JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Force Field Benchmark of Organic Liquids. 2. Gibbs Energy of Solvation

Jin Zhang,[†,‡] Badamkhatan Tuguldur,[§,‡] and David van der Spoel*[,‡]

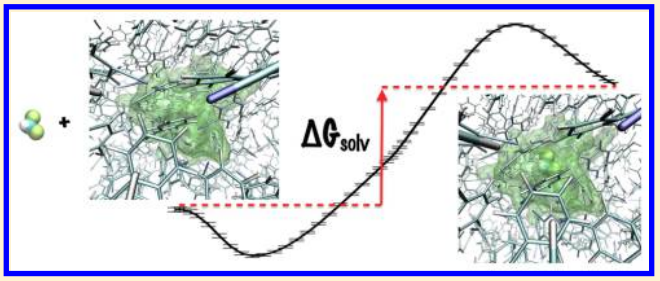
[†]Department of Chemistry, Zhejiang University, Hangzhou 310027, China

[‡]Uppsala Center for Computational Chemistry, Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, Husargatan 3, Box 596, SE-75124 Uppsala, Sweden

[§]Department of Biology, School of Arts and Sciences, National University of Mongolia, Ulaanbaatar 14200, Mongolia

Supporting Information

ABSTRACT: Quantitative prediction of physical properties of liquids is a longstanding goal of molecular simulation. Here, we evaluate the predictive power of the Generalized Amber Force Field (Wang et al. *J. Comput. Chem.* **2004**, *25*, 1157−1174) for the Gibbs energy of solvation of organic molecules in organic solvents using the thermodynamics integration (TI) method. The results are compared to experimental data, to a model based on quantitative structure property relations (QSPR), and to the conductor-like screening models for realistic solvation (COSMO-RS) model. Although the TI calculations yield slightly better correlation to experimental results than the other models, in all fairness we should conclude that the difference between the models is minor since both QSPR and COSMO-RS yield a slightly lower RMSD from that of the experiment (<3.5 kJ/mol). By analyzing which molecules (either as solvents or solutes) are outliers in the TI calculations, we can pinpoint where additional parametrization efforts are needed. For the force field based TI calculations, deviations from the experiment occur in particular when compounds containing nitro or ester groups are solvated into other liquids, suggesting that the interaction between these groups and solvents may be too strong. In the COSMO-RS calculations, outliers mainly occur when compounds containing (in particular aromatic) rings are solvated despite using a ring correction term in the calculations.



INTRODUCTION

Accurate force field parameters play a pivotal role in predicting physical chemistry properties of molecules and liquids using molecular simulation, and therefore the quality of force fields for the prediction of physical properties of molecules and liquids has been under scrutiny for a long while. Due to computational cost not many large scale comparisons of predicted liquid properties to experimental data have been done. Caleman et al. have recently evaluated the performance of the optimized potential for liquid simulations (OPLS/AA)[1] and the Generalized Amber Force Field (GAFF)[2] force fields by simulating approximately 150 liquids and computing the density, enthalpy of vaporization, heat capacity, surface tension, compressibility, expansion coefficient, and the dielectric constant.[3] This benchmark provided insights in strengths and weaknesses of these force fields for use in prediction of liquid properties.

The solvation free energies of small neutral organic molecules are often known experimentally with high accuracy. Therefore, the Gibbs energy of solvation (ΔG_{solv}) is an important target for force field benchmarking[4−8] and development,[9,10] usually with water as the solvent. ΔG_{solv} is also a key property in ligand binding free energy, which is particularly important in drug design.[11] Quantitative free energy prediction requires both force field parameters with sufficient accuracy to represent the experimental system and adequate sampling to establish statistically meaningful measures.[4] Shirts et al. achieved very precise calculations of the hydration free energy of 15 amino acid side chain analogs by performing long simulations.[4] Indeed, it has been confirmed that for simple systems the statistical precision of simulations can be made smaller than the error due to uncertainty in the force field parameters.[12] In addition to the parametrization, the functional form of the potential function puts limits on the accuracy of predictive power of molecular simulation, which is why large efforts have gone into including, e.g., polarizability in force fields.[13,14]

Due to the specific importance of biomolecular properties in aqueous solution, quite a number of studies on hydration free energy for biomolecular force fields are available, including, for instance, GAFF,[7,15−17] CGenFF (CHARMM general force field),[18] and OPLS/AA.[5,19−22] Mobley et al. compared the performance of the GAFF with different methods to derive the partial atomic charges by calculating the hydration free energy of 44 small, neutral molecules in different explicit water models.[15] They found that when a relatively high level of theory was used for deriving the partial atomic charges, the hydration free energies were reproduced marginally better than with lower levels of theory. Similar results were described by Wallin et al.

Received: February 26, 2015
Published: May 26, 2015

ACS Publications © 2015 American Chemical Society

DOI: 10.1021/acs.jcim.5b00106
J. Chem. Inf. Model. 2015, 55, 1192−1201

from binding free energy calculations.[23] A more recent paper by Mobley et al. compared the hydration free energy of 504 neutral small organic molecules with experiments. Here, the authors found good agreement between the simulations and experiments, except for alkynes, for which systematic errors were found.[7] These studies involve complex systems with multiple changed parameters, however, and the interdependency between partial charges and Lennard-Jones parameters was not taken into account. In addition, the amount of sampling influences the calculated free energy, although the error caused by sampling can be made small.[4,12]

Most force field benchmarks and development efforts consider only water as the solvent. An exception to this is the development of the GROMOS 53A6 force field, where cyclohexane was used as a solvent in addition to water[9] and the more recent development of the $53A6_{OXY}$ and $53A6_{OXY+D}$ force fields, where both solvents were considered to reparameterize/extend the nonbonded interaction parameters for functional groups containing oxygen.[24,25] Earlier force field evaluations have shown that the water model plays a significant role in the observed properties of solvated (bio)molecules,[5,26,27] and therefore focusing on water as a solvent only may be misleading. In addition, partition properties between different media play a vital role in biomolecular processes like protein folding, the association of biomolecules, formation of micelles and membranes, and protein−membrane interactions. To faithfully describe the partition properties between different media, it is important to consider the accuracy of reproducing the solvation free energy in other solvents than water in the parametrization of force fields.

In order to test the performance of the GAFF for the prediction of solvation free energy and to avoid systematic errors caused by the water model and protein force field, we here calculated the Gibbs energy of solvation for 228 combinations of (predominantly) organic solvents and solutes and compared our results to experimental data. Each system tested here consists of one solute molecule in a single component organic solvent. The results were compared to predictions using a quantitative structure−property relation (QSPR) model developed by Katritzky et al.[28−30] and the conductor-like screening models for realistic solvation (COSMO-RS), which relies on quantum mechanical calculations and which is considered to be one of the most accurate models for prediction of solvation energies.[31,32]

METHODS

Theory. The experimental solvation free energy used here is the work needed to transfer 1 mol of a solute from the gas phase to solution in equilibrium at room temperature at unit concentration (molarity) of the solutes. It is proportional to log L, the logarithm of the concentration Ostwald solubility coefficient.[33] The relationship between the vapor−liquid equilibrium and the solvation free energy is given by

$$\Delta G_{solv} = -2.3RT \log \frac{C_l}{C_g} \tag{1}$$

where R is the ideal gas constant, T is the temperature in Kelvin, and C_l and C_g are the liquid- and gas-phase molar concentrations of the solute, respectively. Based on eq 1, the experimental equilibrium solubility data, collected by Katritzky et al., and the solubility data predicted by Katritzky et al.[28−30] were converted to solvation free energies.

