JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Application of Molecular Dynamics Simulations in Molecular Property Prediction. 1. Density and Heat of Vaporization

Junmei Wang*,† and Tingjun Hou*,‡

†Department of Pharmacology, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Boulevard, Dallas, Texas 75390-9050, United States

‡Institute of Nano & Soft Materials (FUNSOM) and Jiangsu Key Laboratory for Carbon-Based Functional Materials & Devices, Soochow University, Suzhou, Jiangsu 215123, P. R. China

Supporting Information

ABSTRACT: Molecular mechanical force field (FF) methods are useful in studying condensed phase properties. They are complementary to experiments and can often go beyond experiments in atomic details. Even if a FF is specific for studying structures, dynamics, and functions of biomolecules, it is still important for the FF to accurately reproduce the experimental liquid properties of small molecules that represent the chemical moieties of biomolecules. Otherwise, the force field may not describe the structures and energies of macromolecules in aqueous solutions properly. In this work, we have carried out a systematic study to evaluate the General AMBER Force Field (GAFF) in studying densities and heats of vaporization for a large set of organic molecules that covers the most common chemical functional groups. The latest techniques, such as the particle mesh Ewald (PME) for calculating electrostatic energies and Langevin dynamics for scaling temperatures, have been applied in the molecular dynamics (MD) simulations. For density, the average percent error (APE) of 71 organic compounds is 4.43% when compared to the experimental values. More encouragingly, the APE drops to 3.43% after the exclusion of two outliers and four other compounds for which the experimental densities have been measured with pressures higher than 1.0 atm. For the heat of vaporization, several protocols have been investigated, and the best one, P4/ntt0, achieves an average unsigned error (AUE) and a root-mean-square error (RMSE) of 0.93 and 1.20 kcal/mol, respectively. How to reduce the prediction errors through proper van der Waals (vdW) parametrization has been discussed. An encouraging finding in vdW parametrization is that both densities and heats of vaporization approach their "ideal" values in a synchronous fashion when vdW parameters are tuned. The following hydration free energy calculation using thermodynamic integration further justifies the vdW refinement. We conclude that simple vdW parametrization can significantly reduce the prediction errors. We believe that GAFF can greatly improve its performance in predicting liquid properties of organic molecules after a systematic vdW parametrization, which will be reported in a separate paper.

1. INTRODUCTION

Molecular dynamics simulation has become increasingly important in studying the structures, dynamics, and functions of biomolecules.[1] Nevertheless, it is also an important tool for predicting a variety of molecular properties.[2] The power of MD simulations lies in that it can provide atomic details on the dynamics of a simulation system under conditions that are beyond the scope of experiments. Even for some specific force fields, such as CHARMM,[3,4] AMBER,[5,6] OPLS,[7] and GROMOS,[8,9] which were developed mainly to study the biomolecules, it is often critical for them to reproduce the bulk properties of some small molecules that represent the moieties of biomolecules. With the continually increased demand of "ideal" molecular mechanical models to study ligand−protein and ligand−nucleic acid interactions, it is important to develop a general purpose force field that is totally compatible with one or a set of biomolecular force fields for studying the interactions between biomolecules and organic molecules. The General AMBER force field (GAFF)[10] was developed in this spirit to study biomolecule−ligand interactions in conjunction with the AMBER force fields.[5,6]

$$V_{\text{potential}} = \sum_{\text{bonds}} K_r(r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\varphi - \gamma)] + \sum_{i<j} \left\{ \left[\frac{Ai_j}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} \right] + \frac{q_i q_j}{\varepsilon R_{ij}} \right\} \quad (1)$$

To maintain the compatibility between GAFF and the biomolecular AMBER force fields, GAFF applies the same harmonic function form used by the AMBER additive force fields (eq 1), and it has been developed using the same force field parametrization strategy. The total potential energy is a sum of the intramolecular bonded terms (the first three terms) and intra- and intermolecular nonbonded terms (the last term). In eq 1, the first three terms represent the contribution to the total energy from bond stretching, bond angle bending, and torsion angle twisting, respectively. The last term is the sum of van der Waals

Received: February 24, 2011
Published: May 30, 2011

Table 1. List of the Experimental Densities (g cm^{-3}) and Heats of Vaporization (kcal/mol) for 71 Compounds (Temperatures are in °C)

no.	name	density	$T_{density}$	ΔH_{vap}	T_{vap}	# res[a]	# atom[a]
1	ethane	0.546	−88.63	3.51[c]	−88.63	196	8
2	propane	0.581	−42.07	4.49[d]	−42.07	216	11
3	butane	0.602	−0.5	5.02[e]	−0.5	196	14
4	isobutane	0.551	25	4.57	25	216	14
5	cyclohexane	0.774	25	7.89 ± 0.48[f]	25	180	18
6	propene	0.505[b]	25	3.40	−47.65	324	9
7	*trans*-2-butene	0.598	25	5.15	25	216	12
8	1,3-butadiene	0.615[b]	25	4.99	25	216	10
9	1-butyne	0.678	0	5.58	25	216	10
10	benzene	0.877	20	7.89 ± 0.48[f]	25	216	12
11	water	0.997	25	10.51	25	624	3
12	methanol	0.791	20	8.84 ± 0.48[f]	25	216	6
13	ethanol	0.789	20	10.04 ± 0.48[f]	25	245	9
14	propanol	0.800	25	11.35 ± 0.10[f]	25	216	12
15	2-propanol	0.781	25	10.85	25	180	12
16	2-methyl-2-propanol	0.781	25	11.14	25	180	15
17	phenol	1.055	45	13.82	25	180	13
18	*m*-cresol	1.034	20	14.75	25	180	16
19	methanethiol	0.888	5.96	5.87	5.96	180	6
20	ethanethiol	0.832	25	6.52	25	245	9
21	propanethiol	0.836	25	7.62	25	216	12
22	dimethyl ether	0.735	−24.6	5.14	−24.6	216	9
23	ethyl methyl ether	0.721	7.35	5.91	7.35	196	12
24	diethyl ether	0.714	20	6.48	25	180	15
25	dimethoxymethane	0.854	25	6.90	25	180	13
26	dimethyl sulfide	0.848	20	6.61	25	196	9
27	ethyl methyl sulfide	0.837	25	7.61	25	180	12
28	dimethyl disulfide	1.057	25	9.06[c]	25	180	10
29	acetaldehyde	0.783	18	6.24[g]	25	216	7
30	propanal	0.791	25	7.10 ± 0.05[f]	25	216	10
31	acetone	0.785	25	7.47[c]	25	180	10
32	butanone	0.800	25	8.35[c]	25	216	13
33	acetic acid	1.045	25	12.33[c]	25	180	8
34	propanoic acid	0.988	25	13.15[c]	25	256	11
35	methyl formate	0.971	20	6.93 ± 0.48[f]	25	196	8
36	methylamine	0.656[b]	25	5.59	25	245	7
37	propylamine	0.717	20	7.47	25	147	13
38	dimethyl amine	0.680	0	6.08[c]	25	245	10
39	diethylamine	0.706	20	7.48	25	192	16
40	trimethylamine	0.627[b]	25	5.18	25	180	13
41	triethylamine	0.728	20.000	8.33	25	216	22
42	aniline	1.022	20	12.91 ± 0.96[f]	25	180	14
43	acetonitrile	0.786	20	7.87	25	196	6
44	nitromethane	1.137	20	9.17[c]	25	216	7
45	N-methylformamide	1.011	19	13.43	25	216	9
46	N,N-dimethyl formamide	0.945	25	11.37[h]	25	216	12
47	acetamide	0.981	100	13.40	221.15	245	9
48	N-methyl acetamide	0.894	100	13.30	100	196	12
49	N-ethylacetamide	0.942	4			196	15
50	N,N-dimethylacetamide	0.936	25	11.75	25	216	15
51	N-methylpropanamide	0.931	25	15.50	25	180	15
52	dimethyl sulfoxide	1.101	25	10.30	189	196	10
53	methyl phosphate	1.214	20			216	17

dx.doi.org/10.1021/ct200142z |*J. Chem. Theory Comput.* 2011, 7, 2151–2165

Table 1. Continued

no.	name	density	$T_{density}$	ΔH_{vap}	T_{vap}	# res[a]	# atom[a]
54	triphenyl phosphine oxide	1.212	23			120	34
55	fluorobenzene	1.023	20	8.29[c]	25	180	12
56	trichloromethane	1.479	25	7.48	25	216	5
57	tetrachloromethane	1.594	20	7.75	25	180	5
58	bromomethane	1.676	20	5.45	25	216	5
59	iodomethane	2.279	20	6.68	25	216	5
60	tetrahydrofuran	0.883	25	7.65	25	216	13
61	1,3-dioxolane	1.060	25	8.48[i]	25	216	11
62	pyrrolidine	0.859	20	8.97[c]	25	180	14
63	morpholine	1.001	20	8.87	128	216	15
64	furan	0.951	20	6.62[c]	25	216	9
65	thiophene	1.065	20	8.29	25	180	9
66	pyrrole	0.970	20	10.84[c]	25	216	10
67	pyridine	0.982	20	9.56 ± 0.72[f]	25	180	11
68	pyridazine	1.104	23	12.78	25	216	10
69	4-methyl thiazole	1.112	25	10.48	25	216	11
70	quinoline	1.098	15	14.18[j]	25	180	17
71	dimethyl phosphate	1.323	20			216	14

[a] # res, the number of residues (# res) in a simulation box; # atom, the number of atoms in a residue. [b] Densities were measured with pressures larger than 1 atm (propene, 1,3-butadiene, methylamine, trimethylamine). [c] Ref 23. [d] Ref 24. [e] Ref 25. [f] Average value of several pieces of experimental data, specifically, 10 data points of cyclohexane, 7 for benzene, 6 for methanol, 7 for ethanol, 8 for propanol, 8 for propanal, 6 for methyl formate, 7 for aniline, and 7 for pyridine. The individual data points are adopted from http://webbook.nist.gov/. [g] Ref 26. [h] Ref 27. [i] Ref 28. [j] Ref 29.

and electrostatic energies. A_{ij} and B_{ij} are the Lennard-Jones parameters for repulsion and attraction, which can be expressed in terms of effective van der Waals radii and well depths, R_{ij}^* and ε_{ij} (eqs 2a and 2b), which are further obtained from atomic parameters using mixing rules (eqs 2c and eq 2d). The point charges q_i and q_j are derived to reproduce the *ab initio* electrostatic potential (ESP).

$$A_{ij} = \varepsilon_{ij}(R_{ij}^*)^{12} \quad (2a)$$

$$B_{ij} = 2\varepsilon_{ij}(R_{ij}^*)^6 \quad (2b)$$

$$R_{ij}^* = R_i^* + R_j^* \quad (2c)$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \quad (2d)$$

The most striking difference between AMBER and other force fields is that the atomic partial charges in AMBER are derived to reproduce *ab initio* ESP using the program RESP (Restrained Electrostatic Potential).[11,12] The quantum-mechanics-based charge model is more general and generalizable than a purely empirical approach which automatically includes inherent polarization of the molecules relative to their gas phase charge distribution in an average way. Fortunately, it was found that the 6-31G* basis set enhances the polarity of liquids by about the same magnitude relative to gas phase moments as does the empirical charge models.[13,14] There are two types of charges used in GAFF, the HF/6-31G* RESP[11] and the AM1-BCC.[15,16] The former is the same charge method used to develop AMBER parm94 and parm99 force fields; on the other hand, AM1-BCC, a fast charge method that was parametrized to mimic HF/6-31G* RESP charges, is more suitable for the study of a large number of molecules, such as in virtual screenings. The HF/6-31G* RESP charge model was utilized to develop GAFF in order to maximize its compatibility with AMBER biomolecular force fields.

