JCTC Journal of Chemical Theory and Computation

Modeling Aromatic Liquids: Toluene, Phenol, and Pyridine

Christopher M. Baker[†,‡] and Guy H. Grant*[,‡]

Department of Chemistry, Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford, OX1 3QZ, United Kingdom, and Unilever Centre for Molecular Informatics, The University Chemical Laboratory, Lensfield Road, Cambridge, CB2 1EW, United Kingdom

Received July 4, 2006

Abstract: Aromatic groups are now acknowledged to play an important role in many systems of interest. However, existing molecular mechanics methods provide a poor representation of these groups. In a previous paper, we have shown that the molecular mechanics treatment of benzene can be improved by the incorporation of an explicit representation of the aromatic π electrons. Here, we develop this concept further, developing charge-separation models for toluene, phenol, and pyridine. Monte Carlo simulations are used to parametrize the models, via the reproduction of experimental thermodynamic data, and our models are shown to outperform an existing atom-centered model. The models are then used to make predictions about the structures of the liquids at the molecular level and are tested further through their application to the modeling of gas-phase dimers and cation$-\pi$ interactions.

Introduction

Interactions involving aromatic groups are now recognized as playing a significant role in many important biomolecular processes. Aromatic−aromatic interactions are thought to be important in protein−ligand interactions,[1] DNA structure,[2] and protein structure.[3,4] Computational study of such phenomena therefore requires potential models that can accurately treat interactions involving aromatic groups. Unfortunately, conventional all-atom force fields provide a rather poor description of these interactions. The aromatic groups possess a substantial amount of charge density associated with the aromatic π system that lies above and below the plane of the ring. This charge distribution gives rise to a quadrupole moment perpendicular to the plane of the ring, which cannot be adequately reproduced using a simple atom-centered model. A variety of approaches have been attempted to remedy this problem, including central multipoles,[5] distributed multipoles,[5] and off-atom charges, as suggested by Hunter and Sanders.[6] In a recent paper,[7] the charge-separation model of Hunter and Sanders was combined with an atom-centered model to develop a new force field, based on the OPLSAA force field[8] and named OPLS-CS, for the study of liquid benzene. This new model offered an improved representation of the structure of the liquid relative to the existing all-atom model.

The work on liquid benzene demonstrated the utility of the methodology, but the ultimate aim of this research is to establish a new force field for the modeling of aromatic interactions in biological systems. For such a model to be successful, it is necessary that it be able to treat a variety of aromatic groups, with a variety of different substitution patterns. Here, the method is extended to develop new force field parameters for the benzene derivatives toluene, phenol, and pyridine and to gain new insights into their properties in the liquid phase. As a test of the transferability of these models, they will also be used to investigate the cation$-\pi$ interaction. Cation$-\pi$ interactions are believed to be of fundamental importance in many biological problems and will provide a first test of whether the OPLS-CS model is likely to be of use in the study of biological problems.

Although they have received comparatively little experimental and theoretical attention relative to that of the

* Corresponding author e-mail: ghg24@cam.ac.uk.
† Physical and Theoretical Chemistry Laboratory.
‡ Unilever Centre for Molecular Informatics.

10.1021/ct600218f CCC: $37.00
Published on Web 02/14/2007

archetypical aromatic molecule, benzene, the molecules toluene, phenol, and pyridine play numerous important roles in chemistry. For example, toluene and phenol have been used as analogues for the aromatic amino acids phenylethylamine[9] and tyrosine,[10,11] respectively, in studies on the role of $\pi-\pi$ interactions in proteins,[9] the role of NH$-\pi$ interactions in drug recognition,[10] and the ability of the aromatic amino acids to form different types of hydrogen bonds.[11] In addition, all three chemicals are used in the manufacture of a wide range of products including aviation fuel, paints and dyes, explosives, insecticides, and drugs.

The study of these substituted benzenes is complicated relative to that of benzene itself by the presence of the substituents, which in all cases affect the steric bulk of the molecule and also give rise to a dipole moment where previously none existed. In phenol and pyridine, the OH and N groups can also participate in hydrogen-bonding interactions, increasing the range of structural properties; it has also been shown that methyl groups can act as hydrogen-bond donors to aromatic rings,[12] which may provide a source of structural stabilization in toluene.

Toluene. As with the study of the benzene dimer,[13] there is some controversy over the exact nature of the global energy minimum for the toluene dimer. Gervasio et al.[14] performed molecular mechanics and ab initio calculations which suggested that no "T-shaped" minimum is observed but that, instead, only two stacked minima exist. The work of Chipot et al.,[9] using the AMBER force field,[15] also came to the conclusion that the stacked structure was preferred over the T-shaped, despite failing to consider any displaced structures. Simulations incorporating quadrupolar interactions have identified the T-shaped dimer as an energy minimum[16] but still concluded that stacked structures lie lower in energy.

Spectroscopic results are also ambiguous, with hole-burning experiments[17] having identified the presence of two different dimer structures. These were attributed, on the basis of Schauer and Bernstein's results,[16] to sandwich-shaped and T-shaped structures. Using optical absorption spectroscopy, Law et al.[18] also identified the presence of two structures, which they assigned, by analogy to the case of the benzene−toluene dimer, to parallel displaced and T-shaped arrangements. The most recent spectroscopic work[19] suggests that there might in fact be only one structure present, with the second signal being due to a hot band. This study also proceeded to explore larger clusters of toluene molecules and concluded that they are built from the same basic structure as the dimer, suggesting that an antiparallel stacked arrangement is the most likely.

From a theoretical viewpoint, it is apparent that so-called sandwich structures, which are disfavored for the benzene dimer,[20] should be more stable in all of the aromatic molecules considered in this study because they allow for the possibility of favorable dipole−dipole interactions which would result from antiparallel stacking. Furthermore, in the benzene dimer, the exact energetic preference is known to depend on a subtle balance between the electrostatic and dispersion interactions.[21] The larger size of toluene and phenol would be expected to increase the dispersion interaction in the stacked arrangements.[22] The most accurate theoretical calculations to date on the toluene dimer, at the CCSD(T) level of theory, indeed predict that a face-to-face arrangement of molecules is considerably more favorable in toluene than in benzene.[22,23] They also predict that binding in the T-shaped structure is slightly less favorable in toluene than in benzene, though it still remains as the energetic minimum by 0.28 kcal mol^{-1} (cf. benzene where the difference is 0.82 kcal mol^{-1}).[23]

In the liquid phase, there is almost as little consensus. There are no available experimental studies on liquid toluene, but molecular dynamics (MD)[24,25] have suggested that the structure in the liquid is largely perpendicular. One set of Monte Carlo[26] calculations concluded that liquid toluene was broadly similar to liquid benzene, "with edge-to-face and offset-stacked arrangements predominating". A second Monte Carlo study, however, concluded that stacking interactions occur between nearest neighbors[27] and that the predominant orientational distribution could be described as being "V-shaped". There was also disagreement from Fioroni and Vogt,[28] who used MD simulation with a reparameterization of the GROMOS96 force field[29] to study liquid toluene, concluding that the structure contains "a prevalence of the parallel stacking between dimers of toluene molecules".

The crystal structures of two phases of solid toluene have been determined by X-ray diffraction,[30,31] and in both cases the nearest neighbor arrangement is found to be perpendicular.

Phenol. The study of phenol has received a relatively small amount of interest compared to toluene, but high-level ab initio calculations[32] reveal the presence of a minimum energy structure for the dimer in which the arrangement of molecules is perpendicular, but with one molecule donating an OH−O bond rather than an OH$-\pi$ bond. This structure was also predicted by Sagarik and Asawakun,[33] using a potential based on the test particle model,[34] and is also in good agreement with experimental results.[35−37] A hydrogen-bonded minimum in which both molecules lie approximately in the same plane has also been predicted,[38] though it is the result of Hartree−Fock (HF) calculations which neglect the important role of dispersion in the interactions of aromatic molecules.[21] As in toluene, stacked dimers of phenol are found to be more favorable than those of benzene.[22]

The perpendicular structure also seems to have translated well into the simulation of the liquid phase. Molecular dynamics simulations[39] have found that the structure is largely perpendicular but that molecules "stack over the OH group". Work in the group of Jorgensen using Monte Carlo simulations[26] focused on compounds that are liquid at 298 K and so did not consider phenol, but in the closely related *m*-cresol, they found that a perpendicular structure is preferred but that there is competition between ring stacking and hydrogen bonding.

These results could be called in to question by the work of Thornton et al.,[40] which suggests that all-atom models are insufficient for the modeling of the interactions of phenol. In their work on the interactions between phenol and a carboxylate group, they found that a model constructed using distributed multipole analysis[41] outperformed the CHARMM[42] force field, suggesting that higher-order multipole moments

are required in order to model accurately even those aromatic compounds that possess a dipole moment.

Experimental work has been performed to determine the distribution of tyrosine residues in proteins[43] and found that, while stacked residues were preferred at short separations, as the inter-residue distances increased a T-shaped arrangement became more favorable. It is important to note here that the distribution of aromatic groups in proteins need not necessarily correspond to the minimum energy arrangement. Overstabilization of a protein structure could lead to adverse effects, and for this reason, the structure may not be that which minimizes all possible interactions. In spite of this, it is instructive to consider the biological distribution of the groups as it provides an important insight into the types of interaction that a force field must be able to successfully treat.

The crystal structure of solid phenol has also been determined by X-ray diffraction,[44] and under ambient conditions, it is characterized by "infinite hydrogen-bonded chains, where the molecules are aligned in approximately threefold helices".[45]

Pyridine. As a heteroaromatic molecule, pyridine might be expected to show somewhat different behavior to either toluene or phenol. The presence of the N atom introduces a dipole moment which might be expected to favor the parallel arrangements of the molecules, but the potential dispersion interaction will be smaller than in either toluene or phenol. Mishra and Sathyamurthy[46] have used the MP2 method to investigate the dimer and trimer of pyridine, finding that in both cases an antiparallel displaced arrangement is the most favorable, with the T-shaped structures lying comparatively high in energy. The authors also found that Hartree−Fock and density functional methods predict the dimer and trimer to be unstable, indicating the importance of accounting for the dispersion interaction in these complexes. Piacenza and Grimme[47] attempted to account for the failings of density functional theory methods by introducing an empirical dispersion correction. With this method, they also investigated the structures of the dimer and trimer of pyridine, finding again that the parallel structures are preferred over the T-shaped structure. This result was attributed to the presence of the dipole moment, which was calculated to contribute "about 1 kcal mol^{-1}" to the binding. The pyridine dimer has also been studied by Megiel et al.,[48] using the Roothaan−Hartree−Fock (RHF) method. They identified five dimer minima, all of which are hydrogen-bonded with the global minimum being a planar structure forming a cyclic hydrogen-bonding pattern. However, the fact that RHF neglects dispersion means that this result should be treated cautiously.

The same work by Megiel et al.[48] went on to consider the structure of the liquid via molecular dynamics calculations with the AMBER force field.[15] The authors found that the liquid is dominated by the formation of dimers but that the planar structure identified as the gas-phase energy minimum is now only infrequently observed with the structure instead dominated by three other dimers having hydrogen bonds between the nitrogen atom of the first molecule and either the β or γ hydrogen of the second molecule. The results of these calculations are supported by experimental data from ^{14}N nuclear magnetic resonance. Sagarik and Spohr[49] have studied liquid pyridine using a test particle model[34] in molecular dynamics and Monte Carlo simulations, also concluding that the structure is dominated by the formation of dimers. Monte Carlo simulations using the OPLS all-atom potential[8] have been used by Jorgensen and McDonald[50] to study liquid pyridine. The OPLS model was found to perform well in terms of reproducing thermodynamic data and for the structure of the liquid to predict a largely antiparallel arrangement of the molecules. A molecular dynamics simulation with a simpler potential function (that of Williams and Weller[51]) was performed by Gamba and Klein,[52] who concluded that such a potential was inadequate for the task and decided that the inclusion of the N lone pair is essential for accurate results.

Unlike the other aromatic molecules that have been considered, the solid-phase behavior of pyridine has been difficult to determine. Until 2002, only one crystal structure had been identified,[53] and this was found to be unusually complex. Anghel et al.[54] then used a model derived from distributed multipole analysis[41] to identify 17 structures that are thermodynamically more stable than the observed structure. This result prompted an experimental search for new polymorphs of pyridine, which discovered one new structure, but it was not one of the 17 predicted structures. Five of these 17 structures can be discounted on the grounds of poor mechanical properties, but the reason why the other 12 have not been observed has not been satisfactorily explained.

From the preceding discussion it is clear that all three of the molecules considered in this work have already received a good deal of attention. It is also clear that this attention has thus far failed to yield any unanimous conclusions on the structure of the liquid phase of toluene, phenol, or pyridine. The need for more accurate model potentials to address these issues, as well as many other issues related to the properties of aromatic groups, is pressing.

In this work, we will present new force field models for each of these molecules based on the charge-separation model of Hunter and Sanders.[6] We will parametrize the models by fitting to the available experimental thermodynamic data and then use the resulting models to make predictions on the structure of the liquids, comparing our results to those obtained using a conventional all-atom force field. We will then go on to consider the possibility of using these same parameters to investigate the role of aromatic groups in biological systems. Unfortunately, little experimental data is available for any of these molecules, and in its absence, the OPLSAA model will be used as a reference source where appropriate. OPLSAA is a potential developed around 10 years ago with the goal of being able to model a variety of situations including the liquid phase, gas phase, and aqueous solution. Our objectives here are different: to develop models for the pure liquids, and as such, a failure to reproduce the thermodynamic results obtained with the OPLSAA model would indicate either inappropriate parameter choice or an inappropriate potential model.

