# Evaluation of CM5 Charges for Condensed-Phase Modeling

Jonah Z. Vilseck, Julian Tirado-Rives, and William L. Jorgensen*

Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, United States

S *Supporting Information*

ABSTRACT: The recently developed Charge Model 5 (CM5) is tested for its utility in condensed-phase simulations. The CM5 approach, which derives partial atomic charges from Hirshfeld population analyses, provides excellent results for gas-phase dipole moments and is applicable to all elements of the periodic table. Herein, the adequacy of scaled CM5 charges for use in modeling aqueous solutions has been evaluated by computing free energies of hydration (ΔG_{hyd}) for 42 neutral organic molecules via Monte Carlo statistical mechanics. An optimal scaling factor for the CM5 charges was determined to be 1.27, resulting in a mean unsigned error (MUE) of 1.1 kcal/mol for the free energies of hydration. Testing for an additional 20 molecules gave an MUE of 1.3 kcal/mol. The high precision of the results is confirmed by free energy calculations using both sequential perturbations and complete molecular annihilation. Performance for specific functional groups is discussed; sulfur-containing molecules yield the largest errors. In addition, the scaling factor of 1.27 is shown to be appropriate for CM5 charges derived from a variety of density functional methods and basis sets. Though the average errors from the 1.27*CM5 results are only slightly lower than those using 1.14*CM1A charges, the broader applicability and easier access to CM5 charges via the Gaussian program are additional attractive features. The 1.27*CM5 charge model can be used for an enormous variety of applications in conjunction with many fixed-charge force fields and molecular modeling programs.
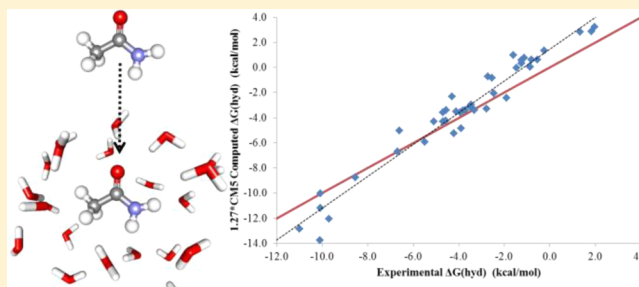
■ INTRODUCTION

Partial atomic charges are an integral part of today's molecular mechanic force fields.[1] Although force fields that involve more rigorous potential energy functions, such as those including polarizability or multipole electrostatics, are growing in popularity and availability,[2,3] fixed-charge force fields continue to receive extensive use for modeling condensed-phase systems.[4−6] The increasing role of computational chemistry in drug discovery and biological modeling often requires the ability to model an arbitrary molecule which has not been explicitly parametrized. While intramolecular and Lennard-Jones parameters are often easily estimated or transferable for new organic functional groups or heterocycles, the assignment of partial atomic charges is not obvious.[7] If experimental data are available, force fields can be parametrized specifically for these molecules, to reproduce liquid-state properties, but this situation is generally rare for many components of drugs such as heterocycles. Alternatively, quantum mechanical (QM) calculations can be performed, whereby atomic charges are derived from the wave functions to reproduce gas-phase dipole moments or electrostatic potentials.[7−10] For use with fixed-charge models in condensed-phase simulations, it is then necessary to scale the gas-phase charges to take into account polarization effects in an average sense.[11−13] Optimization of the scaling factors often focuses on minimizing errors in computed free energies of hydration.

In 2004, Charge Models 1 and 3 (CM1 & CM3)[14,15] were tested for their utility in condensed-phase simulations.[13]

Charges were computed in conjunction with AM1[16] and PM3[17] semiempirical quantum methods (creating the CMxA and CMxP models) and tested for their accuracy to reproduce experimental free energies of hydration for 25 neutral organic molecules. The CM1A and CM3A charge models were found to be preferred. Optimal scaling factors of 1.14 and 1.15 were shown to give the lowest mean unsigned errors (MUEs) of 1.03 and 1.13 kcal/mol for CM1A and CM3A, respectively. Due to the better performance of the 1.14*CM1A charges in reproducing free energies of hydration for amides, this method was taken to be preferred for simulations of biomolecular systems in explicit water.[13]

Since the introduction of the CM1A and CM1P charge models in 1995, successive refinements have been made to produce Charge Models (CMx) 2,[18] 3,[15] 4,[19] and 4M.[20] These charge models transform Mulliken,[21] Löwdin,[22] or redistributed Löwdin[23] population analyses to derive atomic charges capable of accurately reproducing experimental gas-phase dipole moments. Most recently, Charge Model 5 (CM5)[24] was published, which utilizes gas-phase atomic charges from a Hirshfeld population analyses.[25] In that report, Hirshfeld charges were shown to be less sensitive to basis set size than

Löwdin or Mulliken charges and more consistent across a wide range of electronic structure methods. These characteristics are maintained in CM5, allowing the model to consist of only a single set of parameters for various combinations of electronic structure methods and basis sets. CM5 parameters were fit using a test set of 614 neutral molecules containing 26 different elements (H−Ca, Zn, Ge−Br, and I) and were extended to every element in the periodic table via recursive fitting; the mean unsigned error is only 0.26 D. Thus, a new charge model has been created for potential use for an enormous range of molecules and applications.

With the noted success of CM5 for reproducing gas-phase electrostatic properties, the present investigation was initiated to determine its utility in the condensed phase, explicit solvent simulations. Free energies of hydration provide an important measure for testing force field parameters, especially atomic charges, as witnessed by the large number of publications computing this property within the past few years.[26−34] Experimental data are available, and applications to larger systems in chemistry, biology, and medicine are clear.[7] In this report, free energies of hydration are computed initially for an expanded set of 42 neutral, organic molecules to determine an optimal scaling factor for CM5 charges. Subsequent calculations considered an additional 20 small molecules for further testing. Free energies of hydration obtained using the scaled CM5 charges are compared with those derived from other charge methods including 1.14*CM1A and OPLS-AA.

COMPUTATIONAL DETAILS

Atomic Charges. CM1A[14] and CM5[24] partial atomic charges were computed via single point (SP) calculations on geometries optimized using the BOSS program[35] with the OPLS-AA force field.[4,7] This included the use of the nonbonded parameters for thiols and sulfides, which were updated in 2001.[36] CM1A charges were computed from AM1 wave functions by BOSS as before.[7,13] Hirshfeld atomic charges[25] were obtained from a M06-2X[37] density function theory (DFT) single-point calculation using the 6-311+G-(2df,2p) basis set[38] in conjunction with a Hirshfeld population analysis in Gaussian 09, revision A.02.[39] CM5 charges were then calculated with the CM5PAC software, which uses the Hirshfeld charges from the Gaussian output file.[40] It should be noted that in the course of completing this study, Gaussian 09 revision D.01 was released,[41] which automatically incorporates a CM5 charge calculation with every Hirshfeld population analysis. The CM1A and CM5 charges were averaged as usual for equivalent atoms, e.g., hydrogen atoms in a methyl group. This is necessary for force fields to avoid nonphysical asymmetries, such as different energies for identical rotamers.[11,13] Scaled CM1A and CM5 atomic charges for the molecules considered here are recorded in Table S1 of the Supporting Information.

Monte Carlo/Free Energy Perturbations. Metropolis Monte Carlo statistical mechanics[42] in conjunction with Free Energy Perturbation[43] (MC/FEP) calculations were carried out using BOSS to compute relative and absolute free energies of hydration for the 42 molecules. The FEP calculations were based upon the Zwanzig equation (eq 1).[43] The free energy difference between an initial (A) and final state (B) of a system is calculated as an ensemble average of the potential energy difference between states, sampled at the initial state.[1,43−46] To apply FEP theory to chemical equilibria, a thermodynamic cycle is employed (Scheme 1)[44−46] in which the initial and final

Scheme 1

$$\Delta\Delta G(A \rightarrow B) = \Delta G_B - \Delta G_A = \Delta G_2 - \Delta G_1$$
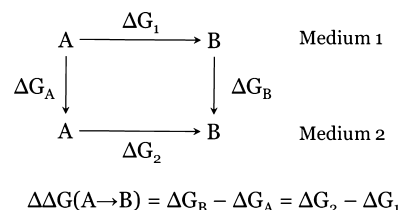
states could be different molecules, conformers, or different force field parameters for the same molecule. Conversion of A to B in two media, ΔG_1 and ΔG_2, provides the difference in free energy of transfer, $\Delta G_B - \Delta G_A$ (Scheme 1). A difference in free energies of hydration can be obtained by taking medium 1 as the gas phase and medium 2 as water. To achieve good convergence, A → B transformations are split into a series of steps, or "windows," categorized by a coupling parameter, λ_i (eq 2), which scales all geometrical and force field parameters, χ_i, as A transforms to B. Such calculations were first performed in 1985.[46]

$$\Delta G(A \rightarrow B)$$
$$= G_B - G_A$$
$$= -k_b T \ln\langle \exp[-(E_B - E_A)/k_b T]\rangle_A \qquad (1)$$

$$\chi_i = \lambda_i \chi_B + (1 - \lambda_i)\chi_A \qquad (2)$$

In BOSS, two sampling schemes are implemented for computing free-energy differences, double-wide sampling (DWS)[44,46] and simple overlap sampling (SOS).[44,47] Both methods deliver accurate results when more than 10 windows are used for molecular transformations involving a single non-hydrogen atom, e.g., $CH_3 \rightarrow H$.[44] In this work, both SOS and DWS have been employed, as discussed below.