As van der Waals parameters generally are dominated by the inner closed shell of electrons, they are typically transferable, and only one set of parameters is needed for different atom types of an element, irrespective of its chemical environment (note that hydrogen is a special case, as it has no inner shell of electrons at all).[13] Therefore, most of the van der Waals parameters in GAFF are adopted from the AMBER biomolecular force fields even for the newly introduced chemical functional groups. For the "hard" force field parameters, the equilibrium bond lengths r_{eq} and bond angles θ_{eq} were obtained through statistical analysis on a large number of optimized structures at the MP2/6-31G* level. The torsional angle potentials were optimized to reproduce both the experimental relative conformational energies and the high-level *ab initio* rotational profiles of model compounds.

GAFF has been extensively tested to predict the structures and energies for a variety of molecules, and an overall satisfactory performance has been achieved. Those tests include a comparison between the GAFF minimized and the crystallographic structures, the interaction energies between nucleic acid base pairs, and the relative energies of a set of conformational pairs. Moreover, GAFF has also been extensively evaluated to calculate the binding free energies by its users.[17,18] Mobley et al. used GAFF to predict the hydration free energies in an explicit solvent for a large set of molecules.[19]

However, GAFF has not been evaluated on predicting the bulk properties of liquids. In this paper series, we set out to systematically study several important molecular properties using GAFF, which include the density, heat of vaporization, isobaric heat capacity, isochoric heat capacity, isothermal compressibility, thermal expansion coefficient, static dielectric constant, free energy of hydration, diffusion constant, rotational correlation time, etc. There are two objectives of this series. First of all, we

hope to establish a set of computational protocols to predict the above molecular properties using GAFF through MD simulations. Second, GAFF will be thoroughly evaluated in this procedure, and the hints of the modification of force field parameters to reduce the prediction errors can guide us to develop a new version of general AMBER force fields.

In this study, we focus on two molecular properties, density and heat of vaporization. Why do we study these two molecular properties? First of all, bulk density and heat of vaporization are almost exclusively associated with the nonbonded interactions.[7,20] The successful optimization of nonbonded force field parameters (as the atomic partial charges in GAFF are obtained in a deterministic fashion, the nonbonded parameters mainly refer to the van der Waals parameters) to reproduce the two condensed phase data sets lays the groundwork for a molecular mechanical force field including those aimed to study biomolecules.

2. METHODS

Data Sources. In Table 1, the compound names and the experimental mass densities (ρ) and heats of vaporization (ΔH_{vap}) are listed. The 71-molecule data set covers diverse functional groups in organic chemistry, which includes hydrocarbons (aliphatic and aromatic, cyclic and acyclic), alcohols, thiols, phenols, ethers, esters, aldehydes, ketones, carboxylic acids, amines, amides, nitriles, nitro-derivatives, disulfides, thiophenes, sulfides, sulfoxides, sulfones, phosphates, halides, and heterocyclic compounds. The structures of the 71 molecules are shown in Figure S1 of the Supporting Information. The experimental values of density are adopted from the CRC Handbook of Chemistry and Physics (edition 86).[21]

Heat of vaporization (ΔH_{vap}) is defined as the enthalpy change in the conversion of 1 mol of liquid to a gas at constant temperature. In experiments, the heat of vaporization can be measured at the boiling point of the neat liquid through calorimetry or determined from the vapor pressure−temperature (P−T) plot using the Clausius−Clapeyron equation.[20] When high-quality P−T data over a wide temperature range are available, it is possible to derive the force field parameters to fit those data.[22] All of the experimental data of ΔH_{vap} are adopted from the CRC Handbook of Chemistry and Physics (edition 86) unless explicitly noted. Other experimental data resources include the publications of Majer and Svoboda,[23] Kemp and Egan,[24] Pedley et al.,[25] Wiberg et al.,[26] Geller,[27] Pihlaja and Heikklia,[28] and Steele et al.[29] When multiple experimental data are available for a compound, the overall guidance is to use the latest one. However, for nine compounds (nos. 5, 10, 12−14, 30, 35, 42, and 67 of Table 1), which have multiple data points (6−10 values) from different sources, the average values are used. For those compounds, the average RMS error is 0.47 kcal/mol and the largest RMS error is 0.96 kcal/mol for aniline.

Similar to QSAR (quantitative structure−activity relationship) studies,[30] it is important to use high-quality experimental data in force field development. Mackerell et al.[20] provided an excellent example on how to cherry-pick the most reliable experimental data. They found that the wide range of heat of vaporization values of N-methylacetamide reported in the literature is due to two basic things: incorrect reporting of the temperatures at which the original values were extracted and limitations in the quality of experimental vapor pressure−temperature data over a wide range of temperatures. We also noticed that the experimental data of ΔH_{vap} for acetic acid and propanoic acid at 25 °C adopted by the CRC Handbook of Chemistry and Physics (edition 86) are much smaller than from the other resources. For acetic acid, the CRC value of 23.36 kJ/mol is less than half of the values reported by Majer and Svoboda (51.6 kJ/mol)[23] and Konicek and Wadso (51.6 ± 1.5 kJ/mol);[31] for propanoic acid, the CRC value of 32.14 kJ/mol is also much smaller than the values reported by Majer and Svoboda (55.0 kJ/mol)[23] and Konicek and Wadso (55.0 ± 2.0 kJ/mol).[31] We adopted the values of Majer and Svoboda in this work, not only because they are relatively newer, but also because the calculated heats of vaporization by both GAFF and OPLS reproduce their experimental values well.

Molecular Mechanical Models. Consistent with the strategy of parametrizing GAFF, the point charges of 71 molecules in the data set were derived by RESP to fit the HF/6-31G* electrostatic potentials which were generated using the Gaussian 03 software package.[32] The other force field parameters come from GAFF in AMBER10.[33] The residue topology files were prepared using the Antechamber module[34] in AMBER 10.[33] For each molecule, an internal program was used to generate a rectangular parallelepiped box filled with multiple copies of the monomer (N_{res} in Table 1), and then the Leap program in AMBER10 was applied to generate the topologies.

Liquid Phase Molecular Dynamics Simulations. All of the liquid phase MD simulations were performed with the periodic boundary condition to produce isothermal−isobaric ensembles using the Sander program of AMBER10.[33] The Particle Mesh Ewald (PME) method[35−37] was used to calculate the full electrostatic energy of a unit cell in a macroscopic lattice of repeating images. Except for TIP3P water, which is described with a special "three-point" algorithm, all degrees of freedom were constrained;[38] the other molecules had all the degrees of freedom free in MD simulations. The nonbonded cutoff of calculating van der Waals and electrostatic energies was set to 9.0 Å, and a continuum model correction term was added to the van der Waals energies. How nonbonded cutoff affects the MD simulations was discussed by Shirts et al.[39]

The integration of the equations of motion was conducted at a time step of 1 fs (except TIP3P water, for which a time step of 2 fs was used). Temperature was regulated using the Langevin dynamics[40] with a collision frequency of 5 ps.[41−43] Pressure regulation was achieved with isotropic position scaling, and the pressure relaxation time was set to 1.0 ps.

There are three phases in a liquid phase MD simulation, namely, the relaxation phase, the equilibrium phase, and the sampling phase. In the relaxation phase, the main chain atoms were gradually relaxed by applying a series of restraints, and the force constants decreased progressively: from 20 to 10, 5, and 1.0 kcal/mol/Å^2. For each force constant, the position-restrained MD simulation was run for 20 ps. In the following equilibrium phase, the system was further equilibrated for 2 ns without any restraint and constraint. In the sampling phases, the snapshots as well as the structural and energetic properties were recorded at an interval of 2 ps. In total, 1500 snapshots were saved for post analysis.

Gas Phase MD Simulations. Gas phase MD simulations were performed for single molecules using four different temperature regulation methods, namely, ntt (the temperature scaling keyword in AMBER) = 0, which corresponds to constant total energy classic dynamics; ntt = 1, which uses the weak-coupling algorithm;[44] ntt = 2, which utilizes the Andersen coupling

scheme;[45,46] and ntt = 3, which applies the Langevin dynamics with a collision frequency of 5 ps^{-1} to scaling temperatures.[40] The systems were equilibrated for 1 ns followed by another 1 ns sampling phase, and in total 500 snapshots were recorded for post analysis.

Hydration Free Energy Calculations. It is a standard practice in today's force field development to test if molecular mechanical models can reproduce the hydration free energies of model compounds. Although GAFF has been extensively evaluated by Mobley et al.[19] for this molecular property using a large data set, we need to find out how the newly derived vdW parameters affect the calculation performance. The hydration free energy of a molecule was calculated using thermodynamic integration (TI). In TI calculations, the system evolves according to a mixed potential, $V(\lambda) = (1-\lambda)^k V_0 + [1-(1-\lambda)^k]V_1$, where λ and k are mixing parameters and V, V_0, and V_1 are the mixed, the unperturbed, and perturbed potentials, respectively. The free energy change, ΔG, is calculated numerically using the following equation: $\Delta G = G_{\lambda=1} - G_{\lambda=0} = \int_0^1 \langle \partial V/\partial\lambda\rangle_\lambda \, d\lambda = \Sigma_i w_i \langle \partial V/\partial\lambda\rangle_i$. The free energy of hydration of a molecule was calculated by summing up the free energy changes in four perturbations, i.e., the gas phase and aqueous phase charge disappearing and the gas phase and aqueous phase atom disappearing. For the charge disappearing, a linear mixing rule was applied (k = 1); for the atom disappearing, k was set to 6, as suggested by Steinbrecher et al.[47] In each free energy perturbation, 12 different free energy simulations with the weights (w_i) fitting the Gaussian quadrature formula were performed to numerically estimate the integral.[48] A similar MD protocol to the above liquid MD simulation was applied to do the free energy simulations, except that the time step of integration was set to 2 fs and the systems were equilibrated for 1 ns followed by production for another nanosecond.

