prepared for the proceedings of the Symposium on Rational Drug Design,
214th Amer. Chem. Soc. Natl. Mtg., Las Vegas, Sept. 7–11, 1997,
to be published in ACS Symposium Series

revised February 12, 1998

New Tools for Rational Drug Design

**Gregory D. Hawkins,[1] Jiabo Li,[1] Tianhai (Tony) Zhu,[1]
Candee C. Chambers,[1,3] David J. Giesen,[1,4] Daniel A. Liotard,[2]
Christopher J. Cramer,[1] and Donald G. Truhlar[1]**

[1] Department of Chemistry and Supercomputer Institute, University of Minnesota, Minneapolis, MN 55455
[2] Laboratoire de Physico-Chimie Theorique, Université de Bordeaux, 351 Cours de la Liberation, 33405 Talence Cedex, France

We have developed two new tools for molecular modeling that can be very useful for computer-aided drug design, namely class IV charges and the SM*x* series of solvation models. This contribution overviews the current status of our efforts in these areas, including the CM2 charge model and the SM5 series of solvation models. The solvation models may be used to estimate partition coefficients for phase transfer equilibria of organic solutes between water and 1-octanol, the most widely used mimic of cellular biophases, and also between water and other solvents that have been used for this purpose, e.g., hexadecane and chloroform.

1. Introduction

The partitioning of an organic solute between an aqueous phase (*aq*) and a nonpolar medium (*np*) is critical for many phenomena in biological and medicinal chemistry. In particular this partitioning can be critical for drug delivery, binding, and clearance. Predictions of the relative free energy of organic molecules in aqueous and nonpolar media can be very useful for predicting the bioavailability of potential drugs. Lipid-like nonpolar media are especially important because they mimic cell membranes, and the lipophilic character of organic compounds is one of the most widely used predictors of their bioactivity. The lipid solubility of a molecule correlates with its ability to enter the brain (i.e., pass the blood-brain barrier) or other parts of the central nervous system and is generally believed to have a large influence on pharmacological properties.

[3] Current address: Departments of Physics and Chemistry, Mercyhurst College, Erie, PA 16504
[4] Current address: Eastman Kodak Company, Rochester, NY 14650

The lipophilic character of a molecule is typically measured quantitatively by its partitioning between an organic phase and water. 1-Octanol is the most widely used solvent for mimicking biophases in this respect, and Hansch and Dunn[1] have attempted to rationalize the success of correlations based on 1-octanol by noting that proteins (with their amide groups) and lipid phases (with their ester and phosphate functionalities) both present accessible hydrogen-bonding opportunities to drug molecules, and the OH functional group of octanol can serve as a hydrogen bond acceptor or donor to mimic such effects, while the molecule is large enough to remain overall hydrophobic. The partitioning coefficient P of organic solutes between water and 1-octanol is widely used in property-activity relationships in rational drug design, and a very large amount of work concerned with the measurement and/or prediction of such partition coefficients has been reported. The reader is referred to representative articles for further references.[1–15]

Hexadecane is another important example of a nonpolar solvent because solute–hexadecane interactions, like solute–1-octanol interactions, are recognized as a surrogate for hydrophobic interactions of molecules with lipid bilayers or other cellular material[16–20] or with the nonpolar active site of an enzyme or receptor. In such models, the partition coefficient of a solute between an alkane solvent and water provides some indication of how likely it is to penetrate the bilayer, skin, brain, central nervous system, or other biophase or to bind to the nonpolar site in (or on) the protein. The difference between log P for an amphiphilic solvent like 1-octanol or 1-hexanol and apolar, aprotic inert solvents like straight-chain alkanes or cyclohexane is generally interpreted as a measure of the hydrogen-bond donor capacity of solutes.[21–24] Furthermore this difference has been used in rational drug design because it correlates with brain/blood and cerebrospinal/blood partitioning equilibria.[25]

Another solvent that has been used for similar purposes as 1-octanol and hexadecane is chloroform. Reynolds[26] has discussed the utility of water/chloroform partition coefficients for correlating membrane permeability and bioactivity properties that depend on such permeability.

The ability to understand the solvation of organic solutes in nonpolar media is also important for conformational analysis of bioactive compounds. A recent example of the importance of solvent effects on conformation is the interpretation of octanol/water and heptane/water partition coefficients for the immunosuppressant cyclosporin A in terms of solvent-dependent conformational changes and of the relationship of these changes to solvent-dependent inhibitory activity.[24]

Historically, most attempts to develop predictive models for solvation free energies or partitioning coefficients have involved multivariate quantitative structure-property relationships (QSPRs).[27–33] More recently, methods for including solvent electrostatic effects self consistently in quantum mechanical solute descriptions have advanced vigorously,[34–45] and such models are preferred for making predictions on molecules outside the QSPR training sets or for transition states. Accurate quantitative predictions must include nonelectrostatic effects as well, and we have developed successful models for quantum mechanical self-

consistent electrostatics in both aqueous solutions[46–59] and organic solvents.[55–65]

An especially important aspect of the framework of our model is that only solute atoms are treated explicitly; the solvent is treated as a continuous fluid. There are three kinds of terms in the solvation free energy: long-range electrostatic contributions (labeled ENP, to denote that they include self-consistent solute electronic and nuclear contributions and solute-solvent electric polarization effects), intermediate-range cavity-structural (CS) contributions, and short-range cavity-dispersion (CD) effects. Hydrogen bonding affects all three terms, ENP, CD, and CS.

The functional forms and parameters of the electrostatic model for organic solvents are identical to those for water except that the dielectric constant, ε, of the organic solvent replaces the dielectric constant of water. The electrostatic treatment involves a three-dimensional integration over the free energy density due to electric polarization of the solvent in the regions of space not occupied by the solute,[38,42,51,66,67] and therefore it reflects the solute shape realistically. The solute electronic wave functions and solute internal energies are calculated with semiempirical molecular orbital theory,[68] *ab initio* Hartree-Fock theory,[69] or density functional theory.[70] The competition between solvent polarization and solute distortion is accounted for by placing solvation terms inside the effective one-electron Hamiltonians for the molecular orbitals.[42,71–73]

The atomic partial charges needed for the electrostatic solvation terms may be calculated by conventional Mulliken analysis or by class IV[50,74,75] charge models. The latter capability is a particular strength of our solvation model since these charges, according to previous validation,[74,75] yield remarkably accurate electrostatic properties, and in addition they are very inexpensive to calculate. Accurate atomic partial charges are of great interest for molecular modeling in general and their usefulness extends beyond solvation modeling.[76] Thus we shall review our recent progress in this area as a separate topic.