Molecular Perturbations. Absolute free energies of hydration can be computed in two ways. Historically, relative molecular perturbations between two molecules (A and B in Scheme 1) have been favored due to their precision and viable computational expense.[46] If changes between A and B are small, then sufficient convergence is achieved using ca. 10 FEP windows with several million Monte Carlo configurations.[13] Relative free energy values can then be added to a reference absolute free energy to derive new absolute free energies. This works well when only a few transformations are needed to connect to the reference molecule, often methane or benzene; however, errors accumulate for longer sequences of transformations. The alternative approach is to perform complete molecular annihilations, which has been increasingly applied in view of the enhancement in computing resources.[26,27,29,31] These calculations make the entire molecule "disappear" from a medium by scaling all nonbonded force-field parameters to zero. While demanding long molecular dynamics or Monte Carlo runs and small FEP steps, these calculations provide a means to calculate an absolute free energy of solvation for any molecule, independent of previous calculations. The statistical uncertainties in the results can also be systematically reduced by increasing the averaging periods. Both types of calculations were performed in this study.

All MC/FEP calculations were carried out with BOSS and closely followed previously reported procedures.[13,44] For aqueous simulations, the TIP4P water model was used.[48] A single solute was placed in a cubic box with 500 water

molecules, and simulations were conducted at 25 °C and 1 atm in the isothermal–isobaric ensemble. Solvent–solvent and solute–solvent cutoff distances, R_c, for calculating nonbonded interactions were set at 10 Å based on all non-hydrogen atom pairs; i.e., if any heavy atom–heavy atom distance is below R_c, the entire intermolecular interaction is evaluated. As before,[4] the intermolecular interactions are quadratically feathered to zero over the last 0.5 Å based on the shortest heavy atom–heavy atom distance, R_m; the scaling factor is $(R_c^2 - R_m^2)/(R_c^2 - R_l^2)$ where $R_l = R_c - 0.5$. All internal degrees of freedom were sampled for the solute, while TIP4P water molecules underwent rigid-body translations and rotations. Solute and volume moves were attempted every 100 and 3125 configurations, respectively. Ranges for translations and rotations of ±0.06 Å and ±6.0° for solutes and ±0.15 Å and ±15.0° for the water molecules were chosen to produce MC acceptance rates of *ca.* 30–50%. Dummy atoms were used when perturbations were performed between molecules with different numbers of atoms.[46] Dummy atoms have zero valued nonbonded parameters, equilibrium bond lengths of 0.30 Å, and force constants for bond stretching and angle bending as in their real counterparts. Statistical uncertainties ($\pm 1\sigma$) were calculated from the batch means procedure using batch sizes of 1 M configurations;[44,49] additional tests of precision are described below. For all ΔG_{hyd} calculations, final uncertainties are below 0.25 kcal/mol, which is comparable to typical experimental uncertainties of 0.3 kcal/mol.[50,51]

This work began by computing ΔG_{hyd} for the 42 molecules using 1.14*CM1A and OPLS-AA charges in the A → B manner. Twenty-one windows of simple overlap sampling (21-SOS) with 5 million configurations of equilibration and averaging (5M/5M) in the gas phase and 15M/30M configurations in water were used. Previously, 10 windows of DWS (10-DWS) were carried out with 10M/20–40M configurations in water to compute ΔG_{hyd} for 25 molecules.[13] In view of the success with SOS procedures,[44] it was decided to recalculate all 1.14*CM1A and OPLS-AA results with this sampling method to further assess precision. In this regard, it is informative to analyze results for closed FEP cycles; the sum of the individual free energy changes for each cycle should be zero. Ten cycles were considered consisting of three or four FEP calculations. As shown in Table 1, the mean signed errors for the cycles using SOS with both charge options are near zero; the mean unsigned errors (MUE) are less than 0.15 kcal/mol, and the maximum individual errors are no more than 0.24 kcal/mol. Absolute free energies of hydration for all 42 molecules were computed in a series of perturbations to methane. Methanol and acetone were used as intermediates for many sequences, and aromatic compounds were perturbed to benzene, which was converted to methane.

In addition to the relative free energy calculations, full molecular annihilations were performed for all molecules using the 1.14*CM1A charges. Annihilations were accomplished by first neutralizing the atomic charges and, in a separate calculation, diminishing the Lennard-Jones parameters to zero. For all annihilations, 21-SOS windows with 8M/8M configurations in the gas phase and 35M/70M configurations in water were used. As λ_i progressed from 0 → 1, each atom was transformed into a corresponding idealized sp² or sp³ dummy atom, and equilibrium bond lengths were perturbed to 0.30 Å. This "shrinks" the molecule over the course of the calculation. Unique to the annihilations, the SOS sampling led to end-point problems when sampling occurred at $\lambda_i = 1$, reminiscent of

Table 1. Free Energy Changes for Closed Thermodynamic Cycles

thermodynamic cycle	hysteresis (kcal/mol)	
	1.14*CM1A	OPLS-AA
CH$_3$OH → CH$_4$ → CH$_3$CH$_3$ → CH$_3$OH	−0.18 ± 0.06	−0.17 ± 0.07
CH$_3$OH → CH$_4$ → CH$_3$NH$_2$ → CH$_3$OH	0.24 ± 0.07	−0.07 ± 0.07
CH$_3$OH → CH$_4$ → CH$_3$CH$_3$ → CH$_3$SH → CH$_3$OH	−0.10 ± 0.06	−0.17 ± 0.06
CH$_3$CH$_3$ → CH$_3$OH → CH$_3$OCH$_3$ → CH$_3$CH$_2$CH$_3$ → CH$_3$CH$_3$	0.16 ± 0.09	0.05 ± 0.09
CH$_3$COCH$_3$ → CH$_3$COOH → CH$_3$CONH$_2$ → CH$_3$COCH$_3$	0.08 ± 0.09	−0.22 ± 0.08
CH$_3$CONH$_2$ → CH$_3$COOH → CH$_3$COOCH$_3$ → (Z)– CH$_3$CONHCH$_3$ → CH$_3$CONH$_2$	0.08 ± 0.14	0.08 ± 0.11
CH$_3$COCH$_3$ → CH$_3$CH$_2$CH$_3$ → CH$_3$CH=CH$_2$ → CH$_3$CHO → CH$_3$COCH$_3$	0.23 ± 0.10	0.05 ± 0.10
PhCl → PhH → PhF → PhCl	−0.11 ± 0.03	0.10 ± 0.04
PhOH → PhH → PhCH$_3$ → PhNH$_2$ → PhOH	0.06 ± 0.08	0.08 ± 0.08
PhSH → PhOH → PhOCH$_3$ → PhSCH$_3$ → PhSH	−0.16 ± 0.09	−0.12 ± 0.10
mean signed error	*0.03*	*−0.04*
mean unsigned error	*0.14*	*0.11*

earlier findings.[52] In this case, solute and solvent molecules can occupy the same space, which introduces infinite nonbonded energies into the energy difference ΔE in eq 1. To correct for this, SOS sampling was performed up to $\lambda_i = 0.99$, at which point, DWS sampling was used to obtain the free-energy result for $\lambda_i = 0.99 → 1$ by running from $\lambda_i = 0.995$ to 0.99 and 1.0. Computer time for a single FEP window averaged 30 min for the relative perturbations and 70 min for the annihilations in water on an Intel Core2 Quad 3.3 GHz processor.

Charge Perturbations. Charge perturbations were then executed to determine the ΔG_{hyd} results using the CM5 charges. 1.14*CM1A charges were first perturbed to unscaled CM5 charges (1.00*CM5) while keeping other force-field parameters constant. Charge perturbations converge rapidly and do not require as extensive sampling. Thus, 10 windows of double-wide sampling were used featuring 5 million configurations for equilibration and averaging for each window in the gas phase and 10M/15M configurations in water. The CM5 charges were subsequently scaled in separate FEP calculations from 1.00*q_{CM5} to 1.50*q_{CM5} in steps of 0.01, where q_{CM5} represents an atomic charge. Absolute free energies of hydration were computed for each scaling, and the optimal scaling factor was chosen to minimize the mean unsigned error for all 42 molecules.