Density Calculations. The average bulk density ρ was computed from the average volume of the simulation box, $\langle V\rangle$ using eq 3, where N_{res} is the number of residues in the simulation box, M is molar mass of the molecule in the study, and N_A is the Avogadro constant. The bulk densities were printed out in the output files of MD simulations by Sander in default.

$$\langle \rho \rangle = \frac{N_{res} M}{N_A \langle V \rangle} \quad (3)$$

Heat of Vaporization. The heat of vaporization or the enthalpy of vaporization ΔH_{vap} can be calculated with eq 4, where $H(p,T)_{gas}$ and $H(p,T)_{liquid}$ are the enthalpies in the gas and liquid phases, respectively. E_{gas} and E_{liquid} are the total energies of the gas and liquid phases, respectively.

$$\Delta H_{vap}(T) = H(p,T)_{gas} - H(p,T)_{liquid}$$
$$= E_{gas}(T) - E_{liquid}(T) + p(V_{gas} - V_{liquid}) \quad (4)$$

V_{liquid} is negligible compared to V_{gas}. Under the assumption that the gas is ideal so that the kinetic energies of a molecule in the gas and liquid phases are identical, eq 4 becomes eq 5, where $E_{gas}^{potential}$ and $E_{liquid}^{potential}$ are the potential energies in the gas and liquid phases, respectively; C is the correction term.

$$\Delta H_{vap}(T) = E_{gas}^{potential}(T) - E_{liquid}^{potential}(T) + RT + C \quad (5)$$

Both of the potential energy terms include the vibrational energies which are obtained through molecular simulations.

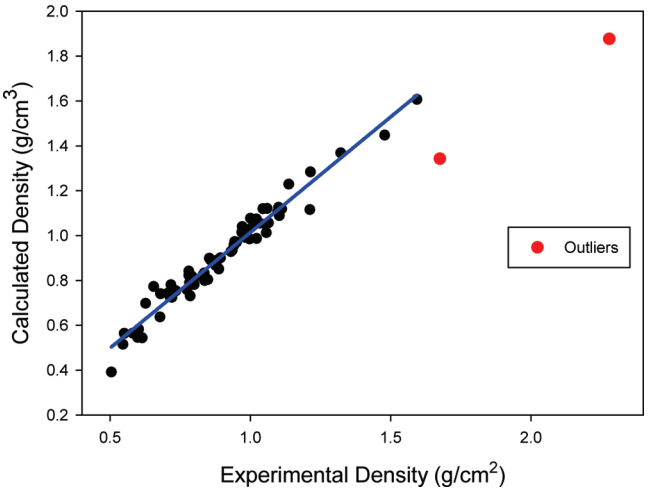

Figure 1. Correlation between experimental and calculated densities for 71 molecules. The two outliers are bromomethane and iodomethane. With the new van der Waals parameters of halides, the experimental densities can be well reproduced: 1.642 versus 1.676 for bromomethane and 2.247 versus 2.279 for iodomethane, respectively.

Note that in MD simulations the actual temperatures T_{MD} may be slightly different from the desirable temperature T. $E_{gas}^{potential}$ can also be estimated using eq 6, where $E_{gas}^{minimized}$ is the minimized energy and N_{atom} and N_{cons} are the number of atoms in the molecules and the number of the constrained degrees of freedom, respectively. Equation 6 has been used by Caldwell and Kollman[13] and Fox and Kollman[14] in calculating the heat of vaporization. $E_{liquid}^{potential}$ consists of two terms, the intramolecular energy E_{liquid}^{intra} and intermolecular energy E_{liquid}^{inter}.

$$E_{gas}^{potential}(T) = E_{gas}^{minimized} + \frac{1}{2}RT(3N_{atom} - 6 - N_{cons}) \quad (6)$$

Equation 5 is the basic equation for calculating the heat of vaporization.[7,49−54] The correction term C in eq 5 accounts for the difference in vibration energies calculated quantum mechanically and classically, as well as the polarization and nonideal gas effects. More details on how to calculate the correction term are presented by Horn et al.[55] In most scenarios, the correction term C is rather small, and it is usually neglected. For example, the quantum correction of the vibrational energies, the major contribution to the correction term C, is estimated to be −0.055 kcal/mol for water.[53]

If the correction term is neglected, and all degrees of freedoms are frozen in MD simulations or the assumption that the intramolecular energy in the liquid phase is the same as that in the gas phase is applied, eq 5 can be further simplified to eq 7. We used eq 7 to calculate the heat of vaporization of TIP3P water.

$$\Delta H_{vap}(T) = E_{liquid}^{inter}(T) + RT \quad (7)$$

In this work, several computation protocols were applied to calculate the heats of vaporization of 67 molecules in the data set. In the first protocol (P1), eq 7 was used to calculate ΔH_{vap}, assuming the energies of the residue are the same in the gas and the liquid phases. In the second protocol (P2), eqs 5 and 6 were applied to calculate ΔH_{vap}. The third protocol, P3, is similar to P2 except that $E_{gas}^{minimized}$ is an average value of minimized energies of multiple conformations sampled in gas phase MD simulations. In the last protocol, P4, ΔH_{vap} was calculated using eq 8, a hybrid

Table 2. Calculation Results of Heats of Vaporization (kcal/mol) for 67 Compounds Using Four Different Basic Protocols (P1−P4)[a]

						P3		P4	
no.	expt	T_{Expt}	T_{MD}[b]	P1	P2	NTT=0	NTT=3	NTT=0	NTT=3
1	3.51	−88.63	−88.31	3.31	2.68	2.68	2.91	2.68	3.31
2	4.49	−42.07	−41.64	4.52	3.43	3.43	3.88	3.43	4.47
3	5.02	−0.5	0.04	5.51	4.42	4.42	5.18	3.84	5.38
4	4.57	25	25.42	5.20	3.48	3.48	4.37	3.48	5.20
5	7.89	25	25.50	8.60	6.49	6.49	7.49	6.49	8.56
6	3.40	−47.65	−47.38	3.90	3.06	3.06	3.39	3.06	3.91
7	5.15	25	25.50	4.60	3.08	3.08	3.83	3.08	4.59
8	4.99	25	25.52	3.97	2.69	2.69	3.32	2.69	3.82
9	5.58	25	25.32	5.02	4.20	4.20	4.84	4.20	5.03
10	7.89	25	25.25	7.59	6.38	6.38	7.08	6.38	7.60
11	10.51	25	25.19	10.76	10.76	10.76	10.76	10.76	10.76
12	8.84	25	25.07	10.76	9.72	9.72	10.11	9.72	10.38
13	10.04	25	25.49	11.85	10.15	10.15	10.54	10.45	11.47
14	11.35	25	25.54	13.94	12.31	11.93	12.69	11.94	13.36
15	10.85	25	25.49	14.28	12.44	12.50	13.29	12.50	13.87
16	11.14	25	25.48	15.15	13.23	13.23	14.08	13.23	14.87
17	13.82	25	25.57	13.27	11.98	11.98	12.65	11.98	13.24
18	14.75	25	25.45	17.25	15.17	15.18	15.90	15.37	16.61
19	5.87	5.96	6.11	4.88	4.26	4.26	4.44	4.26	4.89
20	6.52	25	24.99	6.06	4.95	4.95	5.42	5.00	6.04
21	7.62	25	25.55	7.37	5.72	5.72	6.50	5.88	7.30
22	5.14	−24.6	−24.27	5.55	4.64	4.64	4.97	4.64	5.55
23	5.91	7.35	7.79	6.34	4.81	4.81	5.50	4.99	6.42
24	6.48	25	25.50	7.39	5.24	5.24	6.15	5.55	7.41
25	6.90	25	25.40	8.43	9.28	7.19	8.55	6.49	8.14
26	6.61	25	25.43	5.53	4.44	4.44	4.79	4.44	5.52
27	7.61	25	25.30	6.76	5.23	5.45	6.28	5.33	6.77
28	9.06	25	25.40	7.74	6.64	6.64	7.26	6.64	7.77
29	6.24	25	25.49	7.14	6.57	6.57	6.87	6.58	7.13
30	7.10	25	25.38	8.30	7.36	7.36	7.81	7.36	8.31
31	7.47	25	25.45	8.16	7.20	7.20	7.74	7.20	8.11
32	8.35	25	25.45	8.94	7.57	7.68	8.31	7.71	8.91
33	12.33	25	24.83	14.63	13.38	13.38	13.70	13.38	13.94
34	13.15	25	25.57	16.41	14.95	14.45	15.16	14.59	15.57
35	6.93	25	25.32	9.06	8.19	8.19	8.61	8.19	9.01
36	5.59	25	25.60	8.78	7.47	7.47	7.86	7.47	8.42
37	7.47	25	25.53	11.45	9.25	9.25	9.91	9.28	10.95
38	6.08	25	25.38	7.76	6.37	6.37	7.06	6.37	7.66
39	7.48	25	25.45	9.08	6.90	6.90	7.84	7.04	9.11
40	5.18	25	25.72	6.24	4.60	4.60	5.41	4.60	6.19
41	8.33	25	25.62	9.66	8.44	6.78	8.19	6.75	9.60
42	12.91	25	25.62	13.64	12.16	12.16	12.73	12.16	13.50
43	7.87	25	25.50	7.56	7.19	7.19	7.30	7.19	7.57
44	9.17	25	25.13	12.94	12.66	12.67	12.68	12.67	12.82
45	13.43	25	25.43	13.93	13.28	13.15	13.44	13.16	13.85
46	11.37	25	25.73	12.24	11.48	11.44	11.79	11.44	12.22
47	13.40	221.15	222.09	13.14	11.19	11.19	11.98	11.20	12.76
48	13.30	100	100.82	13.82	12.57	12.58	13.26	12.57	13.78
50	11.75	25	25.56	12.47	11.13	11.11	11.92	11.11	12.46
51	15.50	25	25.55	15.45	14.07	14.12	14.59	14.12	15.36
52	10.30	189	189.74	10.50	8.84	8.84	9.71	8.84	10.47

