Solvation Model for Chloroform Based on Class IV Atomic Charges

David J. Giesen, Candee C. Chambers,† Christopher J. Cramer,* and Donald G. Truhlar*

Department of Chemistry, Supercomputer Institute, and Army High Performance Computing Research Center, 207 Pleasant Street SE, University of Minnesota, Minneapolis, Minnesota 55455-0431

Received: October 4, 1996; In Final Form: December 20, 1996[⊗]

We present a parametrization of the SM5.4 solvation model, previously applied to aqueous solutions and general organic solvents, for predicting free energies of solvation in chloroform. As in all SM5 models, the calculations are based on a set of geometry-based functional forms for parametrizing atomic surface tensions of organic solutes. In particular, the atomic surface tensions depend in some cases on distances to nearby atoms. Combining the atomic surface tensions with electrostatic effects included in a Fock operator by the generalized Born model enables one to calculate free energies of solvation by a quantum mechanical self-consistent reaction field method. Atomic charges are obtained by both the AM1-CM1A and PM3-CM1P class IV charge models, which yield similar results, and hence the same atomic radii and similar surface tension coefficients are used with both charge models. Experimental free energies of solvation and free energies of transfer from aqueous solution are used to parametrize the theory for chloroform. The parametrization is based on a set of 205 neutral solutes containing H, C, N, O, F, S, Cl, Br, and I that we used previously to parameterize a model for general organic solvents plus 32 additional solutes added for this study. For the present parameterization, we used free energies of solvation in chloroform for 88 solutes, free energies of solvation in other solvents for 123 solutes, and free energies of transfer from water to chloroform for 26 other solutes. We obtained a mean unsigned error in the free energies of solvation in chloroform of 0.43 kcal/mol using CM1A atomic charges and 0.34 kcal/mol using CM1P atomic charges.

1. Introduction

The partitioning of an organic solute between an aqueous phase and an organic medium is critical for many phenomena in biological and medicinal chemistry. In particular, this partitioning can impact upon drug delivery, binding, and clearance. The ability to understand the solvation of organic solutes in organic media is also important for conformational analysis in the condensed phase and for the prediction of molecular aggregation. Chloroform is an especially widely used organic solvent, and chloroform/water partition coefficients have been exploited[1] for the prediction of ligand lipophilicity and biological activity of organic ligands.

We have previously developed several solvation models for aqueous and other solvents, culminating in the recent SM5.4 model, which is now parameterized for aqueous solvation[2] and for general organic solvents.[3] The latter parameterization is based on data for 90 different solvents and is intended to be applicable for calculating free energies of solvation in any organic solvent. Examination of the errors and the data available for further refinement led us to the conclusion that only for aromatic hydrocarbon solvents and chloroform solvent would a more specific parameterization effort be worthwhile. The parameterization for aromatic hydrocarbon solvents will be included in ref 3b, and in this paper we present the more specific parameterization effort for chloroform. Two parameter sets are obtained, one for use with the AM1 solute Hamiltonian and one for use with the PM3 solute Hamiltonian. These new parameter sets are called SM5.4/A-chloroform and SM5.4/P-chloroform to distinguish them from the corresponding general organic parameter sets, which are called SM5.4/A-organic and SM5.4/P-organic. We also use a shorthand where SM5.4-water and SM5.4-chloroform (parameterizations for particular solvents) are called just SM5.4, whereas SM5.4-organic is sometimes called OSM5.4.

The new parameterization for calculating free energies of solvation in chloroform, when used along with our previously presented method for calculating free energies of solvation in water, also allows for the quantum mechanical calculation of chloroform/water partition coefficients.

Section 2 summarizes the theory of the SM5.4 models. Sections 3 and 4 summarize the free energy data used for the parameterization of SM5.4 for chloroform and the results we obtained. Section 5 contains discussion and comparison to previous models,[1,4] and section 6 offers concluding remarks.

2. Theory

In all SM*x* models, the standard-state free energy of solvation is written as[2,3,5−12]

$$\Delta G_S^\circ = \Delta G_{ENP} + G_{CDS} \quad (1)$$

where ΔG_{ENP} includes the change in the electronic and nuclear internal energy of the solute and the electric polarization free energy of the solute−solvent system upon insertion of the solute in the solvent, and G_{CDS} is the contribution of first-solvation-shell effects to the standard-state free energy of transfer. Note that S in ΔG_S° stands for solvation, ENP stands for electronic−nuclear-polarization, and CDS stands for cavitation−dispersion−solvent-structure, as explained elsewhere.[3,10] We use a standard state of 298 K and 1 M in both the gas phase and solution.

The ΔG_{ENP} term, which is often called the electrostatic term, can be written as

$$\Delta G_{ENP} = \Delta E_{EN} + G_P \quad (2)$$

where ΔE_{EN} is the change in the electronic and nuclear energy of the solute in going from the gas phase to solution and G_P is the polarization free energy. The form of the quantum me-

† Address as of August 1, 1996: Departments of Physics and Chemistry, Mercyhurst College, 501 East 38th St., Erie, PA 16546.

⊗ Abstract published in *Advance ACS Abstracts,* February 15, 1997.

S1089-5647(96)03080-5 CCC: $14.00

chanical self-consistent reaction field treatment used for this step and all parameters for the polarization free energy in the SM5.4 chloroform model are the same as for the SM5.4 aqueous model[2] and the SM5.4 general organic model,[3] and we note that this treatment in turn is based on several earlier references.[5,6,10,13,14] The dielectric constant (ϵ) for chloroform is taken as 4.71.[15] The solute Hamiltonian is modeled using NDDO molecular orbital theory (AM1[16] and PM3[17]) with class IV atomic charges from the CM1A[18] and CM1P[18] charge models. In the CM1A and CM1P models, charges are obtained from the one-electron density matrix by a semiempirical linear mapping that, as well as possible, makes up both for the deficiencies of the point charge ansatz and for the fact that we are not using converged quantum mechanics. Tests of these charge models indicate that the partial atomic charges obtained this way are competitive with or better than any other model currently extant. Further details of these charge models are presented elsewhere.[18]

The first-solvation-shell term has the form

$$G_{\mathrm{CDS}} = \sum_k \sigma_k A_k(R_{\mathrm{S}}^{\mathrm{CD}}) + \sigma^{\mathrm{CS}} \sum_k A_k(R_{\mathrm{S}}^{\mathrm{CS}}) \tag{3}$$

where k denotes an atom, $A_k(R)$ is the solvent-accessible surface area of atom k calculated with solvent radius R, $R_{\mathrm{S}}^{\mathrm{CD}}$ is the short-range effective solvent radius[3,10] (taken as 1.7 Å), σ_k is the microscopic surface tension of atom k, σ^{CS} is the intermediate-range molecular surface tension, and $R_{\mathrm{S}}^{\mathrm{CS}}$ is the intermediate-range effective solvent radius[3,5] (taken as 3.4 Å).

