GB/SA-Based Continuum Solvation Model for Octanol

Scott A. Best,[‡,§] Kenneth M. Merz, Jr.,[*,§] and Charles H. Reynolds[*,‡]

Rohm and Haas Company, 727 Norristown Road, Spring House, Pennsylvania 19477, and Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania, 16801

Received: April 11, 1997; In Final Form: July 16, 1997[⊗]

A continuum solvation model has been developed for octanol on the basis of the generalized Born/surface area (GB/SA) formalism. This model was parameterized to fit experimental octanol free energies of solvation for a training set of 66 diverse organic solutes. The resulting model is able to reproduce experimental octanol free energies of solution with an unsigned average error of 0.50 kcal/mol when using the recently developed Merck molecular force field (MMFF). Results are poorer when the model is applied to other force fields such as OPLS, where the atomic charges vary significantly from MMFF. The GB/SA octanol model can be used in conjunction with computed or experimental aqueous free energies of solvation to calculate log P_{ow} directly. The average unsigned error in calculated log P_{ow} values for the same set of compounds is 0.37.

Introduction

Solvent effects can play a critical role in determining properties of solutes in solution. Solute properties which may be affected include electronic configuration, conformation, reactivity and free energies of association.[1−6] Numerous approaches have been taken to include the effect of solvation in molecular simulations. Explicit inclusion of solvent in molecular dynamics (MD) and Monte Carlo (MC) simulations have been the most widely used methods for studying solvent properties and their effects.[3−5] However, the use of explicit solvent molecules in these simulations dramatically increases the size of the calculation and creates the dilemma of either running extremely expensive simulations or risking insufficient sampling. Although these types of calculations may be able to reproduce many solution state properties, the associated computational expense has led to the development of other types of solvent models, in particular continuum solvation models.[6−20]

The roots of continuum solvation models can be traced back to the work of Born in 1920.[21] In the years since many continuum approaches to model solvation have been proposed. The most rigorous continuum methods employ numerical solutions to the Poisson−Boltzman (PB) equation that accurately represent the physics of solvation.[20,22−29] These models only compute one, albeit typically the largest, component of the solvation free energy, the electrostatic solvent polarization energy. The PB equation can be adapted into a form that allows calculation of molecular solvation free energies, as evident in the work of Honig and others.[12,24,28,30−33] However, obtaining an accurate solution to the PB equation also incurrs significant computational cost when studying large molecules or assemblies of molecules that contain many atoms. Some recent efforts have focused on developing inexpensive, approximate continuum models which do not sacrifice the accuracy of the PB equation. These approximate continuum models have an advantage over those which employ explicit solvent molecules in that they are less computationally intensive yet can yield accurate results. In these calculations, the solvent is no longer treated explicitly, but as an ensemble averaged continuous medium having properties of the real solvent.[34] Continuum solvation models have been developed for use with both quantum mechanical and empirical force field methods. The semiempirical SMX continuum models developed by Truhlar and Cramer[14,35−41] and the generalized Born/surface area (GB/SA) model of Still and co-workers[13,34,42] are both examples of such models.

Not surprisingly, one of the most extensively studied solvents using continuum methods has been water.[12−14,24,28,30−34,37,38,40,43,44] While a great deal of chemistry occurs in aqueous solution and experimental aqueous solvation data is plentiful, there is also a need for organic continuum solvent models. Examination of compounds in a variety of solvents can provide useful insight into such physical processes as reaction rates and mechanisms. Furthermore, nonpolar solvents play an important role in modeling a wide range of chemical processes including protein ligand interactions, solute partitioning between liquid phases, and transport across lipid bilayers.[30,35] The partition coefficient of a solute between an organic solvent and water is often used as a measure of lipophilicity, and is a leading term in many quantitative structure-activity relationships (QSAR). Experimentally, lipophilicity is usually measured as the log of the equilibrium constant for partitioning a solute between octanol and water and is expressed as log P_{ow}. Due to the wide use of this quantity in predicting the transport and other properties of organic molecules, there has been a long standing interest in methods for computing log P_{ow} directly.[39,45] Hence, a continuum solvation model for octanol would be useful for a wide range of applications.

Octanol presents a variety of unique challenges to the development of a continuum model. Unlike other solvents that have been represented as continuum models such as water, chloroform and hexadecane, octanol exhibits both hydrophilic and hydrophobic characteristics. MD simulations of various solutes in octanol illustrate this amphipathic nature of the solvent.[5,46] For polar regions of the solute, the hydroxyl group of the octanol will tend to associate around these polar regions and form micellar like regions. Conversely, the hydrophobic alkyl portion of the alcohol will preferentially associate with non-polar regions of the solute. Depending on where sampling is occurring, there appear to exist two distinct solvent regions around the solute which may exhibit different dielectrics rather than the accepted uniform dielectric of octanol ($\epsilon = 10.3$).[47] Another way to state this is that both hydrophilic and hydrophobic regions exist within the solvent continuum.

[‡] Rohm and Haas Company.
[§] The Pennsylvania State University.
[⊗] Abstract published in *Advance ACS Abstracts,* November 15, 1997.

S1089-5647(97)01264-9 CCC: $14.00

TABLE 1: Experimental Octanol Free Energies of Solvation for Anhydrous and Water-Saturated Octanol (kcal/mol)

solutes	ΔG_{oct} anhydrous octanol	ΔG_{oct} saturated octanol
n-hexane	−3.32	−2.83
n-heplane	−4.05	−3.73
n-octane	−4.48	−4.16
cyclohexane	−3.72	−3.45
diethylether	−2.95	−2.84
tetrahydrofuran	−3.87	−4.13
propanone	−3.09	−3.57
cyclopentanone	−4.99	−4.76
2-butanone	−3.75	−4.04
3-pentanone	−4.30	−4.76
n-butanol	−5.32	−5.20
t-butanol	−4.54	−4.99
2-butanol	−5.11	−5.41
n-octanol	−8.21	−8.17

A second challenge that arises in developing a continuum model for octanol is related to the limited availability of solvation free energy data for anhydrous octanol. Such data exists only for a relatively small number of solutes. This runs counter to our desire to derive a general method based on a diverse training set. On the other hand, a very large number of experimental log P_{ow} values are available. These can be combined with experimentally determined free energies of solvation in water to back calculate the free energy of solvation in octanol. Besides the additional uncertainty introduced by fitting to a difference of two experimental quantities, there is also the consideration that experimental log P_{ow} values are determined using water saturated octanol rather than neat octanol. Fortunately, a comparison of experimental free energies of solvation in neat octanol with those for water saturated octanol shows that there is very little difference in the two values as illustrated in Table 1.[48,49] In our opinion, the much greater availability of reference data for fitting the model justifies the small increase in uncertainty associated with the octanol free energies of solution derived from experimental log P_{ow} data.