dx.doi.org/10.1021/ct200142z |J. Chem. Theory Comput. 2011, 7, 2151–2165

Table 2. Continued

						P3		P4	
no.	expt	T_{Expt}	T_{MD} [b]	P1	P2	NTT=0	NTT=3	NTT=0	NTT=3
55	8.29	25	25.35	7.80	6.70	6.70	7.21	6.70	7.77
56	7.48	25	25.31	7.03	6.84	6.84	6.88	6.84	7.01
57	7.75	25	25.18	8.09	8.02	8.02	8.03	8.02	8.07
58	5.45	25	25.24	4.39	3.89	3.89	4.06	3.89	4.36
59	6.68	25	25.17	5.14	4.65	4.65	4.98	4.65	5.13
60	7.65	25	25.61	8.79	7.59	7.58	8.28	7.59	8.81
61	8.48	25	25.37	10.66	9.49	9.49	10.03	9.49	10.56
62	8.97	25	25.46	10.54	9.01	8.75	9.51	8.88	10.46
63	8.87	128	128.57	12.24	9.66	9.66	10.68	9.66	11.87
64	6.62	25	25.44	7.22	6.40	6.40	6.90	6.40	7.25
65	8.29	25	25.52	7.73	6.95	6.95	7.37	6.95	7.73
66	10.84	25	25.19	12.46	11.36	11.36	11.93	11.36	12.42
67	9.56	25	25.37	9.87	8.81	8.81	9.24	8.81	9.86
68	12.78	25	25.07	13.24	12.37	12.37	12.91	12.37	13.25
69	10.48	25	25.48	10.97	10.40	10.19	10.68	10.20	10.94
70	14.18	25	25.28	14.74	13.18	13.18	13.92	13.18	14.72

[a] For protocols P3 and P4, only the results of using temperature regulation methods of ntt = 0 and ntt = 3 in gas phase MD are presented. Compounds 49, 53, 54, and 71 are excluded because of a lack of experimental data. [b] T_{MD} is the temperature of MD simulation in the liquid phase (°C)

equation of eqs 5 and 6, where $\Delta T = T_{liquid}^{MD} - T_{gas}^{MD}$ is the difference between the temperatures of MD simulations in the liquid and gas phases. Note that T is the desired temperature, and T_{liquid}^{MD} and T_{gas}^{MD} are the actual temperatures in the liquid phase and gas phase MD simulations.

$$\Delta H_{vap}(T) = E_{gas}^{potential}(T) - E_{liquid}^{potential}(T) + \frac{1}{2}R\Delta T(3N_{atom} - 6 - N_{cons}) + RT \quad (8)$$

With the above-discussed MD protocols, T_{liquid}^{MD} is essentially close to T, but T_{gas}^{MD} may deviate from T for some temperature regulation protocols. For the last two basic protocols, P3 and P4, each has four subprotocols corresponding to the four temperature regulation methods of the gas phase MD simulations (ntt = 0, 1, 2 and 3). For convenience, we use abbreviations to describe the subprotocols in the following text. For example, the abbreviation "P4/ntt0" describes the computation method of ΔH_{vap} using Protocol P4 in conjunction with ntt = 0 in gas phase MD simulations. Other abbreviations, P3/ntt0, P3/ntt1, P3/ntt2, P3/ntt3, P4/ntt1, P4/ntt2, and P4/ntt3, can be interpreted in a similar way.

Statistical Uncertainty Estimation. The density and most energetic terms in heat of vaporization calculations are ensemble averages. The uncertainty of a term (densities, temperatures, and energies) was estimated by the RMS deviation of a series of accumulated means. For the liquid phase terms, the means were calculated using the first 1000, 1025, 1050, 1075, 1100, ...1500 snapshots; for the gas phase terms, the means were calculated using the first 250, 260, 270, ...500 snapshots.

3. RESULTS AND DISCUSSION

Given the fact that GAFF inherited its van der Waals parameters from the AMBER biomolecular force fields, it is important to systematically assess its performance in reproducing some important bulk properties. Rather than studying many molecular properties for a few compounds,[7,13,14,56] we adopt a different presentation style, which studies one or two molecular properties for a large number of compounds. In this work, we focused on density and heat of vaporization, the two most important bulk properties in van der Waals parametrization.

Density Calculations. The calculated densities of 71 molecules are listed in Table S1 of the Supporting Information. The correlation between the experimental and the calculated densities is shown in Figure 1. Bromomethane and iodomethane are recognized as outliers who have the largest prediction errors. This is not a surprise at all since the van der Waals (vdW) parameters of bromide and iodide in GAFF are borrowed from other force fields. The average unsigned error (AUE), root-mean-square error (RMSE), and average percent error (APE) of the whole data set are 0.0436 g/cm^{-3}, 0.0756 g/cm^{-3}, and 4.5%, respectively. The total assigned density difference $\Sigma(\rho_{expt} - \rho_{calc})$ is 0.0169 for 71 molecules, suggesting the symmetrical error in our calculations is minimal.

Exclusion of the two outliers, bromomethane and iodomethane, leads to a significant decrease of AUE, RMSE, and APE, which now are 0.0342, 0.0440, and 4.1%, respectively. It is worth noting that the experimental densities of propene, 1,3-butadiene, methylamine, and trimethylamine were measured with a pressure larger than 1 atm. If we exclude these four molecules, the prediction performance is further improved (AUE = 0.0306, RMSE = 0.0386, APE = 3.430).

Heat of Vaporization. Four basic protocols and 10 subprotocols (P1, P2, P3/ntt0, P3/ntt1, P3/ntt2, P3/ntt3, P4/ntt0, P4/ntt1, P4/ntt2, and P4/ntt3) have been explored to predict the heats of vaporization of 67 compounds. The calculated heats of vaporization of the 10 subprotocols are listed in Table 2 (P1, P2, P3/ntt0, P3/ntt3, P4/ntt0, and P4/ntt3) and Table S2 (Supporting Information; P3/ntt1, P3/ntt2, P4/ntt1, and P4/ntt2). The individual energetic terms as well as the MD temperatures are listed in Table S3 (Supporting Information).

The performance of calculation is summarized in Table 3. Several interesting conclusions can be drawn from this table. First

Table 3. Performance of Heat of Vaporization Calculations[a]

			P3				P4			
protocol	P1	P2	ntt=0	ntt=1	ntt=2	ntt=3	ntt=0	ntt=1	ntt=2	ntt=3
AUE	1.184	1.081	1.061	1.071	1.053	1.069	0.932	2.410	1.019	1.095
RMSE	1.560	1.320	1.285	1.292	1.276	1.288	1.201	3.541	1.303	1.402
R^2	0.864	0.864	0.875	0.878	0.878	0.877	0.880	0.642	0.879	0.877

[a] All energies are in kcal/mol.

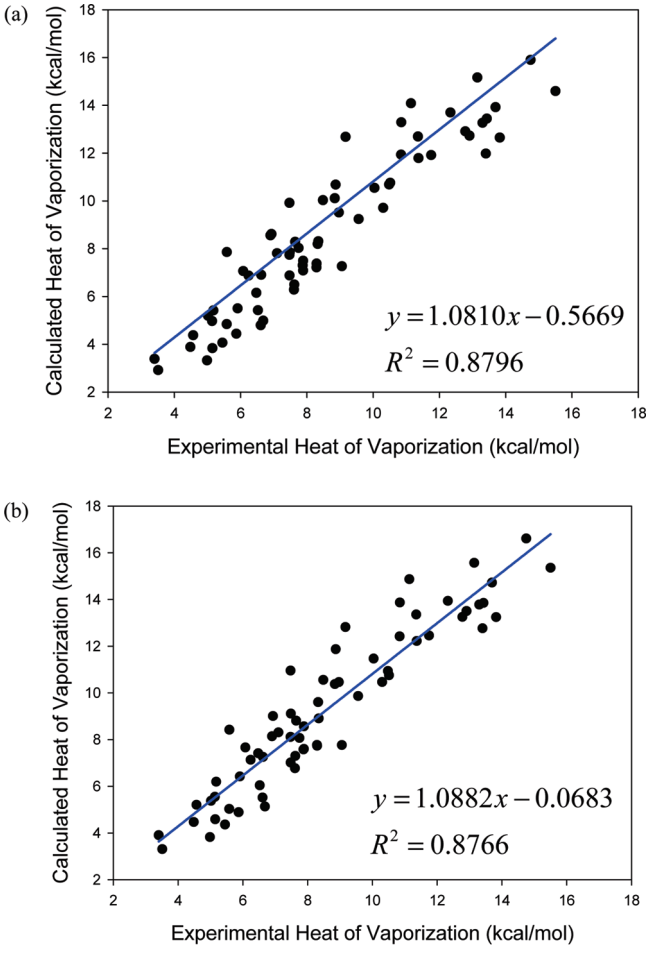

Figure 2. Performance of heats of vaporization calculation using two subprotocols of P4. (a) P4/ntt0, (b) P4/ntt3.

of all, it is a surprise that protocol P4 with the gas phase constant energy classical MD simulations (P4/ntt0) achieves the best performance: the AUE and RMSE are 0.932 and 1.201 kcal/mol, respectively. The total signed ΔH_{vap} difference $\Sigma(\Delta H_{vap}^{expt} - \Delta H_{vap}^{calc})$ is −8.54 for 67 molecules, suggesting that ΔH_{vap} is overestimated by about 0.13 kcal/mol on average. The correlation between the experimental and the calculated heats of vaporization with this model is shown in Figure 2a. The second-best subprotocol in terms of AUE, P4/ntt2, has an AUE of 1.019 kcal/mol. Second, the newly introduced temperature regulation method, Langevin dynamics (ntt = 3),[40] performs slightly worse than the Andersen coupling scheme (ntt = 2) in combination with both the P3 and P4 basic protocols. The correlation between the experimental and the calculated heats of vaporization with P4/ntt3 is shown in Figure 2b. Third, it is not a surprise that all four temperature scaling methods have similar performances in heat of vaporization calculations using protocol 3, since the snapshots sampled by gas MD simulations were further minimized and the means of the minimization energies have much smaller deviations than the MD energies without minimization. Fourth, protocol 4, with the weak coupling algorithm (ntt = 1), performs much worse than the other models, implying that the gas phase energies sampled using ntt = 1 are much different from those from the other temperature regulation methods. Finally, although protocol 1 has the largest approximation in theory, its performance is only marginally worse than the other protocols, except P4/ntt1 (the worst), indicating that the interaction energy makes the largest contribution to the heat of vaporization in most scenarios.