We have previously described the functional forms used for the solute-geometry dependence[2] and solvent-property dependence[3] of σ_k and σ^{CS}. The atomic surface tensions σ_k depend on solute geometry and on the hydrogen-bonding acidity[19] (α), hydrogen-bonding basicity[19] (β), and index of refraction (n) of the solvent. For example,

$$\sigma_k|_{k=\mathrm{H}} = \tilde{\sigma}_{\mathrm{H}} + \sum_{j=\mathrm{C}} \tilde{\sigma}_{\mathrm{HC}} T^{\mathrm{HC}}(R_{kj}) + \sum_{j=\mathrm{O}} \tilde{\sigma}_{\mathrm{HO}} T^{\mathrm{HO}}(R_{kj}) + \sum_{j=\mathrm{N}} \tilde{\sigma}_{\mathrm{HN}} T^{\mathrm{HN}}(R_{kj}) + \sum_{j=\mathrm{S}} \tilde{\sigma}_{\mathrm{HS}} T^{\mathrm{HS}}(R_{kj}) \tag{4}$$

where R_{kj} is the distance between atoms k and j,

$$\tilde{\sigma}_{\mathrm{H}} = \hat{\sigma}_{\mathrm{H}}^{(n)} n \tag{5}$$

$$\tilde{\sigma}_{XY} = \hat{\sigma}_{XY}^{(n)} n + \hat{\sigma}_{XY}^{(\alpha)} \alpha + \hat{\sigma}_{XY}^{(\beta)} \beta \tag{6}$$

$$T^{\mathrm{XY}}(R) = \begin{cases} \exp\left(\dfrac{\Delta R^{\mathrm{XY}}}{R - \Delta R^{\mathrm{XY}} - \bar{R}^{\mathrm{XY}}}\right) & R \le \bar{R}^{\mathrm{XY}} + \Delta R^{\mathrm{XY}} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

ΔR^{XY} and $\bar{R}^{\mathrm{XY}}$ are nonlinear parameters, and the $\hat{\sigma}_X$ and $\hat{\sigma}_{XY}$ are linear parameters (surface tension coefficients) to be optimized. The values of α, β, and n for chloroform are 0.15,[19] 0.02,[19] and 1.4459,[15] respectively. The medium-range molecular surface tension σ^{CS} depends on n and on the macroscopic surface tension γ of the solvent. The latter is 38.4 cal Å^{-2} mol^{-1} for chloroform.[15]

3. Experimental Data, Training Set, and Parametrization

Experimental Data and Training Set. The database of experimental free energies of solvation used to parameterize SM5.4 for chloroform has evolved from data bases used in our group to parameterize solvation models for water[2] and general organic solvents.[3] We started with the data used in the parameterization of the Organic Solvation Model 5.4 (SM5.4-organic).[3] This data had been extracted primarily from the MedChem data base[20] with the rejection of statistical outliers (experimental data that disagree with the mean experimental data by more than two standard deviations) and secondarily from several other sources; all the data referring to chloroform were from the primary source. The majority of the free energies of solvation (which are the same as solvent/air transfer free energies) were obtained from solvent/water partition coefficients. Solvent/water partitioning for a given solute is usually quantified by the partition coefficient P or its logarithm, where

$$P_{\mathrm{Solvent/H_2O}} = \frac{[\mathrm{solute}]_{\mathrm{Solvent}}}{[\mathrm{solute}]_{\mathrm{H_2O}}} \tag{8}$$

$$\log P_{\mathrm{Solvent/H_2O}} = -2.303 \Delta\Delta G_{\mathrm{S}}^{\circ}/RT \tag{9}$$

and

$$\Delta\Delta G_{\mathrm{S}}^{\circ}(\mathrm{Solvent/H_2O}) = \Delta G_{\mathrm{S}}^{\circ}(\mathrm{Solvent}) - \Delta G_{\mathrm{S}}^{\circ}(\mathrm{H_2O}) \tag{10}$$

Here $[x]_y$ denotes the equilibrium concentration of solute x in solvent y, $\Delta G_{\mathrm{S}}^{\circ}(y)$ is the standard-state free energy of solvation of a given solute in solvent y, R is the gas constant, T is the temperature, and $\Delta\Delta G_{\mathrm{S}}^{\circ}(\mathrm{Solvent/H_2O})$ is the transfer free energy of a given solute to the specified solvent from water. Throughout the article, all logarithms are to the base 10. Knowing the solvent/water partition coefficient and the water/air transfer free energy allows calculation of the solvent/air transfer free energy.

A strength of the SM5.4-organic model is that it can predict free energies of solvation in a particular solvent for solutes characterized by functionality for which little or no coverage exists in that solvent's training set *provided* data are available for solutes having that functionality in *other* solvents. The strategy used here is designed to take advantage of this feature as well. The SM5.4-organic training set contains 1784 free energies of solvation for 205 solutes in 90 solvents, and this includes 82 free energies of solvation in chloroform. For these 82 solutes, the chloroform data point was weighted by the number of times the solute appeared in the SM5.4-organic parameterization set, and all other data points for that solute (in other solvents) were eliminated. For example, if the SM5.4-organic training set contains free energies of solvation for a particular solute in five solvents (and one of them is chloroform), the SM5.4-chloroform training set contains the free energy of solvation for that solute in chloroform with a weight of 5. Data for solutes without chloroform free energies of solvation remain with a weight of unity for each free energy.

At this point the data base contains 82 free energies of solvation in chloroform with a total weight of 1399, and 385 free energies of solvation in other solvents, each with a weight of unity. Two other types of data were then added to the SM5.4-chloroform parameterization set. First we added six chloroform free energies of solvation obtained using eqs 8−10. These data were added for diisopropyl ether, methyl benzoate, tetrahydropyran, thiophene, ethoxybenzene, and 1,1-dimethyl-3-phenylurea. These compounds were added because they contain functionalities for which there is reduced representation in the SM5.4-organic training set. These data were given a weight of unity. These six solutes along with the 82 solutes with chloroform/air data in the SM5.4-organic set comprise 88 molecules for which we use experimentally derived free energies of solvation in chloroform, and they will be referred to as the SM5.4-chloroform free energy training set.

TABLE 1: Molar Free Energies of Solvation (kcal) in Wet and Dry Chloroform

solute	wet	dry
toluene	−5.1[a]	−5.5
ethanol	−3.8	−3.9
2-butanone	−5.2	−5.4

[a] There are three chloroform/water partition coefficients for toluene; two are identical (3.41), and the third is quite different (2.48). Our scheme retains all three because they are all within two standard deviations of the mean. If we eliminate the 2.48 number, the molar free energy of solvation of toluene in wet chloroform becomes −5.6 kcal.