■ RESULTS

Absolute free energies of hydration were computed using two CM*x* charge models and the OPLS-AA force field for the initial set of 42 molecules. The absolute free energy of hydration of methane was computed to be 2.51 ± 0.07 kcal/mol via annihilation with the 1.14*CM1A charges. This value is in good agreement with previous results of 2.3–2.6 kcal/mol for OPLS-based models[13,53,54] and experiment (2.01 kcal/mol).[11,50] All other absolute free energies of hydration were anchored by this value via relative FEP transformations. In summary, a mean unsigned error of 1.17 kcal/mol is obtained for the 42 absolute free energies of hydration with the 1.14*CM1A charges. An optimal scaling factor of 1.27 was determined for the CM5

Table 2. Comparison of Previously Published and Recomputed Free Energies of Hydration (kcal/mol) Using 1.14*CM1A Charges

molecule name	molecular formula	10-DWS[a]	21-SOS[b]	exptl.[c]
methanol	CH_3OH	-2.67 ± 0.15	-3.09 ± 0.08	-5.10
chloromethane	CH_3Cl	0.72 ± 0.11	0.87 ± 0.09	-0.55
methylamine	CH_3NH_2	-5.48 ± 0.14	-5.12 ± 0.09	-4.56
ethane	CH_3CH_3	2.55 ± 0.11	2.92 ± 0.09	1.83
methanethiol	CH_3SH	1.24 ± 0.11	0.56 ± 0.08	-1.24
dimethyl ether	CH_3OCH_3	-1.26 ± 0.15	-0.90 ± 0.10	-1.91
acetonitrile	CH_3CN	-3.21 ± 0.14	-2.59 ± 0.10	-3.89
dimethylamine	$(CH_3)_2NH$	-3.34 ± 0.17	-3.20 ± 0.10	-4.30
propane	$CH_3CH_2CH_3$	3.07 ± 0.12	3.30 ± 0.09	1.96
acetone	CH_3COCH_3	-2.96 ± 0.16	-1.91 ± 0.10	-3.81
acetic acid	CH_3COOH	-6.34 ± 0.17	-5.81 ± 0.11	-6.70
methyl acetate	CH_3COOCH_3	-4.35 ± 0.20	-3.27 ± 0.13	-3.32
acetamide	CH_3CONH_2	-10.24 ± 0.19	-10.14 ± 0.12	-9.71^d
(Z)-N-methylacetamide	$CH_3CONHCH_3$	-9.48 ± 0.23	-9.62 ± 0.14	-10.08^d
(E)-N-methylacetamide	$CH_3CONHCH_3$	-9.13 ± 0.21	-8.34 ± 0.14	-10.08^d
N,N-dimethylacetamide	$CH_3CON(CH_3)_2$	-9.30 ± 0.28	-8.18 ± 0.16	-8.55^d
nitroethane	$CH_3CH_2NO_2$	-6.37 ± 0.14	-6.35 ± 0.14	-3.71
benzene	C_6H_6	-1.41 ± 0.23	-1.14 ± 0.17	-0.86
chlorobenzene	C_6H_5Cl	-0.93 ± 0.24	-0.48 ± 0.17	-1.12
toluene	$C_6H_5CH_3$	-1.58 ± 0.24	-1.21 ± 0.17	-0.89
trifluorotoluene	$C_6H_5CF_3$	-0.51 ± 0.24	0.17 ± 0.18	-0.25
phenol	C_6H_5OH	-5.42 ± 0.24	-5.65 ± 0.17	-6.62
anisole	$C_6H_5OCH_3$	-2.70 ± 0.26	-2.81 ± 0.19	-2.46
aniline	$C_6H_5NH_2$	-7.57 ± 0.25	-7.30 ± 0.18	-5.49
pyridine	C_5H_5N	-3.15 ± 0.24	-3.08 ± 0.18	-4.70
mean unsigned error		*1.03*	*1.06*	

[a]From ref 13: 8M/8M gas-phase and 10M/20−40M aqueous-phase configurations. [b]This work: 5M/5M gas-phase and 15M/30M aqueous-phase configurations. [c]ref 50. [d]ref 51.

charges, which yields a MUE of 1.10 kcal/mol for ΔG_{hyd}. Table 2 compares the previous[13] and recomputed 1.14*CM1A results; Table 3 records the full results for the 1.14*CM1A and 1.27*CM5 free energies of hydration. Results for each method are discussed below. A more thorough analysis is made in the Discussion section and focuses on comparing the 1.14*CM1A and 1.27*CM5 results. In addition, results for 44 relative free-energy calculations, $\Delta\Delta G_{hyd}$, which yielded MUEs of 1.08 and 0.66 kcal/mol using the 1.14*CM1A and OPLS-AA charges, are listed in Table S2.

1.14*CM1A Free Energies of Hydration. In 2004, the optimal scaling factor of 1.14 for CM1A charges was determined from fitting to free energies of hydration for 25 molecules using series of FEP calculations with 10 windows of DWS.[13] The CM1A results for these molecules were recalculated in this work with 21 windows of SOS, and then the coverage was expanded to 42 molecules. Table 2 compares the 2004 results with the present ones. Overall, the DWS and SOS results are mostly the same within the statistical limits of about ±0.3 kcal/mol. Improvements of around 0.50 kcal/mol or greater, nearer experimental values, are noted for methanethiol and methyl acetate, while reductions in accuracy of 0.50 kcal/mol or greater are found for acetonitrile, acetone, and (E)-N-methylacetamide. The variations in the results are in-line with previous findings;[44] they are attributed to the differences in numbers of windows and run lengths, which favor the 21-SOS procedure. The net result is an insignificant change in the overall MUE from 1.03 to 1.06 kcal/mol. Some notable discrepancies with experimental values remain. The worst error, 2.64 kcal/mol, is for nitroethane. Additional problem cases are

methanethiol, aniline, and pyridine with errors around 2 kcal/mol.

Expanding the calculations from 25 to 42 molecules increases the MUE from 1.06 to 1.17 kcal/mol using the 1.14*CM1A charges (Table 3). The MUE was not helped by the addition of nitrobenzene and the methylanilines, which continue the pattern of being computed to be too well hydrated with an error now of nearly 4 kcal/mol for nitrobenzene. For the added sulfur-containing molecules, the ΔG_{hyd} for dimethyl sulfide is too positive by 2.79 kcal/mol; however, the errors for thiophenol and thioanisole are less than 0.6 kcal/mol. The errors for most of the other new additions are under 1.5 kcal/mol (propene, methyl formate, fluorobenzene, bromobenzene, benzonitrile, acetophenone, α-methylstyrene, and methyl benzoate). The error for acetaldehyde is 1.57 kcal/mol, and it is 2.77 kcal/mol for dimethyl sulfoxide. Though the error for acetamide is only 0.43 kcal/mol, the newly added benzamide is computed to be too well hydrated by 2.25 kcal/mol. The results are plotted in Figure 1; a linear fit gives an R^2 of 0.87 for the correlation of the computed and experimental data.

1.27*CM5 Free Energies of Hydration. The optimal scaling factor for the CM5 charges turned out to be 1.27 to minimize the MUE for the computed free energies of hydration. The 1.27*CM5 results in Table 3 lead to mean signed and unsigned errors of 0.49 and 1.10 kcal/mol, respectively. The MUE is a little lower than for the 1.14*CM1A results; however, the R^2 for the linear fit to the experimental data in Figure 1 improves to 0.95 from 0.87. The fit with the 1.27*CM5 results is visibly better with fewer strong outliers. It may be noted that the choice of the CM5 scaling