In addition to electrostatics, our solvation models also include non-electrostatic effects in the first solvation shell. These effects are modeled in terms of solvent-accessible surface areas[77,78] and semiempirical atomic surface tensions.[79] The solvent dependence of our predicted free energies of solvation comes from two sources: (i) The electrostatic term contains the factor $(1 - \varepsilon^{-1})$, where ε is the dielectric constant of the solvent. (ii) The atomic surface tensions are determined separately for water and organic solvents, and in the latter case they depend on one or more of the following solvent descriptors: n, the index of refraction; α and β, Abraham's[80–83] hydrogen bond acidity and basicity parameters (converting our notation to his, α is $\Sigma\alpha_2^H$ and β is $\Sigma\beta_2^H$); γ, the macroscopic surface tension of the solvent; and two descriptors which depend upon the fraction of non-hydrogenic atoms within the solvent which are aromatic carbon or electronegative halogen atoms (we define "electronegative halogen atoms" as F, Cl, and Br since these are the halogen atoms that are more electronegative than carbon[84]). A major advantage of using these parameters is that they are

available for almost all possible solvents. Should one desire to treat an unusual solvent for which a and b are not known, three possibilities present themselves. First, they could be determined by generating the kind of partition coefficient data and fits used originally by Abraham.[80–83] Second, they could be determined by correlating them against other acidity or basicity scales[85,86] that *are* known for the solvent of interest. Third, Murray and Politzer[86] have shown that Abraham's single-site hydrogen-bond acidity and basicity parameters (α_2^H and β_2^H) correlate well with maxima and minima of calculated electrostatic potentials on the molecular surface, and these single-site parameters can be used to estimate $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ in most cases.

Section 2 summarizes the current status of class IV charges. Section 3 presents a level chart of SM5 models. Section 4 summarizes the performance of several SM5 models for free energies of solvation in water, 1-octanol, hexadecane, and chloroform.

2. Class IV charges

Partial atomic charges may be classified as follows:[74]

Class I: non-quantum-mechanical charges, for example, the empirical charges in a molecular mechanics force field;

Class II: charges obtained directly from wave functions without calculating physical observables, for example, charges obtained by Mulliken[87] or Löwdin[88] population analysis;

Class III: charges obtained by fitting to electrostatic potentials or multipole moments computed from wave functions, for example, ChElPG[89] charges;

Class IV: charges mapped from class II or class III charges with semiempirical parameters designed to make the mapped charges better reproduce experimental multipole moments or converged quantum mechanical electrostatic potentials or multipole moments.[50,74,75]

We have presented two models for class IV charges: Charge Model 1[50,74] (CM1) and Charge Model 2[75] (CM2).

In the CM1 model, we computed zero-order charges by Mulliken analysis and mapped them as nonlinear functions of calculated bond orders with 15–19 parameters based on data (experimental dipole moments and calculated electrostatic potentials) for compounds containing H, C, N, O, F, Si, S, Cl, Br, and I. Parameters were determined for AM1[91–93] and PM3[94] semiempirical molecular orbital wave functions. We achieved root-mean-square errors in the dipole moments of 0.27 D for maps based on AM1 wave functions and 0.20 D for maps based on PM3 wave functions.[74]

In the CM2 model we computed zero-order charges by Löwdin analysis and mapped them as quadratic functions of calculated bond orders with 20 parameters based on 198 experimental dipole moments for compounds containing H, C, N, O, F, Si, P, S, Cl, Br, and I. Parameters were determined for AM1, for four different basis sets for ab initio Hartree-Fock wave functions (MIDI!,[95] MIDI!(6D), 6-31G*,[69] and 6-31+G*[69]), and for four combinations of basis set (MIDI!, MIDI!(6D), or 6-31G*) with density

Table 1. Partial atomic charges and dipole moments for β-propiolactone[a]

	HF/MIDI!			BPW91/MIDI!		
	Mulliken	Löwdin	CM2	Mulliken	Löwdin	CM2
partial charges:						
O-1	-0.82	-0.46	-0.36	-0.57	-0.33	-0.31
C-2	1.09	0.58	0.56	0.76	0.41	0.51
O (carbonyl on C-2)	-0.69	-0.41	-0.44	-0.50	-0.31	-0.41
C-3	0.09	0.10	0.10	-0.00	0.04	0.02
H-1,2 on C-3	0.20	0.09	0.07	0.22	0.10	0.10
C-4	-0.58	-0.25	-0.20	-0.51	-0.26	-0.26
H-3,4 on C-4	0.25	0.13	0.10	0.19	0.13	0.13
dipole moment (D)	7.69	4.71	4.31	6.02	3.81	4.21

[a]dipole moment from HF/MIDI! density: 4.18 D; from BPW91 density: 3.41 D; from experiment: 4.18 D

functional (BPW91[96,97] or B3LYP[98–100]). We achieved root-mean-square errors in dipole moments in the range 0.17–0.19 D for HF/6-31G*, B3LYP/MIDI!, BPW91/6-31G*, HF/MIDI!, and BPW91/MIDI!, 0.20–0.21 D for two cases with the MIDI!(6D) basis, 0.25 D with AM1, and 0.41 D with HF/6-31+G*, the latter value reflecting the difficulty of obtaining accurate charges from wave functions with diffuse basis functions. On the average, errors in the dipoles computed as expectation values from the full wave functions were about 1.8 times larger than those computed from the CM2 charges.[75]

As an example of the predictions of the CM2 charge model, consider β-propiolactone. The experimental dipole moment is 4.18 D, and the use of BPW91/MIDI! wave functions yields 3.41 D, whereas the CM2 model based on this same BPW91/MIDI! wave function for β-propiolactone yields 4.21 D. The partial charges on the oxygen atoms differ by as much as 0.25 when obtained by Mulliken analysis of HF/MIDI! and BPW91/MIDI! wave functions and by as much as 0.13 for Löwdin analysis. But the mapped charges from these two quite different wave functions agree within 0.05. Full results are given in Table 1.

3. Summary of SM5 models

Aqueous/nonpolar partitioning is usually quantified by the partition coefficient P or its, logarithm ("log P"), where

$$P = \frac{[solute]_{np}}{[solute]_{aq}} \tag{1}$$

Another (equivalent) definition of P is

$$\log P = \Delta\Delta G_S^0/(-2.303RT) \quad (2)$$

where

$$\Delta\Delta G_S^0 = \Delta G_S^0(np) - \Delta G_S^0(aq), \quad (3)$$

$G_S^0(solv)$ is the standard-state free energy of solvation of the solute in solvent *solv*, R is the gas constant, and T is temperature.