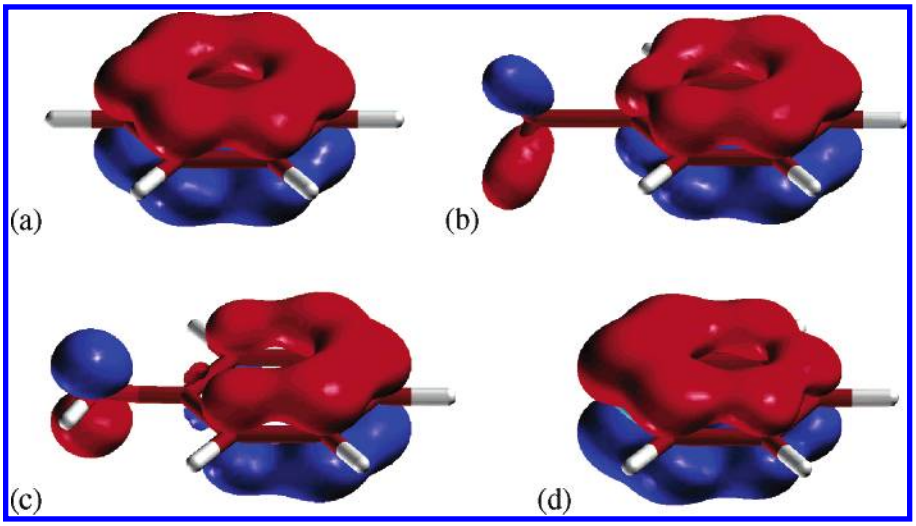

Figure 1. π orbitals for aromatic molecules, calculated at the MP2/6-31+G*//MP2/6-31+G* level of theory. In each case, the lowest-energy orbital in which the aromatic ring lies on the nodal plane is shown. (a) Benzene, (b) toluene, (c) phenol, and (d) pyridine. A contour value of 0.06 is used in all cases.

Cation−π Interactions. Of the many interactions involving aromatic groups that play a role in biology, perhaps the most important is the cation−π interaction, one of the strongest known noncovalent interactions. Energetically, it is on a par with the hydrogen bond,[55] and its role in biology is often just as significant. Cation−π interactions are known to be of importance in determining the structure of small molecules,[56] protein structure,[57,58] protein−protein binding,[59] protein−DNA binding,[60] protein−ligand binding,[61] and protein−lipid interactions.[62]

If a force field aspires to provide an accurate representation of aromatic groups, it is essential that the cation−π interaction be correctly treated. Unfortunately, existing molecular mechanics methodologies are known to provide a rather poor representation of cation−π interactions,[63] the major component of which is electrostatic.[64] As has already been discussed, the electrostatic interactions of aromatic rings are dominated by the quadrupole moment that arises from the aromatic π electrons, which atom-centered force fields fail to account for. As a first step toward the goal of modeling biologically significant aromatic groups, the ability of the OPLS-CS model to treat the cation−π interaction will be investigated.

Many previous studies have employed high levels of theory to identify the preferred binding modes and energetics of cation−π interactions.[65] Unsurprisingly, benzene has been favored, and its interactions with metal ions[66−68] and organic cations[69−72] have been considered. Toluene[73,74] and phenol[75,76] have also been studied in this context, and much is known about the nature of these interactions. The objective of this work is not to gain further insight into the interactions themselves but rather to test whether the OPLS-CS force field is capable of reproducing the results of high-level ab initio calculations which, in the absence of experimental data, can be considered as a benchmark for the calculations.

Methods

Constructing the Potential Model. Conventional all-atom force fields represent the electrostatic interactions of molecules by a series of point charges centered on the nuclear positions. Hunter and Sanders[6] suggested that the modeling of aromatic interactions could be improved by incorporating a number of non-atom-centered point charges into the model, representing the π-electron clouds of the molecule. These point charges lie directly above and below the C atoms, forming two planes that lie above and below the plane of the aromatic ring. This enables the reproduction of the quadrupole moment perpendicular to the plane of the ring, which is absent from an all-atom force field representation. In addition to our work on liquid benzene,[7] such an approach has also been successful in the modeling of a variety of other systems, including porphyrin rings[6] and the aromatic amino acids.[77] In parametrizing their model, Hunter and Sanders assumed that each C atom contributed one π electron to the π system, resulting in each of the 12 π-electron points having a charge of $-0.50e$. These points were then placed at a distance of 0.47 Å from the benzene ring so as to reproduce the component of the gas-phase quadrupole moment in the direction perpendicular to the plane of the benzene molecule.[78]

When considering liquid benzene, this model tended to overestimate the degree of perpendicular ordering in the liquid, and it was concluded that reparameterization of the model was necessary to obtain an accurate description of the liquid.[7] Initial attempts to treat the substituted aromatics using these benzene parameters yielded poor results, indicating that the transferability of these parameters was low. Inspection of the lowest-lying π orbitals of these molecules (Figure 1) suggested that this lack of transferability was not unreasonable; the exact nature of the π orbitals differs considerably throughout the series (as does the component of the quadrupole moment perpendicular to the ring, a key parameter used in fitting of the benzene model). An analysis of the π-electron density (Table 1) also supported this conclusion. Accordingly, it was necessary to develop a new set of parameters for each model.

Computational Details. As in the benzene OPLS-CS simulations,[7] the new models were parametrized by fitting

Table 1. π-Electron Density at Each of the Atomic Sites in the Aromatic Rings, Calculated at the MP2/6-31+G*//MP2/6-31+G* Level of Theory[a]

	C1/N	C2	C2′	C3	C3′	C4
benzene	0.448	0.448	0.448	0.448	0.448	0.448
toluene	0.397	0.449	0.448	0.437	0.435	0.458
phenol	0.291	0.468	0.493	0.418	0.420	0.482
pyridine	0.600	0.406	0.406	0.480	0.480	0.398

[a] The π-electron density is calculated, in each case, by considering the contribution from each of the three molecular orbitals in which the aromatic ring lies on the nodal plane.

Table 2. Calculated Values of Θ_{zz} for Benzene, Evaluated Using a Variety of Different ab Initio Techniques and Basis Sets[a]

	6-31+G**	6-31+G*	6-31G*	6-31G
MP2	−7.05	−7.12	−6.40	−6.73
B3LYP	−6.02	−6.12	−5.34	−5.46
HF	−6.85	−6.89	−6.21	−6.29

[a] All values are in units of ea_0^2.

Table 3. Calculated Values of Θ_{zz} for Pyridine, Evaluated Using a Variety of Different Ab Initio Techniques and Basis Sets

	6-31+G**	6-31+G*	6-31G*	6-31G
MP2	−4.57	−4.61	−4.18	−4.28
B3LYP	−3.64	−3.71	−3.30	−3.20
HF	−4.32	−4.34	−3.95	−3.82

[a] All values are in units of ea_0^2.

the models to reproduce experimental thermodynamic data of the molecular liquids, subject to the constraint that the model must also reproduce the experimental quadrupole moment of the molecule. For toluene and phenol, no experimental quadrupole moment was available, and instead it was necessary to use calculated values for the quadrupole moment. The values used in the parametrization process were calculated ab initio using the GDMA program[80] with the MP2/6-31+G**//MP2/6-31+G** wavefunction. The choice of wavefunction here is arbitrary, and it is necessary to consider whether such a choice is likely to have a significant effect on the calculated quadrupole moment. To assess this, quadrupole moments have been calculated using a variety of different techniques for two aromatic molecules, benzene and pyridine, in which the quadrupole moment is known experimentally (Tables 2 and 3).

For benzene, experimental values for Θ_{zz} range from -6.02 ± 0.25 to -7.42 ± 0.47 ea_0^2.[79] While the values calculated using B3LYP are generally rather low, all of the values calculated with either the MP2 or HF wavefunctions (Table 2) fall comfortably within this range, and there is nothing to suggest that calculating the quadrupole moment at the MP2/6-31+G**//MP2/6-31+G** level of theory would be inappropriate. For pyridine, fewer experimental values of Θ_{zz} are available, but they still range from -4.01 ± 0.45 to -5.13 ± 1.64 ea_0^2.[79] Here, the values calculated using HF methods tend to be slightly low, but all those calculated using MP2 are within the range of experimental values. In both pyridine and benzene, the value calculated using the MP2/6-31+G**//MP2/6-31+G** level of theory provides a reasonable approximation of the experimentally determined quadrupole moment. There is no reason to suppose that, for toluene and phenol, where experimental quadrupole moments are not available, the MP2/6-31+G**//MP2/6-31+G** calculated value of Θ_{zz} will not provide a reasonable approximation to the true value of Θ_{zz}.

In accordance with Hunter and Sanders' model, two new parameters are introduced for each model, q_π, the charge on the π points, and δ, the separation between the nucleus and the π points. Following initial difficulties in parametrizing a model for pyridine, distributed multipole analysis[41] using the program GDMA[80] revealed that the π-electron density is unevenly distributed throughout the molecule, locating preferentially at the nitrogen atom. For this reason, in the pyridine model, an additional parameter, $q_{\pi N}$, was introduced, being the charge on the N π point, and defined such that $q_{\pi N} = 1.75 q_\pi$. In all cases, the bond lengths and angles used are taken directly from the OPLSAA force field.[8]

Because OPLS is an effective potential, interactions that are not explicitly included in the potential will have been "mixed-in" to the model. This means that simply adding charge points to reproduce the quadrupole moment is not an acceptable solution—it would almost certainly have adverse effects on the properties that have been "mixed-in" to the potential. Accordingly, it was deemed necessary to re-examine all of the nonbonding parameters present within the OPLS model. The parametrization scheme followed for toluene and phenol was identical to that used in our previous work on benzene.[7] To the bare molecular frame are added a series of charge and van der Waals parameters. The charge parameters were subject to the constraints that the individual molecules must be charge-neutral, and that $q_\pi = -q_H$, where q_π and q_H are the charges on the π-electron points and hydrogen atoms, respectively. A better fit to the experimental data could probably be obtained by relaxing this constraint and, indeed, by allowing for charge redistribution within the ring. However, such an approach would greatly increase the risk of overfitting,[81] especially given the limited amount of experimental data available, and for this reason was avoided. q_π and δ were then varied systematically along with the van der Waals parameters, σ and ϵ, until the models gave the minimum deviation from experimental thermodynamic results.

Pyridine proved more complicated to parametrize. OPLS assigns different charge values to each of the three different types of C atoms in pyridine, and it was felt necessary to keep this scheme in place when parametrizing the OPLS-CS model. The total charge on each CH unit was determined using the MP2/6-31+G*//MP2/6-31+G* method, and this value was then kept constant throughout the parametrization process. Again, the charge parameters were subject to the constraints that the individual molecules must be charge-neutral, and that $q_\pi = -q_H$, where q_π and q_H are the charges on the π-electron points and hydrogen atoms, respectively. q_π and δ were then varied systematically along with the van

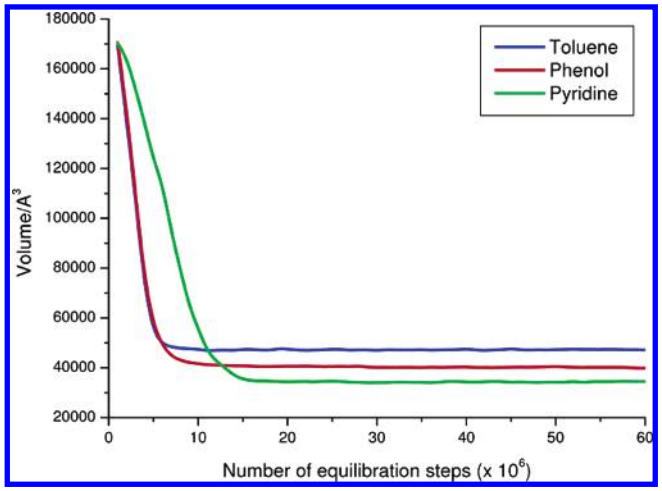

Figure 2. Plot monitoring the volume of the OPLS-CS parametrization simulations over the course of the equilibration period.

der Waals parameters, σ and ϵ, until the models gave the minimum deviation from experimental thermodynamic results.

As noted above, the development of the OPLS-CS models has required wholesale reparameterization of the original OPLS models. If the OPLS-CS models do show an improvement over the OPLS models, the question that then arises is: is the improvement due to the additional point charges or the reparameterization? This is in fact part of a larger issue of how sensitive the obtained results are to the choice of parameters, and whether this is more or less important than the inclusion of the additional point charges. For example, what would happen if the OPLS model compounds had larger bond dipoles? To investigate this issue further, an additional model has been parametrized for phenol. In this model, termed OPLS-Q, the atom-centered representation is retained, but the atomic charge on the aromatic C and H atoms is increased, from $\pm$ 0.115e in the OPLS model to $\pm$ 0.200e in the OPLS-Q model. With these charge values in place, the values of the nonbonding parameters σ and ϵ were then varied until the best set of thermodynamic results were obtained. From the results of the OPLS-CS calculations, it was noted that the calculated values of ΔH_{vap} were somewhat large compared to the experimental values. To investigate the significance of this effect, the OPLS-Q model was parametrized to reproduce the OPLS-CS thermodynamics values rather than the experimental thermodynamic values.

For the purpose of parametrization, a series of Monte Carlo simulations was performed for each molecule. Each simulation included 267 molecules in the constant-temperature, constant pressure (NPT) ensemble, with $P = 1$ atm. For toluene and pyridine, a temperature $T = 298$ K was used, and for phenol, which is solid at 298 K, $T = 333$ K was used. Each simulation consisted of 6.0×10^7 steps of equilibration, sufficient for convergence to have been achieved (Figure 2), followed by 6.0×10^7 steps of averaging. The new parameters obtained from these simulations are given in Table 4. In the simulations of liquid benzene, the molecule was held rigid, and in this case, the rings were also fixed, but the methyl group of toluene and the hydroxyl group of phenol were allowed to rotate. The rotational profile for each of these groups was, however, affected by the addition of the π-electron points, and it was necessary to reparameterize these two dihedral angle terms.