We have chosen to develop an octanol continuum solvation model using the GB/SA formalism and the Merck molecular force field (MMFF).[50] MMFF was chosen as the reference force field for several reasons. First, MMFF has been parameterized for a large number of functional groups and allows for flexibility in choosing compounds for inclusion into the training set. It has also been parameterized to model both simple organic compounds and macromolecular systems such as proteins. Finally, MMFF utilizes an empirical scheme for calculating charges that makes atomic charges generally available. These charges mimic electrostatic potential (ESP) charges calculated quantum mechanically utilizing the HF/6-31G* basis set.[51,52] This provides additional flexibility for modeling compounds with unusual functionality since 6-31G*(ESP) partial charges may be generated for atoms in compounds where the MMFF charges are deficient. This would not be the case for a force field such as OPLS, for example, since the charges in OPLS have been determined empirically to fit solution data.

GB/SA Method

In order to fully describe the development of the solvation model, we must first briefly outline the theory behind the GB/SA method as has been previously described by Still and co-workers.[13,34] The total solvation free energy (G_{sol}) in the GB/SA model can be expressed as the sum of a solvent cavity term (G_{cav}), a solute-solvent van der Waals term (G_{vdW}), and a solute−solvent electrostatic term (G_{pol}):

$$G_{sol} = G_{cav} + G_{vdW} + G_{pol} \tag{1}$$

The first two terms account for the free energy of forming a cavity in the solvent and for the changes in the dispersion interactions that accompany solvation. It has been shown that the nonelectrostatic contributions to G_{sol} are proportional to the solvent-accessible surface area of the solute (SASA).[53−55] Therefore, it follows that the sum of G_{cav} and G_{vdW} can be computed by evaluating the solvent-accessible surface areas for each atom type

$$G_{cav} + G_{vdW} = \sum \sigma_l SA_l \tag{2}$$

where SA_l is the total solvent-accessible surface area of atom type *l*, and σ_l is an empirically derived atomic surface tension term for atom type *l*.

The third term in eq 1, G_{pol}, is the electrostatic free energy. The G_{pol} term is modeled after the generalized Born equation and is modified to approximate the PB equation as shown in eq 3,

$$G_{pol} = -166.0\left(1 - \frac{1}{\epsilon}\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{q_i q_j}{\left(r_{ij}^2 + \alpha_{ij}^2 e^{-D_{ij}}\right)^{0.5}} \tag{3}$$

with $\alpha_{ij} = \sqrt{\alpha_i \alpha_j}$ and $D_{ij} = r_{ij}^2/(2\alpha_{ij})^2$, r_{ij} being the distance between atoms *i* and *j* having charges q_i and q_j, ϵ being the dielectric constant, and α_i being the effective Born radius.

In order to calculate G_{pol}, it is first necessary to calculate a Born radius *(*α*)* for each charged atom in the solute. The Born radius for monatomic spherical ions can be defined simply as the sum of the atomic van der Waals radius plus an offset value that depends on the solvent. The Born equation for spherical ions is

$$G_{pol} = 166.0\left(1 - \frac{1}{\epsilon}\right)\frac{q^2}{\alpha} \tag{4}$$

where the total dielectric polarization free energy is related to the solvent dielectric constant ϵ, the charge *q*, and the Born radius *(*α, the ion's effective radius).[34] In the case of polyatomic molecules the definition of α is not straightforward. The basic approach behind computing α in GB/SA involves breaking the polyatomic solute into its individual atoms and evaluating the contribution to G_{pol} for each atom in the molecule. These atomic contributions ($G_{pol,i}$) are related to the charge on atom *i*, its van der Waals radius, the effective solvent radius, and the degree to which the atom is either exposed to solvent or buried within the solute.

Still and co-workers have recently reported an analytical method to calculate atomic contributions to G_{pol} that includes a series of empirical scaling parameters.[34] These scaling parameters were optimized to reproduce accurately determined atomic $G_{pol,i}$ in water for the individual atoms in a diverse set of molecules calculated using DelPhi. The explicit equation for calculating $G_{pol,i}$ in GB/SA is given in eq 5

$$G_{pol,i} = \frac{-166.0}{R_{vdW-i} + P_1} + \sum_{stretch}\frac{P_2 V_j}{r_{ij}^4} + \sum_{bend}\frac{P_3 V_j}{r_{ij}^4} + \sum_{nonbonded}\frac{P_4 V_j \mathrm{CCF}}{r_{ij}^4} \tag{5}$$

where $G_{pol,i}$ is the polarization energy of atom *i*, r_{ij} is the distance between atoms *i* and *j* (Å), R_{vdW-i} is the van der Waals radius of atom *i* (Å), V_j is the volume of atom *j* (Å^3), P_1 is a single atom scaling factor which also includes the dielectric offset, P_2

is the stretch scaling factor, P_3 is the bend scaling factor, P_4 is the nonbonded scaling factor, P_5 is the soft cutoff parameter (see eq 6 below), and CCF is the close contact function for nonbonded interactions. This CCF function is required due to the deviations that occur when there is significant overlap between nonbonded atom pairs. CCF is expressed as

$$\text{CCF} = 1.0 \qquad \text{if} \qquad \frac{r_{ij}}{R_{\text{vdW}-i} + R_{\text{vdW}-j}} > \frac{1}{P_5}$$

otherwise,

$$\text{CCF} = \left\{0.5\left[1.0 - \cos\left\{\left(\frac{r_{ij}}{R_{\text{vdW}-i} + R_{\text{vdW}-j}}\right)^2 P_5\right\}\right]\right\}^2 \qquad (6)$$

This derivation of the GB/SA model has been used for both water and chloroform, which are the only solvents currently available. In our development of an octanol GB/SA model, we used the same formalism so that our model could be easily implemented alongside the existing continuum solvent models in MacroModel.[56]

Optimization Procedure

In order to obtain the best parameter values ($P_1 - P_5$) for use in eq 5, we followed the precedent of Qiu *et al.*[34] and initially computed $G_{\text{pol},i}$ using a Poisson−Boltzman (PB) calculation as implemented in the DelPhi program.[12,57,58] DelPhi incorporates the PB equation through the use of a finite-difference method in order to provide an approximate but physically complete treatment of electrostatic interactions in solution. This algorithm has been used extensively to predict electrostatic properties in solution and provides an accurate means to estimate $G_{\text{pol},i}$ which is essential in the determination of $P_1 - P_5$.

In computing $G_{\text{pol},i}$ using DelPhi, a unit charge was placed on atom i ($q_i = 1.0$) while the remaining atoms in the polyatomic solute had neutral charge. The solvent dielectric was set to an octanol-like medium ($\epsilon = 10.3$), while the internal solute dielectric constant was set to 1.0. No ionic species were present for these calculations, and so the solution ionic strength was set to zero. The water and chloroform GB/SA models both used a 1.4 Å solvent probe radius, but this would seem to be much too small for a large hydrophobic solvent such as octanol. Instead, a solvent radius of 2.0 Å was used in all calculations. A radius of 2.0 Å was selected because it approximates the van der Waals radius of a united atom CH_2 unit in octanol. This is also consistent with the experimental van der Waals radius for methane, and the probe radius used by Truhlar and Cramer in their recently developed semiempirical quantum mechanical solvation model for *n*-hexadecane.[35] Other alkane solvent continuum models have also used 2.0 Å as a reasonable solvent probe radius.[30] Prior to the DelPhi calculations, each molecule in the training data set was initially minimized in the gas phase using MMFF as incorporated in MacroModel. All compounds that were used in the training set are given in Table 2. Three carboxylic acids are also listed in Table 2, but they were not included in the optimization of electrostatic parameters.