The standard-state free energy of solvation in water is written as

$$\Delta G_S^0 = \Delta G_{ENP} + \sum_\alpha G_{CDS,\alpha} \quad (4)$$

where α denotes one of the atoms of the solute and

$$G_{CDS,\alpha} = A_\alpha \sum_i f_{\alpha i}(R)\tilde{\sigma}_{\alpha i} \quad (5)$$

where $f_{\alpha i}$ is a function of the geometry R of the solute (actually it depends only on selected bond distances, and it has no dependence on bond angles or dihedral angles) and $\tilde{\sigma}_{\alpha i}$ is a surface tension coefficient. The standard-state free energy of solvation in an organic solvent has the same form as for water except that $\tilde{\sigma}_{\alpha i}$ is not a constant but rather depends on solvent descriptors. The solvent descriptors are generally n, α, β, and γ. In some cases (SM5.4 parameterizations) special parameters are used for chloroform, benzene, and toluene; in other cases (SM5.42R, SM5.2R, and SM5.0R parameterizations) two special solvent descriptors are added to the four mentioned in the previous sentence, in particular descriptors computed from the fraction of nonhydrogenic solvent atoms that are aromatic carbons or electronegative halogens. Some $\tilde{\sigma}_{\alpha i}$ values are independent of α and have $f_{\alpha i} = 1$; these are sometimes called the CS terms. The other terms are sometimes called CD terms; however, one should be cautious about physical interpretations of the individual terms.

The actual parameterizaiton is carried out as follows: First the nonlinear parameters are fixed based on a variety of considerations, including trends over solutes and solvents for solvation free energies of neutrals and ions. Then the surface tension coefficients are fit to a large set of data taken chiefly from the tabulation of Cabani *et al.*[101] for ΔG_S^o of neutrals in water and mostly computed from log P values from the MedChem data base[102] for organic solvents.

In the present paper we consider solutes containing H, C, N, O, F, S, Cl, Br, and I. (Some, but not all, models are also parameterized for solutes containing P, but P-containing solutes are not discussed in this chapter.) As an example of the size of the training set, we consider the training set used for solutes with H, C, N, O, F, S, Cl, Br, and I in the SM5.2R model. This training set has data for 43 ions and 248 neutrals in water. It also has 1836 data points for 227 neutrals in 90 organic solvents. The SM5.2R

Table 2. Mean Unsigned Error (kcal/mol) in the Aqueous Solvation Free Energies Predicted by Selected SM*x* Solvation Models.

		SM5.4/		SM5.2R/				SM5.0R
Solute Class	Data Points	AM1	PM3	MNDO(d)	MNDO	AM1	PM3	
Unbranched Alkanes	8	0.6	0.6	0.7	0.7	0.6	0.6	0.5
Branched Alkanes	5	0.7	0.7	0.5	0.5	0.5	0.4	0.3
Cycloalkanes	5	0.2	0.1	0.4	0.4	0.4	0.4	0.5
Alkenes	9	0.5	0.3	0.3	0.3	0.2	0.2	0.2
Alkynes	5	0.2	0.2	0.1	0.1	0.2	0.2	0.1
Arenes	8	0.2	0.2	0.2	0.2	0.2	0.2	0.5
Alcohols	16	0.5	0.4	0.2	0.2	0.2	0.2	0.3
Ethers	12	0.8	0.9	0.6	0.6	0.5	0.6	0.6
Aldehydes	6	0.3	0.4	0.3	0.3	0.3	0.3	0.3
Ketones	12	0.4	0.4	0.5	0.5	0.3	0.4	0.4
Carboxylic Acids	5	0.8	0.8	0.4	0.4	0.4	0.4	0.5
Esters	13	0.5	0.5	0.3	0.3	0.3	0.3	0.3
Bifunctional CHO	5	0.4	0.4	0.5	0.5	0.4	0.4	0.5
Water, Dihydrogen	2	1.6	1.2	0.0	0.0	0.0	0.0	0.9
Aliphatic Amines	15	0.8	0.8	0.5	0.5	0.6	0.5	0.5
Aromatic Amines	10	0.7	0.7	0.8	0.8	0.6	0.7	1.0
Nitriles	4	0.5	0.5	0.5	0.5	0.4	0.3	0.7
Nitrohydrocarbons	6	0.5	0.1	0.1	0.1	0.5	0.4	0.4
Amides & Ureas	4	2.6	1.2	1.4	1.4	1.1	1.1	2.2
Bifunctional HCN and HCNO	5	0.9	1.1	0.9	0.9	0.8	0.9	1.1
Ammonia & Hydrazine	2	2.8	3.1	0.2	0.2	0.4	0.2	1.1
Thiols	4	0.3	0.2	0.6	0.6	0.5	0.6	0.3
Sulfides	6	0.6	0.5	1.2	1.0	1.1	1.0	0.5
Disulfides	2	0.2	0.2	0.1	0.1	0.0	0.0	0.1
Fluorinated Hydrocarbons	6	0.6	0.4	0.7	0.7	0.4	0.5	1.1
Chloroalkanes	13	0.3	0.3	0.3	0.4	0.3	0.8	0.4
Chloroalkenes	5	0.7	0.5	0.6	0.4	0.7	1.0	1.0
Chloroarenes	8	0.2	0.3	0.8	0.9	0.3	0.5	0.3
Brominated Hydrocarbons	14	0.3	0.2	0.2	0.2	0.4	0.2	0.4
Iodinated Hydrocarbons	8	0.3	0.2	0.3	0.3	0.3	0.6	0.3
Other Halo Compounds	25	0.6	0.8	0.7	0.7	0.6	1.0	0.8
All solutes:	248	0.6	0.5	0.5	0.5	0.4	0.5	0.5

Table 3. Mean Unsigned Error (kcal/mol) in the 1-Octanol Solvation Free Energies Predicted by Selected SM*x* Solvation Models.