Table 4. New Parameters Obtained for OPLS-CS Simulations and the Corresponding Parameters Used in the OPLS Simulations[a]

molecule	parameter	OPLS-CS	OPLS
toluene	δ/Å	1.600	n/a
	q_{CA}/e	0.0455	−0.1150
	q_{HA}/e	0.0455	0.1150
	q_{π}/e	−0.0455	n/a
	q_{CT}/e	−0.1345	−0.0650
	q_{HC}/e	0.0600	0.0600
	σ_{CA}/Å	3.5750	3.5500
	σ_{HA}/Å	2.4370	2.4200
	σ_{CT}/Å	3.5000	3.5000
	σ_{HC}/Å	2.5000	2.5000
	ϵ_{CA}/kcal mol^{-1}	0.0630	0.0700
	ϵ_{HA}/kcal mol^{-1}	0.0270	0.0300
	ϵ_{CT}/kcal mol^{-1}	0.0594	0.0660
	ϵ_{HC}/kcal mol^{-1}	0.0270	0.0270
	$\psi_{CACACTHC}$/deg	0.000, −0.435, 0.000, −0.085	0.000, 0.000, 0.000, 0.000
Phenol	δ/Å	1.100	n/a
	q_{C1}/e	0.3480	0.1500
	q_{CA}/e	0.0990	−0.1150
	q_{HA}/e	0.0990	0.1150
	q_{π}/e	−0.0990	n/a
	q_{OH}/e	−0.5850	−0.5850
	q_{HO}/e	0.4350	0.4350
	σ_{C1}/Å	3.5500	3.5500
	σ_{CA}/Å	3.6200	3.5500
	σ_{HA}/Å	2.4677	2.4200
	σ_{OH}/Å	3.0700	3.0700
	σ_{OH}/Å	0.0000	0.0000
	ϵ_{C1}/kcal mol^{-1}	0.0700	0.0700
	ϵ_{CA}/kcal mol^{-1}	0.0700	0.0700
	ϵ_{HA}/kcal mol^{-1}	0.0300	0.0300
	ϵ_{OH}/kcal mol^{-1}	0.1700	0.1700
	ϵ_{HO}/kcal mol^{-1}	0.0000	0.0000
	$\psi_{CACAOHHO}$/deg	0.000, 1.845, 0.000, 0.000	0.000, 1.682, 0.000, 0.000
pyridine	δ/Å	0.600	n/a
	q_N/e	0.0990	−0.6780
	q_{C1}/e	0.1170	0.4730
	q_{C2}/e	0.2750	−0.4470
	q_{C3}/e	−0.0330	0.2270
	q_{H1}/e	0.1000	0.0120
	q_{H2}/e	0.1000	0.1550
	q_{H3}/e	0.1000	0.0650
	q_{π}/e	−0.1000	n/a
	$q_{\pi N}/e$	−0.1750	n/a
	σ_N/Å	3.1585	3.2500
	σ_{CA}/Å	3.4500	3.5500
	σ_{HA}/Å	2.3518	2.4200
	ϵ_N/kcal mol^{-1}	0.1637	0.1700
	ϵ_{CA}/kcal mol^{-1}	0.0674	0.0700
	ϵ_{HA}/kcal mol^{-1}	0.0289	0.0300

[a] Dihedral-angle parameters show the four Fourier coefficients used for energy evaluation. A full list of parameters is given in the Supporting Information.

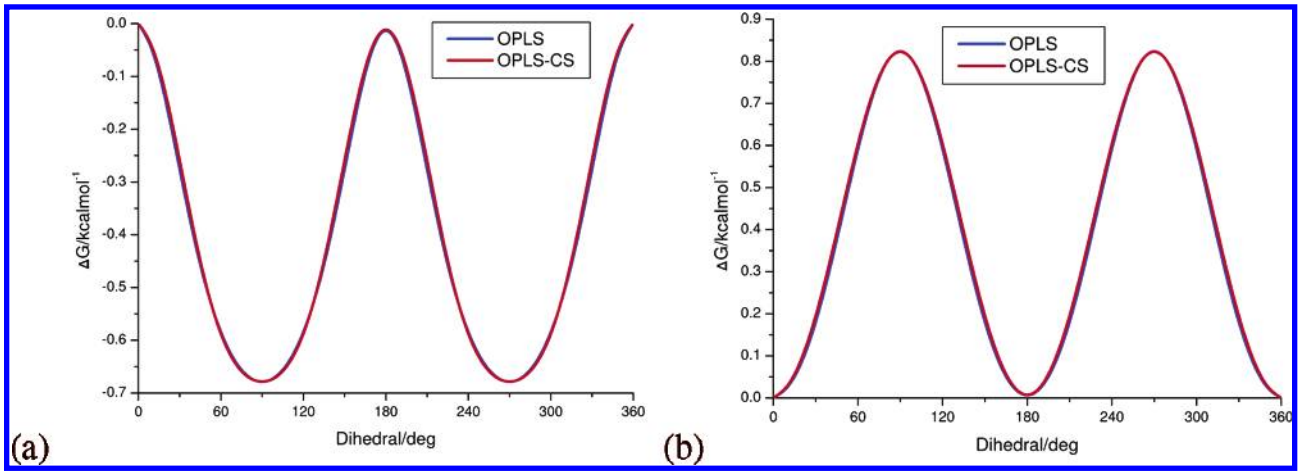

Figure 3. Fittings to OPLS rotational profiles. (a) Toluene and (b) phenol.

This was achieved by fitting to the gas-phase rotational profiles generated using the OPLS potential: for each dihedral angle, the Fourier coefficients were modified until they gave the best fit to the original OPLS-calculated rotational profile. This fitting was performed in the gas phase using simulations at $T = 298$ K with 4×10^5 steps of equilibration and 8×10^5 steps of averaging. The resulting parameters are given in Table 2, and the free-energy profiles are shown in Figure 3.

With the necessary parameters in place, two Monte Carlo simulations were performed for each of toluene and pyridine, and three were performed for phenol. In all cases, the first simulation treated the aromatic molecules using the OPLS all-atom potential[8] (denoted OPLS), the second the OPLS all-atom potential modified in such a way as to incorporate the π-electron points of the charge-separation model (denoted OPLS-CS). In phenol, the third simulation employed the OPLS-Q potential described above. All simulations were performed using BOSS version 4.2;[82] in the OPLS simulations, the standard OPLSAA parameters were used; in the OPLS-CS simulations, the OPLSAA parameters were modified so as to incorporate the π parameters described above. In all cases, the molecules were regarded as being rigid with the exception of the C−O and C−Me bonds in phenol and toluene, respectively, which were allowed to rotate subject to the torsional parameters of the OPLSAA force field in the OPLS and OPLS-Q simulations and the modified OPLSAA parameters in the OPLS-CS simulation. A system consisting of 267 molecules was used in simulations that were run in the NPT ensemble with $P = 1.0$ atm. For toluene and pyridine, a value of $T = 298$ K was used, but for phenol, in which the melting point occurs at 313 K, a value of $T = 333$ K was used. The simulations were begun from a configuration in which all of the molecules were arranged in a parallel fashion, and in all simulations, 4×10^6 equilibration steps were performed followed by 2.5×10^8 steps of averaging.

All molecular graphics were produced using the PyMOL program,[83] except those in Figure 1, which were produced using MOLDEN,[84] and Figures 7b, 10b, and 14b, which were produced using VMD.[85] Where values of the quadrupole moments are given for the force field models, they were evaluated using the BOSS program.[82]

Gas-Phase Dimers. As a first test of the transferability of the OPLS-CS models derived from the liquid-phase calculations, their ability to reproduce the binding energies of the gas-phase dimers of toluene, phenol, and pyridine has been investigated. Dimer structures for each of the molecules were first identified using ab initio theory. Dimer structures obtained from the literature, as discussed above, were constructed and minimized at the MP2/6-31+G* level of theory. The energies of the resulting complexes were then evaluated using the same level of theory with basis set superposition error[86] corrected using the counterpoise method.[87] In this way, a total of 15 minimum energy structures were identified: four for toluene, three for phenol, and eight for pyridine. The energy of each of the dimers was then evaluated using the OPLS and OPLS-CS models (and, in the case of phenol, the OPLS-Q model). Comparison of the energies obtained from the force field and ab initio calculations can give us an insight into whether the charge-separation models are providing a good reproduction of the intermolecular interactions between the molecules.

Modeling Cation−π Interactions. To examine the cation−π interaction, the phenol OPLS-CS model was used to construct an OPLS-CS model of *p*-methylphenol, an analogue of tyrosine (henceforth referred to as Tyr). The toluene OPLS-CS model previously obtained was taken as an analogue of phenylalanine (henceforth referred to as Phe). These models were then used to investigate the interactions of the two molecules with the archetypical organic cation—ammonium. For comparison, equivalent models for Phe and Tyr were constructed using the OPLS models of toluene and phenol. As a reference, ab initio calculations were performed at the MP2/6-311+G** level of theory, with the basis set superposition error[86] corrected using the counterpoise method.[87]

Two sets of calculations were performed. First, an ammonium molecule was placed in a monodentate fashion above the center of the aromatic ring lying in the *xy* plane. The ammonium molecule was then moved in increments of 0.1 Å along the *z* axis, with the energy evaluated via a single point calculation at each increment. This procedure was performed using the ab initio method as well as the OPLS-CS and OPLS models. Second, the separation of the molecules in the *z* direction was fixed at 3 Å while the ammonium molecule was scanned, again with a monodentate

Table 5. Thermodynamic Properties of Liquid Toluene

	OPLS	OPLS-CS	experiment[26]
θ_{zz}/ea$_o^2$	0.45	−6.34	−6.64[a]
dipole/ea$_o$	0.20	0.22	0.14
density/g cm^{-3}	0.857	0.863	0.862
ΔH_{vap}/kcal mol^{-1}	8.89	9.09	9.08
molecular volume/Å^3	178.6	177.3	177.4

[a] Experimental quadrupole moment values for phenol and toluene are not available. The values used in the parametrization process were calculated ab initio using the GDMA program[80] with the MP2/6-31+G**//MP2/6-31+G** wavefunction.

Table 6. Thermodynamic Properties of Liquid Phenol

	OPLS	OPLS-Q	OPLS-CS	experiment[39]
θ_{zz}/ea$_o^2$	0.00	0.00	−6.90	−6.93[a]
dipole/ea$_o$	0.91	0.99	0.79	0.57
density/g cm^{-3}	1.03	1.04	1.04	1.04
ΔH_{vap}/kcal mol^{-1}	12.81	1.49	13.44	13.19

[a] Experimental quadrupole moment values for phenol and toluene are not available. The values used in the parametrization process were calculated ab initio using the GDMA program[80] with the MP2/6-31+G**//MP2/6-31+G** wavefunction.

Table 7. Thermodynamic Properties of Liquid Pyridine

	OPLS	OPLS-CS	experiment[50]
θ_{zz}/ea$_o^2$	0.00	−4.61	−4.61
dipole/ea$_o$	0.94	0.89	0.86
density/g cm^{-3}	0.97	0.97	0.98
ΔH_{vap}/kcal mol^{-1}	9.58	9.66	9.61

geometry, in the x and y directions over the aromatic molecule. The energy was evaluated at 0.2 Å increments in both directions with the same calculations performed using the ab initio, OPLS, and OPLS-CS models. All ab initio calculations were performed using the program Gaussian 98.[88]

Results and Discussion

Aromatic Liquids. From the initial simulations performed, a set of parameters was determined as the "best" OPLS-CS model for each of the aromatic molecules (full details of the parameters are given in the Supporting Information). The comparison with experimental thermodynamic data is shown in Tables 5−7.

Although it should be noted that the model has been parametrized well within experimental uncertainty, the reproduction of the experimental thermodynamic data by the OPLS-CS models is generally good. It offers an improvement over OPLS in all areas, with a significant improvement in the reproduction of the quadrupole moment perpendicular to the plane of the ring in all cases. It is possible to learn something about the behavior of these liquids by consideration of their thermodynamic properties. At constant pressure, the entropy change of vaporization of a system is related to its enthalpy of vaporization via eq 1.

$$\Delta S_{vap} = \frac{\Delta H_{vap}}{T_b} \quad (1)$$

Trouton's rule[89] then tells us that molecular liquids have a standard entropy of vaporization of approximately 85 J K^{-1} mol^{-1} (20 cal K^{-1} mol^{-1}). This observation is based on the fact that the evaporation of any liquid into the gas phase will generate a similar amount of disorder. As a first measure of the level of structuring in the liquids in this study, we can consider the relevant ΔS_{vap} values, which are (calculated via eq 1) 23.68, 29.05, and 24.75 cal K^{-1} mol^{-1} for toluene, phenol, and pyridine, respectively.

It follows that phenol, with the largest deviation from Trouton's rule, is the most structured of the liquids, and toluene the least. This result is unsurprising. Phenol, which contains both hydrogen-bond donors and acceptors, would be expected to be highly hydrogen-bonded. Pyridine will also form hydrogen bonds, but they will be weaker because the only possible donors are the C−H bonds. Toluene, with no scope for hydrogen bonding, will be forced to rely on weaker interactions. Although this is much as would be expected intuitively, it is encouraging. Because the OPLS-CS model has performed better than OPLS in reproducing the experimental ΔH_{vap} values, and these values are an indirect measure of structure in the liquid, it suggests that OPLS-CS should give a better account of liquid structure than does OPLS.