Table 3. Computed Absolute Free Energies of Hydration for 1.14*CM1A and 1.27*CM5 Charge Models (kcal/mol)[a]

molecule name	molecular formula	1.14* CM1A	1.27* CM5	OPLS_2005[b]	AM1-BCC/GAFF[c]	exptl.[d]
methanol	CH_3OH	−3.09	−4.31	−4.33	−3.48	−5.10
chloromethane	CH_3Cl	0.87	0.64	0.23	0.81	−0.55
methylamine	CH_3NH_2	−5.12	−3.39	−2.98	−3.44	−4.56
ethane	CH_3CH_3	2.92	2.88	2.30	2.58	1.83
methanethiol	CH_3SH	0.56	0.35	−0.51	−0.26	−1.24
dimethyl ether	CH_3OCH_3	−0.90	−2.43	−0.75	−0.85	−1.91
dimethyl sulfide	CH_3SCH_3	1.18	0.98	−0.62	0.26	−1.61
acetonitrile	CH_3CN	−2.59	−4.86	−3.70	−1.67	−3.89
dimethylamine	$(CH_3)_2NH$	−3.20	−2.33	−2.47	−3.11	−4.30
propane	$CH_3CH_2CH_3$	3.30	3.25	2.42	2.56	1.96
propene	CH_3CHCH_2	2.76	2.83	1.74	2.44	1.32
acetone	CH_3COCH_3	−1.91	−3.42	−2.64	−3.36	−3.81
acetaldehyde	CH_3CHO	−1.93	−3.12	−2.10	−3.39	−3.50
acetic acid	CH_3COOH	−5.81	−6.69	−5.44	−5.95	−6.70
methyl acetate	CH_3COOCH_3	−3.27	−3.38	−1.54	−3.73	−3.32
methyl formate	$HCOOCH_3$	−3.62	−3.30	−3.46	−3.17	−2.78
acetamide	CH_3CONH_2	−10.14	−12.06	−8.47	−8.62	−9.71[e]
(Z)-N-methylacetamide	$CH_3CONHCH_3$	−9.62	−11.18	−8.47	−8.39	−10.08[e]
(E)-N-methylacetamide	$CH_3CONHCH_3$	−8.34	−10.02			−10.08[e]
N,N-dimethylacetamide	$CH_3CON(CH_3)_2$	−8.18	−8.76	−7.31		−8.55[e]
nitroethane	$CH_3CH_2NO_2$	−6.35	−3.39	−2.75	−1.73	−3.71
dimethyl sulfoxide	CH_3SOCH_3	−12.88	−13.74		−8.32	−10.11
benzene	C_6H_6	−1.14	0.03	−0.06	−0.70	−0.86
fluorobenzene	C_6H_5F	0.01	0.62		−0.07	−0.80
chlorobenzene	C_6H_5Cl	−0.48	0.77	0.29	−0.60	−1.12
bromobenzene	C_6H_5Br	−1.27	−0.01	−0.34	−0.37	−1.46
benzonitrile	C_6H_5CN	−3.41	−5.24	−2.30	−2.74	−4.21
toluene	$C_6H_5CH_3$	−1.21	0.09	−0.74	−0.71	−0.89
trifluorotoluene	$C_6H_5CF_3$	0.17	1.35		−0.55	−0.25
phenol	C_6H_5OH	−5.65	−5.04	−4.64	−5.67	−6.62
anisole	$C_6H_5OCH_3$	−2.81	−2.05	−0.72	−2.30	−2.46
aniline	$C_6H_5NH_2$	−7.30	−5.94	−3.60	−5.92	−5.49
N-methylaniline	$C_6H_5NHCH_3$	−6.79	−4.32		−5.74	−4.69
N,N-dimethylaniline	$C_6H_5N(CH_3)_2$	−5.15	−2.95		−4.70	−3.45
nitrobenzene	$C_6H_5NO_2$	−8.11	−3.49	−2.61	−3.40	−4.12
acetophenone	$C_6H_5COCH_3$	−4.54	−4.25	−3.44	−5.07	−4.58
benzamide	$C_6H_5CONH_2$	−13.26	−12.84		−10.19	−11.01
α-methylstyrene	$C_6H_5CCH_2CH_3$	−0.98	0.58		−1.26	−1.24
methyl benzoate	$C_6H_5COOCH_3$	−4.31	−3.58	−2.44	−5.06	−3.93
thiophenol	C_6H_5SH	−2.32	−0.83	−0.75	−1.43	−2.55
thioanisole	$C_6H_5SCH_3$	−2.16	−0.71	−1.22	−1.21	−2.73
pyridine	C_5H_5N	−3.08	−3.56	−2.68	−3.45	−4.70
mean unsigned error		*1.17*	*1.10*	*1.21*	*0.95*	

[a]Results obtained from a series of FEP/SOS calculations; computed uncertainties are less than 0.25 kcal/mol. [b]ref 31b. [c]ref 26b. [d]ref 50. [e]ref 51.

factor benefited from use of the results for the full set of 42 molecules, while the CM1A scaling factor was derived from the results for the original 25 molecules. For the 25 molecules, the MUE is 0.96 kcal/mol using the 1.27*CM5 charges, which is again a little lower than the MUE of 1.06 kcal/mol with the 1.14*CM1A charges. In comparing several electrostatic-potential and AM1-based charge models, Mobley and co-workers suggested that a root-mean-square error of *ca.* 1.0 kcal/mol may be the current limit for fixed charge models in predicting free energies of hydration.[26] The present results support this notion. It is apparent that some functionality is not well represented by the single charged site per atom model. As discussed more below, the principal problems with the 1.27*CM5 charges occur for sulfur and halogen containing molecules.

The CM*x* results are also compared with published data from molecular dynamics calculations using AM1-BCC/GAFF charges in TIP3P water and OPLS_2005 charges in SPC water in Table 3.[26b,31b] Although results from other methods exist for comparison,[26−34] the chosen ones are viewed as relatively accurate.[26,27,31,34] The mean unsigned errors for the CM*x* models fall between those for the OPLS_2005 and AM1-BCC/GAFF methods. Best fit line statistics are reported in Table S3. Both CM*x* methods have signed errors less positive than the OPLS_2005 and GAFF results. Furthermore, the R^2 values for the OPLS_2005 and GAFF results are 0.96 and 0.92 versus the 0.87 and 0.95 for 1.14*CM1A and 1.27*CM5. The results can be easily biased by selective addition and removal of compounds, e.g., removal of nitro compounds for 1.14*CM1A and DMSO for 1.27*CM5. Completion of the missing
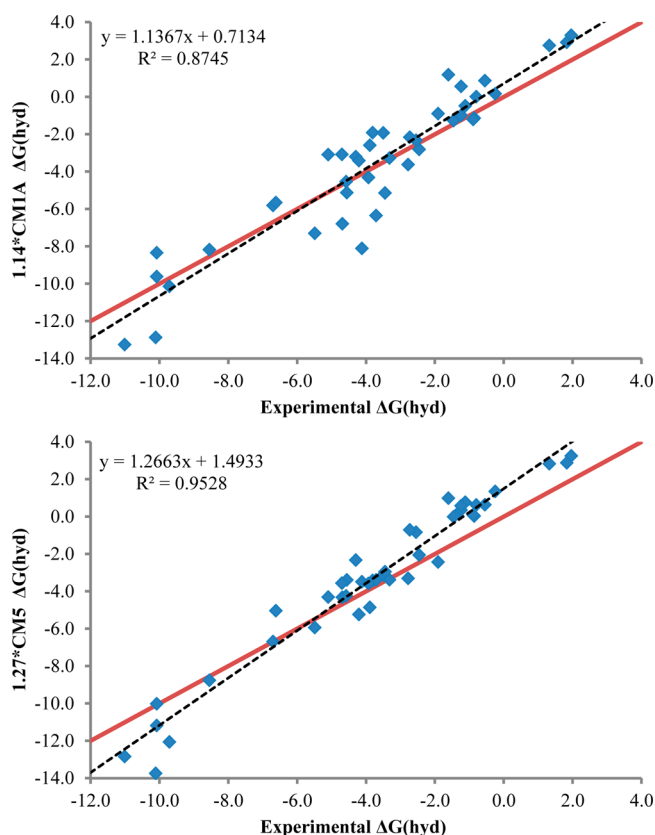
Figure 1. Correlation between experimental and computed free energies of hydration (kcal/mol): 1.14*CM1A results (top), 1.27*CM5 results (bottom). The solid line shows the ideal $y = x$ line; the dashed line represents the best fit line for the computed data.