		SM5.4/		SM5.2R/				SM5.0R
Solute Class	Data Points	AM1	PM3	MNDO(d)	MNDO	AM1	PM3	
Unbranched Alkanes	8	0.4	0.3	0.1	0.1	0.4	0.3	0.1
Branched Alkanes	2	0.1	0.1	0.1	0.1	0.2	0.2	0.1
Cycloalkanes	4	0.6	0.6	0.4	0.4	0.3	0.3	0.4
Alkenes	6	0.6	0.4	0.2	0.2	0.6	0.4	0.2
Alkynes	4	0.3	0.3	0.2	0.2	0.2	0.2	0.1
Arenes	8	0.2	0.2	0.3	0.3	0.3	0.3	0.3
Alcohols	16	0.2	0.3	0.5	0.5	0.5	0.5	0.4
Ethers	11	0.6	0.5	0.5	0.5	0.6	0.5	0.5
Aldehydes	4	0.5	0.5	0.4	0.4	0.4	0.4	0.5
Ketones	10	1.0	1.0	0.8	0.8	0.8	0.8	0.9
Carboxylic Acids	5	0.7	0.7	0.3	0.3	0.4	0.4	0.1
Esters	9	1.2	1.1	0.3	0.3	0.4	0.3	0.6
Bifunctional CHO	4	1.1	1.0	0.8	0.8	0.9	0.8	0.6
Water, Dihydrogen	2	1.2	1.1	0.7	0.7	0.8	0.7	0.5
Aliphatic Amines	9	0.6	0.5	0.4	0.4	0.6	0.5	0.4
Aromatic Amines	7	0.8	0.5	0.5	0.5	0.5	0.5	0.6
Nitriles	4	0.7	0.6	0.2	0.2	0.4	0.2	0.5
Nitrohydrocarbons	6	0.7	0.1	0.2	0.2	0.7	0.4	0.1
Amides & Ureas	1	1.7	0.2	2.6	2.6	3.1	2.1	2.5
Bifunctional HCN and HCNO	3	2.0	1.6	0.7	0.7	0.7	0.6	0.9
Hydrazine	1	2.0	3.3	1.7	1.7	1.7	1.7	1.8
Thiols	2	0.3	0.2	0.4	0.4	0.5	0.4	0.3
Sulfides	3	0.8	0.7	1.0	0.6	0.8	0.6	0.3
Disulfides	1	0.0	0.0	0.1	0.2	0.2	0.3	0.3
Fluorinated Hydrocarbons	2	0.4	0.2	1.2	1.2	0.6	0.3	0.5
Chloroalkanes	7	0.3	0.3	0.3	0.6	0.4	0.5	0.5
Chloroalkenes	3	0.5	0.4	0.8	0.5	0.8	1.1	1.0
Chloroarenes	6	0.7	0.5	0.7	0.9	0.3	0.3	0.3
Brominated Hydrocarbons	12	0.3	0.4	0.3	0.4	0.3	0.3	0.2
Iodinated Hydrocarbons	5	0.2	0.2	0.5	0.6	0.7	0.5	0.6
Other Halo Compounds	15	0.7	0.7	0.7	0.7	0.6	0.9	0.8
All solutes:	180	0.6	0.5	0.5	0.5	0.5	0.5	0.5

Table 4. Mean Unsigned Error (kcal/mol) in the Hexadecane Solvation Free Energies Predicted by Selected SM*x* Solvation Models.

		SM5.4/		SM5.2R/				SM5.0R
Solute Class	Data Points	AM1	PM3	MNDO(d)	MNDO	AM1	PM3	
Unbranched Alkanes	9	0.5	0.5	0.4	0.4	0.2	0.3	0.4
Branched Alkanes	5	0.5	0.5	0.6	0.6	0.4	0.5	0.7
Cycloalkanes	4	1.2	1.2	0.4	0.4	0.2	0.3	0.5
Alkenes	6	0.4	0.2	0.2	0.2	0.4	0.3	0.2
Alkynes	5	0.2	0.1	0.1	0.1	0.1	0.2	0.1
Arenes	9	0.4	0.3	0.4	0.4	0.5	0.5	0.5
Alcohols	17	0.2	0.2	0.2	0.2	0.2	0.3	0.3
Ethers	9	0.3	0.3	0.3	0.3	0.3	0.3	0.2
Aldehydes	7	0.1	0.2	0.3	0.3	0.2	0.3	0.4
Ketones	12	0.3	0.3	0.2	0.2	0.2	0.2	0.3
Carboxylic Acids	5	0.2	0.2	0.4	0.4	0.5	0.4	0.5
Esters	13	0.5	0.4	0.3	0.3	0.3	0.3	0.3
Bifunctional CHO	1	0.2	0.3	0.1	0.1	0.2	0.0	0.4
Water, Dihydrogen	2	0.5	0.5	0.1	0.1	0.1	0.1	0.1
Aliphatic Amines	8	0.3	0.3	0.1	0.1	0.1	0.1	0.1
Aromatic Amines	9	0.4	0.3	0.2	0.2	0.2	0.2	0.2
Nitriles	4	0.2	0.1	0.2	0.2	0.2	0.1	0.2
Nitrohydrocarbons	6	0.5	0.2	0.2	0.2	0.6	0.3	0.1
Amides & Ureas	1	1.2	1.0	0.1	0.1	0.0	0.2	0.4
Bifunctional HCN and HCNO	0	...	...	...	...	...	...	...
Ammonia	1	0.7	0.5	0.4	0.4	0.5	0.3	0.4
Thiols	3	0.5	0.5	0.5	0.5	0.4	0.4	0.4
Sulfides	5	0.2	0.2	0.8	0.6	0.7	0.6	0.5
Disulfides	2	0.1	0.1	0.3	0.3	0.3	0.3	0.3
Fluorinated Hydrocarbons	4	0.7	0.6	0.4	0.4	0.4	0.4	0.6
Chloroalkanes	7	0.2	0.1	0.2	0.2	0.3	0.2	0.2
Chloroalkenes	4	0.4	0.3	0.6	0.5	0.7	0.7	0.5
Chloroarenes	3	0.3	0.2	0.2	0.2	0.4	0.2	0.3
Brominated Hydrocarbons	12	0.1	0.1	0.2	0.2	0.2	0.2	0.4
Iodinated Hydrocarbons	8	0.1	0.2	0.3	0.3	0.3	0.2	0.2
Other Halo Compounds	8	0.4	0.4	0.6	0.7	0.6	0.4	0.3
All solutes:	189	0.3	0.3	0.3	0.3	0.3	0.3	0.3

Table 5. Mean Unsigned Error (kcal/mol) in the Chloroform Solvation Free Energies Predicted by Selected SM*x* Solvation Models.