The solvation free energy was computed using thermodynamic integration (TI):

$$\Delta G_{sim} = \int_0^1 \langle \frac{\partial H}{\partial \lambda} \rangle_\lambda \, d\lambda \tag{2}$$

in which H is the Hamiltonian and where $\lambda = 0$ refers to the coupled state, in which the solute fully interacts with the solvent and $\lambda = 1$ refers to the decoupled state, in which the solute does not interact with solvent. Shirts et al. demonstrated that the difference between ΔG_{solv} and ΔG_{sim} depends on the natural logarithm of the ratio of the mean volume at $\lambda = 1$ and the mean volume of the pure solvent with the same number of solvent molecules as $\lambda = 1$ and showed that this difference can be safely neglected.[4]

The Coulombic interactions between the solute and the solvent molecules were decoupled first and the Lennard-Jones interactions after that. A soft-core potential was incorporated to prevent issues due to strong Lennard-Jones interactions.[34] The soft-core λ power was set to 1 and the parameter α_{LJ} to 0.5, as recommended by Shirts et al.[35] Thus, the nonbonded interaction between the solute and solvent molecules at a particular λ is described by

$$U_{solu-solv}(\lambda_C, \lambda_{LJ}) = \sum_i \sum_j (1 - \lambda_C) \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} + 4\varepsilon_{ij}(1 - \lambda_{LJ}) \times \left[\frac{1}{(0.5\lambda_{LJ} + (r_{ij}/\sigma_{ij})^6)^2} - \frac{1}{0.5\lambda_{LJ} + (r_{ij}/\sigma_{ij})^6} \right] \tag{3}$$

where λ_C and λ_{LJ} are the λ values for the transformation of columbic and van der Waals interactions, respectively, q_i and q_j are the partial charges on atoms i and j, r_{ij} is the distance between atoms i and j, ε_0 is the permittivity of the vacuum, σ_{ij} is the van der Waals radius, and ε_{ij} is the well-depth for atom pair i and j. The Coulombic interaction was decoupled using 25 λ values: 0, 0.02, 0.04, 0.07, 0.1, 0.15, 0.2, ..., 0.8, 0.85, 0.9, 0.93, 0.96, 0.98, 1. The same λ values were used to decouple the Lennard-Jones interaction, and therefore 50 λ simulations were performed for each of the 228 systems, 11 400 simulations in total.

TI Simulation Setup. The GROMACS software suite[36−40] (version 4.5) was used to perform molecular dynamics simulations from which the average $\langle \partial H/\partial \lambda \rangle$ at each λ value was calculated. A total of 49 neutral small organic molecules were used here (Table S2). The force field models of these molecules were taken from the Web server at http://virtualchemistry.org,[3,41] except for water, where the TIP3P model was used.[42] A detailed description of the derivation of the force field files is given by Caleman et al.[3] Briefly, the initial molecule models were built using PRODRG[43] or Molden[44] and optimized using the Gaussian 03 suite of programs[45] at the Hartree−Fock level with a 6-311G**basis set.[46,47] Then the Antechamber software[2,48] was employed for atomic partial charge calculation using the RESP method.[49] The amb2gmx.pl script[50] was used to convert the topologies from AMBER to GROMACS format. Each of the simulation cubic boxes was approximately $4 \times 4 \times 4$ nm^3 large and contained one solute molecule, while the remainder of the box was filled with solvent molecules using the GROMACS tool genbox.

In this work, we followed the protocol of Shirts et al.:[35]

1. Each periodic system was minimized using first steepest-descent minimization and then L-BFGS minimization for 5000 steps or until a maximum force of 100 kJ/(mol^{-1}nm^{-1}) was reached.

DOI: 10.1021/acs.jcim.5b00106
J. Chem. Inf. Model. 2015, 55, 1192−1201

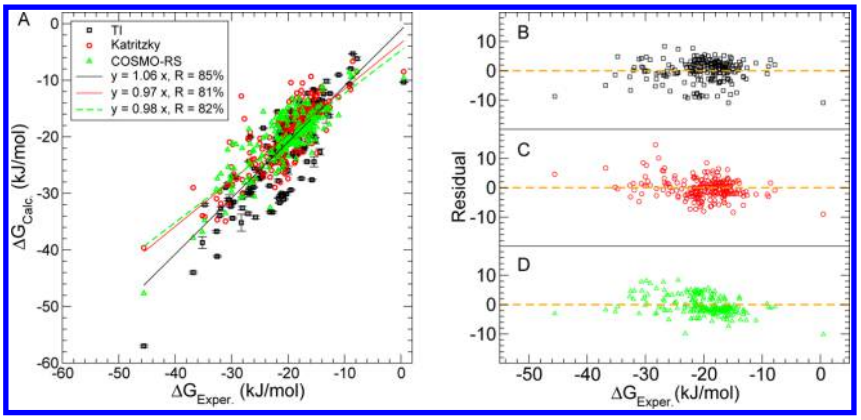


Figure 1. (A) Correlation between experimental and computed ΔG_{solv} from thermodynamics integration methods and (semi)empirical methods. Residual plot for (B) TI, (C) QSPR, and (D) COSMO-RS.

2. Initial velocities were assigned to obtain a Maxwell−Boltzmann distribution corresponding to a temperature of 300 K. Then, the systems were equilibrated for 100 ps using the NVT ensemble at each λ value. A Langevin (stochastic dynamics) integrator[51] was used with a reference temperature of 300 K, a friction coefficient of 2 ps^{-1}, and an integration time step of 2 fs. The LINCS algorighm[52,53] was used to constrain all bonds with an order of 12.

3. Systems were subjected to 100 ps of further equilibration using the NPT ensemble. The Parrinello−Rahman pressure coupling[54] algorithm was used to obtain an NPT ensemble for configurational sampling with a reference pressure of 1 bar, a 5 ps time constant for coupling, and the compressibility set to 5×10^{-5} bar^{-1}.

4. Production simulations were done in order to sample $\langle \partial H/\partial \lambda \rangle_\lambda$ at 50 different λ values according to eq 3. The sampling of $\langle \partial H/\partial \lambda \rangle_\lambda$ was continued until converged results were obtained (no drift in $\langle \partial H/\partial \lambda \rangle_\lambda$). For this reason, the simulation length of each system varyied between 2 and 14 ns (see Supporting Information spreadsheet file).

In all simulations, a neighbor list of 1.0 nm was used and updated every 10 steps to evaluate short-range interaction except for minimization steps, during which the neighbor list was updated every step. The Lennard-Jones interactions were switched off between 0.8 and 0.9 nm, and long-range dispersion corrections[55] were applied to energy and pressure. The Coulombic interaction was calculated using a real space cutoff of 1.0 nm combined with the Particle Mesh Ewald[56] (PME) algorithm for long-range electrostatics with an order of 6 and Fourier spacing of 0.12 nm.

Cubic spline interpolation was performed for integration weighted by the standard deviation in $\langle \partial H/\partial \lambda \rangle_\lambda$. The uncertainty of the calculated free energy ΔG_{sim} is given in the Supporting Information.

COSMO-RS Calculations. The COSMO-RS calculations were performed using the ADF software package.[57] All molecules can be found in the COSMO-RS database,[58−60] and therefore the ADF quantum mechanical calculation using COSMO can be skipped. Default COSMO-RS settings were used to derive the solvation free energy.

Systematic Error Analysis. The approach suggested by Mobley et al. to identify systematic errors caused by specific functional groups[7] was followed here. For this reason, the Boltzmann-enhanced discrimination of receiver-operating characteristic (BEDROC) metric[61] was computed. A labeled list of compounds was made using the unsigned error in ΔG_{solv} for either solvent or solute as the score for BEDROC calculations which were done, using the CROC package.[62] The BEDROC value is higher for functional groups that appear often with high unsigned errors than those with a uniform distribution ($BEDROC_{unif}$). The weighting factor α was set to 1 for the calculation of BEDROC values as discussed by Mobley et al.[7] Given α, the $BEDROC_{unif}$ can be computed with the known fraction of compounds corresponding to the functional group.[61] Bootstrapping with 400 iterations was performed to compute the uncertainty of a BEDROC value for each functional group. The functional groups of each compound were assigned automatically by the Checkmol program,[63] and only those which occurred at least 10 times were retained. The assignments were checked manually and the molecules containing each functional group are listed in Table S3. Finally, a Student's t test was performed to compare the unsigned error for compounds containing a specific functional group, as either solvent or solute based on the unsigned error for the entire set, using the statistics software package R.[64] The Student's t test shows whether the performance of a functional group is significantly different from the whole set as supplement of the BEDROC approach.