It is also possible to learn something about the strength of the interactions between the models. By comparing the ΔH_{vap} value determined experimentally to the ΔH_{vap} value predicted by Trouton's rule (via eq 2), we can make an estimate of the strength of these ordering interactions.

$$\Delta\Delta H = \Delta H_{vap}^{exp} - \Delta H_{vap}^{Trouton} \quad (2)$$

In benzene, $\Delta\Delta H = 0.92$ kcal mol^{-1}, suggesting that the intermolecular interactions are small. In toluene, $\Delta\Delta H = 1.30$ kcal mol^{-1}; in phenol, $\Delta\Delta H = 3.98$ kcal mol^{-1}, and in pyridine, $\Delta\Delta H = 1.73$ kcal mol^{-1}. Again, the relative size of these interactions is much as would be expected, and again the fact that OPLS-CS has reproduced the ΔH_{vap} values more accurately than OPLS suggests that the model should be able to give a better description of the intermolecular interactions within the liquid.

The OPLS-CS model also outperforms the OPLS model in reproducing the density of all three of the liquids. This can again be seen as an encouraging sign since the liquid density provides a fundamental insight into the packing of the molecules in the liquid and, hence, the structure of the liquid. It follows that, with a better reproduction of the liquid densities, we would expect the OPLS-CS model to provide a better representation of the liquid structures.

Structural Information. Toluene. An analysis of $g_{CC}(r)$ (Figure 4a) reveals that the two simulations performed for liquid toluene, using the OPLS and OPLS-CS potentials, give almost identical results. Two small peaks, at around 5.2 Å and 6.2 Å, are identified in each case, with a slight shoulder in the distribution at around 4.3 Å. In the OPLS-CS simulations of liquid benzene,[7] this shoulder appeared as a clear peak and was taken as evidence of a largely perpendicular structure in the liquid. The absence of such a peak in this case indicates that the perpendicular arrangement is not nearly so dominant in toluene. An additional difference between $g_{CC}(r)$ for toluene and benzene is the presence of a small shoulder in the distribution at around 7.5 Å, which is

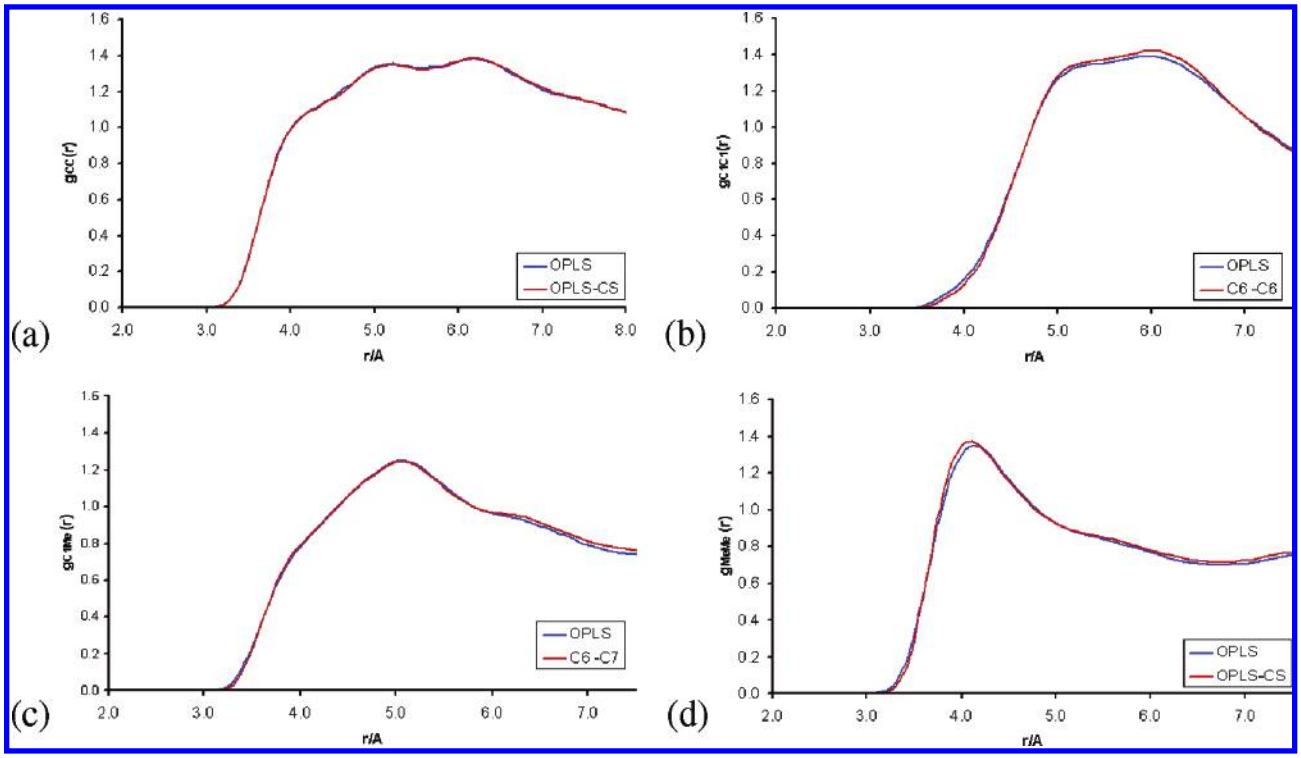

Figure 4. Radial distribution functions for liquid toluene. (a) $g_{CC}(r)$, (b) $g_{C1C1}(r)$, (c) $g_{C1Me}(r)$, and (d) $g_{MeMe}(r)$.

not observed in liquid benzene. Kim and Lee[25] also observed this feature and attributed it to the presence of the methyl group attached to the phenyl ring in a perpendicular arrangement. It is worth noting that the simulation of Kim and Lee in the NPT ensemble[25] heavily overestimated the density of the liquid, and as a result, all four features in $g_{CC}(r)$ were considerably larger than in the current results. A simulation by the same group in the canonical ensemble[24] gave results in much better agreement with our results.

In addition to $g_{CC}(r)$, several other radial distribution functions have been considered (Figure 4). In general, the rather broad and unstructured peaks are similar to those that have been previously observed.[26] The only exception to this rule is in $g_{C1C1}(r)$ (Figure 4b). Previous simulations using the OPLS potential have identified this as a single, broad peak,[26] whereas in the OPLS-CS simulation results, a slight structure is obtained in the form of a splitting into two peaks, the first at around 5.3 Å with a second, larger peak at 6.1 Å. These two distances fit well with two of the minimum energy structures of the toluene dimer (Figure 5), a stacked structure stabilized by a methyl$-\pi$ hydrogen bond[12] and a T-shaped structure. That the second peak, corresponding to a T-shaped structure, is more prominent in the OPLS-CS simulation would be expected; the inclusion of the quadrupole moment arising from the π electrons would be expected to favor the T-shaped orientations. This effect appears to be the only real difference between the simulations performed using the OPLS and OPLS-CS potentials.

The slight preference for the T-shaped structure shown in $g_{C1C1}(r)$ confirms what has been learned from $g_{CC}(r)$: in liquid toluene, a perpendicular arrangement of the molecules is present, but not to such an extent as in benzene, and both parallel displaced and perpendicular structures are well-represented in the liquid phase. This idea is confirmed by

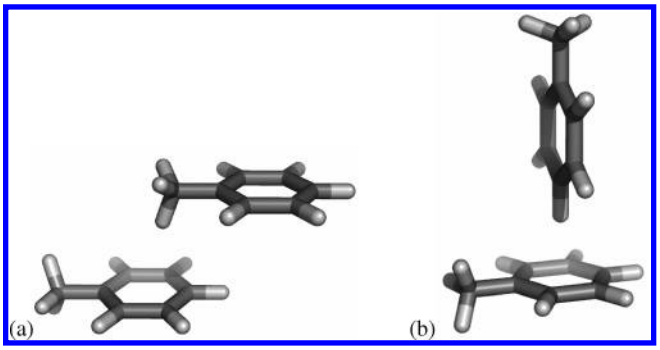

Figure 5. Minimum-energy structures for the toluene dimer. (a) $R_{C1C1} = 5.3$ Å and (b) $R_{C1C1} = 6.1$ Å.

an analysis of the angular distribution function $g(r,\theta)$ (Figure 6), which considers both the relative orientations of the molecules [measured in terms of the angle between the normals to the rings (Figure 6a)] and their separation. The first peak in $g(r,\theta)$ corresponds to the first solvation shell of liquid toluene, and it is clear that both the OPLS and OPLS-CS simulations predict very similar structures; in both cases, the parallel and perpendicular arrangements are approximately equally well-populated.

As has been noted, the molecules in liquid benzene showed a clear preference for orienting in a perpendicular arrangement, while those in liquid toluene show no such preference. This difference must be attributable to the presence of the methyl group in toluene. The methyl group introduces a dipole moment into the molecules, which would be expected to favor parallel (head-to-tail) orientations. In addition, the presence of the methyl group stabilizes parallel displaced interactions such as that shown in Figure 5a by stacking over the center of the ring in an energetically favorable arrangement. The lack of preference for any one orientation can also be seen in the structure of the first solvation shell of a

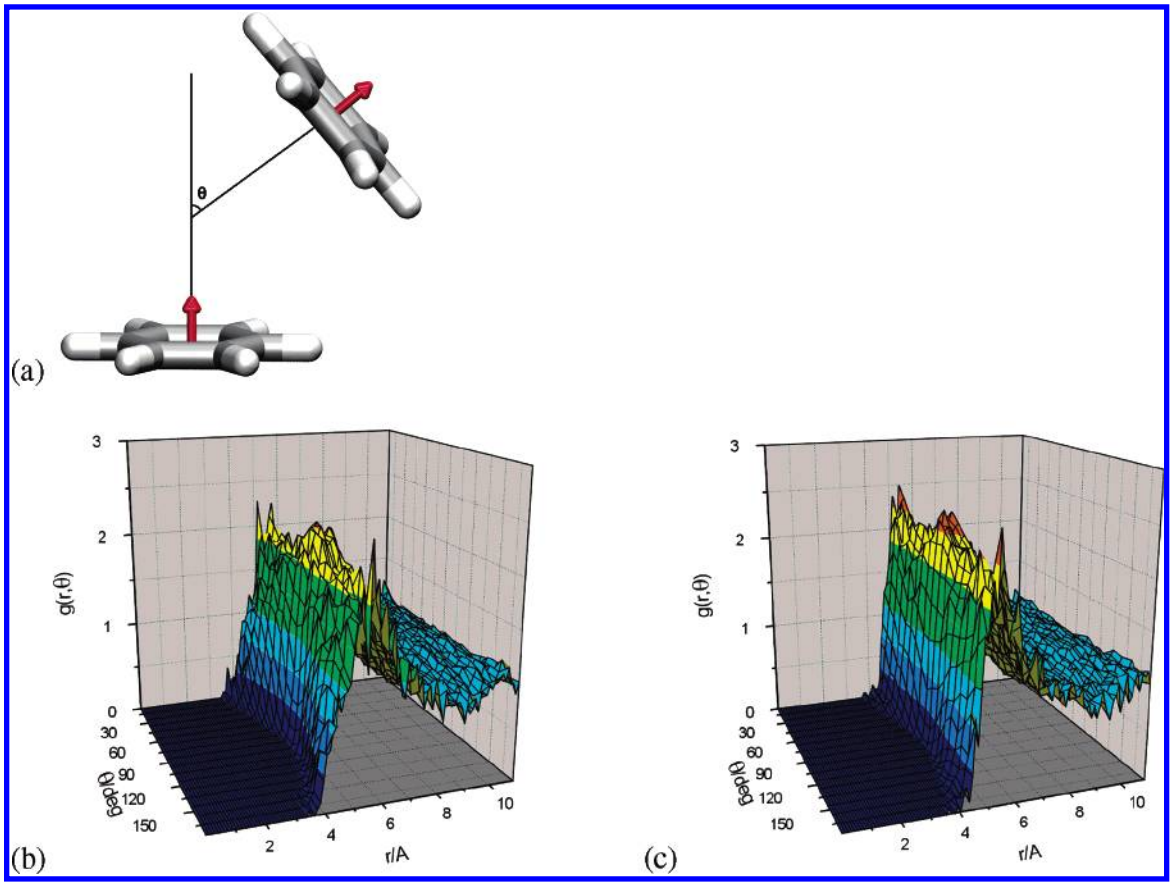

Figure 6. Angular distribution functions for liquid toluene. (a) Definition of θ. (b) $g(r,\theta)$ calculated using OPLS. (c) $g(r,\theta)$ calculated using OPLS-CS.