Finally we added data for 26 additional compounds for which chloroform/water partition coefficients are available, but no experimental water/air data are available with which to calculate air/chloroform transfer free energies. For these 26 compounds, first the water/air transfer free energy was calculated using the SM5.4-aqueous solvation model. This calculated value was used to yield a chloroform/air transfer free energy, which was added to the training set with a weight of 1. Because these data are based on experimental chloroform/water partition coefficients, this set of molecules will be referred to as the SM5.4-chloroform partition coefficient training set. The two training sets combined with the data retained from other solvents will be referred to as the SM5.4-chloroform parameterization set.

The total weight of data points in the SM5.4-chloroform parameterization set is 1816, consisting of 88 free energies of solvation in chloroform with a total weight of 1405, 385 free energies of solvation in other solvents with a total weight of 385, and 26 transfer free energies from water to chloroform with a total weight of 26. (All data have integer weights; we did not attempt to weight data with smaller experimental errors more strongly than other data, except for the statistical outlier step mentioned above.) Two models were developed using this training set. SM5.4/A-chloroform uses the AM1 Hamiltonian for the solute electrons and CM1A atomic charges, and SM5.4/P-chloroform uses the PM3 Hamiltonian and CM1P atomic charges.

Of the 88 free energies of solvation, six (octane, toluene, ethanol, *p*-dioxane, 2-butanone, and nitromethane) come from chloroform/air data, and the other 82 come from combining chloroform/water data with water/air data. The six former data points refer therefore to dry chloroform, and the others to wet chloroform. In three of the cases where we used dry chloroform data, there is also wet chloroform data. The comparison of the free energies of solvation in these three cases is presented in Table 1. We find excellent agreement, even for ethanol with its hydroxyl group. We conclude that the free energy of solvation in neat chloroform equals the free energy of solvation in chloroform saturated with water within experimental error. A further indication that this is the case is provided by the situation for octanol. Although the solubility of water in octanol is much higher than in chloroform, one again finds that the free energies of solvation are essentially the same in the wet and dry solvent.[21]

Parameterization Procedure. Preliminary tests of parameterization with and without the non-chloroform-solvent data indicted that the parameters for which we had reasonable coverage with chloroform data were not significantly affected by deleting the data from other solvents. This other data, however, is essential for the derivation of meaningful values for parameters to which the chloroform data is itself insensitive. Examples of such parameters are the surface tensions that primarily affect disulfides and alkynes.

Four parameters were not optimized but rather were taken over from the corresponding SM5.4-organic (/A or /P) models.

TABLE 2: Microscopic Surface Tensions (cal mol^{-1} Å^{-2}) for SM5.4/A-Chloroform and SM5.4/P-Chloroform by Atom Type and Atom Pair

		SM5.4/A		SM5.4/P	
k		$\tilde{\sigma}_k$	$\tilde{\sigma}_{Hk}$	$\tilde{\sigma}_k$	$\tilde{\sigma}_{Hk}$
H		−75.23		−69.04	
C		−0.83	16.38	−1.95	14.14
O		−88.57	83.00	−94.77	92.00
N		−67.80	10.32	−48.74	10.28
F		−46.09		−41.63	
S		−101.89	81.08	−92.88	64.10
Cl		−68.84		−64.05	
Br		−72.84		−69.38	
I		−77.29		−73.54	
k	k'	$\tilde{\sigma}_{kk'}$	$\tilde{\sigma}^{(2)}_{kk'}$	$\tilde{\sigma}_{kk'}$	$\tilde{\sigma}^{(2)}_{kk'}$
O	O	11.93		2.88	
O	N	64.23		92.42	
O	C	58.59		81.68	
C	C	−45.78	−6.46	−50.70	1.24
S	S	15.71		16.23	
N	C	−16.92		−23.36	
		σ^{CS}		σ^{CS}	
		29.20		26.74	

In particular $\hat{\sigma}^{(n)}_H$ could not be improved since data do not exist for H_2 in chloroform, and $\hat{\sigma}^{(\beta)}_S$, $\hat{\sigma}^{(\beta)}_{HS}$, and $\hat{\sigma}^{(\beta)}_{SS}$ were also not improved due to the low number of sulfur-containing compounds in solvents with nonzero α and β after non-chloroform sulfur data had been removed.

For all parameters, after the surface tension coefficients for the dependencies on n, α, β, and γ were found, they were multiplied by the respective solvent properties of chloroform and added as in eq 6 to obtain the atomic surface tensions of the SM5.4-chloroform models. Table 2 contains the SM5.4/A and SM5.4/P surface tension coefficients for chloroform.

All calculations were carried out with a modified version of the AMSOL code.[22] (The chloroform models presented here will be included in version 6.0 of AMSOL.)

4. Results

Table S1 of the Supporting Information gives free energies of solvation in chloroform calculated by the SM5.4/A-chloroform model and the SM5.4/P-chloroform model along with the values for the ENP and CDS portions of the free energy of solvation for all 88 solutes in the SM5.4-chloroform free energy set. Table 3 presents a subset of Table S1 that is useful for illustrating trends.

Tables 4 and 5 contain the mean signed and mean unsigned errors, respectively, over all functional groups included in the SM5.4-chloroform free energy set for the SM5.4/A, SM5.4/P, OSM5.4/A, and OSM5.4/P models. For informational purposes, we have also included the number of molecules in the SM5.4-chloroform free energy training set that fall into each functional group.

Table S1 of the Supporting Information contains the SM5.4-aqueous and SM5.4-chloroform results for the SM5.4-chloroform partition coefficient training set along with a comparison to experimental chloroform water partition coefficients. Table 6 contains a subset of Table S2 that is useful for illustrating trends.

Table 7 gives results for seven molecules not included in the SM5.4-chloroform parameterization set, but we present them for discussion purposes. Several of these have one or more conformations that lie near the global minimum conformation. For these molecules, the free energy of solvation is calculated using a Boltzmann-averaged population. Since such an average

TABLE 3: Calculated and Experimental[a] Standard State Free Energies of Solvation (kcal/mol) in Chloroform for Selected Molecules in the SM5.4-Chloroform Free Energy Training Set