Using the $G_{\text{pol},i}$ computed from DelPhi, we optimized the electrostatic scaling parameters ($P_1 - P_5$) of eq 5 in order to calculate an accurate $G_{\text{pol},i}$ utilizing the GB/SA model. As was done with the DelPhi calculations, all compounds in the training set were initially minimized in the gas phase using MMFF. The scaling parameters ($P_1 - P_5$) were optimized using the conjugate search optimization routine of Powell adapted for the calculation of $\Delta G_{\text{pol},i}$.[59,60] Powell's method systematically varies the parameters $P_1 - P_5$ of eq 5 without the use of explicit gradients in order to minimize an error function. In this case, the error function was expressed as the rms error between the GB/SA and DelPhi $\Delta G_{\text{pol},i}$ values.

$$\text{ERROR} = \sqrt{\frac{\sum_{i=1}^{N}(G_{\text{pol},i}(\text{eq 5}) - G_{\text{pol},i}(\text{DelPhi}))^2}{N}} \qquad (7)$$

We carried out a systematic search for a minimum in five-dimensional space by creating three different parameter sets for $P_1 - P_5$. Two of the sets consisted of the existing parameter values for chloroform and water as already implemented in the GB/SA model, while the third set was the average of these two parameter sets. Each of the elements in the initial three parameter sets were interchanged resulting in a total of 240 different permutations. Each permutation was used as a starting point for optimization (240 separate parameter optimizations were performed) in order to locate the minimum rms error. This search strategy was chosen with the expectation that the values of $P_1 - P_5$ for octanol would reside somewhere between those of water and chloroform. The parameters that resulted in the lowest rms error are given in Table 3.

The remaining portion of the GB/SA model, which incorporates the energy associated with formation of the solvent cavity, as well as the solute−solvent van der Waals dispersion interactions, was parameterized independently while keeping the optimized electrostatic parameters constant. Recall that this quantity is directly proportional to the SASA and an atomic surface tension parameter σ. In the GB/SA model, the SASA contribution is computed for each individual atom in every compound.[55] Still and co-workers initially assigned the same σ to all atom types. It follows naturally, however, that each atom type should have a distinct surface tension term (σ) associated with it. This approach of allowing atom type dependent σ values was implemented by Truhlar and Cramer in their semiempirical SMx models and has been adopted in a recent version of the aqueous GB/SA model on a more modest scale.[14,34−36,38]

In order to compute σ for the various atom types, experimental solvation free energies for solutes in octanol are required. As discussed previously, we chose to use the octanol solvation free energies derived from experimental log P_{ow} data. These experimental log P_{ow} values can be combined with free energies of solution in water (ΔG_{aq}) to calculate ΔG_{oct}. The relationship can be expressed as

$$\Delta G_{\text{oct}} = \Delta G_{\text{aq}} - 2.30RT\log P_{\text{ow}} \qquad (8)$$

where R is the gas constant and $T = 298$K.

For the parameterization of σ, the following expressions, which are merely expanded versions of eq 1, were implemented

$$\Delta G_{\text{oct},k} - \Delta G_{\text{pol},k} = \Delta G_{\text{cav},k} + \Delta G_{\text{vdW},k} \qquad (9)$$

and

$$\Delta G_{\text{cav},k} + \Delta G_{\text{vdW},k} = \sum_{l=1}^{N} \sigma_{\text{l}} SA_{k,l} \qquad (10)$$

where $SA_{k,l}$ is the sum of all the solvent accessible surface area contributions for atoms of type l in the kth compound, σ_{l} is the derived surface tension term for atom type l, $\Delta G_{\text{oct},k}$ *is* the experimentally determined free energy of solvation for compound k in octanol, and $\Delta G_{\text{pol},k}$ is the electrostatic polarization free energy for compound k computed from eq 5. For example,

TABLE 2: Experimental and Theoretically Predicted Octanol Free Energies of Solvation (kcal/mol)

solutes	ΔG_{oct} (exptl)	ΔG_{oct} (MMFF)	ΔG_{oct} (MMFF) − ΔG_{oct} (exptl)	ΔG_{oct} (OPLS)	ΔG_{oct} (OPLS) - ΔG_{oct} (exptl)
n-butanol	−5.92	−6.44	−0.52	−6.55	−0.63
2-butanol	−5.41	−5.56	−0.15	−5.71	−0.30
t-butanol	−4.99	−5.33	−0.34	−5.51	−0.52
n-octanol	−8.18	−7.66	0.52	−7.77	0.41
n-hexane	−2.83	−2.76	0.07	−2.76	0.07
n-heptane	−3.73	−3.06	0.67	−3.06	0.67
n-octane	−4.16	−3.37	0.79	−3.37	0.79
cyclohexane	−3.45	−2.90	0.55	−2.90	0.55
isobutane	−1.46	−1.82	−0.36	−1.82	−0.36
neopentane	−1.74	−2.03	−0.29	−2.03	−0.29
acetone	−3.57	−4.36	−0.79	−2.31	1.26
2-butanone	−4.03	−4.56	−0.53	−2.49	1.54
cyclopentanone	−4.76	−4.26	0.50	−2.17	2.59
3-pentanone	−4.76	−4.75	0.01	−2.66	2.10
diethylether	−2.84	−4.01	−1.17	−3.67	−0.83
tetrahydrofuran	−4.13	−4.11	0.01	−3.72	0.40
tetrahydropyran	−4.4i	−4.30	0.11	−3.94	0.47
pyridime	−5.59	−6.23	−0.64	−5.12	0.47
4-methylpyridine	−6.56	−6.01	0.55	−4.56	2.00
benzene	−3.80	−5.18	−1.38	−4.62	−0.82
toluene	−4.62	−5.10	−0.48	−4.28	0.34
phenol	−8.59	−8.90	−0.31	−8.35	0.24
anthracene	−10.29	−8.99	1.30	−7.45	2.84
phenanthrene	−10.02	−9.02	1.00	−7.44	2.58
ethylbenzene	−5.08	−5.00	0.08	−4.17	0.91
napthalene	−6.89	−7.21	−0.32	−6.09	0.80
n-butylamine	−5.62	−5.42	0.20	−5.37	0.25
n-propylamine	−5.05	−5.13	−0.08	−5.07	−0.02
piperidine	−6.24	−4.89	1.35	−4.61	1.63
triethylamine	−5.00	−4.20	0.80	−3.62	1.38
acetonitrile	−3.44	−3.50	−0.06	−1.74	1.70
propanenitrile	4.12	−3.84	0.28	−2.12	2.00
acetamide	−7.99	−8.86	−0.87	−8.70	−0.71
N-methylacetamide	−8.67	−8.21	0.46	−7.35	1.32
N,*N*-dimethylacetamide	−7.45	−7.33	0.12	−6.16	1.29
n-propylacetamide	−8.50	−9.08	−0.58	−8.88	−0.38
methyl acetate	−3.55	−3.18	0.37	−2.01	1.54
ethyl acetate	−4.09	−3.69	0.40	−2.52	1.57
methyl propanate	−4.02	−3.32	0.70	−2.14	1.88
ethyl propanate	−4.45	−3.82	0.63	−2.65	1.80
nitroethane[a]	−3.95	−4.84	−0.89		
nitropropane[a]	−4.49	−5.12	−0.63		
nitrobenzene[a]	−6.62	−5.58	1.04		
propyne	−1.58	−1.33	0.25	−1.14	0.44
1-hexyne	−3.42	−2.41	1.01	−2.24	1.18
propene	−1.11	−1.11	0.00	−0.96	0.15
2-methylpropene	−1.98	−1.66	0.32	− 1.49	0.49
ethylene	−0.27	−0.43	−0.16	−0.30	−0.03
diethyl sulfide	−3.96	−4.31	−0.35	−4.41	−0.45
methylethyl sulfide	−3.50	−4.05	−0.55	−4.14	−0.64
benzenethiol	−5.98	−5.47	0.51	−5.76	0.22
chloroethane	−2.55	−2.87	−0.32	−2.53	0.02
1,2-dichloroethane	−3.72	−3.68	0.04	−3.14	0.58
1-chlorobutane	−3.73	−3.49	0.24	−3.15	0.58
1,2-dibromoethane	−4.77	−5.10	−0.33	−4.73	0.04
bromoethane	−2.89	−3.43	−0.54	−3.21	−0.32
1-bromobutane	−4.15	−4.05	0.10	−3.82	0.33
iodoethane	−3.42	−3.37	0.05	−3.13	0.29
2-iodopropane	−4.44	−3.23	1.21	−2.99	1.45
methyl iodide	−2.94	−3.11	−0.17	−2.86	0.08
fluoroform	−0.07	−0.07	−0.00	0.82	0.89
1,1-difluoroethane	−1.12	−1.54	−0.42	−0.66	0.46
trifluoroethanol	−4.86	−4.73	0.13	−4.24	0.62
acetic acid	−6.47	−5.65	0.82	−4.82	1.65
propanoic acid	−6.92	−5.78	1.14	−4.96	1.96
butanoic acid	−7.43	−6.16	1.27	−5.34	2.09
average unsigned error			0.50		0.91