The uncertainty of the heat of vaporization calculations is summarized in Table S4 (Supporting Information). Protocol 1 has the largest average statistical uncertainty of 0.19 kcal/mol for 67 molecules. For the other calculation protocols, the uncertainties are all smaller than 0.085 kcal/mol, indicating that our MD protocols can reliably predict ΔH_{vap} using protocols P2, P3, and P4.

The compounds that have the largest prediction errors via P4/ntt0 are typically those less common compounds in biological systems, such as sulfides (Nos. 26, 27 and 28), nitromethane (No. 44), halides (Nos. 55, 58 and 59), and alkenes (Nos. 7 and 8). This phenomenon suggests that, for those compounds, more effort is needed to tune their van der Waals parameters to reproduce experimental densities and heats of vaporization.

Temperature Dependence of Density and ΔH_{vap} for TIP3P Water. It is important for a molecular mechanical model to accurately predict molecular properties of a broad range of thermodynamic states described by temperature, volume, pressure, etc. Given the fact that water has plenty of experimental data for density and ΔH_{vap} as a function of temperature, we assessed the protocols of MD simulations and ΔH_{vap} prediction using TIP3P water. As shown in Figure 3a, the calculated density of TIP3P decreases more sharply than the experimental value, and the two lines cross around 280 K. At 298 K, the density of TIP3P is 0.985 g/cm^3, about 1.2% lower than the experiment (0.997). For the temperature range from 260 to 310 K, the prediction error is smaller than 2% for TIP3P water in our MD simulations. A $\rho-T$ plot with a similar pattern to ours was reported by Price and Brooks.[54]

It is of interest to locate the temperature in the $\rho-T$ plot where TIP3P water has the maximum density. In a pioneer work, Billeter and co-workers investigated the maximum density temperature of the SPC/E water model using energy-volume correlations.[57] Jorgensen and Jenson also studied the temperatures of maximum density for TIP3P and TIP4P water models.[58] We estimate that the maximum density temperature is around

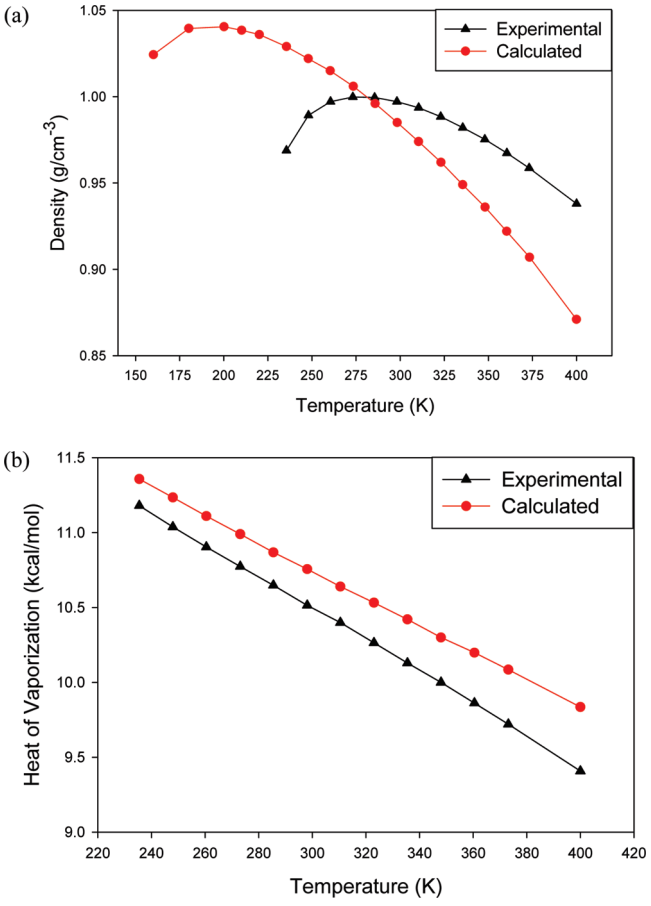

Figure 3. Temperature dependence of density and heat of vaporization of TIP3P water: (a) density, (b) heat of vaporization.

200 K for the TIP3P water. As shown in Figure 3a, when the temperature is above 220 K, the TIP3P model shows monotonically increasing density as the temperature decreases. From 180 to 220 K, there is a flat region, and the maximum density is reached. When the temperature drops below 180 K, the density decreases. This phenomenon is similar to that of TIP4P except TIP4P has its maximum density temperature and flat region around 260 K, as reported by Jorgensen and Jenson.[58] It is clear that the maximum density temperature of TIP4P is much closer to the experimental value (4 °C). Interestingly, Jorgensen and Jenson did not find the maximum density temperature for TIP3P since the temperatures in their Monte Carlo simulations were limited to −50 to 100 °C.[58]

As a special molecule, TIP3P has all its degrees of freedom frozen; as such, all of the protocols of calculating ΔH_{vap} produce the same value. The temperature dependence of heat of vaporization for TIP3P is shown in Figure 3b. It is clear that the calculated ΔH_{vap} is systematically larger than the experimental value at a temperature range of 230−400 K. At 298 K, the calculated ΔH_{vap} is overestimated by about 2.3% (10.51 vs 10.75 kcal/mol). The experimental and calculated data used for plotting Figure 3 are listed in Table S5 of the Supporting Information.

The Major Factors That Affect Density and ΔH_{vap} Calculations. Given the physical limitations of the current nonpolarizabe force fields, it is important to derive force field parameters to reproduce not only the in vacuo ab initio properties but also pure-liquid and hydrated properties. The development of the OPLS-AA force field perfectly demonstrates this philosophy.[7] The quality of molecular mechanical models certainly is one of the major factors that determine the performance of density and ΔH_{vap} calculations. However, it is challenging to conduct a critical assessment of different force fields; hidden flaws (such as wrongly using the mixing rules of van der Waals parameters) and some technical differences could lead to discrepancy. In this work, we only make comparison using the published data.

As a pioneer of developing molecular mechanical force fields through liquid simulations, Jorgensen and co-workers have studied a large set of organic compounds that covers most chemical functional groups. The following is a quick summary of the calculation errors for each compound class using the OPLS-AA force field, and the two values in the appended parentheses represent the average percent errors of density and ΔH_{vap}, respectively: hydrocarbon (3%, 2%);[7] alcohols (1.8%, 2.2%);[7] sulfur compounds (1.8%, 4%);[7] ethers (1.5%, 2.8%);[7] amides (2%, 2.1%);[7] acids (1.9%, 0.7%);[7] aldehydes and ketones (1.0%, 3.7%);[7] pyridine and diazenes (0.8%, 2.7%);[49] pyrrole, furan, diazole, and oxazoles (2%, 9.9%);[50] amines (1.4%, 3.5%);[52] esters (1.4%, 6.2%);[51] nitriles (1.8%, 2.7%);[51] and nitro compounds (1.1%, 2.0).[51] Recently, Price and Brooks further studied a set of 28 mono- and disubstituted benzenes using the OPLS-AA force field.[59] They found that the agreement between OPLS-AA and experimental results is remarkable, with average errors of 1.8% for densities and 0.69 kcal/mol for ΔH_{vap}.

Unlike the remarkable performance of OPLS-AA, the performance of GAFF, with an APE of 3.43% if not counting outliers for density and an AUE of 0.93 kcal/mol (P4/ntt0) for heat of vaporization, is still satisfactory given the following two facts: the partial charges in OPLS-AA are tunable parameters, while GAFF was parametrized using HF/6-31G* RESP charges, and the van der Waals parameters in OPLS-AA were parametrized to reproduce the bulk properties including density and heat of vaporization, while GAFF inherits its vdW parameters directly from the AMBER force fields without further optimization.

Besides the molecular mechanical models, sampling and the other computational techniques could affect the density and ΔH_{vap} calculation. There has been a dramatic change in electrostatic energy calculations of a unit cell since the vdW parameters of GAFF were developed. Nowadays, particle mesh Ewald (PME) has become a standard method in handling the long-range electrostatic interactions in most molecular simulation packages.[35−37] Other methods of calculating long-range electrostatic interactions, such as isotropic periodic sum (IPS),[60−62] begin to emerge. On the other hand, van der Waals parameters implemented in GAFF were developed with short electrostatic cutoffs, and the long-range effect had been embedded in the local interactions in an average sense. It is understandable that van der Waals parameters developed in the old days have some deviations in molecular property calculations using today's simulation techniques (such as PME and Langevin dynamics to regulate temperature).

Exploration of Force Field Parameters. Given the fact that GAFF has not been optimized to reproduce the bulk properties, there should be a lot of room for us to significantly improve the performance of predicting the density and heat of vaporization through proper van der Waals parametrization. Here, we simply carried out an exploring study to prove the concepts, and the systematic van der Waals parametrization for GAFF is the next target in our effort of developing high-quality general purpose force fields.