Solute Class	Data Points	SM5.4/ AM1	SM5.4/ PM3	SM5.2R/ MNDO(d)	SM5.2R/ MNDO	SM5.2R/ AM1	SM5.2R/ PM3	SM5.0R
Unbranched Alkanes	1	0.2	0.3	0.2	0.2	0.1	0.1	0.4
Branched Alkanes	0	...	...	...	...	...	...	...
Cycloalkanes	1	1.1	1.2	0.0	0.0	0.2	0.0	0.1
Alkenes	0	...	...	...	...	...	...	...
Alkynes	0	...	...	...	...	...	...	...
Arenes	6	0.1	0.2	0.2	0.2	0.7	0.5	0.2
Alcohols	14	0.4	0.3	0.3	0.3	0.3	0.3	0.4
Ethers	6	0.4	0.4	0.6	0.6	0.5	0.5	0.5
Aldehydes	3	0.6	0.5	0.9	0.9	0.9	1.0	0.8
Ketones	3	0.2	0.2	0.4	0.4	0.3	0.3	0.4
Carboxylic Acids	5	0.2	0.2	0.1	0.1	0.2	0.2	0.4
Esters	9	0.2	0.2	0.5	0.5	0.4	0.5	0.7
Bifunctional CHO	2	0.8	0.8	0.7	0.7	0.8	0.8	0.8
Water, Dihydrogen	1	1.5	1.2	0.3	0.3	0.1	0.2	0.7
Aliphatic Amines	8	0.4	0.2	0.5	0.5	0.6	0.6	0.5
Aromatic Amines	8	0.3	0.4	0.8	0.8	0.4	0.8	1.0
Nitriles	2	0.2	0.2	0.7	0.7	0.7	0.4	1.0
Nitrohydrocarbons	2	0.1	0.2	0.2	0.2	0.3	0.2	0.5
Amides & Ureas	2	1.5	0.5	2.3	2.3	2.2	2.1	2.5
Bifunctional HCN and HCNO	3	3.0	2.8	1.3	1.3	1.3	1.3	1.0
Ammonia & Hydrazine	2	2.4	2.9	0.8	0.8	0.6	0.8	0.7
Thiols	1	0.7	0.7	0.7	0.6	0.5	0.6	0.8
Sulfides	4	1.0	1.0	1.2	1.2	1.2	1.2	1.2
Disulfides	0	...	...	...	...	...	...	...
Fluorinated Hydrocarbons	1	0.3	0.1	1.0	1.0	0.8	0.8	0.3
Chloroalkanes	0	...	...	...	...	...	...	...
Chloroalkenes	0	...	...	...	....	...	...	...
Chloroarenes	2	0.2	0.2	0.6	0.7	0.6	0.6	0.3
Brominated Hydrocarbons	1	0.0	0.2	0.7	0.7	1.0	0.8	0.4
Iodinated Hydrocarbons	1	0.1	0.2	0.9	1.0	1.4	0.5	0.9
Other Halo Compounds	4	0.6	0.5	0.6	0.5	0.6	0.7	0.7
All solutes:	92	0.5	0.5	0.6	0.6	0.6	0.6	0.6

Table 6. Free Energy of Solvation and Partition Coefficient Results for 1,2 Ethanediol.

Model	ΔG_{ENP}	G_{CDS}	ΔG_S^o	log $P_{org/water}$	
				theory	experiment
		water			
SM5.4/AM1	-6.5	-2.3	-8.8		
SM5.4/PM3	-6.3	-2.9	-9.2		
SM5.2R/MNDO(d)	-1.9	-7.0	-8.9		
SM5.2R/MNDO	-1.9	-7.0	-8.9		
SM5.2R/AM1	-2.8	-6.4	-9.2		
SM5.2R/PM3	-2.2	-7.1	-9.3		
SM5.0R			-8.7		
		1-octanol			
SM5.4/AM1	-5.9	-1.4	-7.2	-1.1	-1.4
SM5.4/PM3	-5.8	-1.9	-7.7	-1.1	
SM5.2R/MNDO(d)	-1.7	-6.3	-8.0	-0.7	
SM5.2R/MNDO	-1.7	-6.3	-8.0	-0.7	
SM5.2R/AM1	-2.5	-5.6	-8.2	-0.7	
SM5.2R/PM3	-2.0	-6.3	-8.2	-0.8	
SM5.0R			-8.1	-0.4	
		hexadecane			
SM5.4/AM1	-3.2	0.1	-3.1	-4.2	-4.8
SM5.4/PM3	-3.2	-0.2	-3.4	-4.3	
SM5.2R/MNDO(d)	-0.9	-2.6	-3.5	-4.0	
SM5.2R/MNDO	-0.9	-2.6	-3.5	-4.0	
SM5.2R/AM1	-1.3	-2.2	-3.5	-4.2	
SM5.2R/PM3	-1.1	-2.6	-3.7	-4.1	
SM5.0R			-3.8	-3.6	
		chloroform			
SM5.4/AM1	-5.0	-0.2	-5.2	-2.6	-2.4
SM5.4/PM3	-5.0	-0.5	-5.5	-2.7	
SM5.2R/MNDO(d)	-1.5	-3.7	-5.2	-2.7	
SM5.2R/MNDO	-1.5	-3.7	-5.2	-2.7	
SM5.2R/AM1	-2.1	-3.3	-5.4	-2.8	
SM5.2R/PM3	-1.7	-3.7	-5.4	-2.9	
SM5.0R			-5.1	-2.6	

parameterizations have 46 surface tension coefficients for organic solvents and 25 for water.

Tables 2–5 show the mean unsigned deviations in standard-state free energies of solvation for various classes of solutes in water and the three organic solvents singled out in the introduction. In each table we show the application of several models to the same set of data, namely our latest and largest training set, excluding phosphorus-containing compounds, except that in Tables 2, 3, and 5, the SM5.4/PM3 results are based on one less data point because hydrazine is excluded when PM3 is used to optimize geometries. Tables 2–5 show that we have uniformly small

Table 7. Free Energy of Solvation and Partition Coefficient Results for Thioanisole.

Model	ΔG_{ENP}	G_{CDS}	ΔG_S^o	$\log P_{org/water}$	
				theory	experiment
		water			
SM5.4/AM1	-3.9	0.7	-3.3		
SM5.4/PM3	-3.0	-0.3	-3.3		
SM5.2R/MNDO(d)	-1.0	-1.7	-2.7		
SM5.2R/MNDO	-1.0	-1.9	-2.9		
SM5.2R/AM1	-3.8	0.8	-3.0		
SM5.2R/PM3	-2.7	-0.3	-3.0		
SM5.0R			-3.4		
		1-octanol			
SM5.4/AM1	-3.5	-3.9	-7.4	3.0	2.7
SM5.4/PM3	-2.7	-4.6	-7.3	2.9	
SM5.2R/MNDO(d)	-0.9	-5.2	-6.1	2.5	
SM5.2R/MNDO	-0.9	-5.5	-6.4	2.6	
SM5.2R/AM1	-3.4	-3.0	-6.4	2.5	
SM5.2R/PM3	-2.4	-4.1	-6.5	2.6	
SM5.0R			-6.4	2.2	
		hexadecane			
SM5.4/AM1	-1.8	-5.1	-6.9	2.6	
SM5.4/PM3	-1.4	-5.7	-7.0	2.7	
SM5.2R/MNDO(d)	-0.5	-5.6	-6.1	2.5	
SM5.2R/MNDO	-0.5	-5.8	-6.3	2.5	
SM5.2R/AM1	-1.8	-4.3	-6.1	2.3	
SM5.2R/PM3	-1.2	-5.0	-6.3	2.4	
SM5.0R			-6.3	2.1	
		chloroform			
SM5.4/AM1	-2.9	-4.8	-7.8	3.3	2.4
SM5.4/PM3	-2.3	-5.7	-8.0	3.4	
SM5.2R/MNDO(d)	-0.8	-6.5	-7.3	3.3	
SM5.2R/MNDO	-0.8	-6.7	-7.4	3.3	
SM5.2R/AM1	-2.9	-4.8	-7.7	3.4	
SM5.2R/PM3	-2.0	-5.7	-7.7	3.4	
SM5.0R			-7.4	2.9	

mean errors. Notice that some solute classes are not well represented in the data sets for specific solvents, and in fact some solute classes are not represented at all in some solvents. The SM5 solvation models are able to treat such cases because all the data for free energies of solvation in organic solvents are fit simultaneously, and the number of solvent descriptors is much smaller than the total number (90) of organic solvents. We believe in this way we have captured all the major physical effects.

