IR spectra for HBP, HBPS, BCA, BIPH, NTAM, and NAPO. The HBPS spectrum was obtained as described in Section 2. The left and right-hand y-axes represent spectral density (computational) and % absorbance/transmittance (experimental), respectively. The magnitude of the peaks has no physical meaning. Rather, the existence of a peak at an appropriate wavenumber, representative of a vibrational phenomenon, is the primary point of comparison.

Table 3

Optimized bond and angle parameters for each molecule. Bonds and angles have been optimized as described here. For the complete set of parameters, please refer to supplemental material.

Angle parameter	K_θ (kcal/mol/rad ²)	θ_0 (°)
(a) HBP		
C2 C1 O1	43.330	110.545
(b) BCA		
C1 C2 C7	36.487	114.086
(c) HBPS		
C7 C8 O1	85.005	118.771
C2 C3 S	15.123	112.738
(d) BIPH		
C7 C8 O1	55.730	115.855
(e) NTAM		
Bond Parameters	K_b (kcal/mol/Å ²)	b_0 (Å)
C3 S	134.889	1.843
N S	126.251	1.765
Angle parameters	K_θ (kcal/mol/rad ²)	θ_0 (°)
C2 C3 S	81.319	125.877
C3 C4 C5	94.337	114.328
C4 C3 S	136.049	113.105
C5 N S	73.526	110.111
S N H	70.303	110.249
C3 S N	199.775	96.373
C3 S O2	105.177	114.081
N S O2	69.221	111.276
(f) NAPO		
Bond parameters	K_b (kcal/mol/Å ²)	b_0 (Å)
C1 S	135.870	1.829
O3 S	88.596	1.765
Angle parameters	K_θ (kcal/mol/rad ²)	θ_0 (°)
C2 C1 S	61.166	121.527
C1 C10 C9	112.940	112.353
C10 C1 S	140.678	108.541
C9 O3 S	62.850	115.994
C1 S O2	75.040	96.537
C1 S O3	212.043	94.171
O2 S O3	101.025	98.555

molecules studied, calculated and experimental spectra were in good agreement in low wavenumber ranges (<2000 cm⁻¹), which were attributed to stretching modes between heavy atoms and C–H bending modes (Table 5). This agreement largely supports vibrational similarity between the parameterized model and the heavy atom structure of these compounds.

However, for the higher frequencies (>2800 cm⁻¹), where C–H and O–H bond stretching appears, the calculated frequencies do not align well with experimental peaks, varying by up to 150 cm⁻¹. The largest mismatches are for the O–H stretching modes. The position of O–H stretching modes are highly sensitive to solvent interactions. In particular, intramolecular and intermolecular hydrogen bonding in protic solvents give distinct shifts. While calculations were completed with an implicit protic solvent, the available experimental data were obtained using aprotic solvents (HBP), solid state (HBPS), and the gas phase (BCA, BIPH); solvation state data for NTAM and NAPO experimental spectra were unavailable. The aromatic C–H stretching calculations were in good agreement (typically <50 cm⁻¹ higher) with the experimental data. These inaccuracies at higher frequencies also arise from anharmonicity effects

Table 4

Optimized dihedral parameters for each molecule. The atom types are given in Table 2. Dihedrals have been optimized as described in Section 2.