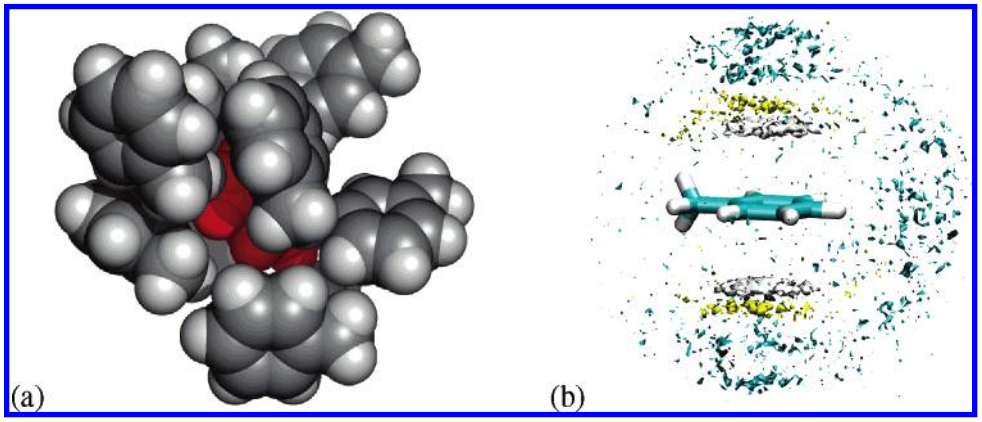

Figure 7. Distribution of molecules in liquid toluene. (a) The first solvation shell around a toluene molecule (in red) taken from the OPLS-CS simulation of liquid toluene. (b) The average atomic density around toluene molecules in the OPLS-CS simulation (cyan = aromatic carbon; yellow = methyl carbon; gray = hydrogen).

toluene molecule (Figure 7a). Examining the average atomic density around toluene molecules (Figure 7b) also leads to the same conclusion. The only real preference we see is for the methyl group to locate above the aromatic ring. Such an arrangement would be expected in either the T-shaped or parallel minima.

Phenol. The use of the OPLS-CS potential had only a small effect on the predicted properties of liquid toluene, but for phenol, its effect is far larger. In the OPLS simulations, $g_{CC}(r)$ appears as a single broad peak, with no discernible structure. In the OPLS-CS simulations, however, the formation of three distinct features within that single peak is observed. Differences between the force fields are even more apparent in the individual $g(r)$ (Figure 8). All of the $g(r)$'s show approximately the same form in both the OPLS and OPLS-CS simulations, but in all cases, the OPLS-CS calculated $g(r)$'s contain smaller peaks. The sizes of these peaks, which focus on the interactions involving the phenol hydroxyl group, can be considered as a measure of the hydrogen bonding in the liquid, and it follows that in the OPLS-CS model it is predicted that less hydrogen bonding occurs. Evidence for hydrogen bonding also comes from the C(O)−H(O) distribution (Figure 8c), which shows the double-peaked structure that is characteristics of alcohols.[90]

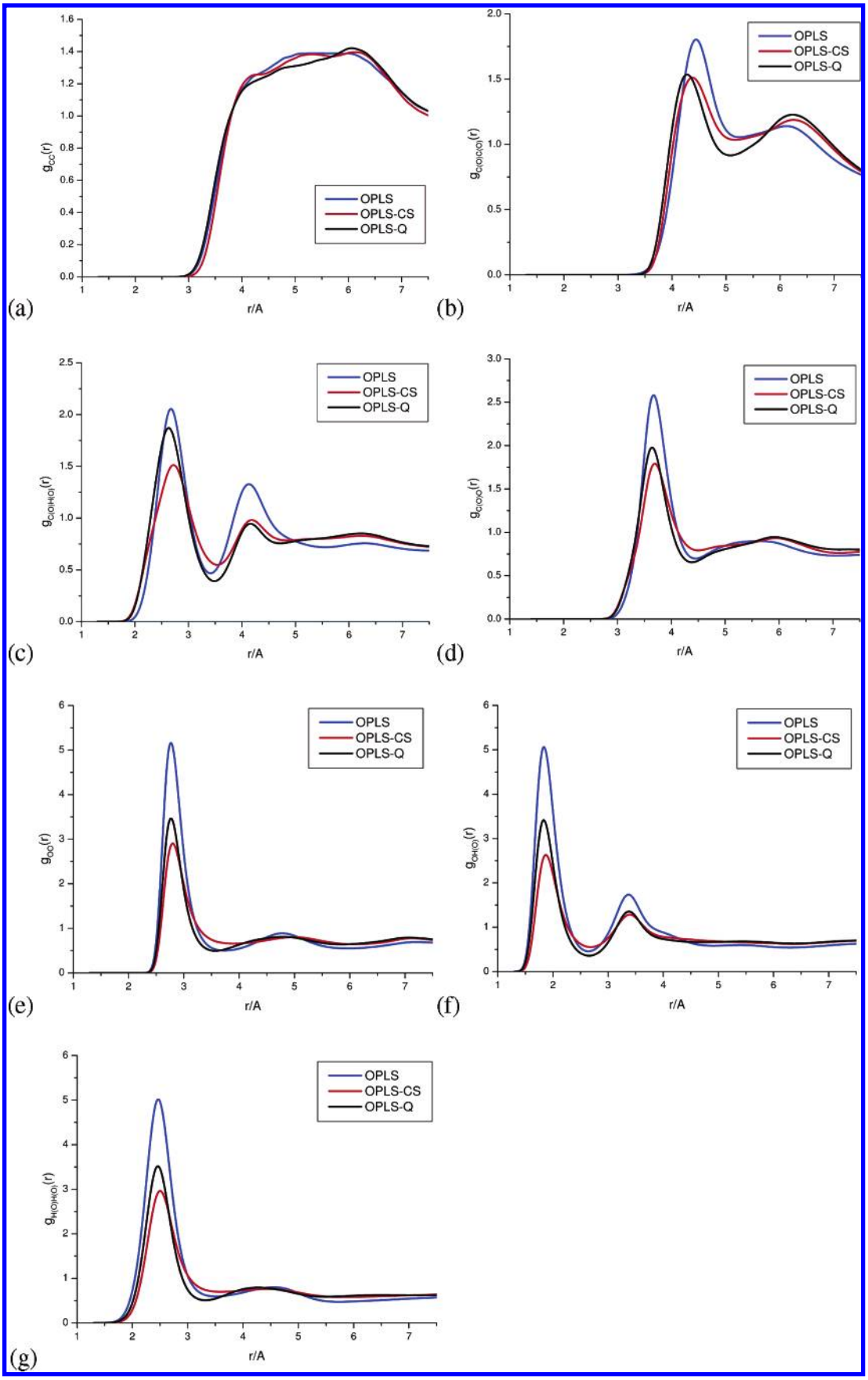

Figure 8. Radial distribution functions for liquid phenol. (a) $g_{CC}(r)$, (b) $g_{C(O)C(O)}(r)$, (c) $g_{C(O)H(O)}(r)$, (d) $g_{C(O)O}(r)$, (e) $g_{OO}(r)$, (f) $g_{OH(O)}(r)$, and (g) $g_{H(O)H(O)}(r)$.

By calculating the areas under the peaks in $g(r)$, we can make an estimate of the average number of hydrogen bonds formed per molecule. Integration of the first peak in $g_{OO}(r)$ to the minimum at 3.7 Å gives a value of 2.6 in the OPLS

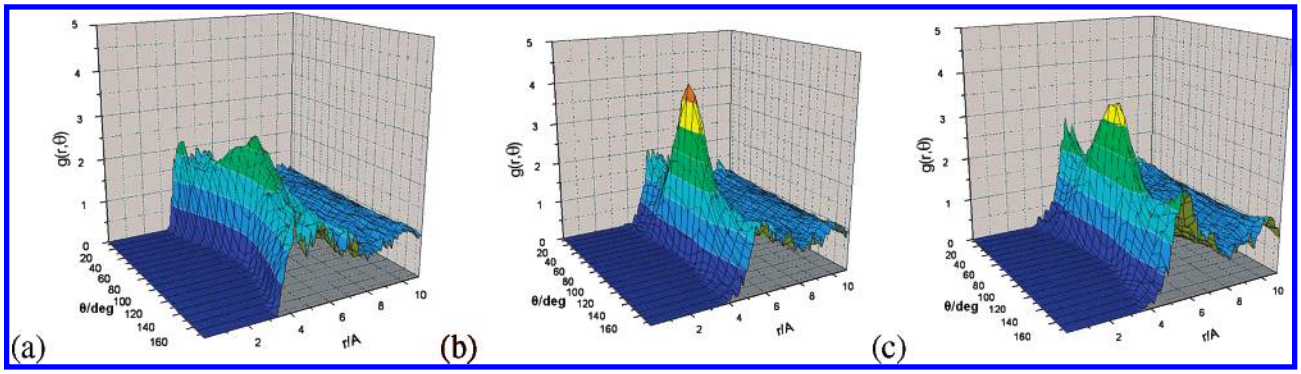

Figure 9. Angular distribution functions for liquid phenol. Calculated using (a) OPLS, (b) OPLS-Q, and (c) OPLS-CS.

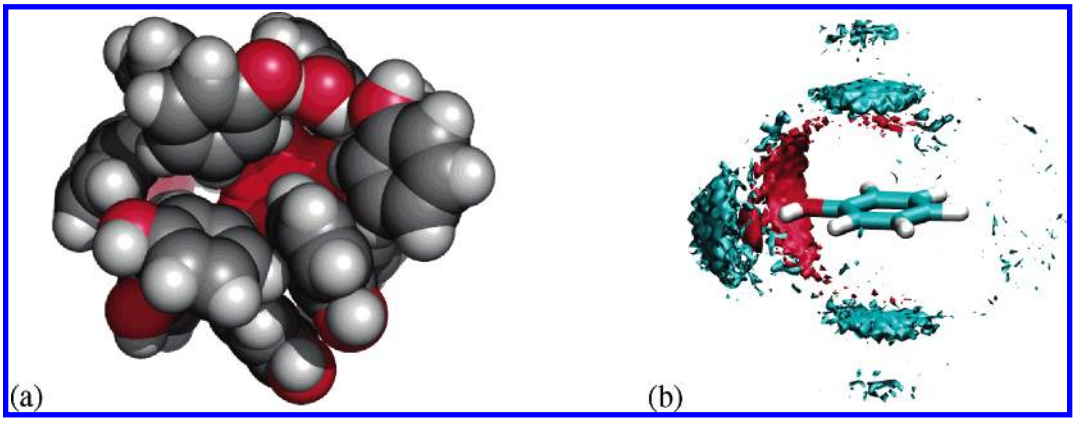

Figure 10. Distribution of molecules in liquid phenol. (a) The first solvation shell around a phenol molecule (in red) taken from the OPLS-CS simulation of liquid phenol. (b) The average atomic density around phenol molecules in the OPLS-CS simulation (cyan = carbon; red = oxygen).

simulations and 1.8 in the OPLS-CS simulations. The OPLS simulation is, therefore, predicting that a molecule forms, on average, 2.5 hydrogen bonds, with the OPLS-CS simulation predicting 1.8. This value is in close agreement with the value of 1.6 predicted by Jorgensen et al.[26] in their simulations of liquid *m*-cresol.

For phenol, a third potential model, OPLS-Q, has also been considered. It is a reparameterization of OPLS in which the OPLS-CS density and ΔH_{vap} are reproduced and the C−H bond dipole is increased. The radial distribution functions obtained using OPLS-Q are shown in Figure 8. In general, the OPLS-Q model seems to be somewhere in between the OPLS-CS and OPLS models, which suggests that a simple reparameterization of the OPLS model might be able to yield a model that is comparable to the OPLS-CS model. Such a conclusion would indicate that the charge-separation strategy might be an overcomplicated method for the development of force field parameters in general, and this idea will be investigated in more detail in the relevant sections of this paper.

Further insight into the structure of the liquid comes from consideration of the angular distribution function $g(r,\theta)$ (Figure 9). Now, we can see that the OPLS-CS simulation predicts a far more perpendicular structure than the OPLS simulation. What is more, the OPLS-CS simulation suggests that the ordering in liquid phenol is far more perpendicular than in liquid toluene, and almost as perpendicular as liquid benzene.[7] Although we have no experimental structural data for liquid phenol, this result seems to be in good agreement with experimental and theoretical studies on the phenol dimer where the minimum energy structure is found to be T-shaped, but with hydrogen bonding occurring between the hydroxyl groups. Examples of such structures can clearly be seen in the first solvation shell of a phenol molecule taken from the OPLS-CS simulation (Figure 10a). Moreover, by considering the average phenol atomic density around phenol molecules in the simulation (Figure 10b), it can clearly be seen that the preferred position for an oxygen atom is adjacent to the oxygen atom of another phenol molecule, either donating or accepting a hydrogen bond. Comparatively little oxygen density is seen above or below the ring, indicating that the formation of OH−π hydrogen bonds is rare, although the carbon density reveals that T-shaped stacking over the ring center does still occur. The angular distribution function obtained from the OPLS-Q simulation (Figure 9b) is surprising. The radial distribution functions suggested that OPLS-Q was somehow intermediate between the OPLS and OPLS-CS simulations. The angular distribution obtained using OPLS-Q favors the perpendicular arrangement even more strongly than the OPLS-CS model does. It is clear from this that the results of the calculations are highly dependent on the choice of parameters. They, along with the type of model to be used, must be carefully considered if a useful potential is to be obtained.

Pyridine. A first measure of the structure of liquid pyridine comes from $g_{NN}(r)$ (Figure 11), which shows two distinct peaks in both the OPLS and OPLS-CS simulations. Jorgensen and McDonald have previously studied liquid pyridine using

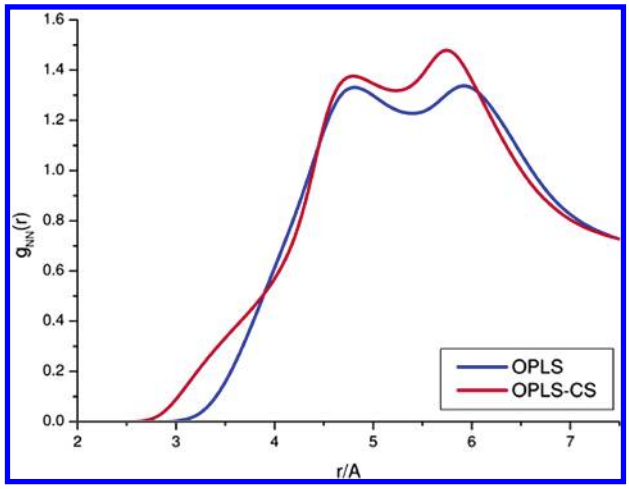

Figure 11. Radial distribution function for liquid pyridine $g_{NN}(r)$.

the OPLS force field[50] and obtained the same $g_{NN}(r)$. They have attributed the two peaks to two different antiparallel arrangements (which correspond to the OPLS calculated dimer minimum), one directly stacked and the other parallelly displaced. However, an analysis of $g(r,\theta)$ (Figure 12) suggests that even with the OPLS potential a perpendicular arrangement of the molecules is slightly favored. In the OPLS-CS simulation, the preference for perpendicular arrangements is even more pronounced.