[a] The nitro functional group is unavailable in OPLS.

one term (l) in eq 10 would relate the contribution of a C(sp^3) atom type to $\Delta G_{cav,k} + \Delta G_{vdW,k}$ for the kth compound while another l represents the contribution of an O(sp^2) atom type. All atom types (Table 4) are represented as a term in the summation. If a compound does not contain a particular atom type, there is no contribution by that type to $\Delta G_{cav,k} + \Delta G_{vdW,k}$.

The σ_l parameters for each atom type were determined using multiple linear regression in order to fit the expression $\Delta G_{oct,k}$

TABLE 3: Optimized Parameters for Calculation of ΔG_{pol} in eq 5

scaling parameter	initial optimized values[a]	final optimized values[b,c]
P_1	0.414	0.614
P_2	1.483	1.483
P_3	6.110	6.110
P_4	31.253	31.253
P_5	0.0034	0.0034

[a] Initial values for the optimized parameters $P_1 - P_5$ derived from comparisons to DelPhi ΔG_{pol} values. [b] Final values for the scaling parameters which are used in the GB/SA octanol model. These scaling parameters were determined from a simple line search in which all of the parameters were kept constant except for P_1. [c] Final parameters include scaling the amide oxygen's radius to 1.28 Å.

TABLE 4: Optimized Values for the Surface Tension Terms

atom functional groups	σ (J/mol) $P_1 = 0.61$	σ^* (J/mol)[a] $P_1 = 0.61$
CH_2 (sp^3)	−35.97	−37.82
CH (sp^3)	−3.99	−12.94
Csp^2 (aromatic)	−50.02	−50.78
Csp^2 (aliphatic)	−0.64	−2.40
Csp	−19.50	−19.87
Osp^3 (hydroxyl)	−159.42	−158.94
Osp^3 (ether)	50.12	53.15
Osp^2	50.12	53.15
Nsp^3 (amines) and Nsp^3 (amides)	−148.99	−116.93
Nsp^2 and Nsp	95.71	94.18
CH_3 and S	−27.83	−26.22
F	43.08	44.77
C1	−8.38	−7.53
Br and I	−49.02	−49.09

[a] σ^* values are calculated with the amide oxygen radius scaled to 1.28 Å.

TABLE 5: Rms Error (kcal/mol) and Corresponding Value of P_1 for Including and Excluding Carboxylic Acids

trial	parameter P_1	rms error (no acids)	rms error (with acids)
1	0.21	0.989	0.991
2	0.41	0.829	0.847
3	0.61	0.765	0.797
4	0.81	0.795	0.820

$- \Delta G_{pol,k}$. This was done utilizing the statistical package JMP.[61] Initially, each atom type was assigned an independent surface tension term. Those σ_1 values computed from the regression analysis that had similar energies were combined into one surface tension parameter, and the regression analysis repeated. The resulting atom types and their corresponding values for σ_1 are given in Table 4.

In order to determine the appropriate balance between the electrostatic and SASA terms, a simple line search was initiated for P_1 in which all of the other scaling factors in eq 5 were kept constant. P_1 was varied in increments of 0.2 Å from 0.21 to 0.81. The $\Delta G_{pol,k}$ and $\Delta G_{cav,k} + \Delta G_{vdW,k}$ were recalculated for the full data set at each value of P_1. Similarly, a new set of σ parameters were derived for each increment of P_1. The line search was evaluated based on the rms errors between calculated and experimental free energies of solvation. The results of the line search are given in Table 5. As depicted in Figure 1, the minimum rms error occurs at $P_1 = 0.61$ Å, which was used as the final P_1 value. Final values for all the electrostatic parameters ($P_1 - P_5$) are listed in Table 3.

Figure 2a illustrates a plot of the theoretically predicted ΔG_{oct} versus the experimental ΔG_{oct}. This plot shows good agreement between the two free energies for almost every compound in the training set. However, there are some exceptions, most notably the amides whose predicted ΔG_{oct}, are significantly lower than their corresponding experimental values. This underestimation was corrected by decreasing the radius of the amide oxygen atom to 1.28 Å. The amide oxygen radius was scaled to 1.28 Å based upon results obtained through varying the value of P_1 in the aforementioned line search, and through an examination of the corresponding atomic charges. Scaling the amide oxygen led to a significant decrease in the rms error

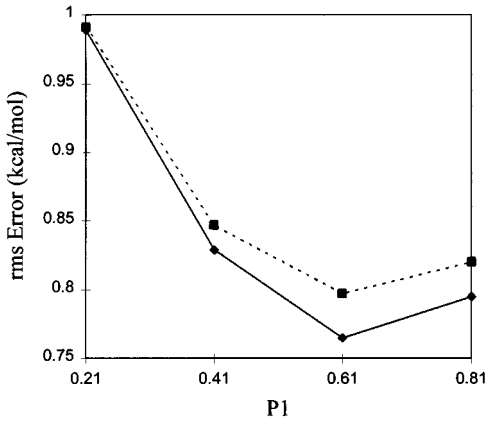

Figure 1. Determination of the scaling parameter P_1 which corresponds to the lowest rms error. The solid line represents the training set that excludes the carboxylic acids while the dashed line represents the training set that includes acids.