Dihedrals parameters	K_ϕ (kcal/mol)	n	δ
(a) HBP			
O C1 C2 C3	2.992	2	180
O C1 C2 C7	3.000	2	180
C2 C1 O H	0.746	2	180
(b) BCA			
O2 C13 C1 C2	1.934	2	180
O1 C13 C1 C2	0.122	2	0
C13 C1 C2 C3	2.727	2	180
C13 C1 C2 C7	2.484	2	180
(c) HBPS			
S C3 C2 C1	1.010	2	180
O1 C8 C7 C12	3.000	2	180
C2 C3 S O2	1.976	2	0
C2 C3 S O2	0.751	4	180
S C3 C2 C7	2.240	2	180
O1 C8 C7 C2	0.549	2	180
C7 C8 O1 H	0.708	2	180
C7 C8 O1 H	0.888	4	0
(d) BIPH			
O1 C8 C7 C12	3.000	2	180
O1 C8 C7 C2	3.000	2	180
C7 C8 O1 H	1.372	2	180
(e) NTAM			
C2 C3 C4 C5	2.970	2	180
C5 N S O2	1.728	3	0
C5 C4 C3 S	0.684	2	0
C4 C5 N S	2.997	3	180
C4 C3 S O2	0.673	6	180
C3 C4 C5 N	0.851	2	0
C2 C3 S N	2.475	2	0
H N S O2	1.056	3	0
H N S C3	1.160	1	0
H N S C3	0.742	2	0
H N S C3	0.662	3	180
C1 C2 C3 S	2.999	2	180
C4 C3 S N	1.485	2	0
C5 N S C3	2.574	1	0
C5 N S C3	1.494	2	180
H C2 C3 S	1.508	2	180
C9 C4 C3 S	0.054	2	180
C6 C5 N S	0.711	2	0
C2 C3 S O2	0.053	6	180
(f) NAPO			
C2 C1 C10 C9	0.847	2	0
C9 C10 C1 S	1.607	2	180
C10 C1 S O2	0.301	6	180
C10 C1 S O3	0.638	2	0
C8 C9 O3 S	1.871	2	180
C3 C2 C1 S	2.935	2	180
C9 C3 S O2	1.926	3	0
C10 C9 O3 S	2.603	2	180
C9 O3 S C1	0.810	3	0
H C2 C1 S	2.340	2	180
C1 C10 C9 O3	2.254	2	180
C5 C10 C1 S	2.207	2	180
C2 C1 S O2	0.013	6	180
C2 C1 S O3	2.103	2	180

resulting from the manner in which theoretical spectra were calculated [43,44]. In the spectra calculations, the system was modeled as harmonic. In the low frequency ranges, the harmonic model is capable of accurately predicting IR spectra by modeling bond stretches as a mass-spring system, satisfying experimental agreement. However, the harmonic mass-spring model begins to break down at higher frequencies, as the relationship between force and displacement is nonlinear. An anharmonic oscillator can be used to represent regions not well described by a harmonic model, but this also nonideal. The anharmonic nonlinear region is not easily calculated. Thus, we applied the harmonic model to calculate

Table 5

Experimental IR spectra frequency values compared with theoretical prediction using MM-computed parameters.

Assignment	Experimental frequency (cm ⁻¹)	Calculated frequency (cm ⁻¹)
HBP		
Aromatic; sp ² ; bending C–H	750	758
Aromatic C–O	1182	1187
Aromatic; stretching C=C	1480	1364
Biphenyl; stretching C–C	1590	1489
Aromatic; sp ² ; overtones C–H	1660–2000	–
Aromatic; sp ² ; stretching C–H	3049 & 3077	3060 & 3110
O–H	3561 & 3611	3473 & 3755
HBPS		
Aromatic; sp ² ; bending C–H	756	772
Aromatic C–O	1002	982
C–S	1296	1290
Stretching S=O	1462	1464
Aromatic; stretching C=C	1460	1432
Biphenyl; stretching C–C	1593	1517
Artifact from diamond ATR surface	1786–2372 ^a	–
Aromatic; sp ² ; stretching C–H	3051	3045 & 3111
O–H	2670–3650	3760
BCA		
Aromatic to Acyl; orthosubstituted C–C	694	627
Aromatic; sp ² ; bending C–H	750	763
Acid; C–O	1346	1262
Aromatic; stretching C=C	1482	1327
Biphenyl; stretching C–C	1550	1447
C=O	1770	1757
Aromatic; sp ² ; stretching C–H	3034 & 3070	3046 & 3178
O–H	3572	3673
BIPH		
Aromatic; sp ² ; bending C–H	750	765
Aromatic C–O	1194	1180
Aromatic; stretching C=C	1474	1432
Biphenyl; stretching C–C	1586	1512
Aromatic; sp ² ; stretching C–H	3046 & 3074	3053 & 3110
O–H	3578	3757
NTAM		
Bending SO ₂	602	705
Stretching C–S	755	755
Stretching S–N	812	826
Ring breathing	1056	998
Symmetric Stretching SO ₂	1142	1177
Stretching C–N	1289	1299
Asymmetric Stretching SO ₂	1363	1347
Aromatic; stretching C=C	1488	1427
Bending N–H	1590	1564
Stretching N–H	3244	3422
Aromatic; sp ² ; stretching C–H	3082	3110
NAPO		
Bending SO ₂	520 & 620	697
Stretching C–S	755	732
Stretching S–O	812	782
Stretching C–O	977	1013
Ring breathing	1034	1045
Symmetric stretching SO ₂	1187	1209
Asymmetric stretching SO ₂	1357	1331
Aromatic; stretching C=C	1485	1502
Aromatic; sp ² ; stretching C–H	3077	3106
Unassigned sultone	3444 (broad peak) [45,46]	–

^a The peaks at 1980, 2130, and 2160 are artifacts from the diamond ATR surface.

the high frequency region, resulting in the energy being coupled to the fundamental vibration frequency, and thus, overestimating wavenumbers.