RESULTS

Overall Correlation and Statistics. The correlation between experiments and calculations for all 228 solvation energies is given in Figure 1A. Residual plots highlighting the deviation from experiment are given in Figure 1B−D for the three methods evaluated. The TI calculations give somewhat better agreement with experiment (higher correlation coefficient) than either the QSPR method[28−30] or the semiempirical COSMO-RS.[31,32] However, a comparison of the correlation between different methods shows that the differences are not significant (Table S1). A breakdown of the root-mean-square deviation (RMSD) and the mean signed error (MSE) per solvent is given in Table 1 for each of the three methods tested. The number of solutes per solvent varies from 5 (*o*-xylene, dibutyl ether) to 29 (1-octanol, chloroform) due to the availability of experimental data. Similarly, a breakdown of the RMSD and MSE per solute is given in Table 2. Here, the number of solvents ranges from 1 to 19 (ethanol) for the same reason. The total RMSD and MSE for all data points and three methods is given at the bottom of Table 1. We find the RMSD in the solvation energy based on QSPR to be 3.4 ± 0.2 kJ/mol (Table 1) which is comparable to a published estimate of 4 kJ/mol.[10] For COSMO-RS, we find the lowest RMSD (3.1 ± 0.2 kJ/mol) and for TI a RMSD

DOI: 10.1021/acs.jcim.5b00106
J. Chem. Inf. Model. 2015, 55, 1192−1201

Table 1. Root Mean Square Deviation (RMSD) and Mean Signed Error (MSE) from the Experimental Solvation Energy (kJ/mol) Per Solvent[a]

			Katritzky		COSMO-RS		TI	
solvent	formula	N	RMSD	MSE	RMSD	MSE	RMSD	MSE
1-octanol	$C_8H_{18}O$	29	**4.4 ± 1.0**	0.6 ± 0.8	2.9 ± 0.3	−0.5 ± 0.5	2.6 ± 0.4	−0.9 ± 0.6
acetone	C_3H_6O	12	3.8 ± 0.8	1.2 ± 1.1	2.6 ± 0.4	−0.8 ± 0.7	3.3 ± 1.1	−1.4 ± 0.9
acetonitrile	C_2H_3N	6	1.6 ± 0.3	0.6 ± 0.6	2.4 ± 0.9	−0.0 ± 0.9	1.1 ± 0.3	0.5 ± 0.5
acetophenone	C_8H_8O	8	2.1 ± 0.5	−0.2 ± 0.7	2.9 ± 0.7	0.5 ± 1.0	**4.2 ± 1.8**	−1.0 ± 1.5
benzonitrile	C_7H_5N	6	1.7 ± 0.5	0.0 ± 0.7	3.1 ± 1.0	1.5 ± 1.1	3.3 ± 1.6	−0.9 ± 1.5
chloroform	$CHCl_3$	29	3.6 ± 0.5	0.2 ± 0.7	**5.1 ± 0.6**	2.5 ± 0.9	**5.0 ± 1.0**	−1.4 ± 0.9
cyclohexanone	$C_6H_{10}O$	9	2.5 ± 0.4	0.6 ± 0.8	2.6 ± 0.5	−0.2 ± 0.9	**4.2 ± 1.4**	−1.0 ± 1.3
dibutyl ether	$C_8H_{18}O$	5	1.9 ± 0.1	−1.1 ± 0.7	**4.1 ± 0.6**	−1.9 ± 1.7	1.5 ± 0.3	1.3 ± 0.4
dimethyl sulfoxide	C_2H_6SO	9	3.2 ± 0.5	−0.4 ± 1.1	2.5 ± 0.4	−0.6 ± 0.8	2.9 ± 0.3	0.3 ± 1.0
ethanol	C_2H_6O	13	3.9 ± 0.8	0.2 ± 1.1	2.3 ± 0.3	−0.1 ± 0.6	3.6 ± 1.1	−1.0 ± 1.0
ethyl acetate	$C_4H_8O_2$	13	**4.1 ± 1.1**	1.1 ± 1.1	1.8 ± 0.4	−0.6 ± 0.5	**4.2 ± 1.2**	−2.1 ± 1.0
methanol	CH_4O	21	3.9 ± 0.5	0.3 ± 0.8	3.3 ± 0.6	−1.0 ± 0.7	**4.2 ± 0.8**	−1.4 ± 0.9
methyl acetate	$C_3H_6O_2$	7	2.4 ± 0.4	1.3 ± 0.8	1.5 ± 0.2	−1.4 ± 0.2	**4.5 ± 1.4**	−2.2 ± 1.5
N-methylformamide	C_2H_5NO	9	2.4 ± 0.4	0.7 ± 0.7	3.2 ± 0.6	0.2 ± 1.1	2.5 ± 0.4	0.9 ± 0.8
nitromethane	CH_3NO_2	10	1.7 ± 0.3	0.3 ± 0.6	2.2 ± 0.8	−0.1 ± 0.7	**4.4 ± 1.2**	0.3 ± 1.4
N,*N*-dimethylacetamide	C_4H_9NO	9	3.3 ± 0.5	0.3 ± 1.1	2.4 ± 0.5	−0.3 ± 0.8	2.2 ± 0.4	−0.3 ± 0.7
N,*N*-dimethylformamide	C_3H_7NO	9	3.9 ± 0.9	−0.1 ± 1.3	2.2 ± 0.4	−0.4 ± 0.7	3.9 ± 1.6	−1.5 ± 1.2
o-xylene	C_8H_{10}	5	**4.3 ± 0.8**	1.2 ± 1.9	1.8 ± 0.4	0.3 ± 0.8	0.7 ± 0.1	−0.3 ± 0.3
pyridine	C_5H_5N	8	1.7 ± 0.7	0.8 ± 0.5	3.5 ± 0.6	−1.6 ± 1.1	**4.0 ± 1.4**	0.1 ± 1.4
toluene	C_7H_8	11	1.6 ± 0.3	−0.6 ± 0.5	2.4 ± 0.5	0.4 ± 0.7	1.9 ± 0.4	−1.3 ± 0.5
total		228	3.4 ± 0.2	0.4 ± 0.2	3.1 ± 0.2	−0.0 ± 0.2	3.7 ± 0.2	−1.0 ± 0.2

[a] N is the number of solutes tested. Numbers in the column Katritzky were taken from refs 28−30, other numbers from this work. Outliers, defined as RMSD > 4 kJ/mol, are marked in bold. Bootstrapping with 1000 iterations were performed using the R program[64] to obtain the uncertainties of the RMSD and the MSE values for each solvent.

of 3.7 ± 0.2 kJ/mol. Despite TI having a slightly larger correlation coefficient R than the other methods (Figure 1), the RMSD is the highest of the three. The uncertainty in the experimental solvation free energy of neutral small organic molecules is not generally available; however, it has been estimated to be 0.8 kJ/mol,[65] which is similar to the statistical uncertainty in the TI calculations (see Supporting Information) but about 5 times smaller than the RMSD in the calculations.

Tables 1 and 2 allow the detection of whether certain solutes or solvents give rise to especially large deviations in TI, which could point to systematic errors in the force field. In order to facilitate such an analysis, we highlighted outliers in bold in the tables. A solvent or solute was defined to be an outlier if the RMSD > 4 kJ/mol. The BEDROC values per solvent and per solute are given in Tables 3 and 4, respectively. The BEDROC value depends on the number of compounds with a particular functional group relative to the number of the entire data points and the α value and is roughly 0.5 for a uniform distribution in our tests. The Student's t-test results are shown in Tables 5 and 6. Functional groups for which the BEDROC value is >0.5 and the p value is <0.05 are highlighted in bold in the respective tables for solvent and solute because for these groups systematic errors may occour.

Outliers in TI. Among the solvents, chloroform has the largest RMSD at 5.0 ± 1.0 kJ/mol (N = 29) followed by methyl acetate (N = 7, RMSD = 4.5 ± 1.4 kJ/mol), nitromethane (N = 10, RMSD = 4.4 ± 1.2 kJ/mol), methanol (N = 21, RMSD = 4.2 ± 0.8 kJ/mol), ethyl acetate (N = 13, RMSD = 4.2 ± 1.2 kJ/mol), and acetophenone (N = 8, RMSD = 4.2 ± 1.8 kJ/mol). The MSE is (by definition) smaller than the RMSD; the largest MSE for solvents are −2.2 ± 1.5 kJ/mol (methyl acetate) and −2.1 ± 1.0 kJ/mol (ethyl acetate). A large MSE signifies that the deviations are not random but rather depend on solvent−solute interactions being either too strong (MSE < 0) or too weak (MSE > 0) compared to solvent−solvent interactions. This notion is consistent with our earlier observations that the solvent−solvent interactions yield a significant contribution to solvation free energies of alkali and halide ions,[66] organic molecules,[67] as well as hydronium and hydroxide ions.[68] Several functional groups (alkyl chlorides and nitro compounds) in the solvents had BEDROC values slightly larger than $BEDROC_{unif}$ (Table 3, note that $BEDROC_{unif}$ differs slightly between functional groups). However, the Student's t test suggests that the difference is not significant.