	SM5.4/A			SM5.4/P			experiment
molecule	G_{CDS}	ΔG_{ENP}	ΔG_S°	G_{CDS}	ΔG_{ENP}	ΔG_S°	ΔG_S°
toluene	−3.0	−2.4	−5.4	−3.9	−1.5	−5.4	−5.5
m-xylene	−3.6	−2.4	−6.1	−4.4	−1.5	−5.9	−5.9
ethanol	−0.1	−3.5	−3.6	−0.4	−3.2	−3.6	−3.9
1-butanol	−2.1	−2.9	−4.9	−2.3	−2.8	−5.1	−5.3
o-cresol	−2.3	−5.2	−7.6	−3.2	−4.0	−7.3	−7.6
1-heptanol	−4.6	−2.8	−7.4	−4.8	−2.8	−7.6	−7.5
tetrahydropyran	−3.3	−2.2	−5.5	−3.7	−1.7	−5.4	−5.8
ethyl phenyl ether	−4.0	−3.2	−7.2	−5.0	−2.1	−7.1	−7.2
ethanal	0.3	−3.6	−3.3	0.4	−3.8	−3.4	−3.7
acetophenone	−3.2	−4.6	−7.9	−4.0	−4.0	−8.0	−7.8
pentanoic acid	−1.0	−5.5	−6.5	−1.2	−5.3	−6.5	−6.6
propyl ethanoate	−2.5	−4.1	−6.5	−2.7	−3.8	−6.4	−6.4
methyl pentanoate	−3.0	−3.6	−6.6	−3.2	−3.3	−6.6	−6.7
p-hydroxybenzaldehyde	−1.8	−7.7	−9.5	−2.5	−6.9	−9.4	−10.3
dimethylamine	−2.5	−1.6	−4.1	−3.3	−0.9	−4.3	−3.7
propylamine	−3.4	−1.4	−4.8	−3.6	−1.1	−4.7	−4.7
trimethylamine	−2.7	−1.6	−4.2	−3.7	−0.5	−4.2	−3.9
piperidine	−4.7	−1.1	−5.8	−5.3	−0.7	−6.1	−6.4
pyridine	−2.7	−3.5	−6.3	−3.7	−2.5	−6.3	−6.5
aniline	−3.5	−3.7	−7.2	−4.2	−3.3	−7.5	−7.3
2-ethylpyrazine	−4.3	−3.5	−7.8	−5.2	−2.8	−8.0	−7.7
ethanonitrile	−0.2	−4.5	−4.7	0.3	−5.0	−4.7	−4.4
nitrobenzene	−1.3	−6.4	−7.7	−2.2	−5.6	−7.7	−7.8
ethanamide	−1.3	−6.0	−7.3	−0.9	−6.2	−7.1	−7.1
diethyl sulfide	−4.0	−1.1	−5.1	−4.1	−1.1	−5.3	−6.4
thioanisole	−5.5	−0.8	−6.3	−5.5	−0.9	−6.5	−6.0
p-dichlorobenzene	−4.4	−1.6	−6.0	−5.1	−1.0	−6.1	−6.3
iodobenzene	−4.7	−2.0	−6.7	−5.6	−1.2	−6.8	−6.6

[a] MedChem database, ref 20.

TABLE 4: Mean Signed Errors (kcal/mol) in Free Energies of Solvation in Chloroform by Functional Group for OSM5.4/A, SM5.4/A, OSM5.4/P, and SM5.4/P Models

functional group	number of molecules	OSM5.4/A	SM5.4/A	OSM5.4/P	SM5.4/P
Compounds Containing at Most C, H, and/or O					
unbranched alkanes	1	0.60	0.18	0.70	0.25
cycloalkanes	1	1.83	1.09	1.97	1.23
arenes	6	0.19	0.00	0.40	−0.01
alcohols	14	0.22	0.14	0.23	0.09
ethers	6	0.17	−0.41	0.40	−0.29
aldehydes	3	−0.37	−0.13	−0.53	−0.28
ketones	3	0.37	0.14	0.33	0.10
carboxylic acids	5	0.39	0.21	0.38	0.13
esters	9	0.38	−0.15	0.47	−0.08
bifunctional	2	0.30	0.68	0.28	0.56
water	1	−2.03	−1.51	−1.77	−1.23
entire subset	51	0.23	0.02	0.30	0.01
Compounds Containing N					
aliphatic amines	8	0.69	−0.08	0.75	−0.08
aromatic amines	8	0.37	0.10	0.56	0.13
nitriles	2	0.15	−0.04	0.21	−0.02
nitrohydrocarbons	2	0.42	0.18	0.12	0.15
amides	1	0.03	−0.25	0.10	−0.05
bifunctional	2	1.46	1.00	0.70	0.01
ammonia, hydrazine	2[a]	2.01	1.48	1.29	0.82
entire subset	25[a]	0.66	0.21	0.58	0.06
Compounds Containing S, H, and/or C					
thiols	1	0.50	0.68	0.72	0.74
organic sulfides, H_2S	4	0.35	0.30	0.28	0.21
entire subset	5	0.38	0.38	0.37	0.32
Compounds Containing Halogens					
fluorinated hydrocarbons	1	0.38	0.30	0.24	−0.07
chloroarenes	2	0.17	0.18	0.25	0.07
brominated hydrocarbons	1	−0.06	0.00	−0.07	−0.19
iodinated hydrocarbons	1	−0.22	−0.07	−0.13	−0.19
other halo molecules	2	−1.32	−0.91	−1.06	−0.74
entire subset	7	−0.31	0.37	−0.22	0.37
entire set	88[a]	0.32	0.08	−0.34	0.02

[a] Hydrazine was not calculated using either the OSM5.4/P or the SM5.4/P method (see ref 2). Therefore, the number of elements for PM3 methods is one less than the value shown in the "number of molecules" column.

TABLE 5: Mean Unsigned Errors (kcal/mol) in Free Energies of Solvation in Chloroform by Functional Group for OSM5.4/A, SM5.4/A, OSM5.4/P, and SM5.4/P Models

functional group	number of molecules	OSM5.4/A	SM5.4/A	OSM5.4/P	SM5.4/P
Compounds Containing at Most C, H, and/or O					
unbranched alkanes	1	0.60	0.18	0.70	0.25
cycloalkanes	1	1.83	1.09	1.97	1.23
arenes	6	0.19	0.10	0.40	0.13
alcohols	14	0.55	0.36	0.36	0.27
ethers	6	0.89	0.53	0.96	0.52
aldehydes	3	0.86	0.62	0.81	0.53
ketones	3	0.37	0.18	0.33	0.22
carboxylic acids	5	0.39	0.21	0.38	0.17
esters	9	0.38	0.23	0.47	0.19
bifunctional	2	0.30	0.68	0.28	0.56
water	1	2.03	1.51	1.77	1.23
entire subset	51	0.56	0.36	0.55	0.32
Compounds Containing N					
aliphatic amines	8	0.69	0.34	0.75	0.21
aromatic amines	8	0.43	0.26	0.59	0.38
nitriles	2	0.19	0.17	0.30	0.20
nitrohydrocarbons	2	0.42	0.18	0.12	0.15
amides	1	0.03	0.25	0.10	0.05
bifunctional	2	1.46	1.64	0.70	0.73
ammonia, hydrazine	2[a]	2.36	2.37	1.29	0.82
entire subset	25[a]	0.71	0.55	0.60	0.32
Compounds Containing S, H, and/or C					
thiols	1	0.50	0.68	0.72	0.74
organic sulfides, H_2S	4	0.63	0.63	0.68	0.64
entire subset	5	0.60	0.64	0.69	0.66
Compounds Containing Halogens					
fluorinated hydrocarbons	1	0.38	0.30	0.24	0.07
chloroarenes	2	0.17	0.18	0.25	0.17
brominated hydrocarbons	1	0.06	0.00	0.07	0.19
iodinated hydrocarbons	1	0.22	0.07	0.13	0.19
other halo molecules	2	1.32	0.93	1.06	0.87
entire subset	7	0.52	0.37	0.44	0.37
entire set	88[a]	0.60	0.43	0.56	0.34

[a] Hydrazine was not calculated using either the OSM5.4/P or the SM5.4/P method (see ref 2). Therefore, the number of elements for PM3 methods is one less than the value shown in the "number of molecules" column.