Table 4. Calculated and Experimental Dipole Moments (D)

molecule	1.14*CM1A	1.27*CM5	1.00*CM5	exptl.[a]
methanol	2.03	2.20	1.67	1.70
chloromethane	2.14	2.27	1.79	1.89
methylamine	1.95	1.74	1.29	1.31
methanethiol	1.74	1.86	1.44	1.52
dimethyl ether	1.88	2.11	1.54	1.30
dimethyl sulfide	1.91	2.01	1.49	1.55
acetonitrile	4.30	4.92	3.88	3.92
dimethylamine	1.81	1.66	1.13	1.01
propene	0.28	0.49	0.38	0.37
acetone	3.29	3.74	2.93	2.88
acetaldehyde	3.01	3.45	2.71	2.75
acetic acid	2.04	2.11	1.67	1.70
methyl acetate	2.23	2.40	1.90	1.72
methyl formate	1.97	2.29	1.83	1.77
acetamide	3.73	4.81	3.71	3.76
(Z)-N-methylacetamide	3.73	4.85	3.77	3.80[b]
(E)-N-methylacetamide	4.30	5.18	4.04	4.17[e]
N,N-dimethylacetamide	4.27	5.22	3.99	3.82
nitroethane	5.25	4.80	3.80	3.23
dimethyl sulfoxide	5.05	5.76	4.46	3.96
fluorobenzene	1.76	2.09	1.65	1.60
chlorobenzene	1.95	2.02	1.59	1.69
bromobenzene	1.82	2.41	1.90	1.70
benzonitrile	4.89	5.84	4.60	4.18
toluene	0.35	0.58	0.46	0.38
trifluorotoluene	2.97	3.32	2.61	2.86
phenol	1.86	1.87	1.28	1.22
anisole	1.78	2.00	1.39	1.38
aniline	1.60	2.36	1.86	1.13
N-methylaniline	1.75	2.60	2.06	
N,N-dimethylaniline	1.89	2.72	2.17	1.68
nitrobenzene	6.57	5.76	4.53	4.22
acetophenone	3.60	4.01	2.99	3.02
benzamide	4.02	4.90	3.53	3.55[c]
α-methylstyrene	0.37	0.56	0.42	
methyl benzoate	2.49	2.44	1.82	[1.94]
thiophenol	1.48	1.51	1.13	1.13[c]
thioanisole	1.92	2.05	1.56	[1.31][d]
pyridine	1.75	2.55	2.01	2.22
mean unsigned error			0.16	

[a]Gas-phase values from ref 55. Brackets indicate dipole moments estimated in solution. [b]Ref 56. [c]Ref 57. [d]Ref58. [e]Ref 24, theoretically estimated.

numbers in Table 3, e.g., for DMSO, the methyanilines, and thiophenol with OPLS_2005, can also be expected to affect the results, adversely in this case. Overall, the results in Table 3 for the different methods can be viewed as similar and reflecting the 1 kcal/mol limit for general charge methods.[26]

CMx Dipole Moments. Dipole moments were calculated for all molecules with the 1.14*CM1A, 1.27*CM5, and 1.00*CM5 charges (Table 4). The charge scaling for neutral molecules is needed to incorporate solute polarization in an average sense for condensed phase modeling with fixed-charge force fields.[7,11] As shown in Table 4, the unscaled CM5 charges yield excellent values for gas-phase dipole moments with a MUE of only 0.16 D. It is noted that some differences occur between 1.00*CM5 dipole moments computed here and those in the original CM5 report.[24] The differences of 0.30 D or more for DMSO, nitroethane, nitrobenzene, and N,N-dimethylaceta-mide can be attributed to the variations in optimized geometries and electronic structure methods used for the CM5 charge computations (see Tables S4 and S5). It is also apparent that in comparing the 1.14*CM1A and 1.27*CM5 dipole moments, the larger scaling factor for the CM5 charges does not always lead to higher dipole moments. The most striking exceptions are for amines and nitro compounds. The lower dipole moments for methylamine and dimethylamine with the 1.27*CM5 charges do correlate with the overly positive computed free energies of hydration (Table 3), while the dipole moments for the nitro compounds are clearly too large with the CM1A charges and contribute to the far too favorable computed ΔG_{hyd} values. The problem with nitro

compounds was previously discussed.[13] For nitroethane, the 1.14*CM1A charges for nitrogen and oxygen are 0.68 and −0.46, while they are 0.54 and −0.37 with OPLS-AA and 0.11 and −0.23 with 1.27*CM5. However, both the CM1 and CM5 approaches lead to overpolarization of the S−O bonds in dimethyl sulfoxide, too large dipole moments, and too negative free energies of hydration.

■ DISCUSSION

CMx Functional Group Trends. Mean unsigned errors in ΔG_{hyd} for compound classes are listed in Table 5 to help identify systematic strengths and weaknesses for the charge models.

The 1.27*CM5 model yields lower average errors than the 1.14*CM1A alternative for most classes except hydrocarbons and halogen and sulfur containing molecules. ΔG_{hyd} for less

2807

dx.doi.org/10.1021/ct500016d | J. Chem. Theory Comput. 2014, 10, 2802−2812

Table 5. Mean Unsigned Errors for Compound Classes (kcal/mol)

	1.14*CM1A	1.27*CM5
all molecules (42)	1.17	1.10
non-aromatic (22)	1.34	1.08
aromatic (20)	0.99	1.12
C, H containing (6)	0.79	1.26
halogen containing (5)	0.70	1.51
C, H, O containing (11)	0.91	0.49
alcohols/ethers (4)	1.09	0.83
carbonyls (7)	0.81	0.29
C, H, N containing (8)	1.37	0.95
amines (6)	1.48	0.93
nitriles (2)	1.05	1.00
C, H, N, O containing (7)	1.70	0.93
amides (5)	1.05	1.11
nitros (2)	3.31	0.48
sulfur containing (5)	1.63	2.31
thiols/sulfides (4)	1.35	1.98
sulfoxide (1)	2.77	3.63

polar benzene derivatives is uniformly too positive by 1−2 kcal/mol with 1.27*CM5 (Table 3), which contributes significantly to the errors for hydrocarbons and halides. The clear failing of the 1.14*CM1A charges is nitro compounds, while both CM*x* approaches show significant deviations for sulfur-containing molecules. Dimethyl sulfoxide (DMSO) is particularly problematic with errors of 2.8 and 3.6 kcal/mol. Unfortunately, additional sulfoxides and sulfones could not be tested due to a lack of experimental data. The 1.27*CM5 method shows higher unsigned errors for the sulfur-containing molecules due to its larger deviations for thiophenol and thioanisole. There is no problem with the computed dipole moments (Table 4), so it is in the details of the charge distributions. For thiophenol, the thiol sulfur and hydrogen have charges of −0.14 and 0.16 with 1.14*CM1A and −0.21 and 0.17 with 1.27*CM5. These differences are not striking, though the larger charge on sulfur would diminish the hydrogen-bond donating ability of the thiol group. However, there is a general pattern that the charges on the carbons and hydrogens in phenyl rings are smaller in magnitude with 1.27*CM5 than with 1.14*CM1A by ca. 0.02 e (Table S1). Remarkably, this translates to less favorable free energies of hydration for benzene, toluene, and α-methylstyrene by 1.2−1.5 kcal/mol with 1.27*CM5 (Table 3) and presumably a similar offset for the other substituted benzenes.

Additional insights can be obtained by considering computed interaction energies for some hydrogen-bonded complexes in the gas phase. Results are reported in Table 6 for nine

Table 6. Computed Optimal Gas-Phase Interaction Energies (kcal/mol)

donor−acceptor	OPLS-AA	1.14*CM1A	1.27*CM5
water−CH₃OH	−6.94	−6.11	−6.80
CH₃OH−water	−5.57	−5.54	−5.72
water−CH₃SH	−3.24	−3.01	−3.38
CH₃SH−water	−2.58	−2.26	−2.65
water−CH₃OCH₃	−5.57	−6.18	−7.09
water−CH₃SCH₃	−4.84	−3.59	−3.82
water−acetamide (1)	−6.92	−6.49	−8.02
water−acetamide (2)	−9.31	−9.25	−10.61
acetamide−water	−6.75	−7.18	−7.41

representative complexes using OPLS-AA and the OPLS/CM*x* models. Images of each complex are provided in Figure S1. OPLS-AA interaction energies correlate well with high-end ab initio results[59] and are given for comparison. The variation of the results for each of the nine complexes is notably small, within 1 kcal/mol for most cases. For both methanol and methanethiol, water is preferred as the hydrogen-bond donor by about 1 kcal/mol, and the interactions are more favorable with the 1.27*CM5 charges than 1.14*CM1A by a few tenths of a kilocalorie per mole. However, in both cases, the free energies of hydration are too positive, though the results are improved with 1.27*CM5 (Table 3). CBS-Q results for the four complexes of methanol and methanethiol in order are −5.69, −5.27, −3.90, and −2.50 kcal/mol, and MP2/6-31++G(2d(X+),p) gives very similar values.[59] Thus, the raw interaction energies are not overly weak, and the conclusion is that the effects of polarization in water are not sufficiently compensated for by the charge scaling in these cases. For dimethyl ether and dimethyl sulfide, the best ab initio and DFT results indicate interaction energies with water of about −5 and −4 kcal/mol,[59] while the scaled CM*x* results in Table 6 are stronger and weaker, respectively. The consequences are reasonably computed ΔG_{hyd} values for dimethyl ether, but overly positive results for dimethyl sulfide by about 2.7 kcal/mol. For dimethyl sulfide, the sulfur atom is not a sufficiently good hydrogen-bond acceptor with the CM*x* charge models. The larger size and increased polarizability of sulfur relative to oxygen likely make its condensed phase electrostatics more difficult to model with fixed charge approximations. Similar difficulties for sulfur and phosphorus have been observed by others.[26]