Inspection of the structures produced in the OPLS-CS simulation suggests that we are not in fact seeing the formation of antiparallel stacked structures, but rather the formation of hydrogen-bonded structures, in which the molecules then adopt an orthogonal arrangement. The peak in $g_{NN}(r)$ at 4.9 Å results from the formation of a structure in which one pyridine molecule donates a hydrogen bond through the hydrogen in the ortho position, with the peak at 5.9 Å arising from a structure in which a pyridine molecule donates a hydrogen bond through the hydrogen in the meta position (Figure 13). The preference for a perpendicular arrangement can also be seen in the first solvation shell of a molecule taken from the OPLS-CS simulation (Figure 14a). Examining the average atomic density around molecules in liquid pyridine (Figure 14b) supports the view that the interactions in the liquid are dominated by the formation of CH−N hydrogen bonds. Nitrogen atom density occurs only in the regions around the hydrogen atoms. Carbon density, however, is seen additionally above the aromatic ring, arising from dimer structures that are stabilized by the formation of a CH−π interaction.

Computational Cost. An important consideration when running any molecular simulation is the relative cost, in terms of computer time, of the calculation. Typically, the need to achieve an acceptable level of accuracy within an acceptable time frame is the determining factor in the choice of the force field model used. With the addition of extra point charges, one would expect the OPLS-CS simulations to take longer than the OPLS simulations, and this is indeed found to be the case. Typically, running on equivalent processors, the OPLS-CS calculations take around 3 times as long as the analogous OPLS calculations. While this represents a significant increase in computational resources, it is still considerably faster than simulations based on distributed multipole models, which have been found to take around 8 times as long as their atom-centered equivalents.[91]

Gas-Phase Dimers. The gas-phase dimer minimum energy structures identified via ab initio calculation are shown in Figure 15, and the corresponding interaction energies are given in Tables 8−10.

For toluene, the OPLS model does well in reproducing the energetic ordering of the dimers, even if the energetic separation between them is underestimated. The OPLS-CS model performs similarly, underestimating the relative stability of tol_b relative to the T-shaped structures tol_c and tol_d but correctly identifying the minimum energy structure tol_a. This information is consistent with that obtained from the simulations on liquid toluene, in which both the OPLS and OPLS-CS models are found to predict very similar structures, but with the OPLS-CS model showing a slight preference for a perpendicular arrangement of molecules.

Interpretation of the results for phenol is hampered by the fact that no perpendicular structures have been identified as minima by the ab initio minimization process. Of the three structures identified as minima, OPLS fails to correctly identify the lowest-energy structure. The reparameterized OPLS model, OPLS-Q, performs better, giving the correct ordering of the minima but heavily underestimating the energy separation between them. OPLS-CS correctly identifies the minimum energy structure and gives the best estimate of its interaction energy. However, it completely fails in its treatment of phe_c, not even identifying it as a minimum.

The relatively simple structure of pyridine means that it has been possible to characterize more fully the minimum energy dimer structures, and it is this molecule that should give the best indication of the ability of the force field models to reproduce dimer interaction energies.

Both the OPLS and OPLS-CS models fail to predict correctly the minimum energy structure of the pyridine dimer. Where the ab initio calculations prefer the formation of a cyclic hydrogen-bonding pattern (pyri_h), OPLS prefers an offset stacked arrangement and OPLS-CS prefers CH−N hydrogen-bonded perpendicular arrangements. Again, these results are consistent with the analysis of the liquid pyridine simulations. For structure pyri_c, the OPLS-CS model performs very poorly, predicting the interaction to be strongly repulsive.

Overall, the results of the calculations on the gas-phase dimers suggest that the OPLS-CS models do not offer a general improvement over the OPLS models. In fact, there are several flaws that can be identified in the OPLS-CS representations. The repulsion in complexes where the molecules stack directly above each other is strongly overestimated, and the relative stability of the T-shaped structures is slightly overestimated relative to the parallel structures. These observations would lead us to suggest that the perpendicular arrangement of molecules is probably overrepresented in the liquid structure predicted using the OPLS-CS models. The absence of perpendicular dimer structures for phenol makes it impossible to truly assess the quality of the OPLS-Q model. If the OPLS-CS model is overestimating

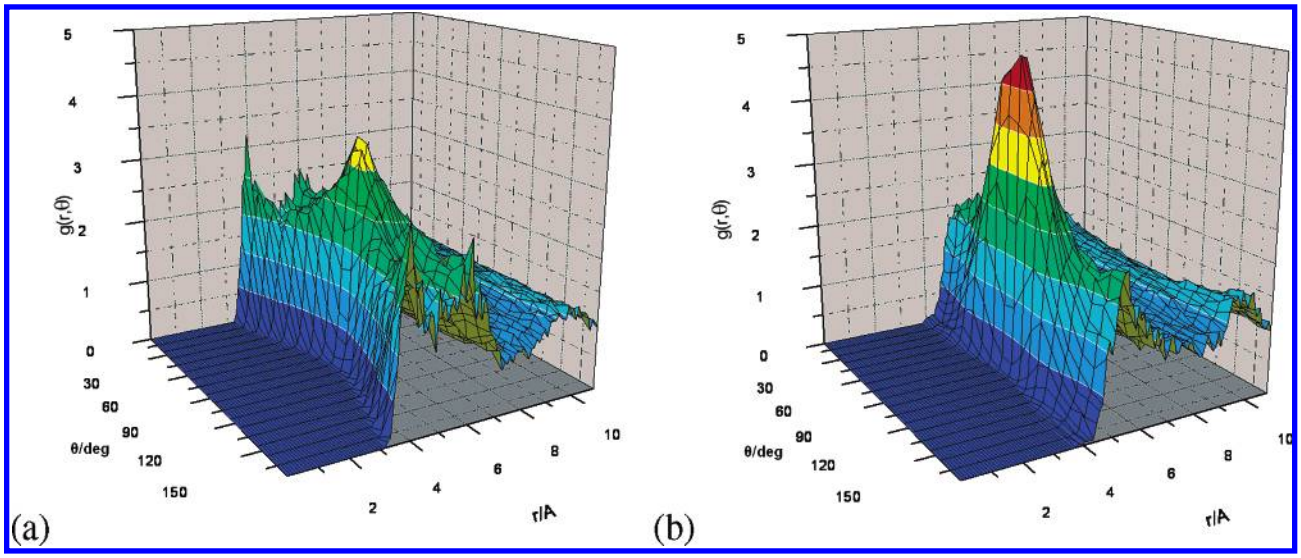

Figure 12. Angular distribution functions for liquid pyridine. Calculated using (a) OPLS and (b) OPLS-CS.

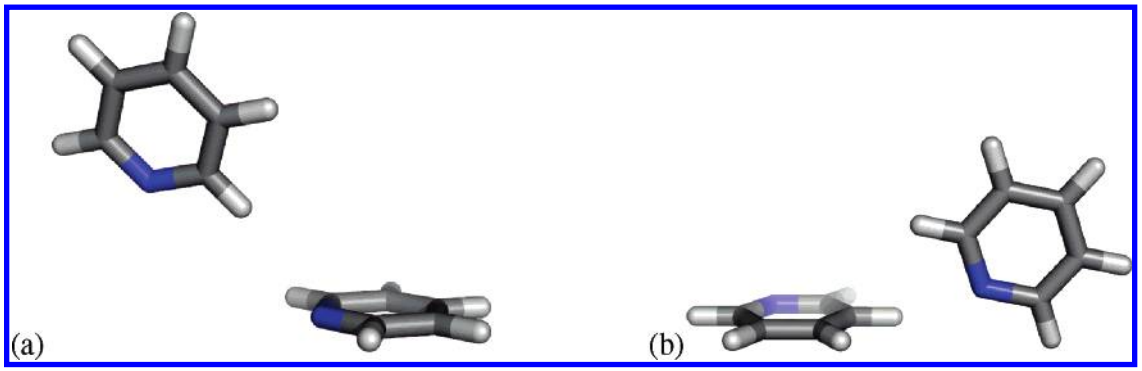

Figure 13. Pyridine−pyridine interactions taken from the OPLS-CS simulation of liquid pyridine. (a) R_{NN} = 4.9 Å and (b) R_{NN} = 5.9 A.

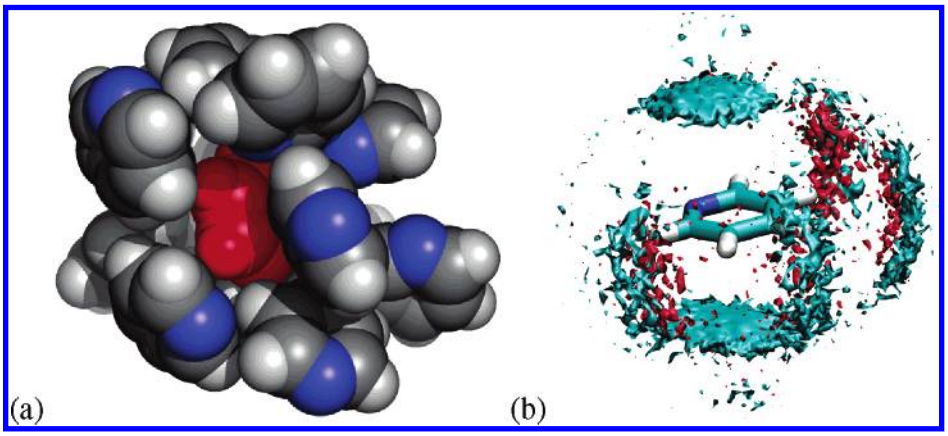

Figure 14. Distribution of molecules in liquid pyridine. (a) The first solvation shell around a pyridine molecule (in red) taken from the OPLS-CS simulation of liquid pyridine. (b) The average atomic density around pyridine molecules in the OPLS-CS simulation (cyan = carbon; red = nitrogen).

perpendicular structure, then the OPLS-Q model is overestimating it even more.

Cation−π Interactions. The models developed thus far have been shown to be successful in the modeling of aromatic liquids. However, the main goal of this work is the development of a force field capable of accurately modeling aromatic groups in biological contexts. The models are not yet sufficiently well-developed that we might be able to consider modeling aromatic−aromatic interactions in proteins, but we can begin to consider isolated examples of the types of interaction that are important in these cases. Such investigations will allow us to test whether the OPLS-CS model is likely to be useful in the study of biological problems and will give an indication of the transferability of the models to situations for which they have not been explicitly parametrized.

As a first test case, the interactions between the amino acid side chains from Phe and Tyr and the ammonium cation

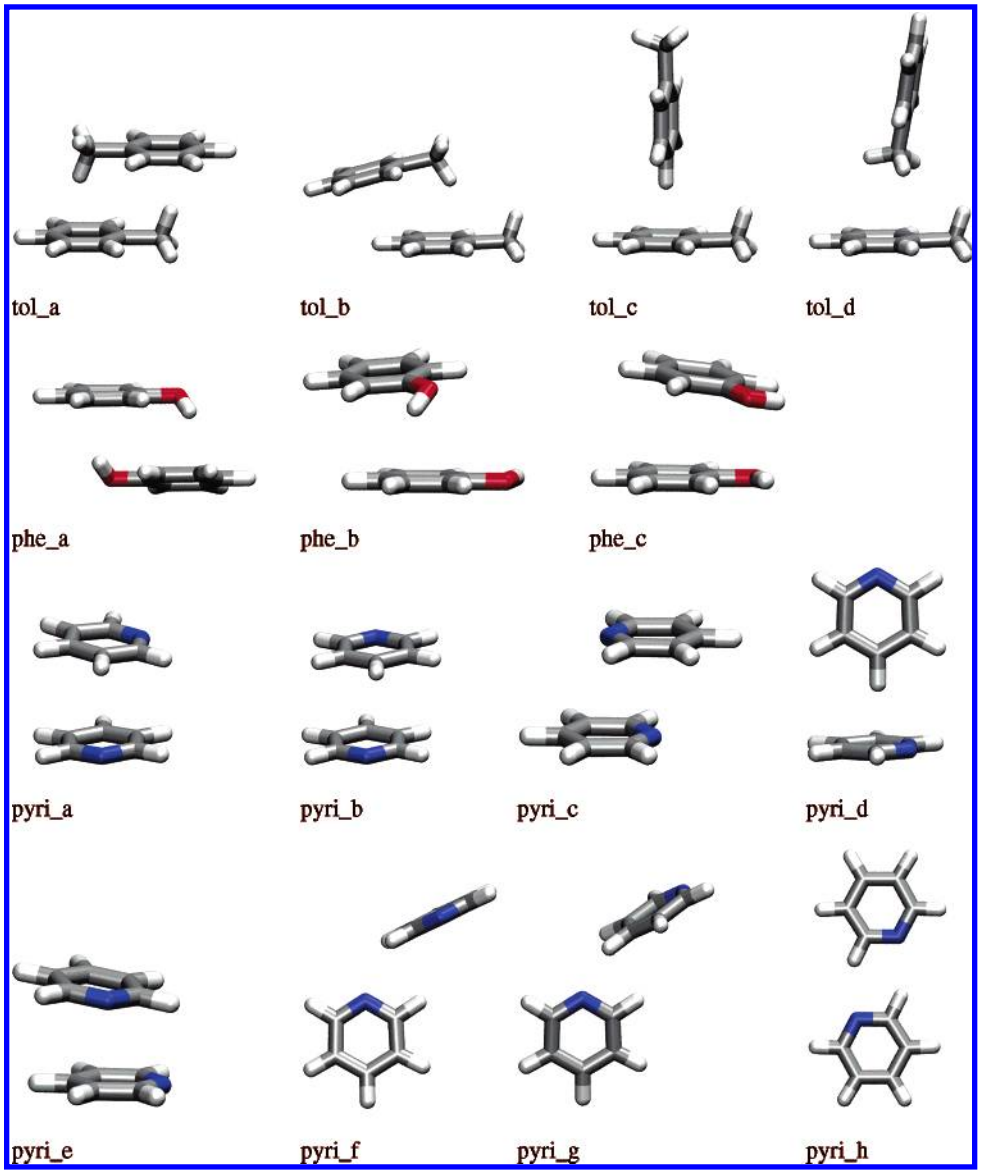

Figure 15. Gas-phase dimer structures for toluene, phenol, and pyridine. Calculated at the MP2/6-31+G*//MP2/6-31+G* level of theory.