TABLE 6: Selected log $P_{c/w}$ Values for the SM5.4-Chloroform Partition Coefficient Training Set As Calculated by the SM5.4 Model and Compared to Experiment

solute	SM5.4/A	SM5.4/P	experiment[a]
malononitrile	0.3	−0.1	−0.5
metrazole	2.2	2.3	1.4
quinoline	2.6	2.7	3.4
quinoxaline	1.4	1.5	2.1
succinimide	−2.0	−1.8	−1.3
N-methyl-2-pyrrolidinone	0.5	0.4	0.3
N-formylpyrrolidine	0.7	1.5	0.7
1-hyroxybenzo-1,2,3-triazole	1.6	1.3	−1.0
1-methylcytosine	−3.1	−2.7	−3.0
thioacetamide	−0.6	−0.4	−1.1
thiourea	−2.3	−2.1	−3.1
2-methylthiopyrimidine	1.2	1.5	2.0
thiobenzamide	1.4	0.3	1.4
5-fluoropyrimidine	−0.1	−0.3	0.9
mean signed error (26 solutes)[b]	0.04	0.04	
mean unsigned error (26 solutes)	0.69	0.75	
root-mean-square error (26 solutes)	0.91	0.93	

[a] Reference 20. [b] See Table S2 of Supporting Information for full set of 26 solutes.

involves not only the Boltzmann-averaged population in the gas phase but also a Boltzmann-averaged population in solution, the free energy of solvation was calculated using[23]

$$\exp[\Delta G_S^\circ/RT] = \sum_C P_C \exp[-\Delta G_S^\circ(C)/RT] \qquad (11)$$

where C runs over all relevant conformations, and P_C is the Boltzmann-averaged mole fraction population of conformation C in the gas phase. The gas phase conformational analysis was performed using geometries and energies computed with the semiempirical Hamiltonian consistent with the solvation model used to calculate the solution geometries and energies. Two molecules, *N*,*N*-dimethyl-*N*′-(3,4-dichloro)phenylurea (diuron) and *N*,*N*-dimethyl-*N*′-(3-trifluoromethyl)phenylurea (fluometuron), have three conformations lying within 2 kcal/mol of the global minimum. The four conformers for each molecule involve combinations of 180° rotations about the C(phenyl)−N′ bond and N′−C(O) bond. One molecule, *N*,*N*-dimethyl-*N*′-(4-chlorophenyl)urea (monuron), has two low-lying conformations which involve 180° rotation about the N′−C(O) bond. The lack of a substituent on the 3-phenyl position causes monuron to have two fewer possible conformations than do either diuron or fluometuron. Finally 3-methylbutanoic acid has four low-lying conformations which involve rotations about the C2−C3 and C3−C4 bonds.

5. Discussion

The parameters in Table 2 all have reasonable magnitudes. In no case does the parameterization lead to unphysical cancellations of large positive and negative contributions. However, the CS terms are always large and positive, and the CD terms are usually even larger, but negative. This situation is entirely consistent with experimentally derived free energies of cavitation and dispersion for nonpolar solutes in alkane solvents.[10,11]

TABLE 7: Calculated and Experimental Free Energies of Solvation in Chloroform, ΔG_S°, (kcal/mol) for Selected Solutes Not Used in the Parametrizations of SM5.4

	number of conformations	OSM5.4/A	SM5.4/A	OSM5.4/P	SM5.4/P	experiment[a]
monuron[b]	2	−11.9	−11.3	−13.2	−13.0	−13.1
diuron[c]	4	−12.7	−11.9	−13.9	−13.6	−14.4
fluometuron[d]	4	−11.0	−10.2	−12.3	−11.9	−13.0
3-methylbutanoic acid	4	−5.9	−5.9	−6.0	−6.1	−6.4
phenanthrene	1	−11.1	−11.1	−10.8	−11.1	−10.9
benzyl alcohol	1	−7.2	−7.3	−7.1	−7.2	−7.7
octanoic acid	1	−8.5	−8.6	−8.4	−8.6	−9.0

[a] Ref 20. [b] *N,N*-dimethyl-*N*′-(4-chlorophenyl)urea. [c] *N,N*-dimethyl-*N*′-(3,4-dichlorophenyl)urea. [d] *N,N*-dimethyl-*N*′-(3-trifluoromethylphenyl)urea.

TABLE 8: Gas Phase Conformational Energies and Free Energies of Solvation (kcal/mol) for Individual Conformations for Molecules in Table 7[a]

	dihedral angle (deg)					
solute	C2−C1−N′−C7	C1−N′−C7−O	AM1 E(g)	SM5.4/A ΔG_S°	PM3 E(g)	SM5.4/P ΔG_S°
monuron	0	0	0.0	−12.0	0.0	−13.5
	0	180	1.1	−10.6	0.3	−11.0
diuron	180	0	0.0	−12.6	0.0	−14.2
	0	0	0.3	−13.0	0.4	−14.3
	0	180	1.3	−11.7	0.3	−12.1
	180	180	1.3	−11.3	0.3	−11.8
fluorometuron	180	0	0.0	−10.7	0.0	−12.3
	0	0	0.9	−11.6	0.7	−12.9
	180	180	1.6	−9.7	0.7	−10.1
	0	180	1.7	−10.2	0.8	−10.5

[a] C1 and C2 are in the phenyl ring; C7 is the carbonyl carbon. Dihedral values are nominal and can vary by as much as 45° in the optimized structure. Note that when the first dihedral changes from 0 to 180°, the second dihedral is affected by that change. Thus a value of 0° for the second dihedral can indicate either *syn* or *anti* to the substituent at the 3 position.

Table S1 of the Supporting Information shows that the actual chloroform/air data is not uniformly spread over the common functional groups. However, since the SM5.4-organic solute set *is* well balanced with respect to solute functional groups, the weighting scheme presented in section 3 ensures that the weights used here do balance the solute functionalities.

Table 3 shows that the SM5.4/P model tends to give a less negative ΔG_{ENP} contribution than does the SM5.4/A model. However, the net predicted solvation energies are more similar than their components.

Table 4 shows that the mean signed errors in ΔG_S° are uniformly small across solute functional groups, with the main exceptions being cyclohexane (the only cycloalkane datum), small inorganics (H_2O, NH_3, N_2H_4), and thiols. For functional groups with three or more representatives, the mean signed errors in the SM5.4 models are smaller than the corresponding OSM5.4 values in 17 out of 20 cases.