Tables 6-8 were included to examine a couple of individual examples, namely, 1,2-ethanediol, thioanisole, and *p*-dichlorobenzene. These tables show the partitioning of the predicted solvation free energy

Table 8. Free Energy of Solvation and Partition Coefficient Results for *p*-Dichlorobenzene.

Model	ΔG_{ENP}	G_{CDS}	ΔG_S^o	log $P_{org/water}$	
				theory	experiment
		water			
SM5.4/AM1	-2.1	1.1	-1.0		
SM5.4/PM3	-1.3	0.2	-1.2		
SM5.2R/MNDO(d)	-1.0	-0.6	-1.6		
SM5.2R/MNDO	-1.4	-0.3	-1.7		
SM5.2R/AM1	-2.5	1.5	-1.0		
SM5.2R/PM3	-1.8	0.2	-1.6		
SM5.0R			-1.0		
		1-octanol			
SM5.4/AM1	-1.9	-3.7	-5.6	3.3	3.4
SM5.4/PM3	-1.2	-4.5	-5.7	3.3	
SM5.2R/MNDO(d)	-0.9	-5.0	-5.9	3.2	
SM5.2R/MNDO	-1.3	-4.8	-6.1	3.2	
SM5.2R/AM1	-2.2	-3.2	-5.5	3.3	
SM5.2R/PM3	-1.6	-4.2	-5.8	3.1	
SM5.0R			-5.6	3.4	
		hexadecane			
SM5.4/AM1	-1.0	-4.7	-5.7	3.4	3.7
SM5.4/PM3	-0.6	-5.1	-5.7	3.3	
SM5.2R/MNDO(d)	-0.5	-5.4	-5.9	3.1	
SM5.2R/MNDO	-0.7	-5.2	-5.9	3.1	
SM5.2R/AM1	-1.2	-4.3	-5.5	3.3	
SM5.2R/PM3	-0.9	-4.9	-5.8	3.1	
SM5.0R			-5.8	3.5	
		chloroform			
SM5.4/AM1	-1.6	-4.3	-5.9	3.6	3.9
SM5.4/PM3	-1.0	-5.1	-6.1	3.6	
SM5.2R/MNDO(d)	-0.8	-6.1	-6.8	3.8	
SM5.2R/MNDO	-1.1	-5.9	-6.9	3.8	
SM5.2R/AM1	-1.9	-4.7	-6.6	4.1	
SM5.2R/PM3	-1.4	-5.5	-6.9	3.9	
SM5.0R			-6.6	4.2	

between the electrostatic (ΔG_{ENP}) and non-electrostatic (G_{CDS}) components as well as the logarithm of the partition coefficient between selected organic solvents and water. The SM5.4 models utilize class IV charges and are designed to optimize solute geometry in the presence of the solvent reaction field. Note that the absolute value of the ΔG_{ENP} term is generally much larger for the SM5.4 parameterizations than for the SM5.2R models which incorporate the less-accurate class II charges. In general, class IV charges lead to greater charge separation within a solute molecule, which results in a larger $|\Delta G_{ENP}|$. Our method of parameterizing the

Table 9. Absolute Value (kcal/mol) of the ΔG_{ENP} and G_{CDS} Terms in Selected SMx Models[a]

Model	$\langle\lvert\Delta G_{ENP}\rvert\rangle$	$\langle\lvert G_{CDS}\rvert\rangle$
	water	
SM5.4/AM1	4.5	1.5
SM5.4/PM3	3.7	1.3
SM5.2R/MNDO(d)	1.7	2.3
SM5.2R/MNDO	1.7	2.3
SM5.2R/AM1	3.2	1.9
SM5.2R/PM3	2.4	2.0
SM5.0R	0.0	3.7
	1-octanol	
SM5.4/AM1	4.0	2.0
SM5.4/PM3	3.4	2.5
SM5.2R/MNDO(d)	1.5	4.2
SM5.2R/MNDO	1.5	4.2
SM5.2R/AM1	2.8	2.9
SM5.2R/PM3	2.1	3.6
SM5.0R	0.0	5.7
	hexadecane	
SM5.4/AM1	2.1	2.4
SM5.4/PM3	1.8	2.8
SM5.2R/MNDO(d)	0.8	3.6
SM5.2R/MNDO	0.8	3.6
SM5.2R/AM1	1.5	2.9
SM5.2R/PM3	1.1	3.3
SM5.0R	0.0	4.5
	chloroform	
SM5.4/AM1	3.4	2.6
SM5.4/PM3	2.9	3.1
SM5.2R/MNDO(d)	1.3	4.4
SM5.2R/MNDO	1.3	4.4
SM5.2R/AM1	2.4	3.5
SM5.2R/PM3	1.8	4.0
SM5.0R	0.0	5.5

[a]Reported averages are for 67 organic solutes for which experimental solvation free energies are available in water, hexadecane, octanol, and chloroform. (A total of 268 data points.)

remaining non-electrostatic term (G_{CDS}) to the experimental solvation free energies allows the diminished electrostatics obtained with the less expensive SM5.2R models to be compensated for by the G_{CDS} term, resulting in fairly accurate absolute solvation free energies and partition coefficients. This approach was taken to the limit in the SM5.0R model

Table 10. Mean Unsigned Errors in Predicted Solvation Free Energies, Organic/Water Partition Coefficients, and Free Energy of Transfer for Selected SMx methods[a]