The picture becomes somewhat clearer when we consider TI outliers in the solutes (Table 2) where several functional groups are found to be problematic. The BEDROC value and the p value of the Student's t test (Tables 4 and 6) highlight that the errors for nitro groups and carboxylic acid derivates are systematic. The ΔG_{solv} for nitromethane as a solute is systematically overestimated by approximately −10 kJ/mol, and a similar result is obtained for nitrobenzene (N = 1, error = −8.6 kJ/mol). Most likely, the charges on the nitro group are too high, which leads to an overestimated ΔG_{solv}, but which also have been suggested to cause too high density and vaporization enthalpy for especially nitromethane, 1-nitropropane, and 2-nitropropane.[3] It was recently concluded in a study of hydration free energy of nitroaromatic compounds that the Lennard-Jones parameters of these compounds need adjusting as well however.[69] Solutes containing ester groups are outliers as well. Ethyl acetate (N = 8, RMSD = 4.8 ± 0.2 kJ/mol) and methyl acetate (N = 4, RMSD = 4.7 ± 0.9 kJ/mol) have a MSE around −5 kJ/mol. For these compounds, the liquid density and enthalpy of vaporization were found to be too high as well.[3] These molecules have a high partial charge on some atoms; for instance, the carbon atom on the ester group of methyl acetate has q = +0.99e. This is known to be an

DOI: 10.1021/acs.jcim.5b00106
J. Chem. Inf. Model. 2015, 55, 1192−1201

Table 2. Root Mean Square Deviation (kJ/mol) and Mean Signed Error (MSE) from the Experimental Solvation Energy (kJ/mol) Per Solute[a]

			Katritzky		COSMO-RS		TI	
solute	formula	*N*	RMSD	MSE	RMSD	MSE	RMSD	MSE
1,2-dichloroethane	$C_2H_4Cl_2$	3	3.1 ± 1.5	1.9 ± 1.5	0.5 ± 0.1	−0.3 ± 0.2	0.7 ± 0.3	−0.5 ± 0.2
1-butanol	$C_4H_{10}O$	6	1.8 ± 0.2	1.0 ± 0.6	2.3 ± 0.7	−0.4 ± 0.9	1.8 ± 0.4	−1.5 ± 0.5
1-chlorobutane	C_4H_9Cl	4	1.2 ± 0.0	−0.0 ± 0.6	1.2 ± 0.2	−1.1 ± 0.2	1.8 ± 0.9	1.4 ± 0.7
1-octanol	$C_8H_{18}O$	1	1.2	1.2	1.6	−1.6	3.5	−3.5
1-pentanol	$C_5H_{12}O$	4	1.3 ± 0.5	0.8 ± 0.5	1.6 ± 0.3	−0.4 ± 0.8	2.7 ± 0.5	−2.8 ± 0.4
2-hexanone	$C_6H_{12}O$	3	2.2 ± 0.5	−0.9 ± 1.1	3.3 ± 0.5	−3.1 ± 0.5	3.8 ± 1.8	−3.3 ± 1.6
2-methyl-2-butanol	$C_5H_{12}O$	1	0.9	0.9	1.5	−1.5	**8.5**	**-8.5**
2-methyl-2-propanol	$C_4H_{10}O$	2	1.7 ± 0.3	0.4 ± 1.2	2.6 ± 0.9	−2.2 ± 0.9	**4.1 ± 0.8**	**-4.4 ± 0.8**
2-methylpyridine	C_6H_7N	1	1.7	1.7	**7.9**	**7.9**	2.0	2.0
3-methylpyridine	C_6H_7N	1	2.3	2.3	**8.4**	**8.4**	2.4	2.4
acetone	C_3H_6O	17	3.0 ± 0.5	0.0 ± 0.7	2.3 ± 0.2	−2.3 ± 0.2	1.6 ± 0.4	−0.9 ± 0.3
acetonitrile	C_2H_3N	12	1.9 ± 0.3	0.9 ± 0.5	2.2 ± 0.5	−1.8 ± 0.4	1.9 ± 0.2	1.1 ± 0.5
acetophenone	C_8H_8O	2	1.2 ± 0.6	0.9 ± 0.6	**4.5 ± 0.7**	**4.4 ± 0.7**	3.7 ± 1.3	−3.0 ± 1.2
anisole	C_7H_8O	2	1.7 ± 0.2	−1.7 ± 0.2	**4.2 ± 0.3**	**4.2 ± 0.3**	2.6 ± 0.8	−1.8 ± 1.3
benzaldehyde	C_7H_6O	1	3.0	3.0	**4.0**	**4.0**	1.9	−1.9
benzonitrile	C_7H_5N	3	1.2 ± 0.2	−0.0 ± 0.7	**4.2 ± 0.5**	**4.1 ± 0.5**	1.7 ± 0.7	−1.3 ± 0.5
chloroform	$CHCl_3$	15	**4.7 ± 0.6**	3.4 ± 0.8	1.3 ± 0.1	−1.1 ± 0.1	2.2 ± 0.2	2.0 ± 0.3
cyclohexanone	$C_6H_{10}O$	1	2.3	2.3	3.7	3.7	1.6	−1.6
dibromomethane	CH_2Br_2	1	**4.6**	**4.6**	0.8	−0.8	0.9	0.9
dibutyl ether	$C_8H_{18}O$	2	1.7 ± 0.6	−0.9 ± 1.1	**4.3 ± 0.1**	**-4.3 ± 0.1**	1.3 ± 0.0	−1.2 ± 0.0
dichloromethane	CH_2Cl_2	14	2.0 ± 0.3	1.4 ± 0.4	1.3 ± 0.1	−1.2 ± 0.1	1.5 ± 0.2	1.4 ± 0.2
diethylamine	$C_4H_{11}N$	3	1.7 ± 0.3	−1.6 ± 0.3	2.4 ± 1.0	1.7 ± 0.9	1.8 ± 0.8	0.7 ± 1.2
diethyl sulfide	$C_4H_{10}S$	1	**7.5**	**7.5**	**5.5**	**5.5**	**6.7**	**6.7**
diisopropyl ether	$C_6H_{14}O$	3	3.3 ± 1.2	−2.6 ± 1.2	**5.9 ± 0.5**	**-5.8 ± 0.5**	**10.3 ± 3.1**	**-8.4 ± 2.5**
dimethyl ether	C_2H_6O	2	2.0 ± 0.8	−1.7 ± 0.7	0.6 ± 0.0	−0.6 ± 0.0	1.1 ± 0.3	1.2 ± 0.2
dimethyl sulfoxide	C_2H_6SO	5	**9.5 ± 2.1**	**8.3 ± 2.1**	1.7 ± 0.4	−0.2 ± 0.7	2.9 ± 1.1	−2.9 ± 1.0
ethanol	C_2H_6O	19	2.0 ± 0.3	0.7 ± 0.4	2.3 ± 0.3	−1.5 ± 0.4	2.2 ± 0.4	−1.0 ± 0.5
ethyl acetate	$C_4H_8O_2$	8	3.6 ± 0.6	−2.1 ± 1.0	3.1 ± 0.4	−2.8 ± 0.4	**4.8 ± 0.2**	**-4.8 ± 0.2**
ethylbenzene	$C_8H_{10}O$	6	**4.3 ± 0.8**	−3.2 ± 1.1	**4.0 ± 0.4**	3.9 ± 0.4	1.9 ± 0.3	0.8 ± 0.8
formaldehyde	CH_2O	1	**9.0**	**-9.0**	**10.1**	**-10.1**	**10.8**	**-10.8**
isopropylbenzene	C_9H_{12}	1	**4.5**	**-4.5**	3.5	3.5	0.5	0.5
methanol	CH_4O	17	1.5 ± 0.1	−0.4 ± 0.4	2.2 ± 0.5	−1.3 ± 0.4	2.1 ± 0.3	−0.8 ± 0.5
methyl acetate	$C_3H_6O_2$	4	2.0 ± 0.4	−0.3 ± 1.0	1.6 ± 0.3	−1.3 ± 0.4	**4.7 ± 0.9**	**-4.7 ± 0.8**
morpholine	C_4H_9NO	1	1.9	−1.9	**6.1**	**6.1**	2.7	−2.7
n-butylamine	$C_4H_{11}N$	3	**5.1 ± 1.4**	**-4.2 ± 1.7**	2.6 ± 0.3	1.4 ± 1.3	0.2 ± 0.4	−0.3 ± 0.4
nitrobenzene	$C_6H_5NO_2$	1	2.1	−2.1	**5.2**	**5.2**	**8.6**	**-8.6**
nitromethane	CH_3NO_2	12	2.3 ± 0.4	−0.1 ± 0.7	0.6 ± 0.1	−0.2 ± 0.2	**9.7 ± 0.2**	**-9.7 ± 0.2**
N-methylformamide	C_2H_5NO	2	2.8 ± 1.2	−1.6 ± 1.6	1.7 ± 0.9	−1.2 ± 0.8	2.2 ± 0.3	2.3 ± 0.2
N,*N*-dimethylacetamide	C_4H_9NO	1	**9.4**	**9.4**	2.5	2.5	0.8	−0.8
N,*N*-dimethylformamide	C_3H_7NO	6	**5.9 ± 0.7**	**5.7 ± 0.7**	1.4 ± 0.3	1.2 ± 0.3	3.2 ± 0.6	−3.2 ± 0.6
o-xylene	C_8H_{10}	3	**4.2 ± 1.4**	−1.8 ± 2.2	3.7 ± 0.9	3.4 ± 0.9	0.8 ± 0.3	−0.0 ± 0.5
phenol	C_6H_6O	1	1.5	1.5	**8.8**	**8.8**	2.9	2.9
pyridine	C_5H_5N	3	1.8 ± 1.0	−0.9 ± 0.9	**6.2 ± 1.1**	**5.9 ± 1.0**	2.0 ± 0.5	1.9 ± 0.4
tetrahydrofuran	C_4H_8O	5	1.4 ± 0.2	0.0 ± 0.6	2.5 ± 0.3	2.4 ± 0.3	1.4 ± 0.3	−1.2 ± 0.3
thiophene	C_4H_4S	1	2.1	−2.1	**8.8**	**8.8**	**5.9**	**5.9**
toluene	C_7H_8	17	2.9 ± 0.4	−1.5 ± 0.6	**4.7 ± 0.2**	**4.6 ± 0.2**	1.9 ± 0.2	1.8 ± 0.2
triethylamine	$C_6H_{15}N$	3	**5.7 ± 1.4**	−0.2 ± 3.2	2.2 ± 0.7	−0.6 ± 1.2	**4.9 ± 1.5**	3.6 ± 1.6
triethyl phosphate	$C_6H_{15}PO_4$	1	**5.9**	**5.9**	2.1	−2.1	**11.4**	**-11.4**
water	H_2O	2	3.1 ± 0.7	−1.0 ± 2.0	**6.6 ± 3.1**	**-5.1 ± 3.0**	3.0 ± 1.1	2.5 ± 0.9