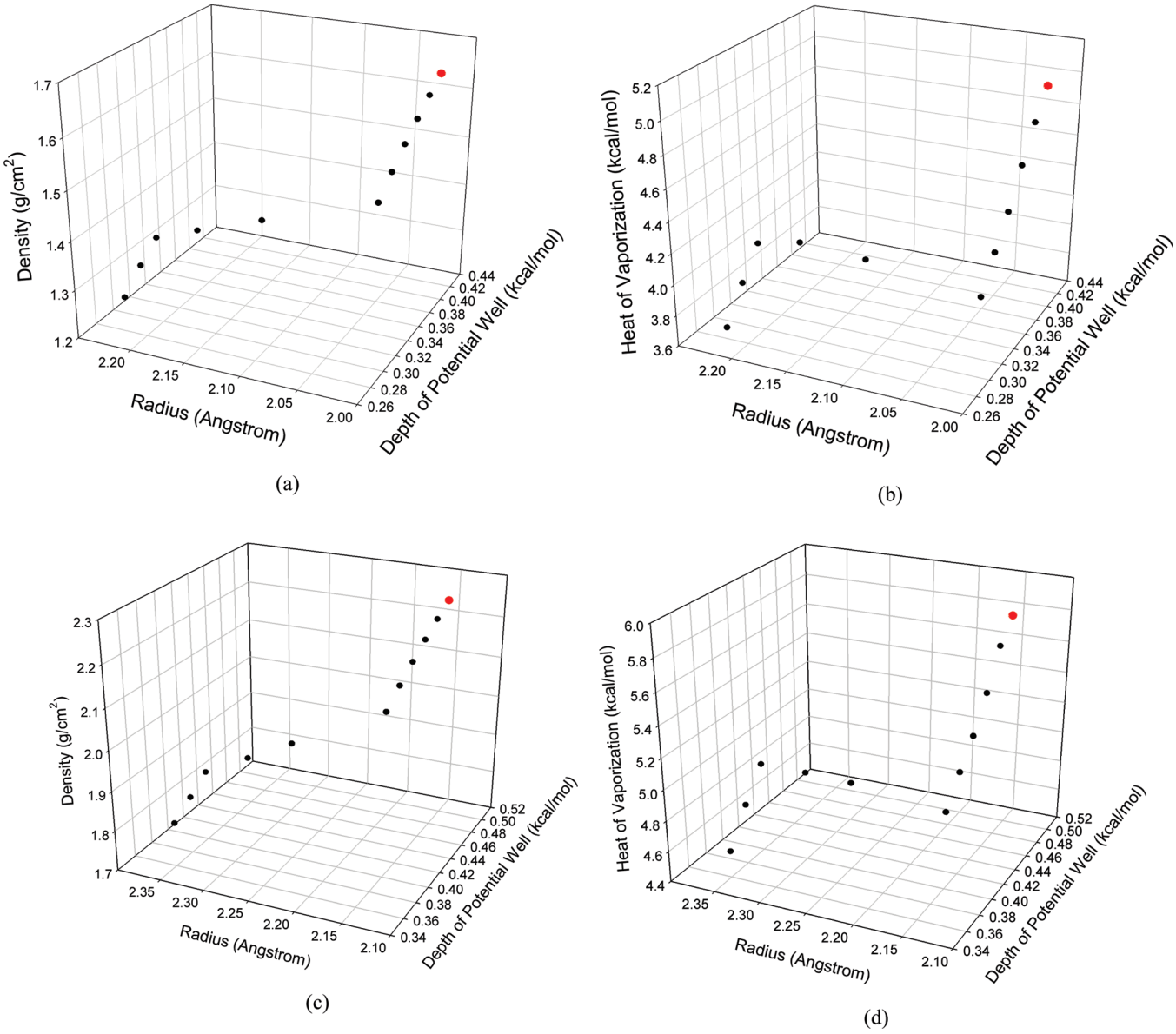

Figure 4. How the calculated density and heat of vaporization change as a function of van der Waals parameters. (a) Density of bromomethane. (b) Heat of vaporization of bromomethane. (c) Density of iodomethane. (d) Heat of vaporization of iodomethane.

First of all, we optimized the vdW parameters of bromide and iodide. There are no parameters for bromide and iodide in the AMBER biomolecular force fields. The vdW parameter of bromide was borrowed from MM3,[63] and that of iodine was from Weiner et al.[64] The ideal parameters of radius R and depth of potential well ε were located using systematic searches. How the density and ΔH_{vap} change as a function of R and ε is shown in Figure 4. Comparing Figure 4a and b for bromomethane and 4c and d for iodomethane, we can reach the following conclusion: the patterns of density and heat of vaporization change as a function of R and ε are essentially similar, and they almost reach the points of experimental values synchronously. In summary, the vdW parameter of bromide changes from R = 2.22 and ε = 0.32 to R = 2.02 and ε = 0.42, and the density and ΔH_{vap} of the new parameter are 1.642 g cm^{-3} and 4.952 kcal/mol (P4/ntt = 0), respectively. In contrast, the density and ΔH_{vap} of the old parameter are 1.343 g cm^{-3} and 3.89 kcal/mol, respectively (the experimental values are 1.676 and 5.45, respectively). As to iodide, the vdW parameter changes from R = 2.35 and ε = 0.40 to R = 2.15 and ε = 0.50). Similarly, both density and ΔH_{vap} can much better reproduce the experimental values (ρ = 2.279 g cm^{-3} and ΔH_{vap} = 6.68) using the new parameter (ρ = 2.247 and ΔH_{vap} = 5.782) than the original one (ρ = 1.877 and ΔH_{vap} = 4.65). The prediction errors of densities and ΔH_{vap} are summarized in Table 4.

In the next example, we attempted to improve the density and ΔH_{vap} of aromatic compounds by tuning the vdW parameter of aromatic carbon. Benzene (No. 10) was selected as the model compound for the vdW parametrization. We found that the two properties of benzene are not sensitive to the radius parameter R; therefore, only the depth of potential well ε of aromatic carbon was scanned using the original radius parameter R = 1.908 Å. As shown in Figure 5, the density and ΔH_{vap} are proportional to ε, and the ideal value of ε ranges from 0.090 to 0.095 kcal/mol.

Table 4. Performance of Density and Heat of Vaporization Calculation Using the Original and the Revised GAFF

		original GAFF				revised GAFF			
		density (g/cm^3)		ΔH_{vap} (kcal/mol)		density (g/cm^3)		ΔH_{vap} (kcal/mol)	
molecular class	#mols	AUE	RMSE	AUE	RMSE	AUE	RMSE	AUE	RMSE
bromide	1	0.333	0.333	1.56	1.56	0.034	0.034	0.50	0.50
iodide	1	0.402	0.402	2.03	2.03	0.032	0.032	0.90	0.90
benzene derivatives	5	0.023	0.051	0.72	1.61	0.020	0.045	0.48	1.07
amine	3	0.041	0.044	1.28	1.41	0.026	0.026	0.44	0.70

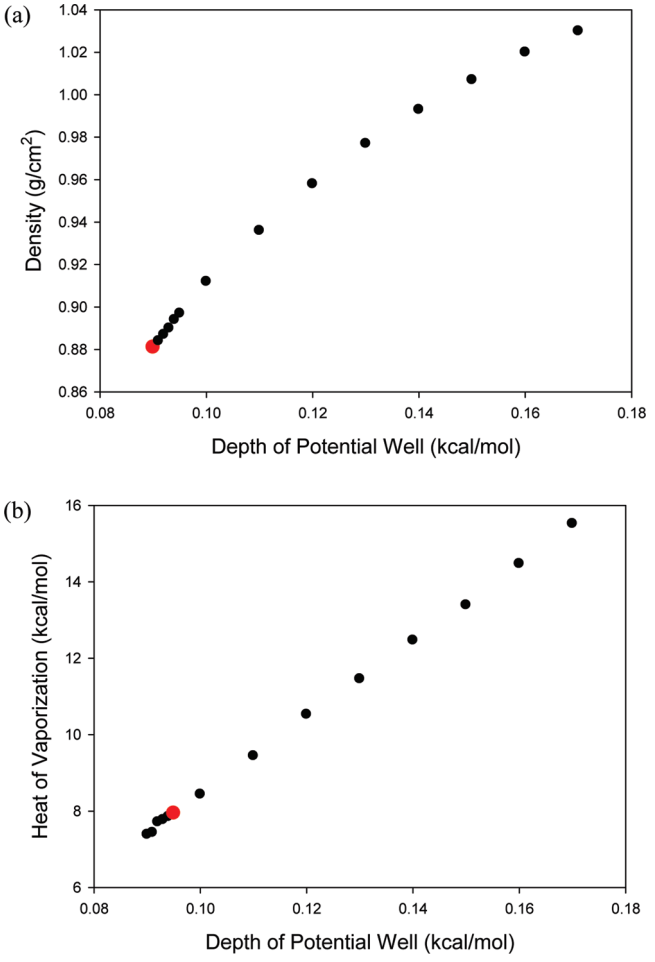

Figure 5. How the density and heat of vaporization of benzene change as a function of the depth of potential well of the aromatic carbon: (a) density, (b) heat of vaporization. The red dots indicate which van der Waals parameter best reproduces the experimental values.

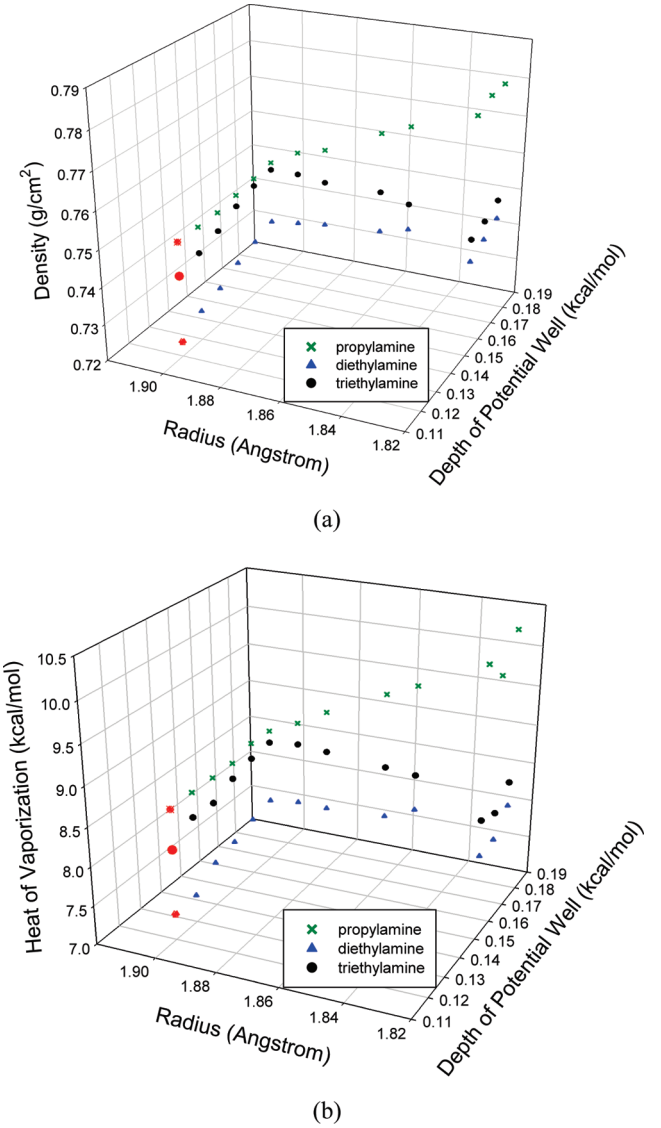

Figure 6. How the density and heat of vaporization of three amines change as a function of the van der Waals parameters of amine nitrogen: (a) density, (b) heat of vaporization. The red symbols indicate which van der Waals parameter best reproduces the experimental values.