Table 5 gives a side-by-side comparison of the vibrational phenomena observed in the experimental and calculated IR spectra. We discuss only the low frequency range phenomena given the

limitations of theoretical spectra calculations described above. Within the low frequency range, the biphenyl compounds all exhibit aromatic and biphenyl carbon-carbon stretching and aromatic carbon-hydrogen stretching. In all compounds, carbon-hydrogen bond stretching in the aromatic rings is reasonably well predicted by calculation. Carbon-carbon bond stretching, both aromatic and biphenyl, is perhaps the least accurate vibrational phenomena in the low frequency region. As the bond and angle parameters primarily defining aromatic carbon-carbon stretching have not been modified from the CGenFF parameters, this small discrepancy is likely common to all aromatic molecules modeled with the CGenFF force field. Otherwise, the biphenyl compound parameters largely reproduce the vibrational phenomena observed in the experimental spectra. The planar naphthenic parameters are capable of reproducing the low frequency ring-based, carbon-sulfur vibrational phenomena, and SO₂ asymmetric and symmetric stretching and bending. Overall, the optimized parameter set adequately predicts primary vibrational phenomena in classical molecular simulation.

4. Conclusions

Here, we have determined CHARMM-compatible molecular mechanics force fields for HBP, HBPS, BIPH, BCA, NTAM, and NAPO using the fTK optimization tools available in VMD. The force field parameters are provided in the Supplementary Data as individual files, including a .psf file with the partial atomic charges and the bond, angle, and dihedral parameters in a .par file for each molecule. Explicit solvent molecular simulation confirmed the ability of the implemented parameter force constants to maintain the optimized geometry equilibrium values within the 95% confidence interval. The optimized parameters were used to calculate theoretical IR spectra, which were then compared to experimental spectra from literature or generated as part of this study. In general, the theoretical spectra were in good agreement with experiment at lower frequencies, but calculated spectra were not particularly accurate in the high frequency ranges as a result of both solvent effects and anharmonicity not captured by harmonic models. Overall, the parameters sufficiently capture the equilibrium geometry and vibrational phenomena of a subset of biphenyl and planar naphthenic molecules that naturally occur in unrefined crude oil. Thus, these parameters can be implemented in biomolecular simulations of 4S pathway enzymes such as DszB to understand the molecular-level interactions driving substrate binding and product inhibitions.

Acknowledgments

The computational resources for this work were provided by the Extreme Science and Engineering Discovery Environment (XSEDE) [47], which is supported by National Science Foundation grant number ACI-1053575. Additional computational resources for testing and data analysis were made available by the Center for Computational Sciences DLX cluster at the University of Kentucky. Acknowledgement is made to the Donors of the American Chemical Society Petroleum Research Fund for support of this research (53861-DNI4). We also thank Mingyang Sun for his technical support in graphical analysis.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jmglm.2016.12.005>.