[a]*N* is the number of solvents tested. Numbers in the column Katritzky were taken from refs 28−30, other data from this work. Outliers, defined as RMSD > 4 kJ/mol, are marked in bold. Bootstrapping with 1000 iterations were performed using the R program[64] to obtain the uncertainties of the RMSD and the MSE values for each solute.

artifact of fitting point charges to reproduce the molecular electrostatic potential (ESP) determined from quantum chemistry. The ESP is highly correlated, and atoms far from the ESP data points can get large values because they become an "error sink" for the fitting procedure.[70−73] Even though most applications use the *ad-hoc* method known as restrained ESP (RESP),[49] this does not really resolve the problem.[74] In addition to outliers containing the functional groups mentioned above, triethyl phosphate (N = 1, error = −11.4 kJ/mol), diisopropyl ether (N = 3, RMSD = 10.3 ± 3.1 kJ/mol, MSE = −8.4 ± 2.5 kJ/mol),

DOI: 10.1021/acs.jcim.5b00106
J. Chem. Inf. Model. 2015, 55, 1192−1201

Table 3. BEDROC Values by Functional Group for the Different Functional Groups of Solvents, Compared to What Would Be Expected for a Uniform Distribution of the Same Number of Compounds ($BEDROC_{unif}$)[a]

functional group	N	$BEDROC_{unif}$	Katritzky	COSMO-RS	TI
carbonyl compound	29	0.44	0.40 ± 0.06	0.43 ± 0.05	0.41 ± 0.05
alcohol	63	0.46	**0.53 ± 0.04**	0.47 ± 0.04	0.47 ± 0.05
alkyl chloride	29	0.44	**0.51 ± 0.05**	**0.68 ± 0.06**	**0.59 ± 0.06**
carboxylic acid deriv.	47	0.45	0.47 ± 0.05	0.34 ± 0.04	0.44 ± 0.05
nitrile	12	0.43	0.26 ± 0.07	0.33 ± 0.10	0.26 ± 0.08
nitro compound	10	0.43	0.27 ± 0.07	0.23 ± 0.09	**0.58 ± 0.10**
aromatic compound	38	0.45	0.31 ± 0.05	0.41 ± 0.05	0.33 ± 0.05

[a]*N* represents the times the functional group occurred in our tests. The functional groups with BEDROC > 0.5 are marked in bold.

Table 4. BEDROC Values by Functional Group for the Different Functional Groups of Solutes, Compared to What Would Be Expected for a Uniform Distribution of the Same Number of Compounds ($BEDROC_{unif}$)[a]

functional group	N	$BEDROC_{unif}$	Katritzky	COSMO-RS	TI
carbonyl compound	25	0.44	0.48 ± 0.06	**0.56 ± 0.05**	0.34 ± 0.06
alcohol	50	0.45	0.28 ± 0.03	0.34 ± 0.04	039 ± 0.04
ether	15	0.43	0.32 ± 0.06	**0.63 ± 0.08**	0.37 ± 0.07
amine	10	0.42	**0.54 ± 0.10**	0.45 ± 0.10	0.37 ± 0.10
alkyl chloride	36	0.44	0.45 ± 0.06	0.17 ± 0.02	0.31 ± 0.04
carboxylic acid deriv.	21	0.43	**0.63 ± 0.07**	0.38 ± 0.06	**0.69 ± 0.05**
nitrile	15	0.43	0.29 ± 0.06	0.40 ± 0.07	0.34 ± 0.05
nitro compound	13	0.43	0.34 ± 0.08	0.13 ± 0.06	**0.98 ± 0.01**
aromatic compound	43	0.45	0.47 ± 0.05	**0.89 ± 0.02**	0.39 ± 0.04
heterocyclic compound	12	0.43	0.29 ± 0.06	**0.76 ± 0.08**	0.36 ± 0.07

[a]*N* represents the times the functional group occurred in our tests. The functional groups with BEDROC > 0.5 are marked in bold.

Table 5. Statistics from Student's *t* Test of the Unsigned Error Per Functional Group of Solvent Compounds and Mean Signed Error (MSE) from the Experimental Solvation Energy (kJ/mol) Per Functional Group[a]

		Katritzky			COSMO-RS			TI		
functional group	N	*t* value	*p* value	MSE	*t* value	*p* value	MSE	*t* value	*p* value	MSE
carbonyl compound	29	−4.0e-01	6.9e-01	0.6	4.6e-01	6.5e-01	−0.3	−7.4e-02	9.4e-01	−1.0
alcohol	63	−1.1e-01	9.2e-01	0.4	1.3e+00	2.1e-01	−0.6	1.3e+00	2.1e-01	−1.7
alkyl chloride	29	2.9e-01	7.8e-01	0.2	−3.0e+00	**5.9e-03**	2.5	6.7e-01	5.1e-01	−1.6
carboxylic acid deriv.	47	−5.8e-01	5.7e-01	0.7	1.2e+00	2.5e-01	−0.5	−1.4e-02	9.9e-01	−1.0
nitrile	12	1.3e-01	9.0e-01	0.3	−9.4e-01	3.6e-01	0.8	−7.7e-01	4.5e-01	−0.3
nitro compound	10	1.5e-01	8.8e-01	0.3	7.4e-02	9.4e-01	−0.1	−9.8e-01	3.5e-01	0.5
aromatic compound	38	5.7e-01	5.7e-01	0.1	−3.4e-01	7.3e-01	0.1	−5.3e-01	6.0e-01	−0.7

[a]*N* is the number of compounds in each functional group.

Table 6. Statistics from Student's *t* Test of the Unsigned Error Per Functional Group of Solute Compounds and Mean Signed Error (MSE) from the Experimental Solvation Energy (kJ/mol) Per Functional Group[a]

		Katritzky			COSMO-RS			TI		
functional group	N	*t* value	*p* value	MSE	*t* value	*p* value	MSE	*t* value	*p* value	MSE
carbonyl compound	25	7.7e-01	4.5e-01	−0.2	2.5e+00	**1.8e-02**	−1.7	1.5e+00	1.5e-01	−1.9
alcohol	50	8.8e-03	9.9e-01	0.4	3.7e+00	**3.7e-04**	−1.3	1.4e+00	1.6e-01	−1.6
ether	15	3.1e+00	**4.9e-03**	−1.2	2.7e-02	9.8e-01	−0.1	1.1e+00	2.8e-01	−2.0
amine	10	1.8e+00	1.0e-01	−2.0	−1.5e+00	1.6e-01	1.4	−1.8e+00	1.0e-01	0.7
alkyl chloride	36	−3.3e+00	**1.5e-03**	2.1	4.8e+00	**2.9e-06**	−1.1	−8.1e+00	**1.9e-13**	1.6
carboxylic acid deriv.	21	−6.6e-01	5.2e-01	1.0	1.9e+00	6.5e-02	−1.0	3.9e+00	**5.2e-04**	−3.6
nitrile	15	−6.5e-01	5.2e-01	0.7	7.9e-01	4.4e-01	−0.6	−3.8e+00	**1.0e-03**	0.9
nitro compound	13	9.0e-01	3.8e-01	−0.2	−5.0e-01	6.2e-01	0.2	2.8e+01	**2.2e-16**	−9.6
aromatic compound	43	3.5e+00	**7.8e-04**	−1.3	−1.5e+01	**2.2e-16**	4.8	−4.5e+00	**2.6e-05**	0.9
heterocyclic compound	12	1.1e+00	3.0e-01	−0.2	−6.2e+00	**3.9e-05**	5.1	−2.4e+00	**3.3e-02**	0.7

[a]*N* is the number of compounds in each functional group.

