

MODELING THE EFFECT OF SOLVATION ON STRUCTURE, REACTIVITY, AND PARTITIONING OF ORGANIC SOLUTES: UTILITY IN DRUG DESIGN

CANDEE C. CHAMBERS*, DAVID J. GIESEN†, GREGORY D. HAWKINS‡,
CHRISTOPHER J. CRAMER‡, DONALD G. TRUHLAR‡, AND
WOUTER H.J. VAES§

Abstract. The SM x family of quantum mechanical solvation models accounts for electric and electronic polarization via the generalized Born model and for non-electrostatic components of solvation by microscopic surface tensions. The SM5.4 model, which is the most physical member of the SM x family, has been parameterized using two electronic Hamiltonians, AM1 and PM3. For both Hamiltonians, solvation parameters are obtained for water and for any organic solvent for which certain macroscopic data are available, in particular, the index of refraction, bulk surface tension, dielectric constant, and hydrogen bonding acidity and basicity as measured by the Abraham empirical α_2^H and β_2^H scales. For neutral solutes, the mean unsigned errors for aqueous and non-aqueous free energies of solvation are both 0.5 kcal/mol based on 215 and 1786 data points, respectively (for either Hamiltonian). By adding solvation effects to the gas-phase Hamiltonian, it is possible to model the effects of solvent on conformational analysis, molecular recognition, reaction kinetics, etc. The SM5.4 model is also useful for the calculation of solute partitioning between two solvents. Moreover, it is possible to generalize the SM5.4 model to media that are less well characterized than homogeneous solvents—an example is presented here for the case of bilayers of phosphatidyl choline—in order to model partitioning between biophases.

1. Introduction. By their very nature as molecular bullets aimed at (usually macromolecular) targets in condensed-phase systems, drugs, or more accurately their structure and reactivity, can only be realistically modeled within the context of their surroundings. For instance, it is well appreciated within the pharmaceutical industry that *in vitro* activity will not correlate with *in vivo* efficacy if drug delivery is stymied by insolubility in water. Efficiency of transport from one biophase to another, e.g., blood to cell membrane, can also play a critical role in drug bioactivity, and this efficiency is fundamentally related to the ability of the two different phases to differentially solvate the drug [1, 2]. Another example of the role of a drug's surroundings is that interaction of a drug with its target necessitates the desolvation of each over some fraction of their molecular surface areas. If those surfaces interact favorably with the surrounding biophase, the loss of this interaction must reduce binding efficiency.

In principle, all of these effects can be modeled from knowledge of the

*Departments of Physics and Chemistry, Mercyhurst College, 501 E. 38th St., Erie, PA 16546.

†Eastman Kodak Co., 1999 Lake Avenue, Rochester, NY 14650.

‡Department of Chemistry and Supercomputer Institute, 207 Pleasant St. SE, Minneapolis, MN 55455-0431.

§Research Institute of Toxicology, Utrecht University, P.O. Box 80176, NL-3508 TD Utrecht, The Netherlands.

free energies of solvation of the system at relevant sites (e.g., the drug in different phases or the isolated drug and receptor vs. the complex of the two). This chapter will focus on the use of the semiempirical SM5 family of quantum mechanical solvation models to calculate solvation and partitioning free energies for organic solutes. A special advantage of quantum mechanical solvation models is that they can predict the dependence of solvation free energy on conformation. Such information is important for modeling partitioning and binding phenomena, but only rarely is it available from experiment.

Section 2 reviews the functional forms of the models and documents their performance, Section 3 illustrates application of the models to two conformational analysis problems in solution, Section 4 describes the application of the models to compute partition coefficients between water and chloroform and water and a phospholipid bilayer, and Section 5 offers final remarks.

2. The SM5 models. Full descriptions of the SM5 models are presented elsewhere [3–8]. This section reviews those features that are most critical as background for the issues addressed in this chapter. The SM5 models are quantum statistical continuum solvation models. The solvent is modeled as a homogeneous surrounding medium that affects the quantum mechanical self-consistent-field equations for the solute orbitals, and the model includes additional terms designed to account for specific interactions between the solute and its surrounding first solvent shell. For all SM5 models (and for all previous SM x models), the standard-state free energy of solvation is partitioned into two contributions [9–13]

$$(1) \quad \Delta G_S^0 = \Delta G_{ENP} + G_{CDS}$$

where ΔG_{ENP} is the electrostatic component of the solvation free energy, i.e., the change in the electronic and nuclear internal energy of the solute and the electric polarization free energy of the solute-solvent system upon insertion of the solute in the solvent, and G_{CDS} is the contribution of non-electrostatic short- and medium-range effects to the standard-state free energy of transfer. The S in ΔG_S^0 stands for solvation, ENP stands for electronic-nuclear-polarization, and CDS stands for cavitation-dispersion-solvent-structure. All SM5 models use a standard state of 298 K and 1 M in both the gas phase and solution.