References

- [1] M. Carrales, R.W. Martin, Sulfur Content of Crude Oils, US Department of the Interior, Bureau of Mines, 1975.
- [2] B.L. McFarland, Biodesulfurization, *Curr. Opin. Microbiol.* 2 (3) (1999) 257–264.
- [3] T. Kabe, A. Ishihara, H. Tajima, Hydrodesulfurization of sulfur-containing polyaromatic compounds in light oil, *Ind. Eng. Chem. Res.* 31 (6) (1992) 1577–1580.
- [4] P. Leprince, BST Petroleum Refining. Conversion Processes, Vol. 3, Editions Technip, 2001.
- [5] R. Javadli, A. de Klerk, Desulfurization of heavy oil, *Appl. Petrochem. Res.* 1 (1–4) (2012) 3–19.
- [6] M.S. Rana, V. Samano, J. Ancheyta, J. Diaz, A review of recent advances on process technologies for upgrading of heavy oils and residua, *Fuel* 86 (9) (2007) 1216–1231.
- [7] T. Swaty, Global refining industry trends: the present and future, *Hydrocarbon Process.* 84 (9) (2005) 35.
- [8] M.R. Gray, A.R. Ayasse, E.W. Chan, M. Veljkovic, Kinetics of hydrodesulfurization of thiophenic and sulfide sulfur in Athabasca bitumen, *Energy Fuels* 9 (3) (1995) 500–506.
- [9] J.J. Kilbane, Microbial biocatalyst developments to upgrade fossil fuels, *Curr. Opin. Biotechnol.* 17 (3) (2006) 305–314.
- [10] S. Le Borgne, R. Quintero, Biotechnological processes for the refining of petroleum, *Fuel Process. Technol.* 81 (2) (2003) 155–169.
- [11] T. Ohshiro, Y. Izumi, Microbial desulfurization of organic sulfur compounds in petroleum, *Biosci. Biotechnol. Biochem.* 63 (1) (1999) 1–9.
- [12] K. Kodama, K. Umehara, K. Shimizu, S. Nakatani, Y. Minoda, K. Yamada, Identification of microbial products from dibenzothiophene and its proposed oxidation pathway, *Agric. Biol. Chem.* 37 (1) (1973) 45–50.
- [13] J.R. Gallagher, E.S. Olson, D.C. Stanley, Microbial Desulfurization of Dibenzothiophene: A Sulfur-Specific Pathway, 1993, February.
- [14] C. Oldfield, O. Pogrebinsky, J. Simmonds, E.S. Olson, C.F. Kulpa, Elucidation of the metabolic pathway for dibenzothiophene desulfurization by *Rhodococcus* sp. strain IGTS8 (ATCC 53968), *Microbiology* 143 (9) (1997) 2961–2973.
- [15] L.M. Watkins, R. Rodriguez, D. Schneider, R. Broderick, M. Cruz, R. Chambers, E. Ruckman, M. Cody, G.T. Mrachko, Purification and characterization of the aromatic desulfinase, 2-(2'-hydroxyphenyl)benzenesulfinate desulfinase, *Arch. Biochem. Biophys.* 415 (1) (2003) 14–23.
- [16] A. Caro, K. Boltes, P. Leton, E. Garcia-Calvo, Description of by-product inhibition effects on biodesulfurization of dibenzothiophene in biphasic media, *Biodegradation* 19 (4) (2008) 599–611.
- [17] N. Nakayama, T. Matsubara, T. Ohshiro, Y. Moroto, Y. Kawata, K. Koizumi, Y. Hirakawa, M. Suzuki, K. Maruhashi, Y. Izumi, R. Kurane, A novel enzyme, 2-hydroxybiphenyl-2-sulfinate desulfinase (DszB), from a dibenzothiophene-desulfurizing bacterium *Rhodococcus erythropolis* KA2-5-1: gene overexpression and enzyme characterization, *Biochim. Biophys. Acta* 1598 (12) (2002) 122–130.
- [18] K.A. Gray, O.S. Pogrebinsky, G.T. Mrachko, L. Xi, D.J. Monticello, C.H. Squires, Molecular mechanisms of biocatalytic desulfurization of fossil fuels, *Nat. Biotechnol.* 14 (13) (1996) 1705–1709.
- [19] B.R. Folsom, D.R. Schieche, P.M. DiGrazia, J. Werner, S. Palmer, Microbial desulfurization of alkylated dibenzothiophenes from a hydrodesulfurized middle distillate by *Rhodococcus erythropolis* I-19, *Appl. Environ. Microbiol.* 65 (11) (1999) 4967–4972.
- [20] K.A. Gray, G.T. Mrachko, C.H. Squires, Biodesulfurization of fossil fuels, *Curr. Opin. Microbiol.* 6 (3) (2003) 229–235.
- [21] B.R. Brooks, C.L. Brooks, A.D. Mackerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodosek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus, CHARMM: the biomolecular simulation program, *J. Comput. Chem.* 30 (10) (2009) 1545–1614.
- [22] A.D. Mackerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B* 102 (18) (1998) 3586–3616.