2-methyl-2-butanol (N = 1, error = −8.5 kJ/mol), as well as 2-methyl-2-propanol (N = 2, RMSD = 4.1 ± 0.8 kJ/mol, MSE = −4.4 ± 0.8 kJ/mol) might be problematic for the same reason. Mobley et al. found that the assigned partial charges of triethyl phosphate do significantly influence its hydration free energy.[15]

DOI: 10.1021/acs.jcim.5b00106
J. Chem. Inf. Model. 2015, 55, 1192−1201

All sulfide-containing solutes are outliers in the TI calculations. Diethyl sulfide has an error of 6.7 kJ/mol, and its liquid density and vaporization enthalpy are underestimated as well,[3] as was noticed by Wang et al.[75] Polarizability of the sulfur atom could play an important role in interactions, the lack of which is not compensated for by Lennard-Jones interactions.

QSPR. The QSPR model derived by Katritzky et al.[28−30] is an efficient method to evaluate the solvation free energy, which performs almost as well as the other two physics-based methods in our tests. There are outliers among the solvents (Table 1), 1-octanol (N = 29, RMSD = 4.4 ± 1.0 kJ/mol), ethyl acetate (N = 13, RMSD = 4.1 ± 1.1 kJ/mol), and *o*-xylene (N = 5, RMSD = 4.3 ± 0.8 kJ/mol), but their MSE is not significant, most likely because of the fitting procedure applied to derive the QSPR model. A number of solutes are outliers according to our criteria, but most of these outliers were used as solutes only for a small number of solvents. Considering the BEDROC metric (Tables 3 and 4) and Student's *t*-test results (Tables 5 and 6), it seems that none of the error of the functional groups is systematic.

COSMO-RS. Of the solvents used in the COSMO-RS method, chloroform has the largest RMSD at 5.1 ± 0.6 kJ/mol (N = 29) followed by dibutyl ether (4.1 ± 0.6 kJ/mol, N = 5). Virtually all outliers in the COSMO-RS ΔG_{solv} calculations contain rings, in particular aromatic groups. The results seem to be systematic since the MSE for aromatic solutes is positive, which means ΔG_{solv} is too low. There are five solutes containing a pyridine moiety with a RMSD of ∼7 kJ/mol. Even a common aromatic solute such as toluene (N = 17) performs relatively poorly (RMSD = 4.7 ± 0.2 kJ/mol). It was somewhat surprising to note that water is an outlier (N = 2, RMSD = 6.6 ± 3.1 kJ/mol, MSE = −5.1 ± 3.0 kJ/mol), mainly because the solvation free energy of water in methanol solvent was poorly predicted (−9.34 kJ/mol deviation from the expriment). The BEDROC values and p values of Student's t test of aromatic compounds and heterocyclic compounds confirms that their errors are systematic (Tables 4 and 6). Although the BEDROC metric and Student's t test suggests that the results for chloroform as a solvent are systematically off, this is likely caused by the solutes, 13 of which have (predominantly aromatic) rings. Once the compounds with rings are removed from the analysis, the BEDROC value for chloroform reduces from 0.68 to 0.51 (N = 16) and the p value increases to 0.89. Overall, COSMO-RS predicted the solvation free energy of the tested system with the lowest RMSD and the MSE as close to 0 in most cases. Compared with the other two methods, it performs better for the problematic solutes chloroform, *N*,*N*-dimethylformamide, and especially nitromethane.

DISCUSSION

Force fields are a critical component for the prediction of properties of materials using molecular simulation. Their accuracy and applicability depends on the availability of good reference data, and considerable effort is still required to validate and improve force fields. In a recent paper Caleman et al. compared the performance of OPLS/AA[1] and GAFF[2] by computing properties such as liquid density, enthalpy of vaporization, heat capacity, surface tension, compressibility, expansion coefficient, and the dielectric constant.[3] The solvation free energy is important for biomedical application, and a lot of literature reports the performance of force fields on hydration free energy. To our knowledge only a few large scale tests of solvation free energies of organic molecules in media other than water have been done.[76] Many force fields, like GAFF, were originally introduced for modeling drug-like small molecules in combination with the Amber force field.[77] Since biomolecules are composed of functional groups similar to the organic molecules tested in our work, the performance of force fields for computing the solvation free energy in organic solvent is instructive even in the context of biomolecular applications. The solvation free energy of molecules in organic solvents is interesting for applications in the chemical industry as well. Therefore, computationally efficient methods such as COSMO-RS and QSPR have been developed and used for ΔG_{solv} predictions in organic solvents. In particular, COSMO-RS can be used to predict multiple component liquid properties.[31,32]

The Generalized Amber force field is attractive for force field calculations since derivation of new models is almost completely automated using the Antechamber package,[48] facilitating high-throughput computational predictions. We also notice STaGE, a newly developed program that can generate GROMACS topology for force fields such as CGenFF and OPLS-AA automaticlly,[78] would faciliate a large scale force field free energy benchmark in the future. Here, we tested the performance of the GAFF for predicting the solvation free energy of organic molecules in organic solvents and compared the results to experimental data and to the two methods mentioned above.

The TI calculations presented here yield a slightly higher correlation coefficient (Figure 1) than either the QSPR method[28−30] or the semiempirical COSMO-RS,[31,32] at a significantly increased computational cost (see Figure S1). However, the RMSD from experiment for TI is somewhat higher than for the other methods, meaning there is no sigificant difference between the predictive power of the methods. The overall performance of GAFF is good considering that it was not aimed directly at organic liquids. All three methods achieved a RMSD within 4 kJ/mol in our tests, on par with our previous estimates for accuracy in the enthalpy of vaporization[3] and with Mobley's hydration free energy calculations (5 kJ/mol[7]). Even so, there is still considerable room for improvement in force fields like GAFF. Although a ring correction[79] was applied to the COSMO-RS calculations, molecules with rings still have large deviations (4−9 kJ/mol) from the experiment using this method. It should be noted that the ring correction term is not understood completely.[32]

In the evaluation of ΔG_{solv} presented here based on thermodynamics integration calculations, outliers are found for molecules continaing nitro, ester, and sulfoxide groups. Caleman et al. reported that the liquid density and vaporization enthalpy are reproduced poorly for such molecules as well.[3] Mobley et al. paid particular attention to the partial charge influence on hydration free energy of small molecules and found that the assigned partial charges influenced the results somewhat.[15] It seems that a different charge assignment scheme may be needed for polar molecules than the RESP method that is used in GAFF.[74] It should be kept in mind however that there is an interdependency between the Lennard-Jones parameters and partial charges, and hence it is important to tune the parameters for van der Waals interactions as well. In this light it is encouraging that the densities and enthalpies of vaporization predicted by GAFF can be improved by tuning the Lennard-Jones parameters.[75] Force field parametrization must be carried out focusing on pure liquids and solvation properties at the same time. This is crucial, since focusing on only liquid properties or solvation properties could result in distinct and incompatible force fields that work well only for the "training objective," as it happened for GROMOS 53A5 and 53A6.[9] The recently

DOI: 10.1021/acs.jcim.5b00106
J. Chem. Inf. Model. 2015, 55, 1192−1201

developed GROMOS $53A6_{OXY}$,[24] $53A6_{OXY+D}$,[25] and $53A6_{OXY+A}$[80] force fields tackled this problem by using both liquid and solvation properties in the calibration.