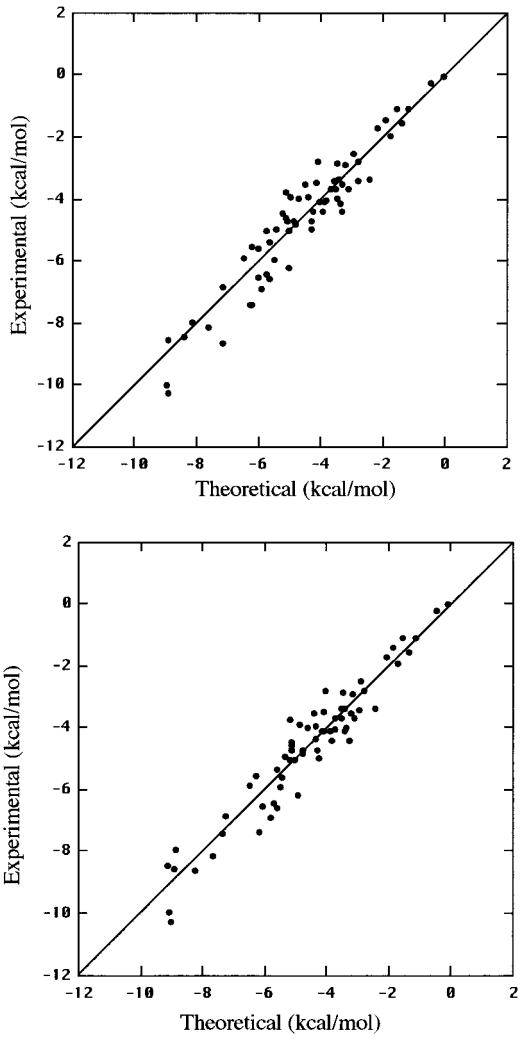

Figure 2. Comparison of theoretically computed ΔG_{oct} and experimental ΔG_{oct} for (a) the optimized model that does not scale the amide oxygen radius and (b) the optimized model that scales the amide oxygen radius.

TABLE 6: Free Energies of Solvation (kcal/mol) for Carboxylic Acids Calculated Using MMFF

compound	ΔG_{oct} (exptl)	ΔG_{oct}[a] (no acids)	ΔG_{oct}[b] (with acids)
acetic acid	−6.47	−4.57	−5.65
propanoic acid	−6.92	−4.71	−5.78
butanoic acid	−7.43	−5.11	−6.16

[a] Model excluding carboxylic acids from the training set. [b] Model including carboxylic acids in the training set.

for the training set from 0.770 to 0.736 kcal/mol (Figure 2b). The final surface tension terms are given in Table 4 for both the scaled (σ*) and unscaled (σ) oxygen radius, with all subsequent calculations incorporating σ*.

In order to examine differences in the electrostatic properties of the different force fields, the atomic charges generated by the MMFF and OPLS force fields were compared to those calculated quantum mechanically for various classes of molecules in the training set. Semiempirical calculations were initially carried out using MOPAC 6.0.[62] The geometries of the structures were optimized using a restricted Hartree−Fock calculation and the PM3 Hamiltonian.[63] These optimized structures were used in the subsequent single-point Hartree−Fock quantum mechanical calculations. The atomic charges for the molecule were obtained by using the electrostatic potential (ESP) charges calculated using the 6-31G* basis set[51,52] and rapid electrostatic fitting.[64]

Results and Discussion

The MMFF GB/SA computed ΔG_{oct} values for the organic compounds in our training set are summarized in Table 2. A wide variety of molecules were included in the training set so that the optimization would not be biased toward any one particular class of compounds. Each generic class (i.e., alcohol) was represented by at least three molecules, initially excluding organic acids as well as compounds which contained phosphorous. Both mono- and multifunctional compounds were included. The various atom types that exist in MacroModel were also taken into consideration in the selection of molecules for the training set.[56] The compounds not only had to represent different functionalities, but also had to be representative of these various atom types. The 66 compounds included in the training set fulfilled both requirements.

There is very little experimental solvation data for compounds containing phosphorous other than a few phosphate esters. These esters are not useful for deriving surface tension parameters for phosphorous because the phosphorous atom is buried deep withing the structure and has no surface area contribution. In the absence of data for compounds with more exposed phosphorous atoms, we simply assigned phosphorous the same σ as CH_3.

Carboxylic acids were initially omitted from the parameterization because of concerns about their behavior in hydrophobic solution. In nonpolar solvents the acids may dimerize leading to significant errors in their experimental free energies of solvation. This phenomenon as well as problems inherent in measuring log P values for neutral acids create additional uncertainty in ΔG_{oct}. Table 6 compares predicted and experimental ΔG_{oct} data for three carboxylic acids, and this comparison shows that the predicted ΔG_{oct} for the three acids were underestimated by approximately 2.0 kcal/mol. Assuming that the experimental data is reasonable, such a large difference is clearly not desirable since correctly modeling acid groups is essential in many areas of interest including modeling polymeric acids such as poly(acrylic acid). In order to improve upon the results, these compounds were included in the training set for the SASA dependent terms and the model was reoptimized. The parameterization of ΔG_{pol} was not repeated since the oxygen parameters were already well represented in the training set. The surface tension terms were optimized as previously described, and the line search was also repeated through variation of the scaling parameter P_1. The rms error of the regression analysis was only slightly greater with the inclusion of the acids into the training set (0.797 kcal/mol as compared to 0.765 kcal/mol). As seen in Figure 1, the minimum rms error still corresponded to $P_1 = 0.61$. The error in ΔG_{oct} calculated for the acids using the newly optimized parameters decreased by approximately 1 kcal/mol as shown in Table 6. Hence, including the organic acids did not significantly alter the original model yet allowed solution properties of this class of compounds to be predicted with a higher degree of certainty. Tables 3 and 4 provide a listing of the parameters used in the optimized continuum model.

TABLE 7: Free Energies of Solvation (kcal/mol) for Amides Calculated Using MMFF

compound	ΔG_{oct}[a]	ΔG_{oct}[b]	ΔG_{oct}[c]	ΔG_{oct} (exptl)
acetamide	−8.12	−8.86	−9.52	−7.99
N-methylacetamide	−7.12	−8.21	−8.73	−8.67
N,*N*-dimethylacetamide	−6.16	−7.33	−7.24	−7.45
n-propylacetamide	−8.34	−9.08	−8.94	−8.50

[a] Oxygen radius is unscaled (1.48 Å). [b] Oxygen atom is scaled to 1.28 Å. [c] Oxygen unscaled with 6-31G* ESP charges.