When ε = 0.095, ΔH_{vap} has the best value (7.944 by P4/ntt0) compared to the experimental result (8.086 kcal/mol), while a smaller ε is required to achieve the best density (ρ = 0.877 when ε = 0.091). Finally, we set the depth of the potential well of aromatic carbon to 0.092 kcal/mol as a compromise. The calculated densities and ΔH_{vap} for five aromatic compounds (Nos. 10, 17, 18, 42, and 55) using the new vdW parameter are listed in Table S6 (Supporting Information). Although only the depth of potential well ε of aromatic carbon was adjusted, the performance of the density and ΔH_{vap} prediction is significantly improved: the AUE, RMSE, and APE of density calculations using the new parameter are 0.02, 0.045, and 1.83%, respectively, in comparison to AUE = 0.023, RMSE = 0.051, and APE = 2.07% using the original parameter. More significant improvement was

dx.doi.org/10.1021/ct200142z |J. Chem. Theory Comput. 2011, 7, 2151–2165

Table 5. List of the Electrostatic Part (Charge Disappearing), the van der Waals Part (Atom Disappearing), and the Calculated Free Energies (kcal/mol) of Hydration for 10 Representative Molecules

			original GAFF			revised GAFF		
no.	name	exptl[a]	elec.	vdW	calcd	elec.	vdW	calcd
1	bromomethane	−0.82	−1.40	2.10	0.70	−1.55	1.46	−0.10
2	iodomethane	−0.89	−1.43	1.98	0.55	−1.54	1.38	−0.16
3	*propylamine*	*−4.39*	*−7.26*	*1.84*	*−5.42*	*−6.54*	*2.02*	*−4.53*
4	*diethylamine*	*−4.07*	*−3.97*	*2.33*	*−1.63*	*−3.64*	*2.19*	*−1.45*
5	*triethylamine*	*−3.22*	*−0.92*	*2.56*	*2.37*	*−0.79*	*2.25*	*2.20*
6	benzene	−0.86	−2.69	1.77	−0.92	−2.70	1.37	−1.33
7	phenol	−6.61	−5.92	1.48	−4.44	−5.88	1.15	−4.73
8	*m*-cresol	−5.49	−7.52	1.47	−6.06	−7.41	1.43	−5.98
9	aniline	−5.49	−5.92	1.30	−4.62	−5.87	1.11	−4.76
10	fluobenzene	−0.80	−2.28	2.02	−0.26	−2.25	2.05	−0.19

[a] The experimental data are from ref 19.

achieved for ΔH_{vap}, and the AUE, RMSE, and APE of the new parameter are 0.477, 1.067, and 4.63%, respectively, in comparison to AUE = 0.720, RMSE = 1.610, and APE = 6.99% of the original parameter of the aromatic carbon. We believe that the prediction performance can be further improved when more atom types are tuned simultaneously. The new depth of potential well of aromatic carbon, 0.092, is much larger than the original value, 0.086. Interestingly, we found that a larger value of depth of potential well of carbon can significantly reduce the prediction errors of interaction energies for a large set of amino acid analog pairs.[65,66]

In the last example, the vdW parameter of amine carbon was optimized using three different types of amines, namely, propylamine (No. 37), diethylamine (No. 39), and triethylamine (No. 41). How the densities and heats of vaporization change as a function of the radius parameter R and depth of potential well ε is shown in Figure 6. Encouragingly, the best prediction values of densities and heats of vaporization for the three compounds are obtained synchronously. The optimized parameter of amine nitrogen (R = 1.900 and ε = 0.140) has a larger radius and smaller depth of potential well compared to the original one (R = 1.824 and ε = 0.170). The calculated densities and ΔH_{vap} for three amines using the new vdW parameter are listed in Table S7 (Supporting Information). Again, the new parameter significantly reduces the prediction errors of both molecular properties: for density, the AUE, RMSE, and APE of the new parameter are 0.026, 0.026, and 3.57%, respectively, in comparison to AUE = 0.041, RMSE = 0.044, and APE = 5.67% of the original parameter. As to heat of vaporization, the improvement is even more significant: the AUE, RMSE, and APE of the new parameter are 0.441, 0.699, and 5.89%, respectively, in comparison to AUE = 1.276, RMSE = 1.408, and APE = 16.34% of the original parameter. The performance of the original and revised GAFF in calculating density and ΔH_{vap} is summarized in Table 4.

It is worth noting that tuning the partial charges can also improve the molecular property prediction. For example, in OPLS-AA after turning both the partial charges and Lennard-Jones parameters, Jorgensen and Rizzo could reproduce the densities, the heats of vaporization, and hydration free energies of pure liquid amines very well.[52] In another example, with the adjustment of several bond charge correction parameters, AM1-BCC can better reproduce the absolute solvation energies.[15,16] Certainly, it is also possible for us to tune BCC parameters so that the calculated densities, heats of vaporization, and free energies of hydration agree better with experimental results. However, in this study, the HF/6-31G* RESP charges were used since GAFF was originally parametrized with this charge method.

In summary, the prediction errors of density and heat of vaporization can be significantly reduced by tuning one or two van der Waals parameters. Density and ΔH_{vap} approach their ideal values synchronously in all three cases, although the best parameters for density and ΔH_{vap} may be slightly different. It should be pointed out that although the new vdW parameters were optimized using Protocol P4/ntt0 to calculate ΔH_{vap}, the calculation errors of other protocols (such as P2, P4/ntt3) are also reduced (data not shown). Certainly, the optimum vdW parameter sets are slightly different from one protocol to another.

Hydration Energies Calculations. As the hydration free energy of a molecule is also directly affected by the vDW parameters. It is important for us to check whether the newly developed vDW parameters deteriorate the hydration energy calculations or not. Ten molecules, which are directly affected by the newly developed vDW parameters, were selected to conduct hydration free energies using TI. Encouragingly, the revised GAFF significantly improves the hydration energy calculations, as shown in Table 5: the AUE is reduced from 1.62 to 1.38 kcal/mol and RMSE from 2.21 to 2.05 kcal/mol. It is pointed out that most errors come from three amines, i.e., propylamine, diethylamine, and triethylamime. It is widely known that the hydration energies of amine and amide series are difficult to predict, and the previous studies using free energy perturbation with AMBER force fields strongly overestimate the free energy changes from methylamine to dimethylamine and from dimethylamine to trimethylamime.[52] If the three amines are excluded, the AUE and RMSE of the revised GAFF become 0.80 and 0.92 kcal/mol, respectively. All three compound classes in Table 5 achieve a better performance of hydration free energy calculation using the revised GAFF. Specifically, for the two halides, both the AUE and RMSE are reduced by 0.75 kcal/mol. For the three amines, the AUE and RMSE are reduced by 0.3 and 0.1 kcal/mol, respectively. For five benzene derivatives, the AUE and RMSE are reduced by 0.01 and 0.11 kcal/mol, respectively. In summary, although the prediction of hydration free energies was not taken into account when we optimized vDW parameters, the revised GAFF achieved better performance in hydration energy calculations using TI. The statistical uncertainties of individual energy

terms of TI calculations are summarized in Table S8 of the Supporting Information.

4. CONCLUSIONS

This is the first paper in a series of predicting molecular properties using the General AMBER Force Field (GAFF). The bulk densities and heats of vaporizations of 71 organic molecules that cover the most common chemical functional groups have been predicted through molecular dynamics simulations. The overall performance of the prediction is satisfactory: the average unsigned error (AUE) of the density of the whole molecular set is 0.0436 g cm^{-3}, and the AUE of the heat of vaporization calculation using Protocol P4/ntt0 is 0.932 kcal/mol. In total, 10 different heat of vaporization subprotocols (P1, P2, P3/ntt=0, P3/ntt=1, P3/ntt=2, P3/ntt=3, P4/ntt=0, P4/ntt=1, P4/ntt=2, P4/ntt=3) have been investigated, and P4/ntt0 outperforms the other subprotocols. van der Waals parametrization has been conducted for three molecular subsets in order to reduce the prediction errors of both density and heat of vaporization. Encouragingly, both density and heat of vaporization can much better reproduce the experimental data after adjusting one or two van der Waals parameters. Moreover, the free energies of hydration for 10 molecules also achieve a much better performance for the revised GAFF. Given the fact that GAFF inherited van der Waals parameters from the AMBER biomolecular force fields without further optimization, it is very likely that GAFF can be significantly improved on predicting molecular properties after a systematic van der Waals parametrization.

■ ASSOCIATED CONTENT

Supporting Information. The calculated densities of 71 molecules are listed in Table S1. The structures of the 71 molecules are shown in Figure S1. The calculated heats of vaporization using Protocols P3/ntt1, P4/ntt1, P3/ntt2, and P4/ntt2 are summarized in Table S2. The individual energetic terms and statistical uncertainties are listed in Tables S3 and S4, respectively. The data used to prepare Figure 3 are listed in Table S5. In Tables S6−S7, the calculated densities and heats of vaporization using new van der Waals parameters of aromatic carbon (Table S6) and amine nitrogen (Table S7) are listed. The statistical uncertainties of TI calculations are listed in Table S8. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*Tel.: (214)-645-5966 (J.W.). E-mail: junmei.wang@utsouthwestern.edu, tjhou@suda.edu.cn.

■ ACKNOWLEDGMENT

We are grateful for the research support from NIH (R01GM79383, Y. Duan, P.I.) and Natural Science Foundation of China (No. 20973121, T. Hou, P.I.) and the TeraGrid for the computational time (TG-CHE090098, J. Wang, P.I.).

■ ABBREVIATIONS

GAFF, General AMBER Force Field; MD, Molecular Dynamics; vdW, van der Waals; ρ, density; ΔH_{vap}, heat of vaporization; ε, depth of potential well; AUE, average unsigned errors; RMSE, root-mean-square errors; APE, average percent errors; R^2, correlation coefficient square

■ REFERENCES

(1) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652.

(2) van Gunsteren, W. F.; Mark, A. E. Validation of molecular dynamics simulation. *J. Chem. Phys.* **1998**, *108*, 6109–6116.

(3) Foloppe, N.; MacKerell, A. D. J. All-atom empirical force field for nucleic acids. 1) Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **2000**, *21*, 86–104.

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

(5) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(6) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21*, 1049–1074.

(7) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(8) Oostenbrink, C.; Soares, T. A.; van der Vegt, N. F.; van Gunsteren, W. F. Validation of the 53A6 GROMOS force field. *Eur. Biophys. J.* **2005**, *34*, 273–284.