A force field consists of a potential energy function and a set of parameters. The GAFF[81] used here does not include explicit polarization, which is known to play a role in the transfer of molecules between different environments. This omission makes it almost impossible to find a set of nonbonded parameters that accurately reproduces liquid properties and solvation properties in all kinds of solvents simultaneously.[82] Hydrogen bonds, which are crucial in many compounds including biomolecules, are difficult to treat with fixed point charges for example,[24] while the ubiquitous Lennard-Jones potential was suggested to be an important source of error for studies of mixing water and alcohol.[83] Recent efforts to develop complete biomolecular force fields including polarization[13,14] are a significant step toward phase-transferable models. In the case of the AMOEBA force field,[13] the Lennard-Jones potential was replaced by a more physical function.[84] We have previously demonstrated that modern polarizable potentials can yield a new level of insight for free energy calculations of ions in water.[66,68] Further studies of polarizable organic liquids will need to be done to show whether these can improve the accuracy of predictions compared to the simpler force field used in this work. We hope that this work contributes to building improved force field models and to encouraging other large-scale systematic comparisons of force field predictions to more computation-efficient models.

■ ASSOCIATED CONTENT

Supporting Information

Detailed figures showing the correlation between experimental data and calculated per solvent are shown. All $\partial H/\partial\lambda$ curves are given. The experimental data and predicted data by three methods of each system are provided as well in a portable (semicolon-separated values) spreadsheet file. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.5b00106.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +46 18 4714205. E-mail: david.vanderspoel@icm.uu.se.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

J.Z. was supported in part by a scholarship from Zhejiang University. B.T. was supported by the Erasmus Mundus MAHEVA scholarship program. The Swedish research council is acknowledged for financial support to D.v.d.S. (grant 2013-5947), and for a grant of computer time (SNIC2013-26-6) through the High Performance Computing Center North in Umeå, Sweden.

■ REFERENCES

(1) Jorgensen, W. L.; Tirado-Rives, J. Potential Energy Functions for Atomic-Level Simulations of Water and Organic and Biomolecular Systems. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6665−6670.

(2) 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.

(3) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Compressibility, Expansion Coefficient and Dielectric Constant. *J. Chem. Theory Comput.* **2012**, *8*, 61−74.

(4) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. Extremely Precise Free Energy Calculations of Amino Acid Side Chain Analogs: Comparison of Common Molecular Mechanics Force Fields for Proteins. *J. Chem. Phys.* **2003**, *119*, 5740−5761.

(5) Hess, B.; van der Vegt, N. F. A. Hydration Thermodynamic Properties of Amino Acid Analogues: A Systematic Comparison of Biomolecular Force Fields and Water Models. *J. Phys. Chem. B* **2006**, *110*, 17616−17626.

(6) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Dill, K. A. Predictions of Hydration Free Energies from All-Atom Molecular Dynamics Simulations. *J. Phys. Chem. B* **2009**, *113*, 4533−4537.

(7) 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.

(8) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. Correction to Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations. *J. Chem. Theory Comput.* **2015**, *11*, 1347−1347.

(9) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656−1676.

(10) Hansen, N.; van Gunsteren, W. F. Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.* **2014**, *10*, 2632−2647.

(11) Hetenyi, C.; Paragi, G.; Maran, U.; Timar, Z.; Karelson, M.; Penke, B. Combination of a Modified Scoring Function with Two-Dimensional Descriptors for Calculation of Binding Affinities of Bulky, Flexible Ligands to Proteins. *J. Am. Chem. Soc.* **2006**, *128*, 1233−1239.

(12) Cailliez, F.; Pernot, P. Statistical Approaches to Forcefield Calibration and Prediction Uncertainty in Molecular Simulation. *J. Chem. Phys.* **2011**, *134*, 054124.

(13) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.

(14) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, J.; Alexander, D. Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput* **2013**, *9*, 5430−5449.

(15) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. Comparison of Charge Models for Fixed-Charge Force Fields: Small-Molecule Hydration Free Energies in Explicit Solvent. *J. Phys. Chem. B* **2007**, *111*, 2242−2254.

(16) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. Comparison of Charge Models for Fixed-Charge Force Fields: Small Molecule Hydration Free Energies in Explicit Solvent. *J. Phys. Chem. B* **2011**, *115*, 1329−1332.

(17) Martins, S. A.; Sousa, S. F.; Ramos, M. J.; Fernandes, P. A. Prediction of Solvation Free Energies with Thermodynamic Integration Using the General Amber Force Field. *J. Chem. Theory Comput.* **2014**, *10*, 3570−3577.

(18) Knight, J. L.; Yesselman, J. D.; Brooks, C. L. Assessing the Uality of Absolute Hydration Free Energies Among CHARMM-Compatible Ligand Parameterization Schemes. *J. Comput. Chem.* **2013**, *34*, 893−903.

(19) Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. Free Energies of Hydration and Pure Liquid Properties of Hydrocarbons from the OPLS All-Atom Model. *J. Phys. Chem.* **1994**, *98*, 13077−13082.

(20) Udier-Blagovic, M.; Tirado, P. M. D.; Pearlman, S. A.; Jorgensen, W. L. Accuracy of Free Energies of Hydration Using CM1 and CM3 Atomic Charges. *J. Comput. Chem.* **2004**, *25*, 1322−1332.

(21) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 1509−1519.

(22) Shivakumar, D.; Harder, E.; Damm, W.; Friesner, R. A.; Sherman, W. Improving the Prediction of Absolute Solvation Free Energies Using