- [23] A.D. Mackerell, Empirical force fields for biological macromolecules: overview and issues, *J. Comput. Chem.* 25 (13) (2004) 1584–1604.
- [24] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A.D. Mackerell, CHARMM General Force Field (CGenFF): a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, *J. Comput. Chem.* 31 (4) (2010) 671–690.
- [25] W. Humphrey, A. Dalke, K. Schulten, VMD – Visual molecular dynamics, *J. Mol. Graph.* 14 (1996) 33–38.
- [26] C.G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid, J.C. Gumbart, Rapid parameterization of small molecules using the Force Field Toolkit, *J. Comput. Chem.* 34 (32) (2013) 2757–2770.
- [27] M. ElGamacy, L. Van Meervelt, A fast topological analysis algorithm for large-scale similarity evaluations of ligands and binding pockets, *J. Cheminform.* 7 (1) (2015) 42.
- [28] K. Vanommeslaeghe, A.D. Mackerell Jr., Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing, *J. Chem. Inf. Model.* 52 (12) (2012) 3144–3154.
- [29] W. Yu, X. He, K. Vanommeslaeghe, A.D. Mackerell, Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations, *J. Comput. Chem.* 33 (31) (2012) 2451–2468.
- [30] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, A. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian09 Revision D. 01, Gaussian Inc., Wallingford, CT, 2009.
- [31] O. Guvench, A.J. Mackerell, Automated conformational energy fitting for force-field development, *J. Mol. Model.* 14 (8) (2008) 667–679.
- [32] J.A. Nelder, R. Mead, A simplex method for function minimization, *Comput. J.* 7 (4) (1965) 308–313.
- [33] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes, 1990.
- [34] M. Feig, J. Karanicolas, C.L. Brooks Iii, MMTSB tool set: enhanced sampling and multiscale modeling methods for applications in structural biology, *J. Mol. Graph. Model.* 22 (5) (2004) 377–395.
- [35] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, *J. Comput. Chem.* 26 (16) (2005) 1781–1802.
- [36] S.E. Stein, Infrared Spectra in **NIST Chemistry WebBook, NIST Standard Reference Database Number 69**, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899, <http://webbook.nist.gov>, (retrieved October 26, 2016).
- [37] D.E. Tanner, K.-Y. Chan, J.C. Phillips, K. Schulten, Parallel generalized born implicit solvent calculations with NAMD, *J. Chem. Theory Comput.* 7 (11) (2011) 3635–3642.
- [38] A. Almenningen, O. Bastiansen, L. Fernholt, B.N. Cyvin, S.J. Cyvin, S. Samdal, Structure and barrier of internal rotation of biphenyl derivatives in the gaseous state: Part 1. the molecular structure and normal coordinate analysis of normal biphenyl and pedesterated biphenyl, *J. Mol. Struct.* 128 (1) (1985) 59–76.
- [39] M. Perrin, K. Bekkouch, A. Thozet, Structure of 2-hydroxybiphenyl, *Acta Crystallogr. Sect. C: Cryst. Struct. Commun.* 43 (5) (1987) 980–982.
- [40] X. Chen, G. Luo, M. Tong, Z. Zhou, 2,2'-biphenol monohydrate, *Acta Crystallogr. Sect. C: Cryst. Struct. Commun.* 52 (7) (1996) 1727–1729.
- [41] D. Cruickshank, A detailed refinement of the crystal and molecular structure of naphthalene, *Acta Crystallogr.* 10 (8) (1957) 504–508.
- [42] V. Balasubramanian, Peri interaction in naphthalene derivatives, *Chem. Rev.* 66 (6) (1966) 567–641.
- [43] M.C. Holthausen, W. Koch, Chemist's Guide to Density Functional Theory, 2nd ed., Wiley-VCH, New York, 2001.
- [44] T.I. Morrow, E.J. Maginn, Molecular dynamics study of the ionic liquid 1-n-butyl-3-methylimidazolium hexafluorophosphate, *J. Phys. Chem. B* 106 (49) (2002) 12807–12813.
- [45] H. Li, J. Zhou, Synthesis of 1,3-propane sultone, *J. Chem. Sci. Technol.* 2 (2013) 209–212.
- [46] SDBSWeb, <http://sdb.sdb.aist.go.jp> (National Institute of Advanced Industrial Science and Technology), (SDBS No.:5331).
- [47] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G.D. Peterson, et al., XSEDE: accelerating scientific discovery, *Comput. Sci. Eng.* 16 (5) (2014) 62–74.