The free energies of hydration of aliphatic amides, a functional group particularly important for biologically related studies, are well reproduced by the 1.14*CM1A model (Table 3). With the 1.27*CM5 charges, the ΔG_{hyd} values for acetamide and *Z*-NMA are too favorable by 2.4 and 1.1 kcal/mol. For the water−acetamide complexes in Table 6, the hydrogen bonds to the carbonyl oxygen are significantly stronger with 1.27*CM5 than with OPLS-AA or 1.14*CM1A. If the scaling factor for the CM5 charges were reduced to 1.20, the ΔG_{hyd} errors would decrease to 0.58 kcal/mol for both acetamide and *Z*-NMA. Benzamide turns out to be too well hydrated for both charge models with errors of 2.3 and 1.8 kcal/mol (Table 3). In general, the 1.14*CM1A charges for amide nitrogens seem large in magnitude at ca. −1.2 e, whereas the values are ca. −0.7 e with 1.27*CM5 and −0.76 and −0.50 e for primary and secondary amides with OPLS-AA. It is also noted that both charge models predict poorer hydration of the *E* conformer of NMA by about 1.2 kcal/mol, while the available experimental data indicate there is no *E/Z* differential.[51] This discrepancy is a recurrent one going back to the earliest FEP calculations for amide *E/Z* interconversions.[60] The experimental measurements are challenging owing to the low population of the *E* conformer.

Molecular Annihilations. As an additional check of the reported free energies of hydration in Table 3, full molecular annihilations were performed for all molecules. For molecules with up to 10 heavy atoms, sufficient convergence required sampling roughly three-times the number of configurations for a typical single heavy atom perturbation. The free energy results from the sequential and annihilation procedures with the 1.14*CM1A charges are compared in Figure 2; a similar plot for the 1.27*CM5 charges is shown in Figure S2. Excellent
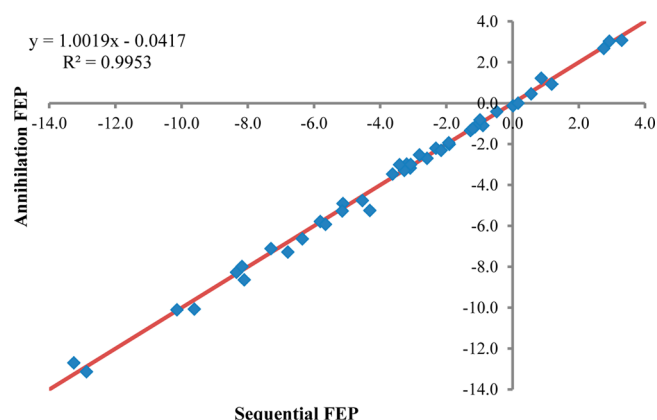
Figure 2. Comparison of computed free energy of hydration (kcal/mol) using a series of FEP calculations or annihilation with 1.14*CM1A charges. The solid line is the ideal $y = x$ line.

correlations are observed for both CM*x* methods; the linear fits have slopes of 1.0, intercepts near 0.0, and R^2 values of 0.995. The computed absolute free energies of hydration are listed in Table S6. Average deviations between the sequential and annihilation techniques are *ca.* 0.06 kcal/mol. The consistency confirms the precision of the results in Table 3, and it also provides validation of the annihilation methodology itself. This permits more automated calculations of free energies of hydration without the need to devise a mutation sequence.

Further Testing. An additional 20 neutral, organic molecules were studied by molecular annihilation to ascertain the validity and transferability of the 1.27 scale factor. Molecules were chosen to exhibit molecular diversity and for which independent free energies of hydration have been previously computed. Five new heterocycles were also included. The OPLS_2005 and AM1-BCC/GAFF methods were again chosen for comparison. Both E and Z isomers have been

included for *N*-methylformamide, although no AM1-BCC/GAFF data are available for this molecule. The free energies of hydration are reported in Table 7; scaled CM*x* charges are reported in Table S7. The MUE with the 1.27*CM5 charges is 1.31 kcal/mol, while it is again higher, 1.47 kcal/mol, for the 1.14*CM1A charges. AM1-BCC/GAFF has the lowest MUE for this set, though based on the error of 1.7 kcal/mol for *N*-methylacetamide in Table 3, inclusion of results for NMF could increase the MUE by ca. 0.2 kcal/mol. The results indicate that average errors of 1.0−1.5 kcal/mol can be expected for application of these charge models to other small molecules. For the 1.27*CM5 charges, larger errors again occur for hydrocarbons and sulfur and halogen containing molecules. Aliphatic amines are under-hydrated with all methods, while the performance for amides in Tables 3 and 7 is notably improved with the CM*x* charges. The cumulative errors for polypeptides can be expected to be large with many charge models, which has implications for modeling protein stability and folding.

Transferability of the CM5 Scaling Factor. A promising attribute of Charge Model 5 is the ability to provide atomic charges for any electronic structure method in conjunction with a Hirshfeld population analysis.[24] For the present context, tests were performed to analyze the transferability of the 1.27 scale factor for alternative density functional methods. Specifically, M06,[37] M06-L,[37] B3LYP,[61] and mPW1PW91[62] were utilized to derive CM5 charges for a subset of 10 molecules. The same procedure for obtaining the CM5 charges was followed as before, only the M06-2X method was replaced by the alternatives. Absolute free energies of hydration were determined by charge perturbations from 1.14*CM1A to the new 1.27*CM5 charges for each molecule in the gas phase and in water. The computed $\Delta\Delta G_{hyd}$ was added to the ΔG_{hyd}(1.14*CM1A) for each molecule to determine the new ΔG_{hyd}(1.27*CM5). The final results are summarized in Table 8 including the standard deviation for the five computed values

Table 7. Computed Absolute Free Energies of Hydration for 1.14*CM1A and 1.27*CM5 Charge Models (kcal/mol) for an Additional 20 Molecules[a]

molecule name	1.14*CM1A	1.27*CM5	OPLS_2005[b]	AM1-BCC/GAFF[c]	exptl.[d]
2,2-dimethylbutane	4.00	3.87	2.51	2.53	2.51
Z-2-pentene	2.59	2.72	2.31	2.55	1.31
cyclohexane	2.54	2.45	1.64	1.67	1.23
1-chloropropane	1.53	1.40	0.96	0.92	−0.33
hydrogen sulfide	1.43	1.23	1.01	−1.17	−0.70[b]
o-xylene	−1.09	0.42	−1.54	−0.52	−0.90
triethylamine	−1.40	−0.78	−1.07	−1.83	−3.22
propionitrile	−2.21	−4.56	−3.39	−1.27	−3.85
benzaldehyde	−4.46	−3.87	−2.97	−4.99	−4.02
2,2,2-trifluoroethanol	−1.27	−1.57	−3.81	−3.95	−4.31
2-propen-1-ol	−2.61	−3.00	−4.55	−3.23	−5.03
1,4-dioxane	−4.43	−5.23	−2.60	−4.35	−5.06
2-methylpyrazine	−3.07	−3.83	−3.33	−6.10	−5.51
azetidine	−4.62	−2.64	−3.45	−3.41	−5.56[b]
3-methylindole	−9.04	−5.78	−4.73	−6.55	−5.91
butanoic acid	−5.68	−6.08	−5.32	−5.46	−6.36
4-bromophenol	−5.74	−5.49	−5.56	−5.47	−7.12
morpholine	−5.86	−6.42	−5.54	−6.28	−7.17
Z-N-methylformamide	−9.67	−11.19	−6.63		−10.00[b]
E-N-methylformamide	−9.02	−10.43			−10.00[e]
mean unsigned error	***1.47***	***1.31***	***1.33***	***1.02***	

[a]Results obtained from a FEP annihilations; computed uncertainties are less than 0.25 kcal/mol. [b]Ref 31b. [c]Ref 26b. [d]Ref 50. [e]Ref 51.