TABLE 8: Calculated Atomic Partial Charges for the Amide Nitrogens and the Carbonyl Oxygens

	HF/6-31G* ESP		MMFF	
compound	nitrogen	oxygen	nitrogen	oxygen
acetamide	−0.86	−0.65	−0.80	−0.57
N-methylacetamide	−1.02	−0.63	−0.73	−0.57
N,*N*-dimethylacetamide	−0.35	−0.60	−0.66	−0.57
n-propylamide	−1.00	−0.61	−0.80	−0.57

Experimental ΔG_{oct} are plotted against the predicted values in Figure 2. The predicted solvation energies are strongly correlated with experiment giving an $r^2 = 0.92$ and a theoretical slope of 1.0. The errors between predicted and observed ΔG_{oct} are centered around zero and have an rms error of 0.72 kcal/mol (unsigned av = 0.50 kcal/mol). These errors relative to experiment are less than those reported by Qiu *et al.* for the GB/SA aqueous solvation free energies,[13,34] albeit their reported error is for a much larger dataset. These results are also comparable to the errors in calculated free energies of solvation in octanol reported by Giesen *et al.* using OSM5.4/A. This quantum mechanically based method gives an average unsigned error of 0.6 kcal/mol across 14 alcohol solvents including octanol.[41] The largest errors that occur in our model correspond to compounds which contain amide, acid, and nitro functionalities. The larger error found between experimental and predicted values for the acids has been addressed already. The differences for the nitro compounds may be a result of atom typing problems in MacroModel. A comparison of the atomic charges calculated by the MMFF and the calculated ESP charges show a strong correlation. However, in order to obtain these charges, the nitro functionality must be described in MacroModel as an ion with double bonds to both of the oxygen atoms, rather than some resonance hybrid.

In order to correctly model peptides and much larger biological molecules such as proteins and enzymes, solution properties for compounds which contain the amide functionality must be reproduced accurately. An examination of Table 7 illustrates that the ΔG_{oct} for secondary and tertiary amides are underestimated by more than 1 kcal/mol. Because of the

TABLE 9: Experimental and Predicted log P_{ow} Values[a]

solutes	ΔG_{aq} (exptl)	ΔG_{aq} (MMFF)	log P_{ow} (exptl)	log P_{ow} (predicted)[a]	log P_{ow}[c] (predicted)
n-butanol	−4.72	−4.65	0.88	1.31	1.26
2-butanol	−4.51	−3.25	0.61	1.70	0.77
t-butanol	−4.58	−3.04	0.35	1.68	0.55
n-octanol	−4.09	−3.80	3.00	2.83	2.62
n-hexane	2.48	2.47	3.90	3.84	3.85
n-heptane	2.62	2.67	4.66	4.21	4.17
n-octane	2.89	2.88	5.18	4.59	4.60
cyclohexane	1.23	1.79	3.44	3.44	3.03
isobutane	2.30	2.17	2.76	2.93	3.03
neopentane	2.50	2.41	3.11	3.26	3.33
acetone	−3.90	−7.47	−0.24	−2.28	0.34
2-butanone	−3.64	−6.36	0.29	−1.32	0.68
cyclopentanone	−4.54	−7.35	0.16	−2.27	−0.21
3-pentanone	−3.41	−5.29	0.99	−0.40	0.98
diethyl ether	−1.63	−3.02	0.89	0.73	1.75
tetrahydrofuran	−3.50	−4.36	0.46	−0.18	0.45
tetrahydropyran	−3.12	−3.51	0.95	0.58	0.87
pyridine	−4.70	−6.89	0.65	−0.48	1.12
4-methylpyridine	−4.90	−6.47	1.22	−0.34	0.82
benzene	−0.90	−2.61	2.13	1.89	3.14
toluene	−0.90	−2.16	2.73	2.16	3.08
phenol	−6.60	−7.58	1.46	0.97	1.69
anthracene	−4.23	−5.02	4.45	2.92	3.50
phenanthrene	−3.95	−5.15	4.46	2.84	3.72
ethylbenzene	−0.80	−1.77	3.15	2.37	3.08
napthalene	−2.39	−3.95	3.30	2.39	3.54
n-butylamme[d]	−4.30	−2.55	0.97	2.11	0.82
n-propylamine[d]	−4.40	−2.76	0.48	1.74	0.54
piperidine[d]	−5.10	−2.20	0.84	1.98	−0.15
triethylamine[d]	−3.04	−1.33	1.44	2.11	0.85
acetonitrile	−3.90	−4.71	−0.34	−0.89	−0.29
propanenitrile	−3.90	−4.13	0.16	−0.21	−0.04
acetamide	−9.71	−11.39	−1.26	−1.86	−0.62
N-methylacetamide	−10.10	−9.67	−1.05	−1.07	−1.39
N,*N*-dimethylacetamide	−8.50	−7.87	−0.77	−0.40	−0.86
n-propylacetamide	−9.40	−10.47	−0.66	−1.02	−0.23
methyl acetate	−3.30	−4.85	0.18	−1.23	−0.09
ethyl acetate	−3.10	−4.41	0.73	−0.53	0.43
methyl propanate	−2.90	−3.96	0.82	−0.47	0.31
ethyl propanate	−2.80	−3.52	1.21	0.22	0.75
nitroethane	−3.70	− 12.80	0.18	−5.84	0.84
nitropropane	−3.30	−12.32	0.87	−5.29	1.34
nitrobenzene	−4.10	−10.50	1.85	−3.61	1.09
propyne	−0.30	−0.18	0.94	0.84	0.76
1-hexyne	0.30	0.67	2.73	2.26	1.99
propene	1.30	0.01	1.77	0.82	1.77
2-methylpropene	1.20	0.46	2.34	1.56	2.10
ethylene	1.27	−0.31	1.13	0.09	1.25
diethylsulfide	−1.30	−1.23	1.95	2.26	2.21
methylethylsulfide	−1.40	−1.77	1.54	1.67	1.95
benzenethiol	−2.55	−2.56	2.52	2.14	2.14
chloroethane	−0.60	−1.60	1.43	0.93	1.67
1,2-dichloroethane	−1.70	−4.61	1.48	−0.68	1.45
1-chlorobutane	−0.13	−1.04	2.64	1.80	2.47
1,2-dibromaethane	−2.10	−4.01	1.96	0.80	2.20
bromoethane	−0.7	−1.19	1.61	1.64	2.00
1-bromobutane	−0.41	−0.68	2.75	2.47	2.67
iodoethane	−0.70	−1.09	2.00	1.67	1.96
2-iodopropane	−0.50	−0.56	2.89	1.96	2.00
methyl iodide	−0.89	−1.55	1.51	1.15	1.63
fluoroform	0.80	−3.95	0.64	−2.85	0.64
1,1-difluoroethane	−0.10	−3.00	0.75	−1.07	1.06
trifluoroethanol	−4.30	−8.62	0.41	−2.86	0.32
acetic acid	−6.70	−9.08	−0.17	−2.51	−0.77
propanoic acid	−6.47	−8.16	0.33	−1.75	−0.51
butanoic acid	−6.36	−8.00	0.79	−1.35	−0.15
average unsigned error	1.16	0.37			

[a] All free energies of solvation are expressed in kcal/mol. [b] log *P* is calculated from predicted free energies of solvation using MMFF. [c] log *P* is calculated from the predicted free energies of solvation for octanol and experimental aqueous free energies of solvation. [d] Experimental log *P* values were measured at pH = 13.