Table 5 shows that there is also considerable across-the-board improvement in the mean unsigned errors. The most significant improvements in mean unsigned errors are for ethers, esters, and aliphatic amines. The new parameters do not improve the bifunctional solutes, but there are so few of these that we do not regard the comparison as significant. More importantly, even for the bifunctional solutes, the absolute errors obtained with the new parameters are small. A detailed discussion of the application of the present model to partitioning of substituted purine and pyrimidine bases between water and chloroform will be presented elsewhere;[24] the new parameters do better than the general organic ones for these multifunctional solutes, which involve functionality not present in the data set for absolute free energies of solvation.

The 26 solutes in the partition coefficient training set have 13 positive values of log K_{CH_3Cl/H_2O} and 13 negative values (Table S2 of the Supporting Information and Table 6). The SM5.4/A-chloroform model gets the correct sign in 23 of the 26 cases, and the SM5.4/P-chloroform model gets the correct sign in 24. Both models predict the wrong sign for 5-fluoropyrimidine and 1-hydroxybenzo-1,2,3-triazole. For the latter, the electrostatics are about 2 kcal/mol more negative in water, but the CDS contribution is much more negative in chloroform.

The additional tests in Table 7 are primarily encouraging in that the four different models all predict quite similar results, despite the conformational complexity of these solutes. Table 8 shows the gas phase energies and free energies of solvation for individual conformers of the first three molecules in Table 7. We note that the lowest energy gas phase conformation is best solvated for monuron, and the two lowest gas phase conformations are the ones that are best solvated for both diuron and flurometuron. For 3-methylbutanoic acid, the solvation energies of the four conformations are all the same within 0.1 kcal.

Recently, another quantum mechanical solvation model for chloroform appeared in the literature.[4] Two parameterizations have been presented; we will call these B-MST/A and B-MST/6-31G* since they are part of the Barcelona project for parameterizing the formulation of Miertius, Scrocco, and Tomasi (MST) using the AM1 and HF/6-31G* levels of theory, respectively. As compared to the present models, these models use different treatments of the electrostatics and different functional forms for the surface tensions. Table 9 provides the results for our models and the models of Luque and co-workers over their smaller parameterization set (27 solutes). With the exception of the two nitrophenols, all of their molecules were also included in the SM5.4-chloroform free energy set, where they had a total weight of 664 out of 1816. Both models show comfortably low errors which begin to approach experimental accuracy (which we and others[25] estimate as being typically ±0.2 kcal).

It is also of interest to compare intermediate quantities and other characteristics of our new solvation model to the results

TABLE 9: Calculated (SM5.4/A, SM5.4/P, and Two Models of Luque and Orozco[a]) and Experimental[b] Free Energies of Solvation in Chloroform, ΔG_S°, (kcal/mol)

			B-MST[a]		
	SM5.4/A	SM5.4/P	6-31G*	AM1	expt[b]
methanol	−2.6	−2.7	−3.3	−2.5	−3.3
ethanol	−3.6	−3.6	−3.7	−3.5	−3.9
1-propanol	−4.3	−4.2	−4.5	−4.2	−4.4
1-butanol	−4.9	−5.1	−5.1	−5.0	−5.3
1-pentanol	−5.7	−5.8	−5.9	−5.8	−5.9
phenol	−7.2	−7.1	−7.3	−7.2	−7.1
1-hexanol	−6.6	−6.9	−6.6	−6.6	−6.7
o-cresol	−7.6	−7.3	−7.5	−7.6	−7.6
p-cresol	−7.9	−7.7	−7.9	−7.8	−7.6
1-heptanol	−7.4	−7.6	−7.4	−7.5	−7.5
propanone	−4.4	−4.4	−5.1	−4.9	−4.4
acetophenone	−7.9	−8.0	−8.5	−8.8	−7.8[c]
ethanoic acid	−4.7	−4.8	−5.0	−5.0	−4.7
propanoic acid	−5.2	−5.3	−5.5	−5.5	−5.4
butanoic acid	−5.7	−5.8	−6.2	−6.2	−6.0
methyl ethanoate	−4.8	−4.8	−5.3	−5.4	−4.9
ethyl ethanoate	−5.9	−5.8	−5.9	−5.0	−5.6
ethylamine	−4.2	−4.0	−4.0	−4.1	−4.0
propylamine	−4.8	−4.7	−4.6	−4.9	−4.7
trimethylamine	−4.2	−4.2	−3.6	−3.9	−3.9
butylamine	−5.5	−5.4	−5.4	−5.7	−5.4
diethylamine	−4.7	−5.0	−4.4	−4.7	−5.2
pyridine	−6.3	−6.3	−6.3	−6.4	−6.5
aniline	−7.2	−7.5	−7.4	−6.4	−7.3
p-hydroxybenzaldehyde	−9.5	−9.4	−10.3	−9.6	−10.3
m-nitrophenol	−9.9	−9.8	−10.0	−10.5	−10.5
p-nitrophenol	−10.6	−10.3	−10.1	−10.6	−11.0
mean signed error	0.14	0.13	0.01	0.06	
mean unsigned error	0.25	0.23	0.25	0.32	
root-mean-square error	0.32	0.32	0.36	0.43	

[a] Reference 4. [b] Reference 20. [c] Luque *et al.* (ref 4) give −8.4 kcal/mol as the experimental value for acetophenone.

of Luque et al. For example, they reported electrostatic polarization energies for several cases calculated under the assumption of unperturbed solute; in our code, AMSOL, this is denoted the NOPOL option. (In a NOPOL polarization energy, the solvent is polarized but the solute is not.) The comparison is shown in Table 10. On average, our values are 1−2 kcal more negative. In the first place, we point out that the division of free energy of solvation effects into electrostatic and "other" is not unique, either experimentally or by simulation. However, having said that, we also point out that we have developed the class IV charge model to make our atomic charges as realistic as possible, and we have attempted to choose atomic radii for the solvation calculation that make the electrostatic term as physical as possible.

Although our electrostatic contributions are much larger in magnitude than those computed with Luque et al.'s final parameters, they also presented a second set of results, which are equally interesting to compare to ours. The second set of results will be labeled B-MST/1.2A and B-MST/1.25−6-31G*, because of the way that they differ from the parameterization that Luque et al. accepted. This difference is in the choice of scaling factor employed for the atomic radii. Their final choice for this scaling factor is 1.6 in the B-MST/A and B-MST/6-31G* models, but it is reduced to 1.2 and 1.25 in the B-MST/1.2A and B-MST/1.25−6-31G* models, respectively. The last two columns of Table 10 present the NOPOL polarization energies for the B-MST models with smaller atomic radii, and they are typically larger than ours. Our results would thus agree best with theirs if they chose a scaling factor between 1.25 and 1.6. These differences have important implications for the semiempirical surface tensions. As an example, we point out that our model yields positive CDS terms for water, methanol, and acetic acid, which have polarizabilities of 5.1, 1.5, and 3.3 Å^3, respectively, all smaller than the polarizability of chloroform, which is 9.5 Å^3.