Table 8. Toluene Dimer Interaction Energies (in kcal mol^{-1})

	tol_a	tol_b	tol_c	tol_d
MP2/6-31+G*	−3.69	−2.90	−2.02	−1.29
OPLS	−3.14	−2.24	−2.16	−1.95
OPLS-CS	−2.44	−1.28	−1.96	−1.71

Table 9. Phenol Dimer Interaction Energies (in kcal mol^{-1})

	phe_a	phe_b	phe_c
MP2/6-31+G*	−5.33	−3.48	−3.59
OPLS	−2.19	−2.65	−3.11
OPLS-Q	−2.41	−2.23	−2.24
OPLS-CS	−6.26	−5.25	0.91

have been investigated. The results of scanning the ammonium molecule along the *z* axis above the aromatic ring center (Figure 16) reveal that both the OPLS and OPLS-CS models perform reasonably well in reproducing the general shape of the ab initio calculated curves. For the Phe−ammonium interaction, OPLS-CS does slightly better in reproducing the position of the energy minimum while OPLS does better in reproducing the well depth. For Tyr−ammonium, the situation is more clear-cut; OPLS-CS performs much better than OPLS in reproducing both the position and depth of the curve.

While the *z*-scan results do not definitively demonstrate that OPLS-CS outperforms OPLS, the results obtained by scanning the ammonium molecules in the *xy* plane above the aromatic ring (while keeping *z* fixed at 3 Å) are far more conclusive (Figures 17 and 18). For both Phe and Tyr, the OPLS model fails completely to reproduce the general shape of the potential energy surface, severely overestimating the repulsion that occurs when the cation is above the atomic sites. The OPLS-CS models give a much better reproduction of the ab initio data. The well depth is exaggerated, but the shape of the potential energy surface is close to that obtained from the ab initio calculations. The same set of calculations has also been performed using the OPLS-Q model (results

Table 10. Pyridine Dimer Interaction Energies (in kcal mol^{-1})

	pyri_a	pyri_b	pyri_c	pyri_d	pyri_e	pyri_f	pyri_g	pyri_h
MP2/6-31+G*	−3.02	−2.94	−1.85	−1.58	−2.93	−2.57	−2.89	−3.31
OPLS	−3.02	−3.33	−1.84	−2.08	−2.99	−2.62	−2.77	−2.27
OPLS-CS	−1.22	−0.09	4.41	−2.76	−1.43	−2.72	−2.93	−1.00

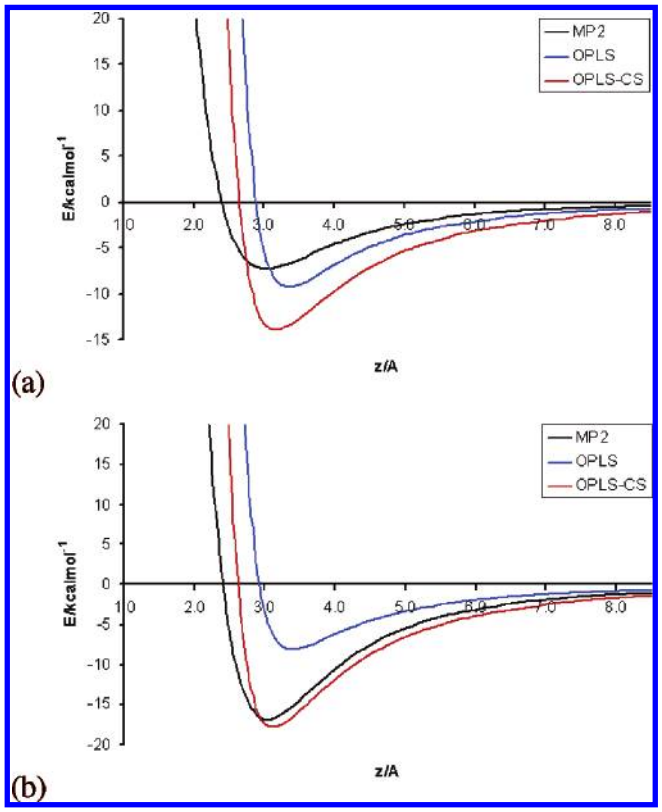

Figure 16. Energy profiles obtained from scanning an ammonium cation in the *z* direction above the aromatic ring center of (a) Phe and (b) Tyr.

not shown). In this case, the results are found to be similar to those obtained using the OPLS-CS model. As was found to be true in the discussion of the liquid-phase results, it seems likely that results equivalent to those obtained using a charge-separation model could be obtained using a standard OPLS model with appropriate reparameterization.

An accurate reproduction of the potential energy surface for cation−π interactions is likely to be essential for the accurate modeling of a variety of biological systems. For example, Thornton and Singh have performed a detailed analysis of protein side-chain interactions,[92] finding that both Arg and Lys are observed experimentally interacting with both Phe and Tyr in a way that places the charged groups in the region above the ring atomic sites rather than above the ring center. Where the OPLS model would fail to find such interactions, predicting them to be repulsive, the OPLS-CS model would be expected to perform much better.

The lack of quantitative agreement between the OPLS-CS and ab initio results might be considered a problem. Indeed, a force field that aspires to provide the most accurate possible account of aromatic interactions must reproduce accurately the true interaction energies. It must be acknowledged that there is still a considerable amount of work to be done before the OPLS-CS model can treat perfectly cation−π interactions. Learning from other cases, it is highly likely that a "correct" representation of these interactions will not be approached without the use of more sophisticated potential functions incorporating, for example, explicit polarization. This work can be seen as a step in the right direction rather than the ultimate solution to the problem of modeling cation−π interactions.

In spite of these caveats, it is possible to say that the OPLS-CS potential, previously derived for the modeling of aromatic liquids, offers a considerable improvement over the OPLS model for the modeling of cation−π interactions.

Conclusions

Monte Carlo calculations have been used to parametrize new OPLS-CS force field models for the aromatic molecules toluene, phenol, and pyridine. Using these models, an accurate reproduction of the available experimental data has been obtained.

The models developed have then been used to examine the structure of these aromatic liquids at the molecular level. Predictions of the structures have been made, but unfortunately, the lack of experimental structural data for these systems makes corroboration of these predictions impossible. Comparison with an existing all-atom potential (OPLS) reveals both similarities and differences. In toluene, both models predict an approximately equal distribution of parallel and perpendicular arrangements of molecules, the parallel structures probably stabilized by a combination of large dispersion interactions and the presence of a dipole moment.

In phenol and pyridine, the OPLS-CS model predicts a far more perpendicular structure than does the OPLS model. Both liquids are characterized by the formation of hydrogen bonds.

To test the transferability of the models developed, they have been used to evaluate the energies of a number of dimeric structures for toluene, phenol, and pyridine. It is found that the OPLS-CS model often overestimates the stability of the perpendicular dimer structures relative to the parallel dimer structures. If this energetic preference is carried over into the liquid, it is likely that the OPLS-CS model is overestimating the extent of perpendicular structure.

As a first test of the ability of the OPLS-CS model to model aromatic interactions in biology, the OPLS-CS models of toluene and phenol have been used to construct models of the aromatic amino acids phenylalanine and tyrosine, respectively. The interactions between these amino acids and the ammonium cation have then been investigated, and it has been shown that the OPLS-CS model considerably outperforms the OPLS model in reproducing the ab initio calculated potential energy surface for this interaction, most significantly in regions that are known experimentally to be

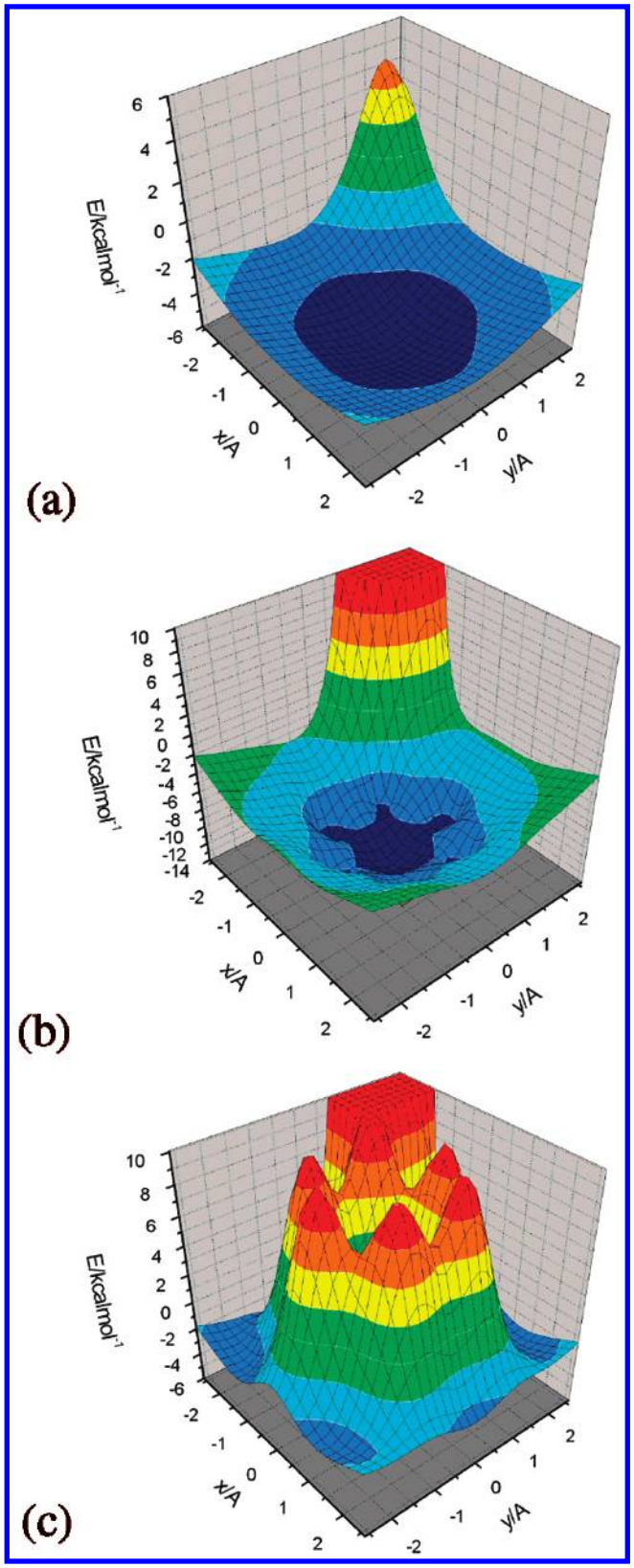

Figure 17. Potential-energy surface for Phe−ammonium interactions calculated using (a) MP2/6-311+G**, (b) the OPLS-CS model, and (c) the OPLS model. In all cases, $z =$ 3.0 Å.