Table 8. Computed Free Energies of Hydration using 1.27*CM5 Charges from Different Density Functional Methods (kcal/mol)[a]

	M06-2X	M06	M06-L	mPW1PW91	B3LYP	exptl.[b]	std. dev.
methanol	−4.31 ± 0.08	−4.36 ± 0.08	−4.08 ± 0.08	−4.19 ± 0.08	−4.01 ± 0.08	−5.10	0.15
ethane	2.88 ± 0.09	2.90 ± 0.09	2.90 ± 0.09	2.88 ± 0.09	2.90 ± 0.09	1.83	0.01
dimethyl sulfide	0.98 ± 0.10	1.15 ± 0.10	1.12 ± 0.10	0.91 ± 0.10	1.07 ± 0.10	−1.61	0.10
acetone	−3.42 ± 0.10	−3.62 ± 0.10	−3.09 ± 0.10	−3.37 ± 0.10	−3.21 ± 0.10	−3.81	0.20
acetamide	−12.06 ± 0.13	−12.03 ± 0.13	−11.65 ± 0.13	−12.44 ± 0.13	−11.92 ± 0.13	−9.71[c]	0.28
benzene	0.03 ± 0.17	0.31 ± 0.17	0.42 ± 0.17	0.00 ± 0.17	0.25 ± 0.17	−0.86	0.18
toluene	0.09 ± 0.18	0.40 ± 0.17	0.57 ± 0.17	0.10 ± 0.17	0.38 ± 0.17	−0.89	0.21
phenol	−5.04 ± 0.18	−4.94 ± 0.18	−4.70 ± 0.18	−5.18 ± 0.18	−4.80 ± 0.18	−6.62	0.19
aniline	−5.94 ± 0.19	−6.00 ± 0.18	−5.91 ± 0.18	−6.36 ± 0.18	−5.86 ± 0.18	−5.49	0.20
acetophenone	−4.25 ± 0.19	−4.47 ± 0.19	−3.81 ± 0.19	−4.44 ± 0.19	−3.96 ± 0.19	−4.58	0.29
average							*0.18*

[a]All methods used the 6-311+G(2df,2p) basis set. [b]Ref 50. [c]Ref 51.

Table 9. Computed Free Energies of Hydration using 1.27*CM5 Charges with Different Basis Sets (kcal/mol)[a]

	6-311+G(2df,2p)	6-311G(d,p)	6-31+G(d,p)	6-31G(d)	exptl.[b]	std. dev.
methanol	−4.31 ± 0.08	−4.68 ± 0.08	−4.65 ± 0.08	−4.99 ± 0.08	−5.10	0.28
ethane	2.88 ± 0.09	2.89 ± 0.09	2.89 ± 0.09	2.88 ± 0.09	1.83	0.00
dimethyl sulfide	0.98 ± 0.10	0.94 ± 0.10	0.87 ± 0.10	0.85 ± 0.10	−1.61	0.06
acetone	−3.42 ± 0.10	−3.12 ± 0.10	−3.40 ± 0.10	−3.16 ± 0.10	−3.81	0.16
acetamide	−12.06 ± 0.13	−12.46 ± 0.13	−12.49 ± 0.13	−12.65 ± 0.13	−9.71[c]	0.25
benzene	0.03 ± 0.17	−0.01 ± 0.17	−0.01 ± 0.17	−0.17 ± 0.17	−0.86	0.09
toluene	0.09 ± 0.18	0.01 ± 0.17	0.08 ± 0.17	−0.07 ± 0.17	−0.89	0.07
phenol	−5.04 ± 0.18	−5.49 ± 0.18	−5.32 ± 0.18	−5.83 ± 0.18	−6.62	0.33
aniline	−5.94 ± 0.19	−6.63 ± 0.18	−6.35 ± 0.18	−6.83 ± 0.18	−5.49	0.39
acetophenone	−4.25 ± 0.19	−4.22 ± 0.19	−4.24 ± 0.19	−4.26 ± 0.19	−4.58	0.02
average						*0.17*

[a]All calculations used the M06-2X method. [b]Ref 50. [c]Ref 51.

for each molecule. The effects of changing the basis set on the ΔG_{hyd} results were also examined with the M06-2X functional. Three additional basis sets, 6-311G(d,p), 6-31+G(d,p), and 6-31G(d),[38,63] were used to compute new 1.27*CM5 charges. The absolute free energies of hydration for the 10 molecules were again computed by charge perturbations. The results are listed in Table 9.

Overall, little effect is found on the computed free energies of hydration from changes in the density functional method or basis set. The standard deviations average below 0.20 kcal/mol, which is near the level of the statistical uncertainties in the results. The conclusion is that the scale factor of 1.27 for CM5 charges is valid for a wide range of DFT methods.

Polarization of CM5 Charges by Implicit Solvation. The ease of access to CM5 charges from the Gaussian 09 program coupled with its implementation of continuum solvation models raised the possibility of combining these methods to compute solvent-polarized charges. Such charges might be appropriate for direct use in molecular mechanics based fluid simulations. Single-point calculations were carried out for the 42 molecules using M06-2X/6-311+G(2df,2p) and the conductor-like polarizable continuum model (CPCM)[64] with the dielectric value for water, 78.3. The Hirshfeld charges were computed and processed with CM5PAC to derive new CM5 charges (CM5/CPCM). Perturbations from the 1.14*CM1A to the CM5/CPCM charges were then performed to obtain the ΔG_{hyd} values for the CM5/CPCM model. The resultant ΔG_{hyd} MUE for the CM5/CPCM approach is 2.11 kcal/mol. This represents improvement from the MUE of 3.08 kcal/mol with unscaled CM5 charges; however, it is not

competitive with the MUE of 1.10 kcal/mol using the 1.27*CM5 charges (Table 3). Table S8 reports the CM5/CPCM enhanced dipole moments and resultant free energies of hydration. CM5/CPCM dipole moments are significantly different from the Hirshfeld/CPCM dipole moments due to empirical adjustments by CM5. Although the CM5/CPCM dipole moments are enhanced by about 20% over CM5 and gas-phase experimental values, this is insufficient for optimal modeling of chemistry in aqueous solution with fixed-charge force fields.

CM1A or CM5. Since the average errors for free energies of hydration with the 1.14*CM1A and 1.27*CM5 methods are similar and near the 1 kcal/mol level, both approaches are reasonable for use in simulations of aqueous systems. A simple Welch Two-Sample t test comparing CM1A and CM5 signed errors of the training set produces a p-statistic of 0.328 and a t-statistic of 0.985. While no statistical significance can be made for the training set based on this t test, correlation with experimental data and precision are improved with the 1.27*CM5 charges (Figure 1, Table S3, Figure S3). A smaller standard deviation of unsigned errors for 1.27*CM5 versus 1.14*CM1A, 0.78 vs 0.89 kcal/mol, respectively, and a repeated reduced mean unsigned error in the validation set, 1.31 vs 1.47 kcal/mol, respectively, support this conclusion and highlight CM5 improvements as important.

There are several additional advantages to the 1.27*CM5 approach. First, the CM1A method is wed to semiempirical AM1 wave functions, while the CM5 charges can be derived from Hirshfeld charges obtained from a wide variety of DFT methods and basis sets.[24] The Hirshfeld charges, CM5 charges,

and, as shown here, the resultant ΔG_{hyd} values are insensitive to reasonable choices for the DFT method and basis set. Second, the variety of elements is limited with AM1, while the CM5 model has been parametrized for all atoms of the periodic table.[24] Third, the CM5 method yields more reasonable charges for buried atoms, e.g., amide nitrogen atoms as noted above, and less dependence on conformation than alternative charge models.[24] Finally, incorporation of CM5 charges into molecular mechanics software is straightforward given access to Gaussian 09. For BOSS, the subroutine that requests quantum mechanics calculations was modified to invoke an external call to a linker script, which interfaces with Gaussian. The script (i) creates a Gaussian input file using solute coordinates and variables specifying the calculation type from BOSS, (ii) initializes and executes a single-point calculation in Gaussian, and (iii) reads the output and loads the total energy and CM5 charges back into BOSS. Though the present testing has only considered a few elements, optimism can be expressed for broader utility given the excellent performance of CM5 for dipole moments for a wide range of molecules.[24]

CONCLUSION

The recently reported CM5 charge model was tested for performance in computing free energies of hydration in conjunction with the OPLS-AA force field for the remaining force-field parameters. An optimal scaling factor of 1.27 for the CM5 charges was determined to minimize the error in the computed free energies of hydration for 42 neutral organic molecules. The mean unsigned error of 1.10 kcal/mol from the 1.27*CM5 calculations compares favorably with the error of 1.17 kcal/mol from the well-established 1.14*CM1A alternative. Both approaches give larger errors for sulfur-containing molecules, and the 1.14*CM1A method makes nitro groups much too polar. Additional attractive features of the CM5 approach include the transferability of the 1.27 scale factor to a wide range of density functional methods and basis sets, coverage of all elements of the periodic table, and the availability of CM5 charges from the Gaussian program. Use of the scaled CM5 charges in conjunction with fixed-charge force fields should provide a viable approach to modeling an enormous range of molecular systems in solution.