We have negative CDS terms of magnitude less than 1.0 kcal/mol for acetonitrile, ethanol, and acetone, which have intermediate-size polarizabilities of 4.4, 5.3, and 6.4 Å^3, respectively, and we have more negative CDS contributions for ethylamine and *N*-methylacetamide, which have polarizabilities of 7.1 and 7.8 Å^3, respectively. We think it is not unreasonable to have positive CDS terms for the case where the solute polarizability of the solute is less than the solvent. Dissolving such a solute involves replacing chloroform−chloroform dispersion interactions by weaker solute−chloroform interactions, and the positive cavitation terms might dominate in such a case. Further work will be necessary to decide which parameterization is more "physical." (We stress again that the division of free energy of solvation effects into electrostatic and "other" is not unique, either experimentally or by simulation; the essential issue for an accurate semiempirical model is that any terms added to the electrostatic effects—these are the CDS terms in our models—must be parameterized *consistently* with the necessarily somewhat arbitrary choices made for the atomic radii, which are artificial concepts and hence not systematically improvable. However, solute properties like dipole moments are more sensitive to details of the model, because the nonelectrostatic effects are not treated self-consistently. As a practical matter, it appears to be easier to parameterize a model by underestimating the bulk electrostatic effects and including the deficit with first-solvation effects than by overestimating them and trying to cancel the overestimate with CDS terms.)

Luque et al. also reported the change in solvation free energy when the system is allowed to relax, i.e., when the electronic structure of the solute is allowed to change self-consistently in solution. Their results for this quantity are compared to ours

TABLE 10: NOPOL Electrostatic Polarization Energies (kcal/mol) for Selected Solutes in Chloroform

	B-MST/			B-MST/	
solute	6-31G*	A	SM5.4/A	1.2A	1.25−6-31G*
water	−2.8	−1.8	−5.8	−5.7	−6.0
ammonia	−1.9	−1.4	−2.5	−5.2	−4.8
methanol	−1.8	−1.2	−3.1	−4.0	−4.3
acetonitrile	−2.4	−1.2	−3.4	−2.7	−4.4
ethylamine	−1.3	−1.0	−1.4	−4.5	−4.0
ethanol	−1.7	−1.1	−3.1	−4.2	−4.1
acetic acid (*cis*)	−2.4	−2.0	−5.3	−6.1	−6.0
acetic acid (*trans*)	−4.2	−3.3	−7.3	−8.2	−8.3
acetone	−1.9	−1.5	−2.8	−4.1	−4.3
anti-*N*-methylacetamide	−3.0	−2.2	−4.1	−6.5	−6.6
trans-*N*-methylacetamide	n.a.[a]	n.a.[a]	−4.2	n.a.[a]	n.a.[a]

[a] Not available.

TABLE 11: Energy Change (kcal/mol) upon Relaxation of Solute in Chloroform

	B-MST/			B-MST/	
solute	6-31G*	A	SM5.4/A	1.2A	1.25−6-31G*
water	−0.1	−0.1	−0.7	−0.4	−0.4
ammonia	−0.1	0.0	−0.3	−0.1	−0.2
methanol	−0.1	0.0	−0.4	−0.3	−0.3
acetonitrile	−0.2	−0.1	−1.1	−0.5	−0.6
ethylamine	0.0	0.0	−0.2	−0.1	−0.2
ethanol	−0.1	0.0	−0.4	−0.3	−0.2
acetic acid (*cis*)	−0.1	−0.1	−1.0	−0.5	−0.4
acetic acid (*trans*)	−0.3	−0.3	−2.3	−1.3	−1.0
acetone	−0.2	−0.1	−0.9	−0.7	−0.5
anti-*N*-methylacetamide	−0.3	−0.2	−1.1	−1.2	−0.8
trans-*N*-methylacetamide	n.a.[a]	n.a.[a]	−1.2	n.a.[a]	n.a.[a]

[a] Not available.

TABLE 12: Dipole Moments (Debyes) in the Gas Phase and in Chloroform Solution

	Luque et al.			present	
solute	gas	B-MST/A	B-MST/1.2A	gas	SM5.4/A
water	1.86	1.94	2.09	2.03	2.29
ammonia	1.85	1.87	1.95	1.79	1.92
methanol	1.62	1.71	1.89	1.63	1.94
acetonitrile	2.89	3.20	3.67	3.80	4.99
ethylamine	1.55	1.63	1.78	1.38	1.61
ethanol	1.55	1.67	1.89	1.56	1.95
acetic acid (*cis*)	1.89	2.06	2.34	1.97	2.61
acetic acid (*trans*)	4.39	4.84	5.48	4.53	6.15
acetone	2.85	3.17	3.71	2.95	4.11
anti-*N*-methylacetamide	3.51	3.98	4.81	3.71	4.99
trans-*N*-methylacetamide	n.a.[a]	n.a.[a]	n.a.[a]	2.92	4.20

[a] Not available.

TABLE 13: log $P_{c/w}$ Calculated Using SM5.4 Methods and Values Calculated by Reynolds[a]

	log $P_{c/w}$			
	SM5.4/A	SM5.4/P	Reynolds	expt
benzene	2.7	2.8	4.3	2.8
toluene	3.3	3.3	5.0	3.1
ethylbenzene	3.9	3.9	5.5	3.7
methanol	−2.0	−2.0	−1.5	−1.3
ethanol	−0.9	−0.8	−0.5	−0.9
1-pentanol	1.1	1.0	1.1	1.1
phenol	0.4	0.5	1.0	0.4
1-hexanol	1.6	1.5	1.6	1.7
ethanal	−0.7	−0.6	−0.6	0.1
propanone	0.1	0.1	1.6	0.4
acetophenone	2.3	2.4	3.5	2.4
ethanoic acid	−1.3	−1.3	−1.4	−1.4
butanoic acid	0.2	0.3	−0.1	−0.3
hexanoic acid	1.6	1.5	1.0	1.0
diethylamine	1.4	1.4	2.8	0.9
pyridine	0.9	1.0	2.1	1.3
benzonitrile	3.0	2.8	3.1	2.3
2-propen-1-ol	−0.4	−0.3	0.1	−0.5
ethanamide	−2.2	−2.0	−3.6	−2.0
chlorobenzene	3.1	3.1	4.4	3.2
bromobenzene	3.2	3.3	4.3	3.4
p-bromophenol	1.0	1.1	0.8	1.1
2-methylpropan-1-ol	0.7	0.7	1.0	0.3
1-naphthol[b]	2.2	2.1	2.3	1.2
4-methylbenzoic acid	1.8	1.8	1.4	1.4
methylamine	−1.5	−1.2	−0.6	−1.0
4-hexylpyridine	4.8	4.7	5.4	5.0
propanamide	−1.2	−1.1	−2.7	−1.4
m-chlorophenol[b]	0.9	0.9	0.9	1.0
p-chlorophenol	0.9	0.9	0.9	1.0
mean signed error	0.03	0.08	0.44	
mean unsigned error	0.30	0.27	0.73	
root-mean-square error	0.39	0.36	0.95	

[a] Reference 1. [b] Conformationally averaged for present calculations.

in Table 11. We see that the effect of relaxation is considerably larger in SM5.4/A than in their models, and in most cases this remains true even when they use the smaller cavities.