Model	MUE ΔG_S^o	MUE $\Delta\Delta G^o_{org/water}$	MUE $\log P_{org/water}$
	water		
SM5.4/AM1	0.54		
SM5.4/PM3	0.47		
SM5.2R/MNDO(d)	0.45		
SM5.2R/MNDO	0.44		
SM5.2R/AM1	0.43		
SM5.2R/PM3	0.39		
SM5.0R	0.46		
	1-octanol		
SM5.4/AM1	0.56	0.63	0.46
SM5.4/PM3	0.51	0.54	0.40
SM5.2R/MNDO(d)	0.43	0.38	0.28
SM5.2R/MNDO	0.42	0.38	0.28
SM5.2R/AM1	0.47	0.42	0.31
SM5.2R/PM3	0.41	0.37	0.27
SM5.0R	0.40	0.40	0.38
	hexadecane		
SM5.4/AM1	0.29	0.49	0.36
SM5.4/PM3	0.29	0.49	0.36
SM5.2R/MNDO(d)	0.27	0.45	0.33
SM5.2R/MNDO	0.26	0.45	0.33
SM5.2R/AM1	0.27	0.51	0.38
SM5.2R/PM3	0.26	0.47	0.35
SM5.0R	0.30	0.30	0.43
	chloroform		
SM5.4/AM1	0.32	0.46	0.33
SM5.4/PM3	0.27	0.39	0.28
SM5.2R/MNDO(d)	0.45	0.45	0.33
SM5.2R/MNDO	0.46	0.45	0.33
SM5.2R/AM1	0.47	0.45	0.33
SM5.2R/PM3	0.44	0.41	0.30
SM5.0R	0.50	0.50	0.41

[a]Reported averages are for 67 organic solutes from our training set for which experimental solvation free energies are available in water, hexadecane, octanol, and chloroform. (A total of 268 data points.)

which contains no explicit electrostatic or SCF treatment. Although it is likely that the very inexpensive SM5.0R approach will have difficulty predicting solvation free energies in cases where the charge distribution within a given solute differs significantly from the implicit distributions parameterized into the model, SM5.0's predictions for the example molecules and overall training set are reasonably similar to those predicted by SM5 models with more rigorous electrostatic treatments.

compounds through cells, and drug activity. We hope the models will be useful for a variety of purposes in the humanistic endeavor of designing better drugs.

5. Acknowledgments

This work was supported in part by the National Science Foundation.

References

1. Hansch, C.; Dunn, W. J. III *J. Pharm. Sci.* **1972**, *61*, 1.
2. Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.* **1964**, *86*, 5175.
3. Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525.
4. Leo, A.; Jow, P. Y. C.; Silipino, C.; Hansch, C. *J. Med. Chem.* **1975**, *18*, 865.
5. Franke, R.; Dove, S.; Kuhne, B. *J. Med. Chem.* **1977**, **14**, 363.
6. Funisaki, N.; Hada, S.; Neya, S. *J. Phys. Chem.* **1985**, *89*, 3046.
7. Camilleri, P.; Watts, S. A.; Boraston, J. A. *J. Chem. Soc. Perkin Trans. II* **1988**, *188*, 1699.
8. Lewis, D. F. V. *J. Comput. Chem.* **1989**, *10*, 145.
9. Dunn, W. J. III: Nagy, P. I.; Collantes, E. R. *J. Am. Chem. Soc.* **1991**, *113*, 7898.
10. Suzuki, T. *J. Comput.-Aided Molec. Des* **1991**, *5*, 149.
11. Richards, N. G. J.; Williams, P. B.; Tute, M. S. *Int. J. Quantum Chem. Quantum. Biol. Symp.* **1991**, *18*, 299.
12. Dallas, A. J.; Carr, P. W. *J. Chem. Soc. Perkin Trans. 2*, **1992**, 2155.
13. Leo, A. *J. Chem. Rev.* **1993**, *93*, 1281.
14. Viswanadhan, V. N.; Reddy, M. R.; Bacquet, R. J.; Erion, M. D. *J. Comp. Chem.* **1993**, *14*, 1019.
15. Slater, B.; McCormack, A.; Avdeef, A.; Comer, J. E. A. *J. Pharm. Sci.* **1994**, *83*, 1280.
16. Collander, R. *Physiol. Plant.* **1954**, *7*, 420
17. Mayer, S.; Maickel, R. P.; Brodie, B. B. *J. Pharmacol. Exptl. Therap.* **1959**, *127*, 205.
18. Tute, M. S. *Adv. Drug Res.* **1971**, *6*, 1.
19. Venable, R. M.; Zhang, Y.; Hardy, B. J.; Pastor, R. W. *Science* **1993**, *262*, 223.
20. Solé-Violan, L.; Devallez, B.; Postel, M.; Riess, J. G. *New J. Chem.* **1993**, *17*, 581.
21. Sieler, P. *Eur. J. Med. Chem.* **1974**, *9*, 473.
22. Abraham, M. H.; Chadha, H. S.; Mitchell, R. C. *J. Pharm. Sci.* **1994**, *83*, 1257.
23. Tayar, N. E.; Tsai, R.-S.; Testa, B.; Carrupt, P.-A.; Leo, A. *J. Pharm. Sci.* **1991**, *80*, 590.
24. Tayar, N. A.; Mark, A. G.; Vallat, P.; Brunne, R. M.; Testa, B.; van Gunsteren, W. F. *J. Med. Chem.* **1993**, *36*, 3757.
25. Young, R.-C.; Mitchell, R. C.; Brown, T. H.; Ganellin, C. R.; Griffiths, R.; Jones, M.; Rana, K. K.; Saunders, D.; Smith, I. R.; Sore, N. E.; and Wilks, T. J. *J. Med. Chem.* **1988**, *31*, 656.