(9) Soares, T. A.; Hunenberger, P. H.; Kastenholz, M. A.; Krautler, V.; Lenz, T.; Lins, R. D.; Oostenbrink, C.; van Gunsteren, W. F. An improved nucleic acid parameter set for the GROMOS force field. *J. Comput. Chem.* **2005**, *26*, 727–737.

(10) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(11) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges- the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(12) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. Application of the multimolecule and multiconformational RESP methodology to biopolymers- charge derivation for DNA, RNA, and proteins. *J. Comput. Chem.* **1995**, *16*, 1357–1377.

(13) Caldwell, J. W.; Kollman, P. A. Structure and Properties of Neat Liquids Using Nonadditive Molecular-Dynamics - Water, Methanol, and N-Methylacetamide. *J. Phys. Chem.* **1995**, *99*, 6208–6219.

(14) Fox, T.; Kollman, P. A. Application of the RESP methodology in the parametrization of organic solvents. *J. Phys. Chem. B* **1998**, *102*, 8070–8079.

(15) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast efficient generation of high quality atomic charges. AM1-BCC model: I. method. *J. Comput. Chem.* **2000**, *21*, 132–146.

(16) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641.

(17) Kuhn, B.; Gerber, P.; Schulz-Gasch, T.; Stahl, M. Validation and use of the MM-PBSA approach for drug discovery. *J. Med. Chem.* **2005**, *48*, 4040–4048.

(18) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69–82.

(19) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations. *J. Chem. Theory Comput.* **2009**, *5*, 350–358.

(20) MacKerell, A. D.; Shim, J. H.; Anisimov, V. M. Re-evaluation of the reported experimental values of the heat of vaporization of N-methylacetamide. *J. Chem. Theory Comput.* **2008**, *4*, 1307–1312.

(21) Physical constants of organic compounds. In *CRC Handbook of Chemistry and Physics*, 86th ed.; Lide, D. R. E., Eds.; CRC Press: Boca Raton, FL, 2005; pp 1−3.

(22) Siepmann, J. I.; Karaborni, S.; Smit, B. Simulating the Critical-Behavior of Complex Fluids. *Nature* **1993**, *365*, 330–332.

(23) *Enthalpies of Vaporization of Organic Compounds: A Critical Review and Data Compilation*; Majer, V., Svoboda, V., Eds.; Blackwell Scientific Publications: Oxford, U. K., 1985; pp 1−304.

(24) Kemp, J. D.; Egan, C. J. Hindered Rotation of the Methyl Groups in Propane. The Heat Capacity, Vapor Pressure, Heats of Fusion and Vaporization of Propane. Entropy and Density of the Gas. *J. Am. Chem. Soc.* **1938**, *60*, 1521–1525.

(25) Pedley, J. B.; Naylor, R. D.; Kirby, S. P. *Thermochemical Data of Organic Compounds*, 2nd ed.; Chapman and Hall: New York, 1986; pp 1−792.

(26) Wiberg, K. B.; Crocker, L. S.; Morgan, K. M. Thermochemical Studies of Carbonyl-Compounds 0.5. Enthalpies of Reduction of Carbonyl Groups. *J. Am. Chem. Soc.* **1991**, *113*, 3447–3450.

(27) Geller, B. E. Properties of the dimethylformamide water system. I. Thermochemical properties. *Russ. J. Phys. Chem.* **1961**, *35*, 542–564.

(28) Pihlaja, K.; Heikklia, J. Enthalpies of formation of cyclic acetals. 1,3-dioxolane, 2-methyl-1,3-dioxolane, and 2,4-dimethyl-1,3-dioxolanes. *Acta Chem. Scand.* **1969**, *23*, 1053–1055.

(29) Steele, W. V.; Archer, D. G.; Chirico, R. D.; Collier, W. B.; Hossenlopp, I. A.; Nguyen, A.; Smith, N. K.; Gammon, B. E. The thermodynamic properties of quinoline and isoquinoline. *J. Chem. Thermodyn.* **1988**, *20*, 1233–1264.

(30) Wang, J.; Hou, T.; Xu, X. Aqueous solubility prediction based on weighted atom type counts and solvent accessible surface areas. *J. Chem. Inf. Model.* **2009**, *49*, 571–81.

(31) Konicek, J.; Wadso, I. Enthalpies of Vaporization of Organic Compounds 0.7. Some Carboxylic Acids. *Acta Chem. Scand.* **1970**, *24*, 2612–2626.

(32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.

(33) Case, D. A.; Darden, T. A.; Cheatham, I.; Simmerling, C.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER10*; University of California at San Francisco: San Francisco, CA, 2008.

(34) Wang, J. M.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Model.* **2006**, *25*, 247–260.

(35) Darden, T.; Perera, L.; Li, L.; Pedersen, L. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* **1999**, *7*, R55–60.

(36) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(37) Sagui, C.; Pedersen, L. G.; Darden, T. A. Towards an accurate representation of electrostatics in classical force fields: Efficient implementation of multipolar interactions in biomolecular simulations. *J. Chem. Phys.* **2004**, *120*, 73–87.

(38) Miyamoto, S.; Kollman, P. A. Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952–962.

(39) Shirts, M. R.; Mobley, D. L.; Chodera, J. D.; Pande, V. S. Accurate and efficient corrections for missing dispersion interactions in molecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 13052–13063.

(40) Uberuaga, B. P.; Anghel, M.; Voter, A. F. Synchronization of trajectories in canonical molecular-dynamics simulations: observation, explanation, and exploitation. *J. Chem. Phys.* **2004**, *120*, 6363–6374.

(41) Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D. Langevin stabilization of molecular dynamics. *J. Chem. Phys.* **2001**, *114*, 2090–2098.

(42) Larini, L.; Mannella, R.; Leporini, D. Langevin stabilization of molecular-dynamics simulations of polymers by means of quasisymplectic algorithms. *J. Chem. Phys.* **2007**, *126*, 104101.

(43) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin Dynamics of Peptides - the Frictional Dependence of Isomerization Rates of N-Acetylalanyl-N′-Methylamide. *Biopolymers* **1992**, *32*, 523–535.

(44) Harvey, S. C.; Tan, R. K. Z.; Cheatham, T. E. The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J. Comput. Chem.* **1998**, *19*, 726–740.

(45) Andersen, H. C. Molecular-Dynamics Simulations at Constant Pressure and-or Temperature. *J. Chem. Phys.* **1980**, *72*, 2384–2393.

(46) Andrea, T. A.; Swope, W. C.; Andersen, H. C. The Role of Long Ranged Forces in Determining the Structure and Properties of Liquid Water. *J. Chem. Phys.* **1983**, *79*, 4576–4584.

(47) Steinbrecher, T.; Mobley, D. L.; Case, D. A. Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations. *J. Chem. Phys.* **2007**, *127*, 214108.

(48) Hummer, G.; Szabo, A. Calculation of free-energy differences from computer simulations of initial and final states. *J. Chem. Phys.* **1996**, *105*, 2004–2010.

(49) Jorgensen, W. L.; McDonald, N. A. Development of an all-atom force field for heterocycles. Properties of liquid pyridine and diazenes. *THEOCHEM* **1998**, *424*, 145–155.

(50) McDonald, N. A.; Jorgensen, W. L. Development of an all-atom force field for heterocycles. Properties of liquid pyrrole, furan, diazoles, and oxazoles. *J. Phys. Chem. B* **1998**, *102*, 8049–8059.

(51) Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L. Gas-phase and liquid-state properties of esters, nitriles, and nitro compounds with the OPLS-AA force field. *J. Comput. Chem.* **2001**, *22*, 1340–1352.

(52) Rizzo, R. C.; Jorgensen, W. L. OPLS all-atom model for amines: Resolution of the amine hydration problem. *J. Am. Chem. Soc.* **1999**, *121*, 4827–4836.

(53) Yu, H. B.; Hansson, T.; van Gunsteren, W. F. Development of a simple, self-consistent polarizable model for liquid water. *J. Chem. Phys.* **2003**, *118*, 221–234.

(54) Price, D. J.; Brooks, C. L. A modified TIP3P water potential for simulation with Ewald summation. *J. Chem. Phys.* **2004**, *121*, 10096–10103.

(55) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved

four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.

(56) Yu, H. B.; Geerke, D. P.; Liu, H. Y.; van Gunsteren, W. E. Molecular dynamics simulations of liquid methanol and methanol-water mixtures with polarizable models. *J. Comput. Chem.* **2006**, *27*, 1494–1504.

(57) Billeter, S. R.; King, P. M.; Vangunsteren, W. F. Can the Density Maximum of Water Be Found by Computer-Simulation. *J. Chem. Phys.* **1994**, *100*, 6692–6699.

(58) Jorgensen, W. L.; Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *J. Comput. Chem.* **1998**, *19*, 1179–1186.

(59) Price, D. J.; Brooks, C. L. Detailed considerations for a balanced and broadly applicable force field: A study of substituted benzenes modeled with OPLS-AA. *J. Comput. Chem.* **2005**, *26*, 1529–1541.

(60) Wu, X.; Brooks, B. R. Isotropic periodic sum: a method for the calculation of long-range interactions. *J. Chem. Phys.* **2005**, *122*, 44107.

(61) Klauda, J. B.; Wu, X.; Pastor, R. W.; Brooks, B. R. Long-range Lennard-Jones and electrostatic interactions in interfaces: application of the isotropic periodic sum method. *J. Phys. Chem. B* **2007**, *111*, 4393–400.

(62) Takahashi, K.; Yasuoka, K.; Narumi, T. Cutoff radius effect of isotropic periodic sum method for transport coefficients of Lennard-Jones liquid. *J. Chem. Phys.* **2007**, *127*, 114511.

(63) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular Mechanics - the MM3 Force-Field for Hydrocarbons. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.

(64) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force-Field for Simulations of Proteins and Nucleic-Acids. *J. Comput. Chem.* **1986**, *7*, 230–252.

(65) Wang, J. M.; Cieplak, P.; Li, J.; Hou, T. J.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations I: Parameterization of Atomic Polarizability. *J. Phys. Chem. B* **2011**, *115*, 3091–3099.

(66) Wang, J. M.; Cieplak, P.; Li, J.; Wang, J.; Cai, Q.; Hsieh, M. J.; Lei, H. X.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations II: Induced Dipole Models Significantly Improve Accuracy of Intermolecular Interaction Energies. *J. Phys. Chem. B* **2011**, *115*, 3100–3111.

NOTE ADDED AFTER ASAP PUBLICATION

This manuscript was originally published on May 30, 2011 with the given name and surname of one of the authors transposed. The corrected manuscript was published on June 22, 2011.