Table 12 presents a comparison of dipole moments. The dipole moments from the present calculations are based on the CM1A partial charges, whereas their dipole moments are not calculated from partial charges at all, but rather from the wave function. Their dipole moments increase about 8−10% (0.1−0.3 D) when the solute is dissolved according to the B-MST/A model, whereas ours change about 10−40%. This increase correlates more closely with the B-MST/1.2A model. We note that calculated solute properties, like solution phase dipole moments, are sensitive to the division of solvation effects into electrostatic and "other" because it is only the former that are treated self-consistently in terms of the solute electronic structure.

Finally, another kind of method for calculating chloroform/water partition coefficients was developed by Reynolds.[1] His calculations are based on the GB/SA model of Still et al.,[14] which is a nonpolarizable molecular mechanics model with fixed charges and a single molecular surface tension. Table 13 compares the performance of our SM5.4-aqueous and SM5.4-chloroform models to that of Reynolds' calculations for the training set he used to parameterize a subsequent quantitative structure−activity relationship for lipophilicity. Two molecules in Table 13 were conformationally averaged using the same method as described above for Table 7. The two molecules are 1-naphthol and *m*-chlorophenol, both of which have two conformations which differ by a 180° rotation of the hydroxyl group. The present polarizable-solute model performs considerably better than the model of ref 1 for the 30 molecules considered by Reynolds. Of these 30 molecules, the first 22 in Table 13 were included in the SM5.4-chloroform parameterization set with a total weight of 482. The remaining eight

molecules not included in the parameterization set are treated with equivalent accuracy.

6. Concluding Remarks

Lack of reliable information about solvation free energies prevents reliable molecular modeling of many interesting processes that take place in solutions, and the work presented here is part of a systematic effort to address this gap in the knowledge base. In this paper we have presented two new parameterizations (one based on AM1 and one based on PM3) of the SM5.4 model[2] for calculating free energies of solvation for organic solutes in chloroform. The parameterizations are based on the general organic parameterizations of the SM5.4 model,[3] and using this approach allows them to encompass a wider variety of functional groups than we could consider if we restricted ourselves only to solutes with experimental data in chloroform. The new parameterizations predict solvation free energies that begin to approach experimental accuracy for the 88 solutes in the SM5.4-chloroform free energy training set.

Combining the parameterizations of the SM5.4 model for water and chloroform now allows one to calculate accurate chloroform/water partition coefficients.

Acknowledgment. D.J.G. gratefully acknowledges a Kodak Graduate School Fellowship. This work was supported in part by the National Science Foundation under Grant No. CHE94-23927 and by the Army High Performance Computing Research Center under the auspices of the Department of the Army, Army Research Laboratory. The content does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

Supporting Information Available: Tables S1 and S2 providing the calculated and experimental free energies of solvation in chloroform for all 88 solutes in the SM5.4-chloroform free energy set and the log $P_{c/w}$ values for the 26 solutes in the SM5.4-chloroform partition coefficient training set (7 pages). Ordering information is given on any current masthead page.

References and Notes

(1) Reynolds, C. H. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 738.

(2) SM5.4/A and SM5.4/P for aqueous solution: Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996,** *100,* 16385.

(3) (a) Communication, SM5.4/A-organic: Giesen, D. J.; Gu, M. Z.; Cramer, C. J.; Truhlar, D. G. *J. Org. Chem.* **1996**, *61*, 8720 (communication). (b) Full paper, S5.4/A and SM5.4/P for general organic solvents and for aromatic hydrocarbon solvents: Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. Full paper in preparation.

(4) Luque, F. J.; Zhang, Y.; Alemán, C.; Bachs, M.; Gao, J.; Orozco, M. *J. Phys. Chem.* **1996**, *100*, 4269.

(5) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc*. **1991**, *113*, 8305.

(6) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213.

(7) Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **1992**, *13*, 1089.

(8) Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aid. Mol. Des.* **1992**, *6*, 629.

(9) Liotard, D. A.; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **1995**, *16*, 442.

(10) Giesen, D. J.; Storer, J. W.; Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1995**, *117*, 1057.

(11) Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1995**, *99*, 7137.

(12) Storer, J. W.; Giesen, D. J.; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; Truhlar, D. G.; Liotard, D. A. In *Structure, Energetics, and Reactivity in Aqueous Solution: Characterization of Chemical and Biological Systems*; Cramer, C. J., Truhlar, D. G., Eds.; American Chemical Society: Washington, DC, 1994; p 24.

(13) Tapia, O.; In *Quantum Theory of Chemical Reactions;* Daudel, R., Pullman, A., Salem, L., Veillard, A., Eds.; Reidel: Dordrecht, 1980; Vol. 2, p 25.

(14) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc*. **1990**, *112*, 6127.

(15) *CRC Handbook of Chemistry and Physics*, 75th ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, FL, 1994.

(16) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. E.; Stewart, J. J. P. *J. Am. Chem. Soc*. **1985**, *107*, 3902. Dewar, M. J. S.; Zoebisch, E. G. *J. Mol. Struct. (THEOCHEM)* **1988**, *180*, 1. Dewar, M. J. S.; Yate-Ching, Y. *Inorg. Chem.* **1990**, *29*, 3881.

(17) Stewart, J. J. P. *J. Comput. Chem*. **1989**, *10*, 221.

(18) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aid. Mol. Des.* **1995**, *9*, 87.

(19) Abraham, M. H. *Chem. Soc. Rev.* **1993**, *1993*, 73.

(20) Masterfile (1994) from MedChem Software, BioByte Corp., P.O. 517, Claremont, CA 91711-0157.

(21) Dallas, A. J.; Carr. P. W. *J. Chem. Soc., Perkin Trans. 2* **1992**, 2155.

(22) Cramer, C. J.; Hawkins, G. D.; Lynch, G. C.; Giesen, D. J.; Rossi, I.; Storer, J. W.; Truhlar, D. G.; Liotard, D. A. *QCPE Bull.* **1995**, *15*, 41.

(23) Cramer, C. J.; Truhlar, D. G. *Rev. Comput. Chem.* **1995**, *6*, 1.

(24) Giesen, D. J.; Chambers, C. C.; Cramer, C. J.; Truhlar, D. G. To be published.

(25) Abraham, M. H.; Whiting, G. S.; Ruchs, R.; Chambers, E. J. *J. Chem. Soc., Perkin Trans. 2* **1990**, 291.