DOI: 10.1021/acs.jcim.5b00106
J. Chem. Inf. Model. 2015, 55, 1192−1201

the Next Generation OPLS Force Field. *J. Chem. Theory Comput.* **2012**, *8*, 2553−2558.
(23) Wallin, G.; Nervall, M.; Carlsson, J.; Åqvist, J. Charges for Large Scale Binding Free Energy Calculations with the Linear Interaction Energy Method. *J. Chem. Theory Comput.* **2009**, *5*, 380−395.
(24) Horta, B. A. C.; Fuchs, P. F. J.; van Gunsteren, W. F.; Hunenberger, P. H. New Interaction Parameters for Oxygen Compounds in the GROMOS Force Field: Improved Pure-Liquid and Solvation Properties for Alcohols, Ethers, Aldehydes, Ketones, Carboxylic Acids, and Esters. *J. Chem. Theory Comput.* **2011**, *7*, 1016−1031.
(25) Fuchs, P. F. J.; Hansen, H. S.; Hünenberger, P. H.; Horta, B. A. C. A GROMOS Parameter Set for Vicinal Diether Functions: Properties of Polyethyleneoxide and Polyethyleneglycol. *J. Chem. Theory Comput.* **2012**, *8*, 3943−3963.
(26) van der Spoel, D.; van Buuren, A. R.; Tieleman, D. P.; Berendsen, H. J. C. Molecular Dynamics Simulations of Peptides from BPTI: A Closer Look at Amide-Aromatic Interactions. *J. Biomol. NMR* **1996**, *8*, 229−238.
(27) van der Spoel, D.; Lindahl, E. Brute-Force Molecular Dynamics Simulations of Villin Headpiece: Comparison with NMR Parameters. *J. Phys. Chem. B* **2003**, *107*, 11178−11187.
(28) Katritzky, A. R.; Oliferenko, A. A.; Oliferenko, P. V.; Petrukhin, R.; Tatham, D. B.; Maran, U.; Lomaka, A.; Acree, W. E. A general Treatment of Solubility. 1. The QSPR Correlation of Solvation Free Energies of Single Solutes in Series of Solvents. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1794−1805.
(29) Katritzky, A. R.; Oliferenko, A. A.; Oliferenko, P. V.; Petrukhin, R.; Tatham, D. B.; Maran, U.; Lomaka, A.; Acree, W. E. A General Treatment of Solubility. 2. QSPR Prediction of Free Energies of Solvation of Specified Solutes in Ranges of Solvents. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1806−1814.
(30) Katritzky, A. R.; Tulp, I.; Fara, D. C.; Lauria, A.; Maran, U.; Acree, W. E. A General Treatment of Solubility. 3. Principal Component Analysis (PCA) of the Solubilities of Diverse Solutes in Diverse Solvents. *J. Chem. Inf. Model.* **2005**, *45*, 913−923.
(31) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99*, 2224−2235.
(32) Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. Refinement and Parametrization of COSMO-RS. *J. Phys. Chem. A* **1998**, *102*, 5074−5085.
(33) Battino, R. The Ostwald coefficient of gas solubility. *Fluid Phase Equilib.* **1984**, *15*, 231−240.
(34) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding Singularities and Numerical Instabilities in Free Energy Calculations Based on Molecular Simulations. *Chem. Phys. Lett.* **1994**, *222*, 529−539.
(35) Shirts, M. R.; Pande, V. S. Comparison of Efficiency and Bias of Free Energies Computed by Exponential Averaging, the Bennett Acceptance Ratio, and Thermodynamic Integration. *J. Chem. Phys.* **2005**, *122*, 144107.
(36) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(37) Lindahl, E.; Hess, B.; Van Der Spoel, D. GROMACS 3.0: a Package for Molecular Simulation and Trajectory Analysis. *J. Mol. Model* **2001**, *7*, 306−317.
(38) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(39) Hess, B.; Kutzner, C.; Van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(40) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: a High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845−854.
(41) van der Spoel, D.; van Maaren, P. J.; Caleman, C. GROMACS Molecule & Liquid Database. *Bioinformatics* **2012**, *28*, 752−753.
(42) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(43) Schuettelkopf, A. W.; van Aalten, D. M. F. PRODRG - a Tool for High-Throughput Crystallography of Protein-Ligand Complexes. *Act. Crystallogr. D* **2004**, *60*, 1355−1363.
(44) Schaftenaar, G.; Noordik, J. H. Molden: a Pre- and Post-Processing Program for Molecular and Electronic Structures. *J. Comput. Aid. Mol. Des.* **2000**, *14*, 123−134.
(45) Frisch, M. J. et al. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(46) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent Molecular-orbital Methods. 20. Basis Set for Correlated Wave-functions. *J. Chem. Phys.* **1980**, *72*, 650−654.
(47) McLean, A. D.; Chandler, G. S. Contracted Gaussian-basis Sets for Molecular Calculations. 1. 2nd Row Atoms, Z=11−18. *J. Chem. Phys.* **1980**, *72*, 5639−5648.
(48) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Antechamber, an Accessory Software Package for Molecular Mechanical Calculations. *J. Comput. Chem.* **2005**, *25*, 1157−1174.
(49) 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.
(50) Mobley, D. L.; Chodera, J. D.; Dill, K. A. On the Use of Orientational Restraints and Symmetry Corrections in Alchemical Free Energy Calculations. *J. Chem. Phys.* **2006**, *125*, 84902.
(51) van Gunsteren, W. F.; Berendsen, H. J. C. A Leap-Frog Algorithm for Stochastic Dynamics. *Mol. Sim.* **1988**, *1*, 173−185.
(52) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(53) Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 116−122.
(54) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(55) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford Science Publications: Oxford, 1987.
(56) 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*, 8577−8592.
(57) Pye, C. C.; Ziegler, T.; van Lenthe, E.; Louwen, J. N. An Implementation of the Conductor-Like Screening Model of Solvation within the Amsterdam Density Functional Package — Part II. COSMO for Real Solvents 1. *Can. J. Chem.* **2009**, *87*, 790−797.
(58) Mullins, E.; Oldland, R.; Liu, Y. A.; Wang, S.; Sandler, S. I.; Chen, C.-C.; Zwolak, M.; Seavey, K. C. Sigma-Profile Database for Using COSMO-Based Thermodynamic Methods. *Ind. Eng. Chem. Res.* **2006**, *45*, 4389−4415.
(59) Mullins, E.; Liu, Y. A.; Ghaderi, A.; Fast, S. D. Sigma Profile Database for Predicting Solid Solubility in Pure and Mixed Solvent Mixtures for Organic Pharmacological Compounds with COSMO-Based Thermodynamic Methods. *Ind. Eng. Chem. Res.* **2008**, *47*, 1707−1725.
(60) Phillips, K. L.; Sandler, S. I.; Greene, R. W.; Toro, D. M. D. Quantum Mechanical Predictions of the Henry's Law Constants and Their Temperature Dependence for the 209 Polychlorinated Biphenyl Congeners. *Environ. Sci. Technol.* **2008**, *42*, 8412−8418.
(61) Truchon, J.-F.; Bayly, C. I. Evaluating Virtual Screening Methods: Good and Bad Metrics for the "Early Recognition" Problem. *J. Chem. Inf. Model.* **2007**, *47*, 488−508.
(62) Swamidass, S. J.; Azencott, C.-A.; Daily, K.; Baldi, P. A CROC Stronger than ROC: Measuring, Visualizing and Optimizing Early Retrieval. *Bioinformatics* **2010**, *26*, 1348−1356.

DOI: 10.1021/acs.jcim.5b00106
J. Chem. Inf. Model. 2015, 55, 1192−1201

(63) Haider, N. Functionality Pattern Matching as an Efficient Complementary Structure/Reaction Search Tool: an Open-Source Approach. *Molecules* **2010**, *15*, 5079−5092.
(64) R Core Team, *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2015.
(65) Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. Extension of the Platform of Applicability of the SM5.42R Universal Solvation Model. *Theor. Chem. Acc.* **1999**, *103*, 9−63.
(66) Caleman, C.; Hub, J. S.; van Maaren, P. J.; van der Spoel, D. Atomistic Simulation of Ion Solvation in Water Explains Surface Preference of Halides. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6838−6842.
(67) Hub, J. S.; Caleman, C.; van der Spoel, D. Organic Molecules on the Surface of Water Droplets an Energetic Perspective. *Phys. Chem. Chem. Phys.* **2012**, *14*, 9537−9545.
(68) Hub, J. S.; Wolf, M. G.; Caleman, C.; van Maaren, P. J.; Groenhof, G.; van der Spoel, D. Thermodynamics of Hydronium and Hydroxide Surface Solvation. *Chem. Sci.* **2014**, *5*, 1745−1749.
(69) Ahmed, A.; Sandler, S. I. Hydration Free Energies of Multifunctional Nitroaromatic Compounds. *J. Chem. Theory Comput.* **2013**, *9*, 2774−2785.
(70) Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984**, *5*, 129−145.
(71) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11*, 431−439.
(72) Francl, M. M.; Carey, C.; Chirlian, L. E.; Gange, D. M. Charges Fit to Electrostatic Potentials. 2. Can Atomic Charges be Unambiguously Fit to Electrostatic Potentials? *J. Comput. Chem.* **1996**, *17*, 367−383.
(73) Francl, M. M.; Chirlian, L. E. The Pluses and Minuses of Mapping Atomic Charges to Electrostatic Potentials. *Rev. Comput. Chem.* **2000**, *14*, 1−31.
(74) Sigfridsson, E.; Ryde, U. Comparison of Methods for Deriving Atomic Charges from the Electrostatic Potential and Moments. *J. Comput. Chem.* **1998**, *19*, 377−395.
(75) Wang, J.; Tingjun, H. Application of Molecular Dynamics Simulations in Molecular Property Prediction. 1. Density and Heat of Vaporization. *J. Chem. Theory Comput.* **2011**, *7*, 2151−2165.
(76) Duffy, E.; Jorgensen, W. Prediction of Properties from Simulations: Free Energies of Solvation in Hexadecane, Octanol, and Water. *J. Am. Chem. Soc.* **2000**, *122*, 2878−2888.
(77) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(78) Lundborg, M.; Lindahl, E. Automatic GROMACS Topology Generation and Comparisons of Force Fields for Solvation Free Energy Calculations. *J. Phys. Chem. B* **2015**, *119*, 810−823.
(79) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. New Model for Calculation of Solvation Free Energies: Correction of Self-Consistent Reaction Field Continuum Dielectric Theory for Short-Range Hydrogen-Bonding Effects. *J. Phys. Chem.* **1996**, *100*, 11775−11788.
(80) Horta, B. A. C.; Lin, Z.; Huang, W.; Riniker, S.; van Gunsteren, W. F.; Hünenberger, P. H. Reoptimized Interaction Parameters for the Peptide-Backbone Model Compound N-methylacetamide in the GROMOS Force Field: Influence on the Folding Properties of Two Beta-Peptides in Methanol. *J. Comput. Chem.* **2012**, *33*, 1907−1917.
(81) Wang, W.; Skeel, R. D. Fast evaluation of polarizable forces. *J. Chem. Phys.* **2005**, *123*, 164107.
(82) Yu, H. B.; van Gunsteren, W. F. Accounting for polarization in molecular simulation. *Comput. Phys. Commun.* **2005**, *172*, 69−85.
(83) Wensink, E. J. W.; Hoffmann, A. C.; van Maaren, P. J.; van der Spoel, D. Dynamic Properties of Water/Alcohol Mixtures Studied by Computer Simulation. *J. Chem. Phys.* **2003**, *119*, 7308−7317.
(84) Halgren, T. A. The representation of van der Waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters. *J. Am. Chem. Soc.* **1992**, *114*, 7827−7843.

DOI: 10.1021/acs.jcim.5b00106
J. Chem. Inf. Model. 2015, 55, 1192−1201