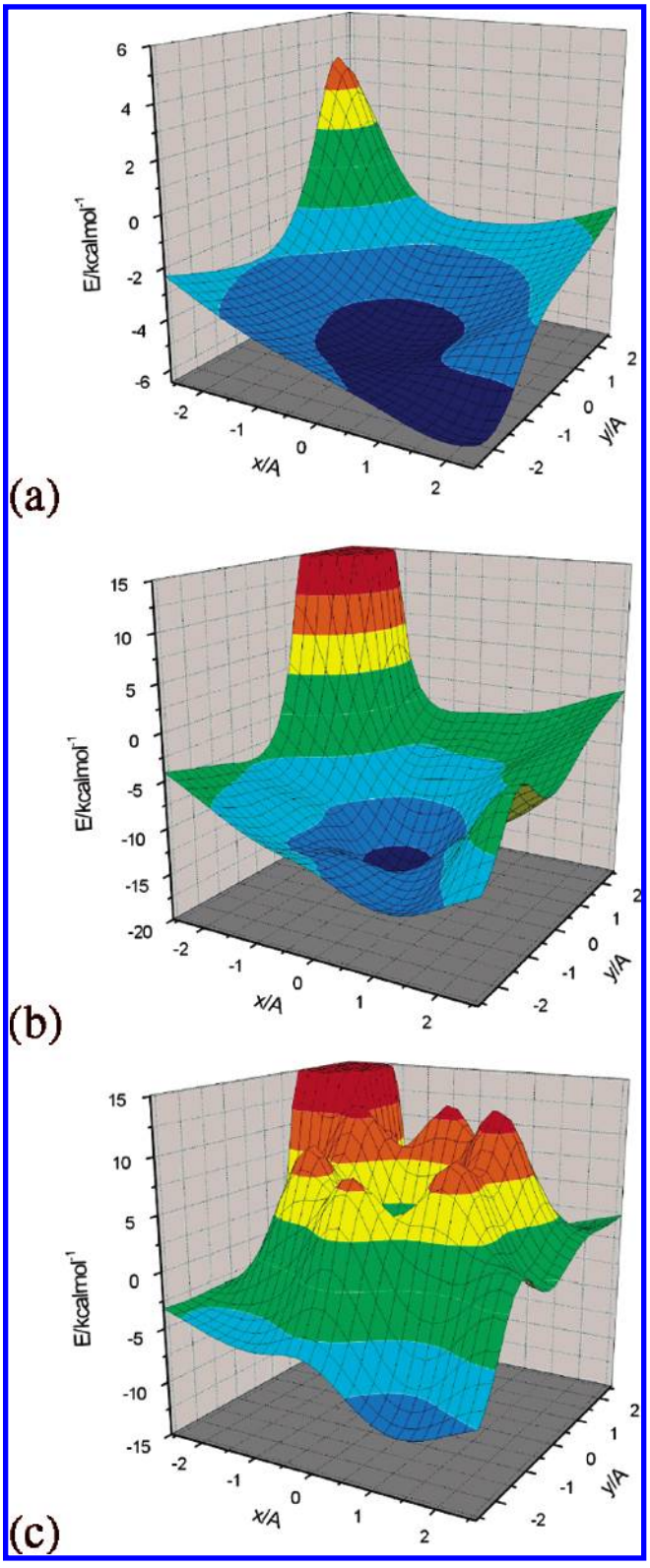

Figure 18. Potential-energy surface for Tyr−ammonium interactions calculated using (a) MP2/6-311+G**, (b) the OPLS-CS model, and (c) the OPLS model. In all cases, $z =$ 3.0 Å.

of biological significance. That the OPLS-CS model outperforms the atom-centered method in this case suggests that it might well find use in the modeling of aromatic interactions in biology, though further development will be necessary before this goal can be realized.

In tandem with the development of the OPLS-CS potential for phenol, an additional model has been developed as a reparameterization of the basic OPLS model. When this

reparameterized model is used, it is possible to obtain results similar to those obtained using the OPLS model—it is likely that the quality of the parametrization process is as important as the model used.

Overall, it has been shown that there are situations in which the OPLS-CS models can outperform the OPLS models. However, this is not always the case, and as with all empirical force fields, it is essential to choose the appropriate model to tackle the problem in question.

Acknowledgment. The authors thank Prof. W. L. Jorgensen for his generous provision of the BOSS program. C.M.B. thanks the National Foundation for Cancer Research for funding, and G.H.G. thanks Boehringer Ingelheim for a fellowship.

Supporting Information Available: Force field parameters for the OPLS-CS models of toluene, phenol, and pyridine. This information is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Sussman, J. L.; Harel, M.; Frolow, F.; Oefner, C.; Goldman, A.; Toker, L.; Silman, I. *Science* **1991**, *253*, 872.

(2) Hunter, C. A. *Philos. Trans. R. Soc. London, Ser. A* **1993**, *345*, 77.

(3) Vondrášek, J.; Bendová, L.; Klusák, V.; Hobza, P. *J. Am. Chem. Soc.* **2005**, *127*, 2615.

(4) Burley, S. K.; Petsko, G. A. *Science* **1985**, *229*, 23.

(5) Fowler, P. W.; Buckingham, A. D. *Chem. Phys. Lett.* **1991**, *176*, 11.

(6) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525.

(7) Baker, C. M.; Grant, G. H. *J. Chem. Theory Comput.* **2006**, *2*, 947.

(8) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(9) Chipot, C.; Jaffe, R.; Maigret, B.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 11217.

(10) Cheney, J.; Cheney, B. V.; Richards, W. G. *Biochim. Biophys. Acta* **1988**, *954*, 137.

(11) Scheiner, S.; Kar, T.; Pattanayak, J. *J. Am. Chem. Soc.* **2002**, *124*, 13257.

(12) Hobza, P.; Špirko, V.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem. A* **1998**, *102*, 2501.

(13) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.

(14) Gervasio, F. L.; Chelli, R.; Procacci, P.; Schettino, V. *J. Phys. Chem. A* **2002**, *106*, 2945.

(15) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. C.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(16) Schauer, M.; Bernstein, E. R. *J. Chem. Phys.* **1985**, *82*, 3722.

(17) Ishikawa, S.; Ebata, T.; Ishikawa, H.; Inoue, T.; Mikami, N. *J. Phys. Chem.* **1996**, *100*, 10531.

(18) Law, K. S.; Schauer, M.; Bernstein, E. R. *J. Chem. Phys.* **1984**, *81*, 4871.

(19) Musgrave, A.; Wright, T. G. *J. Chem. Phys.* **2005**, *122*, 074312.

(20) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 4209.

(21) Tsuzuki, S.; Honda, K.; Uchimara, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2002**, *124*, 104.

(22) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2003**, *107*, 8377.

(23) Sinnokrot, M. O.; Sherrill, C. D. *J. Am. Chem. Soc.* **2004**, *126*, 7690.

(24) Kim, J. H.; Lee, S. H. *Bull. Korean Chem. Soc.* **2002**, *23*, 441.

(25) Kim, J. H.; Lee, S. H. *Bull. Korean Chem. Soc.* **2002**, *23*, 447.

(26) Jorgensen, W. L.; Laird, E. R.; Nguyen, T. B.; Tirado-Rives, J. *J. Comput. Chem.* **1993**, *14*, 206.

(27) Kumar, A. V. A.; Yashonath, S.; Chaplot, S. L. *J. Chem. Phys.* **2000**, *113*, 8070.

(28) Fioroni, M.; Vogt, D. *J. Phys. Chem. B* **2004**, *108*, 11774.

(29) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596.

(30) Andre, D.; Fourme, R.; Bruneaux-Poulle, J.; Bosio, L. *J. Mol. Struct.* **1982**, *81*, 253.

(31) Anderson, M.; Bosio, L.; Bruneaux-Poulle, J.; Fourme, R. *J. Chim. Phys. Phys. Chim. Biol.* **1977**, *74*, 68.

(32) Hobza, P.; Riehn, C.; Weichert, A.; Brutschy, B. *Chem. Phys.* **2002**, *283*, 331.

(33) Sagarik, K.; Asawakun, P. *Chem. Phys.* **1997**, *219*, 173.

(34) Böhm, H.-J.; Ahlrichs, R. *J. Chem. Phys.* **1982**, *77*, 2028.

(35) Hartland, G. V.; Henson, B. F.; Venturo, V. A.; Felker, P. M. *J. Phys. Chem.* **1992**, *96*, 1164.

(36) Ebata, T.; Watanabe, T.; Mikami, N. *J. Phys. Chem.* **1995**, *99*, 5761.

(37) Connell, L. L.; Ohline, S. M.; Joireman, P. W.; Corcoran, T. C.; Fleker, P. M. *J. Chem. Phys.* **1992**, *96*, 2585.

(38) Parthasarathi, R.; Subramanian, V.; Sathyamurthy, N. *J. Phys. Chem. A* **2005**, *109*, 843.

(39) Mooney, D. A.; Müller-Plathe, F.; Kremer, K. *Chem. Phys. Lett.* **1998**, *294*, 135.

(40) Thornton, J. M.; MacArthur, M. W.; McDonald, I. K.; Jones, D. T.; Mitchell, J. B. O.; Nandi, C. L.; Price, S. L.; Zvelebil, M. J. J. M. *Philos. Trans. R. Soc. London, Ser. A* **1993**, *345*, 113.

(41) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233.

(42) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(43) Chelli, R.; Gervasio, F. L.; Procacci, P.; Schettino, V. *J. Am. Chem. Soc.* **2002**, *124*, 6133.

(44) Zavodnik, V. E.; Bel'skii, V. K.; Zorkii, P. M. *J. Struct. Chem.* **1987**, *28*, 793.

(45) Allan, D. R.; Clark, S. J.; Dawson, A.; McGregor, P. A.; Parsons, S. *Acta. Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **2002**, *58*, 1018.

(46) Mishra, B. K.; Sathyamurthy, N. *J. Phys. Chem. A* **2005**, *109*, 6.

(47) Piacenza, M.; Grimme, S. *Chem. Phys. Chem.* **2005**, *6*, 1554.

(48) Megiel, E.; Kasprzycka-Guttman, T.; Jagielska, A.; Wróblewska, L. *J. Mol. Struct.* **2001**, *569*, 111.

(49) Sagarik, K.; Spohr, E. *Chem. Phys.* **1995**, *199*, 73.

(50) Jorgensen, W. L.; McDonald, N. A. *THEOCHEM* **1998**, *424*, 145.

(51) Williams, D. E.; Weller, R. R. *J. Am. Chem. Soc.* **1983**, *105*, 4143.

(52) Gamba, Z.; Klein, M. L. *Chem. Phys.* **1989**, *130*, 15.

(53) Mootz, D.; Wussow, H.-G. *J. Chem. Phys.* **1981**, *75*, 1517.

(54) Anghel, A. T.; Day, G. M.; Price, S. L. *CrystEngComm* **2002**, *4*, 348.

(55) Dougherty, D. A. *Science* **1996**, *271*, 163.

(56) Urban, J. J.; Cramer, C. J.; Famini, G. R. *J. Am. Chem. Soc.* **1992**, *114*, 8226.

(57) Gallivan, J. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *17*, 9459.

(58) Tsou, L. K.; Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 14917.

(59) Scrutton, N. S.; Raine, A. R. C. *Biochem. J.* **1996**, *319*, 1.

(60) Wintjens, R.; Lievin, J.; Rooman, M.; Buisine, E. *J. Mol. Biol.* **2000**, *302*, 395.

(61) Zacharias, N.; Dougherty, D. A. *Trends Pharmacol. Sci.* **2002**, *23*, 281.

(62) Petersen, F. N. R.; Jensen, M. O.; Nielsen, C. H. *Biophys. J.* **2005**, *89*, 3985.

(63) Caldwell, J. W.; Kollmann, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4177.

(64) Mecozzi, S.; West, A. P., Jr.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10566.

(65) Mo, Y.; Subramanian, G.; Gao, J.; Ferguson, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 4832.

(66) Tsuzuki, S.; Uchimaru, T.; Mikami, M. *J. Phys. Chem. A* **2003**, *107*, 10414.

(67) Alberti, M.; Castro, A.; Lagana, A.; Moix, M.; Pirani, F.; Cappelletti, D.; Liuti, G. *J. Phys. Chem. A* **2005**, *109*, 2906.

(68) Rocha-Rinza, T.; Hernandez-Trujillo, J. *Chem. Phys. Lett.* **2006**, *422*, 36.

(69) Zhu, W.-L.; Tan, X.-J.; Puah, C. M.; Gu, J.-D.; Jiang, H.-L.; Chen, K.-X.; Felder, C. E.; Silman, I.; Sussman, J. L. *J. Phys. Chem. A* **2000**, *104*, 9573.

(70) Liu, T.; Zhu, W.; Gu, J.; Shen, J.; Luo, X.; Chen, G.; Puah, C. M.; Silman, I.; Chen, K.; Sussman, J. L.; Jiang, H. *J. Phys. Chem. A* **2004**, *108*, 9400.

(71) Xu, Y.; Shen, J.; Zhu, W.; Luo, X.; Chen, K.; Jiang, H. *J. Phys. Chem. B* **2005**, *109*, 5945.

(72) Lee, J. Y.; Lee, S. J.; Choi, H. S.; Cho, S. J.; Kim, K. S.; Ha, T.-K. *Chem. Phys. Lett.* **1995**, *232*, 67.

(73) Reddy, A. S.; Sastry, G. N. *J. Phys. Chem. A* **2005**, *109*, 8893.

(74) Chipot, C.; Maigret, B.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 2998.

(75) Tan, X. J.; Jiang, H. L.; Zhu, W. L.; Chen, K. X.; Ji, R. Y. *J. Chem. Soc., Perkin Trans. 2* **1999**, 107.

(76) Vaden, T. D.; Lisy, J. M. *J. Chem. Phys.* **2005**, *123*, 074302.

(77) Hunter, C. A.; Singh, J.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218*, 837.

(78) Price, S. L.; Stone, A. J. *J. Chem. Phys.* **1987**, *86*, 2859.

(79) Doerksen, R. J.; Thakkar, A. J. *J. Phys. Chem. A* **1999**, *103*, 10009.

(80) Stone, A. J. *J. Chem. Theory Comput.* **2005**, *1*, 1128.

(81) Hawkins, D. M. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1.

(82) Jorgensen, W. L. *BOSS 4.2*; Yale University: New Haven, CT, 2001.

(83) DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific: San Carlos, CA, 2002.

(84) Schaftenaar, G.; Noordik, J. H. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 123.

(85) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.

(86) Ransil, B. J. *J. Chem. Phys.* **1961**, *34*, 2109.

(87) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(88) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98, Rev. A.7*; Gaussian, Inc.: Pittsburgh, PA, 1998.

(89) Atkins, P. W. *Physical Chemistry*, 6th ed.; Oxford University Press: Oxford, U. K., 1998; p 106.

(90) Jorgensen, W. L. *J. Phys. Chem.* **1986**, *90*, 1276.

(91) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933.

(92) Singh, J.; Thornton, J. M. *Atlas of Protein Side-Chain Interaction*; IRL Press: Oxford, United Kingdom, 1992; pp 94−105, 494−505.

CT600218F