26. Reynolds, C. H. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 738.
27. Hine, J.; Mookerjee, P. K. *J. Org. Chem.* **1975**, *40*, 287.
28. Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *50*, 563.
29. Abraham, M. H.; Whiting, G. S.; Fuchs, R.; Chambers, E. J. *J. Chem. Soc. Perkin Trans. II* **1990**, 291.
30. Famini, G. R.; Penski, C. A.; Wilson, L. Y. *J. Phys. Org. Chem.* **1992**, *5*, 395.
31. Leo, A. *J. Chem. Rev.* **1993**, *93*, 1281.
32. Cramer, C. J.; Famini, G. R.; Lowrey, A. H. *Acc. Chem. Res.* **1993**, *26*, 599.
33. Bodor, N.; Huang, M.-J.; Harget, A. *J. Mol. Struct. (Theochem)* **1994**, *390*, 259.
34. Tapia, O. *J. Math. Chem.* **1992**, *10*, 139.
35. Olivares del Valle, F. J.; Aguilar, M. A. *J. Mol. Struct. (Theochem)* **1993**, *280*, 25.
36. Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.
37. Cramer, C. J.; Truhlar, D. G. In *Quantitative Treatments of Solute/Solvent Interactions*, Politzer, P., Murray, J. S., Eds.; Elsevier: Amsterdam, 1994; p. 9.
38. Cramer, C. J.; Truhlar, D. G. *Rev. Comp. Chem.* **1995**, *6*, 1.
39. Rivail, J.-L.; Rinaldi, D. In *Computational Chemistry, Review of Current Trends, Vol. 1*; Leszczynski, J., Ed.; World Scientific: Singapore, 1996; p. 139.
40. Orozco, M., Alhambra, C.; Barril, X.; Lopez, J. M.; Busquets, M. A.; Luque, F. J. *J. Mol. Modeling* **1996**, *2*, 1.
41. Gao, J. *Rev. Comp. Chem.* **1996**, *7*, 119.
42. Cramer, C. J.; Truhlar, D. G. In *Solvent Effects and Chemical Reactivity*; Tapia, O., Bertrán, J., Eds.; Kluwer: Dordrecht, 1996; p. 1.
43. Takahashi, O.; Sawahata, H.; Ogama, Y.; Kikuchi, O. *J. Mol. Struct. (Theochem)* **1997**, *393*, 141.
44. Baldridge, K.; Klamt, A. *J. Chem. Phys.* **1997**, *106*, 6622.
45. Zhan, C.-G.; Bentley, J.; Chipman, D. M. *J. Chem. Phys.* **1998**, *108*, 177.
46. Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8305.
47. Cramer, C. J.; Truhlar, D. G. *Science* (Washington, DC) **1992**, *256*, 213.
48. Cramer, C. J.; Truhlar, D. G. *J. Comp. Chem.* **1992**, *13*, 1089.
49. Cramer, C. J.; Truhlar, D. G. *J. Comp.-Aided Mol. Des.* **1992**, *6*, 629.
50. Storer, J. W.; Giesen, D. J.; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; Truhlar, D. G. In *Structure and Reactivity in Aqueous Solution*, Cramer, C. J., Truhlar, D. G., Ed.; American Chemical Society: Washington, DC, 1994; p. 24.
51. Liotard, D. A.; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; Truhlar, D. G. *J. Comp. Chem.* **1995**, *16*, 422.
52. Barrows, S. E.; Dulles, F. J.; Cramer, C. J.; Truhlar, D. G.; French, A. D. *Carb. Res.* **1995**, *276*, 219.
53. Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 16385.
54. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824.

55. Giesen, D. J.; Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. In *Computational Thermochemistry: Prediction and Estimation of Molecular Thermodynamics*; Irikura, K., Frurip, D. J., Eds.; American Chemical Society Symposium Series 677: Washington D.C., in press.
56. Chambers, C. C.; Giesen, D. J.; Hawkins, G. D.; Vaes, W. H. J.; Cramer, C. J.; Truhlar, D. G. In *Computer-Assisted Drug Design*; Truhlar, D. G.; Howe, W. J.; Blaney, J.; Hopfinger, A. J.; Dammkoehler, R. A., Eds.; Springer: New York, in press.
57. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* to be published (SM5.2R model).
58. Zhu, T.; Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. to be published (SM5.42R/DFT parameterizations).
59. Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* to be published (first SM5.42R/HF paramterization).
60. Giesen, D. J.; Storer, J. W.; Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1995**, *117*, 1057.
61. Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1995**, *99*, 7137.
62. Giesen, D. J.; Gu, M. Z.; Cramer, C. J.; Truhlar, D. G. *J. Org. Chem.* **1996**, *61*, 8720.
63. Giesen, D. J.; Chambers, C. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys Chem.* **1997**, *101*, 2061.
64. Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acct.* **1997**, *98*, 85.
65. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. to be published (SM5.0R-organic model).
66. Bucher, M.; Porter, T. L. *J. Phys.Chem.* **1986**, *90*, 3406.
67. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6129.
68. Zerner, M. C. *Rev. Comp. Chem.* **1991**, *2*, 313.
69. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.
70. Laird, B. B.; Ross, R. B.; Ziegler, T., Eds. *Chemical Applications of Density Functional Theory*; American Chemical Society: Washington, DC, 1996.
71. Yomosa, S. *J. Phys. Soc. Japan* **1973**, *35*, 1738.
72. Rivail, J.-L.; Rinaldi, D. *Chem. Phys.* **1976**, *18*, 233.
73. Tapia, O. *Theor. Chem. Acta* **1978**, *47*, 157.
74. Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comp.-Aided Mol. Des.* **1995**, *9*, 87.
75. Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* in press.
76. Bachrach, S. M. *Rev. Comp. Chem.* **1993**, *5*, 171.
77. Lee, B.; Richards, R. M. *J. Mol. Biol.* **1971**, *55*, 379.
78. Hermann, R. B. *J. Phys. Chem.* **1972**, *76*, 2754.
79. Eisenberg, D.; McLachlan, A. D. *Nature* **1986**, *319*, 199.
80. Abraham, M. H. *Chem. Soc. Rev.* **1993**, 73.
81. Abraham, M. H. *et al. J. Chem. Soc. Perkin Trans. II* **1989**, 699.
82. Abraham, M. H. *J. Phys. Org. Chem.* **1993**, *6*, 660.

83. Abraham, M. H. In *Quantitative Treatments of Solute/Solvent Interactions*; Politzer, P., Murray, J. S.; Eds.; Elsevier: Amsterdam, 1994; p. 83.
84. Pauling, L. *The Nature of the Chemical Bond*; 3rd ed.; Cornell University Press: Ithaca, 1960; p. 93.
85. Reichardt, C. In *Solvents and Solvent Effects in Organic Chemistry*; VCH: New York, 1990.
86. Gutmann, V.; Resch, G. *Lecture Notes on Solution Chemistry*; World Scientific: Singapore, 1995.
87. Murray, J. S.; Brinck, T.; Lane, P.; Paulsen, K.; Politzer, P. *J. Mol. Struct. (Theochem)* **1994**, *307*, 55.
88. Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.
89. Löwdin, P.-O. *J. Chem. Phys.* **1950**, *18*, 365.
90. Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1989**, *11*, 361.
91. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
92. Dewar, M. J. S.; Zoebisch, E. G. *J. Mol. Struct. (Theochem)* **1988**, *180*, 1.
93. Dewar, M. J. S.; Yuan, Y.-C. *Inorg. Chem.* **1990**, *29*, 3881.
94. Stewart, J. J. P. *J. Comp. Chem.* **1989**, *10*, 221.
95. Easton, R. E.; Giesen, D. J.; Welch, A.; Cramer, C. J.; Truhlar, D. G. *Theo. Chim. Acta* **1996**, *93*, 281.
96. Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
97. Perdew, J. P.; Burke, K.; Wang, Y. *Phys. Rev. B* **1996**, *54*, 6533.
98. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
99. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
100. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
101. Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563.
102. Leo, A. J. *Pamona Medchem Database*; BioByte Corp.: Claremont, CA, 1994.