biological significance of these compounds decreasing the magnitude of this error was a priority. The source of the discrepancy was a deficiency in the polarization free energy for these molecules which is directly proportional to the atomic charges. Quantum mechanically calculated ESP atomic charges were determined for a variety of amides and were compared to

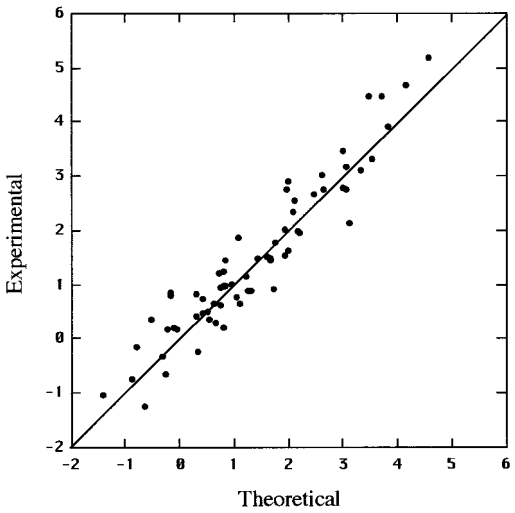

Figure 3. Values of log P_{ow} were calculated using the ΔG_{oct} values predicted by our solvation model and experimental ΔG_{aq} data.

the charges generated by MMFF. As seen in Table 8, the ESP atomic charges varied depending on the substitution at nitrogen, unlike the charges generated by MMFF. The magnitude of these charges were also greater than the MMFF atomic charges. In order to compensate for the too positive polarization free energy, the van der Waals radius for the amide's carbonyl oxygen atom was scaled from 1.48 to 1.28 Å, resulting in a more negative ΔG_{pol}, consequently resulting in more accurate ΔG_{oct} for these compounds as illustrated in Table 7. The amide oxygen was chosen for scaling since the MMFF charges are consistently more positive than the ESP derived HF/6-31G* charges for this atom (Table 8). The average unsigned error in predicting ΔG_{oct} for the compounds in the training set was calculated to be 0.65 and 0.50 for the unscaled and scaled oxygen radius respectively. Thus, scaling the amide oxygen led to a much improved model, and the scaled amide oxygen van der Waals radii have been included in the preferred parameter set.

In developing the dielectric continuum model for octanol, it was hoped that the model would be general, and could be used in conjunction with other commonly available force fields besides MMFF. Unfortunately, this is not the case as evidenced by the data in Table 2 for the prediction of ΔG_{oct} using the OPLS force field. Although a substantial number of the free energies of solvation are reproduced satisfactorily, numerous compounds are underestimated by more than 1.0 kcal/mol. Most of these compounds contain carbonyl functionality, such as ketones, esters, amides, and acids. A comparison of the OPLS atomic charges, the MMFF atomic charges, and the quantum mechanically derived ESP charges reveals that the OPLS charges for these classes of compounds are consistently lower than the quantum mechanically generated charges.[50] A similar trend exists when comparing OPLS to the MMFF charges as well. This illustrates quite well the dominant role atomic charges play in solvation thermodynamics and the need for high quality and consistent atomic charges (*e.g.*, HF/6-31G* ESP). Small differences between atomic charges may result in large errors in the free energies of solvation. The practical result is that a different set of parameters would be necessary in order to obtain results comparable to MMFF using OPLS or AMBER.

Although the optimized octanol continuum model is parameterized only for MMFF, the ΔG_{oct} predicted by the model are quite good when using MMFF. Most compounds are strongly correlated with experimental results as illustrated in Figure 2b ($r^2 = 0.92$). On the basis these results, it would appear that the model has overcome the initial concerns associated with developing an octanol continuum model. One of the major concerns was the existence of both polar and non-polar regions in using an amphipathic solvent like octanol. Although there most likely exist regions within the solvent that have either a greater or lesser dielectric than seen for neat octanol ($\epsilon = 10.3$),[47] it appears that an averaging of these regions occurs and the continuum model overcomes this inhomogeneity of the solvent. It is also likely that the atom-type specific σ parameters absorb some of these "local" solvent interactions.

As mentioned previously, there is a significant need for organic continuum models for a variety of possible applications. Such applications include modeling polymers in solution, analyzing molecular conformations in solution and estimating lipophilicity for use in QSAR studies. Estimation of lipophilicity utilizing the thermodynamic approach expressed in eq 11 has significant advantages. This approach is relatively general and can be applied to virtually any molecule where high quality atomic charges are available. Thus, in contrast to fragment and molecular property based approaches, it is possible to estimate log P_{ow} for classes of compounds where no previous experimental log P_{ow} data is available.

$$\log P_{ow} = (\Delta G_{aq} - \Delta G_{oct})/2.30RT \qquad (11)$$

In order to further test the model, partition coefficients were estimated for the 66 compounds in the training set as given in Table 9. The experimental log P_{ow} values were estimated through eq 11 using ΔG_{oct} and ΔG_{aq} which were both calculated using the MMFF. The results were not as accurate as expected, with an average unsigned error of 1.16. However, the source of the large errors lies in the accuracy of ΔG_{aq} not ΔG_{oct}. Experimental ΔG_{aq} are given in Table 9 along with the predicted ΔG_{aq}, and comparison reveals large discrepancies. This is a result of using the aqueous GB/SA model that was derived for the OPLS force field with MMFF. As we observed for octanol, the error arrises because of differences between the charges used by the two force fields, resulting in significant differences in ΔG_{aq}. Partition coefficients estimated from experimental ΔG_{aq} and calculated ΔG_{oct} using the octanol continuum model are excellent. These are plotted against experimental log P_{ow} in Figure 3. The predicted log P_{ow} are strongly correlated with experiment with $r^2 = 0.90$. The average unsigned error for these computed values of log P_{ow} is 0.37.

Conclusion

We have developed a GB/SA continuum solvation model for octanol that is able to reproduce experimental free energies of solution remarkably well. The average unsigned error in ΔG_{oct} is 0.50 kcal/mol for 66 organic solutes chosen to span a wide variety of size and functionality. As is the case with other solvents, the calculated ΔG_{oct} is very sensitive to the quality of the atomic partial charges. One practical result of this sensitivity, is that the accuracy of our octanol model is seriously degraded when using molecular force fields other than the MMFF, at least unless charges comparable to HF/6-31G*(ESP) are used.

As is the case with previous GB/SA models for water and chloroform, the octanol solvation energies can be calculated at little additional cost making it possible to model polymers, biopolymers and other macromolecular systems in octanol. Partition coefficients estimated from the calculated ΔG_{oct} and experimental ΔG_{aq} accurately reproduce experimental log P_{ow} values. This is one promising application for the GB/SA octanol model. It provides a rapid and relatively general method to estimate lipophilicity in molecules for which no experimental

data exists, as long as accurate partial charges are available. In addition, this model should complement the already existing chloroform and water GB/SA models, since most organic compounds are at least somewhat soluble in octanol, and since octanol has a dielectric that is intermediate between water and chloroform.