ASSOCIATED CONTENT

Ⓢ Supporting Information

Scaled 1.14*CM1A and 1.27*CM5 charges are provided for all molecules along with results from free energy difference calculations ($\Delta\Delta G_{hyd}$). Best-fit line statistics for CM1A, CM5, OPLS_2005, and AM1-BCC/GAFF methods are provided for the 42 molecule training set. Dipole and geometry differences between OPLS/CM5 and M06-L/CM5 techniques are tabulated. Structures of the complexes from Table 6 are shown. ΔG_{hyd} results from the molecular annihilations and a plot of sequential FEP vs annihilation results with the 1.27*CM5 charges are provided. Atomic charges for the validation set and CM5/CPCM dipole moments and free energies of hydration are reported. Finally, a box plot of signed errors for 1.14*CM1A and 1.27*CM5 is drawn. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: william.jorgensen@yale.edu.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Leach, A. R. *Molecular Modelling: Principles and Applications*, 2nd ed.; Pearson Education Limited: Essex, England, 2001; pp 181−203.
(2) Some review articles: (a) Marshall, G. R. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 107−114. (b) Cieplak, P.; Dupradeau, F.-Y.; Duan, Y.; Wang, J. *J. Phys.: Condens. Matter* **2009**, *21*, 333102. (c) Rick, S. W.; Stuart, S. J. *Rev. Comput. Chem.* **2002**, *18*, 89−146.
(3) (a) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. B* **1999**, *103*, 4730−4737. (b) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. A* **2004**, *108*, 621−627. (c) Ren, P.; Ponder, J. W. *J. Comput. Chem.* **2002**, *23*, 1497−1506. (d) Ponder, J. W.; Case, D. A. *Adv. Protein Chem.* **2003**, *66*, 27−85. (e) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933−5947. (f) 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.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549−2564. (g) Patel, S.; Brooks, C. L., III. *J. Comput. Chem.* **2004**, *25*, 1−16.
(4) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(5) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(6) (a) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr. *J. Comput. Chem.* **2010**, *31*, 671−690. (b) Momany, F. A.; Rone, R. *J. Comput. Chem.* **1992**, *13*, 888−900.
(7) Jorgensen, W. L.; Tirado-Rives, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6665−6670.
(8) Francl, M. M.; Chirlian, L. E. *Rev. Comput. Chem.* **2000**, *14*, 1−31.
(9) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361−373.
(10) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(11) Carlson, H. A.; Nguyen, T. B.; Orozco, M.; Jorgensen, W. L. *J. Comput. Chem.* **1993**, *14*, 1240−1249.
(12) Kaminski, G.; Jorgensen, W. L. *J. Phys. Chem. B* **1998**, *102*, 1787−1796.
(13) Udier-Blagovic, M.; Morales De Tirado, P.; Pearlman, S. A.; Jorgensen, W. L. *J. Comput. Chem.* **2004**, *108*, 16264−16270.
(14) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput-Aided Mol. Des.* **1995**, *9*, 87−110.
(15) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **2003**, *24*, 1291−1304.
(16) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(17) (a) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220. (b) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221−264.
(18) Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 1820−1831.
(19) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133−1152.
(20) Olson, R. M.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 2046−2054.
(21) (a) Mulliken, R. S. *J. Chem. Phys.* **1935**, *3*, 564−572. (b) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833−1840. (c) Mulliken, R. S. *J. Chem. Phys.* **1962**, *36*, 3428−3439.

(22) (a) Löwdin, P. O. *J. Chem. Phys.* **1950**, *18*, 365−375. (b) Baker, J. *Theor. Chim. Acta* **1985**, *68*, 221−229.

(23) Thompson, J. D.; Xidos, J. D.; Sonbuchner, T. M.; Cramer, C. J.; Truhlar, D. G. *Phys. Chem. Comm.* **2002**, *5*, 117−134.

(24) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2012**, *8*, 527−541.

(25) (a) Hirshfeld, F. L. *Theor. Chim. Acta* **1977**, *44*, 129−138. (b) Ritchie, J. P. *J. Am. Chem. Soc.* **1985**, *107*, 1829−1837. (c) Ritchie, J. P.; Bachrach, S. M. *J. Comput. Chem.* **1987**, *8*, 499−509.

(26) (a) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Dill, K. A. *J. Phys. Chem. B* **2009**, *113*, 4533−4537. (b) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. *J. Chem. Theory Comput.* **2009**, *5*, 350−358. (c) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. *J. Phys. Chem. B* **2007**, *111*, 2242−2254. (d) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. *J. Phys. Chem. B* **2011**, *115*, 1329−1332.

(27) (a) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119*, 5740−5761. (b) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.

(28) Jorgensen, W. L.; Ulmschneider, J. P.; Tirado-Rives, J. *J. Phys. Chem. B* **2004**, *108*, 16264−16270.

(29) Vilseck, J. Z.; Sambasivarao, S. V.; Acevedo, O. *J. Comput. Chem.* **2011**, *32*, 2836−2842.

(30) Xu, Z.; Luo, H. H.; Tieleman, D. P. *J. Comput. Chem.* **2007**, *28*, 689−697.

(31) (a) Shivakumar, D.; Deng, Y.; Roux, B. *J. Chem. Theory Comput.* **2009**, *5*, 919−930. (b) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. *J. Chem. Theory Comput.* **2010**, *6*, 1509−1519.

(32) Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. *J. Med. Chem.* **2008**, *51*, 769−779.

(33) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.

(34) Shivakumar, D.; Harder, E.; Damm, W.; Friesner, R. A.; Sherman, W. *J. Chem. Theory Comput.* **2012**, *8*, 2553−2558.

(35) (a) Jorgensen, W. L.; Tirado-Rives, J. *J. Comput. Chem.* **2005**, *26*, 1689−1700. (b) Jorgensen, W. L. *BOSS*, version 4.8; Yale University: New Haven, CT, 2011.

(36) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.

(37) (a) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101. (b) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157−167. (c) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215−241.

(38) (a) Raghavachari, K.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650−54. (b) McLean, D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639−5648. (c) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294−301. (d) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265−3269.

(39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; 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.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.

(40) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *CM5PAC*, version 2011; University of Minnesota: Minneapolis, MN, 2011.

(41) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; 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.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2013.

(42) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087−1092.

(43) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.

(44) Jorgensen, W. L.; Thomas, L. L. *J. Chem. Theory Comput.* **2008**, *4*, 869−876.

(45) Acevedo, O.; Jorgensen, W. L. *Acc. Chem. Res.* **2010**, *43*, 142−151.

(46) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050−3056.

(47) Lu, N.; Kofke, D. A.; Woolf, T. B. *J. Comput. Chem.* **2004**, *25*, 28−39.

(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(49) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Clarendon: Oxford, 1987; pp 191−198.

(50) Abraham, M. H.; Andonian-Haftvan, J.; Whiting, G. S.; Leo, A.; Taft, R. S. *J. Chem. Soc., Perkin Trans. 2* **1994**, 1777−1791.

(51) (a) Wolfenden, R. *Biochemistry* **1978**, *17*, 201−204. (b) Radzicka, A.; Pedersen, L.; Wolfenden, R. *Biochemistry* **1988**, *27*, 4538−4541.

(52) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529−539.

(53) (a) Jorgensen, W. L.; Blake, J. F.; Buckner, J. K. *Chem. Phys.* **1989**, *129*, 193−200. (b) Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. *J. Phys. Chem.* **1994**, *98*, 13077−13081.

(54) Gallicchio, E.; Kubo, M. M.; Levy, R. M. *J. Phys. Chem. B* **2000**, *104*, 6271−6285.

(55) Lide, D. R. *CRC Handbook of Chemisry & Physics*, 93rd ed.; CRC Press: Boca Raton, FL, 2012.

(56) Pralat, K.; Jadzyn, J.; Balanicka, S. *J. Phys. Chem.* **1983**, *87*, 1385−1390.

(57) McClellan, A. L. *Tables of Experimental Dipole Moments*; W.H. Freeman & Company: San Francisco, CA, 1963.

(58) Aroney, M. J.; Le Fèvre, R. J. W.; Pierens, R. K.; Mida, G. N. *J. Chem. Soc. B* **1971**, 1132−1135.

(59) Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. *J. Phys. Chem. A* **1998**, *102*, 3782−3797.

(60) Jorgensen, W. L.; Gao, J. *J. Am. Chem. Soc.* **1988**, *110*, 4212−4216.

(61) (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789. (c) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652. (d) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

(62) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664−675.

(63) (a) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724−728. (b) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261. (c) Hariharan, P. C.; Pople, J. A. *Theor. Chem. Acc.* **1973**, *28*, 213−22. (d) Hariharan, P. C.; Pople, J. A. *Mol. Phys.* **1974**, *27*, 209−14. (e) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654−3665.

(64) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669−681.