Acknowledgment. Financial support was provided by the Research Fellows Organization of Rohm and Haas Company (graduate research assistantship for S.A.B.) and the office of Naval Research. We also thank D. Qiu, P. S. Shenkin, F. P. Hollinger, and W. C. Still for kindly providing us a preprint of their manuscript, as well as Dr. Lee Lynn for programming assistance with MacroModel.

References and Notes

(1) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(2) Reichardt, C. *Solvent Effects in Organic Chemistry*; Verlag Chemie: Weinheim, 1979.
(3) Jorgensen, W. L.; Ravimohan, C. J. *J. Chem. Phys.* **1985**, *83*, 3050−3054.
(4) Rossky, P. J.; Karplus, M. *J. Am. Chem. Soc.* **1979**, *101*, 1913−1936.
(5) DeBolt, S. E.; Kollman, P. *A. J. Am. Chem. Soc.* **1995**, *117*, 5316−5340.
(6) Cramer, C. J.; Truhlar, D. G. Continuum Solvation Models: Classical and Quantum Mechanical Implementations. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D., Ed.; VCH: New York, 1995; Vol. 6, pp 1−72.
(7) Eisenberg, D.; McLachlan, A. D. *Nature* **1986**, *319*, 199−203.
(8) Ooi, T.; Oobatake, M.; Nemethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086−3090.
(9) Scrocco, E.; Tomasi, J. *Top. Curr. Chem.* **1973**, *42*, 381.
(10) Kang, Y. K.; Nemethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1987**, *91*, 4105−4108, 4109−4117, 4118−4120.
(11) Warshel, A.; Russel, S. T. *Rev. Biophys.* **1984**, *17*, 283.
(12) Gilson, M.; Honig, B. *Proteins* **1988**, *4*, 7−18.
(13) Still, W. C.; Tempczyk, A.; Hawley, R.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(14) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8305−8311.
(15) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027−2094.
(16) Tomasi, J.; Bonaccorsi, R.; Cammi, R.; Valle, F. J. O. d. *J. Mol. Struct. (THEOCHEM)* **1991**, *234*, 401−424.
(17) Tawa, G. J.; Martin, R. L.; Pratt, L. R.; Russo, T. V. *J. Phys. Chem.* **1996**, *100*, 1515−1523.
(18) Chen, J. L.; Noodleman, L.; Case, D. A.; Bashford, D. *J. Phys. Chem.* **1994**, *98*, 11059−11068.
(19) Rashin, A. A.; Young, L.; Topol, I. A. *Biophys. Chem.* **1994**, *51*, 359−374.
(20) Davis, M. E.; McCammon, J. A. *Chem. Rev.* **1990**, *90*, 509−524.
(21) Born, M. Z. *Phys.* **1920**, *1*, 45.
(22) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *157*, 671−679.
(23) Klapper, I.; Hangstrom, R.; Fine, R.; Sharp, K.; Honig, B. *Proteins* **1986**, *1*, 1986.
(24) Gilson, M. K.; Sharp, K. A.; Honig, B. *J. Comput. Chem.* **1988**, *9*, 327−335.
(25) Sharp, K.; Jean-Charles, A.; Honig, B. *J. Phys. Chem.* **1992**, *96*, 3822−3828.
(26) Rashin, A. A. *J. Phys. Chem.* **1989**, *93*, 4664−4669.
(27) Juffer, A. H.; Botta, E. F. F.; van Keulen, B. A. M.; van der Plog, A.; Berendsen, H. J. C. *J. Comput. Phys.* **1991**, *97*, 144−171.
(28) Zauher, R. J.; Morgan, R. S. *J. Comput. Chem.* **1988**, *9*, 171−187.
(29) You, T. J.; Harvey, S. C. *J. Comput. Chem.* **1993**, *14*, 484−501.
(30) Sitkoff, D.; Ben-Tal, N.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 2744−2752.
(31) Gilson, M. K.; Rashin, A.; Fine, R.; Honig, B. *J. Mol. Biol.* **1985**, *88*, 503−516.
(32) Zauher, R. J.; Morgan, R. S. *J. Mol. Biol.* **1985**, *186*, 815−820.
(33) Luty, B. A.; Davis, M. E.; McCammon, J. A. *J. Comput. Chem.* **1992**, *13*, 768−771.
(34) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005−3014.
(35) Giesen, D. J.; Storer, J. W.; Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1995**, *117*, 1057−1068.
(36) Giesen, D. G.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1995**, *99*, 7137−7146.
(37) Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 16385−16398.
(38) Cramer, C. J.; Truhlar, D. G. *J. Comput. Aided Mol. Des.* **1992**, *6*, 629−666.
(39) Giesen, D. J.; Chambers, C. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, *101*, 2061−2069.
(40) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824−19839.
(41) Giesen, D. J.; Gu, M. Z.; Cramer, C. J.; Truhlar, D. G. *J. Org. Chem.* **1996**, *61*, 8720−8721.
(42) Edinger, S. R.; Cortis, C.; Shenkin, P. S.; Friesner, R. A. *J. Phys. Chem. B* **1997**, *101*, 1190−1197.
(43) Schmidt, A. B.; Fine, R. M. *Molecular Simulation* **1994**, *13*, 347−365.
(44) Klamt, A. *J. Phys. Chem.* **1995**, *99*, 2224−2235.
(45) Reynolds, C. H. J. *Chem. Inf. Comput. Sci.* **1995**, 738−742.
(46) Best, S. A.; Merz, K. M., Jr.; Reynolds, C. H. 1997. In preparation.
(47) West, R. C. *CRC Handbook of Chemistry and Physics, 62nd ed.;* West, R. C., Ed.; CRC Press, Inc.: Boca Raton, FL, 1981.
(48) Bernazzani, L.; Cabani, S.; Conti, G.; Mollica, V. *J. Chem. Soc. Faraday Trans.* **1995**, *91*, 649−655.
(49) Berti, P.; Cabani, S.; Conti, G.; Mollica, V. *J. Chem. Soc., Faraday Trans. 1* **1986**, *82*, 2547−2556.
(50) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490−512, 520−552, 553−586, 587−615, 616−641.
(51) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129−145.
(52) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431−439.
(53) Shrake, A.; Rupley, J. A. *J. Mol. Biol.* **1973**, *79*, 351−371.
(54) Richards, F. M. *Ann. Rev. Biophys. Bioeng.* **1977**, *6*, 151−176.
(55) Hasel, W.; Hendrickson, T. F.; Still, W. C. *Tetrahedron Computer Methodology* **1988**, *1*, 103−116.
(56) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, 11, 440.
(57) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435−445.
(58) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144−1149.
(59) Powell, M. J. D. *Comput. J.* **1969**, *7*, 155.
(60) Hestenes, M. *Conjugate Direction Methods in Optimization*; Springer-Verlag: New York, 1980.
(61) *JMP* Version 3.15; SAS Institute Inc., 1989−1995.
(62) *MOPAC* Version 6.0; Quantum Chemistry Program Exchange, Program 455; Indiana University: Bloomington, IN, 1990.
(63) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220, 221−264.
(64) Bayley, C. I.; Cieplak, P.; Cornell, W. D. *J. Phys. Chem.* **1993**, *97*, 10269−10280.