The remainder of Section 2 describes the steps employed to parameterize an SM5 model. That process involves both the ENP and the CDS terms, and we will discuss the procedure in that order. It is worth emphasizing here, however, that the characteristics that distinguish SM5 models from previous SM x solvation models [9, 14–21] are the use of fixed atomic radii for defining molecular cavities and surfaces of all atoms except (sometimes) hydrogens, the avoidance of strictly empirical modifications of the electrostatics (thereby placing explicitly empirical terms in the CDS term

rather than the ENP term of the equation above), and the use of geometry-dependent atomic surface tensions, as described more fully below.

2.1. Electrostatic components of solvation. The ΔG_{ENP} term can be further partitioned as

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

where ΔE_{EN} is the change in the electronic and nuclear energy of the solute in going from the gas phase to solution, and G_P is the electric polarization free energy of the solvent, due to solute-solvent interactions and the change in solvent-solvent interactions upon inserting the solute. In the SM5 models parameterized so far, the terms in equation 2 are calculated using neglect-of-diatomic-differential-overlap molecular orbital theory with either the Austin Model 1 [22–24] (AM1) or Parameterized Model 3 [25] (PM3) Hamiltonian; these choices are distinguished by the notations SM5.../AM1 and SM5.../PM3, respectively, where the ellipses represent additional notation required to fully specify the model (vide infra).

G_P is calculated from the partial atomic charges q_k according to the Generalized Born equation [9, 10, 17, 26–36]

$$(3) \quad G_P = -\frac{1}{2} \left(1 - \frac{1}{\varepsilon} \right) \sum_{k,k'} q_k q_{k'} \gamma_{kk'}$$

where ε is the solvent dielectric constant, k and k' label atoms, and $\gamma_{kk'}$ is a Coulomb integral that accounts for either the self-energy of a charge in a dielectric medium ($k = k'$) or the dielectric screening of the Coulomb interaction of two charges ($k \neq k'$). Additional details of the computation of $\gamma_{kk'}$ are provided in Section 2.1.2, but we note here that a very critical aspect of the screening algorithm is that it allows other parts of the solute to descreen each partial charge in terms of the solvent volume that they displace. Since the partial atomic charges used in equation 3 are derived from the wave function, the Fock operator must be modified to make the total energy, which includes this term, variational [3, 9–13, 20, 30–33, 37]. This is accomplished by including a self-consistent reaction field (SCRf) in the Fock operator of the Hartree-Fock equations. The Fock operator is modified by the dependence of the partial atomic charges on density matrix elements, and the next section discusses this dependence for the different SM5 models.

2.1.1. Partial atomic charges. We have previously identified four classes of partial atomic charges that find use in diverse modeling efforts [16, 38]. These classes are defined as: (I) charges postulated without reference to quantum mechanics, e.g., one might calculate partial charges for a diatomic molecule by dividing its dipole moment by its bond length, (II) charges extracted from a quantum mechanical wave function by analyzing the wave function using *ad hoc* prescriptions—of these, Mulliken population

analysis [39] is perhaps the most widely used, (III) charges extracted from a quantum mechanical wave function by analyzing a physical observable *predicted from the wave function*, e.g., the electrostatic potential [40–43] and (IV) methods that map Class II or Class III charges in a manner designed to improve agreement with a physical observable, e.g., the dipole moment [38], *as measured by experiment or converged quantum mechanics*. Different SM5 models are designed to use different classes of partial atomic charges, as further described in the next three sections.

2.1.1.1. SM5.4. The addition of “.4” to “SM5” indicates that the charge model employed for the electrostatic component of the free energy of solvation is a Class IV charge model. All SM5.4 calculations described in this paper are carried out by the SM5.4/AM1 model, in which the CM1A charge model [38], which maps AM1 zero-differential-overlap Mulliken charges to Class IV charges, is employed. (The SM5.4/PM3 solvation model uses the PM3-based charge model, CM1P [38], but is not discussed further here). The CM1 charge models were designed to deliver charges that reproduce gas-phase dipole moments for small neutral solutes as closely as possible [38]. Note that for polar neutral molecules the electrostatic potential at large distances is determined entirely by the dipole moment, so this is a special case of electrostatic potential fitting. Since the charge distribution of a large molecule is determined even at shorter distances primarily by bond dipoles and individual functional groups, a model that works well for dipole moments of small molecules should also work well for charge distributions of large molecules, if the model is properly formulated. The SM5.4 models have been parameterized for water [3, 5] and for any organic solvent for which five readily available descriptors are known, as described in more detail in Section 2.2 below.

In additional work not yet available at the time that this chapter is being prepared, our group is creating a series of 5.42R models based on AM1, PM3, *ab initio* Hartree-Fock, and density functional theory. These models involve a new class IV charge model called CM2, and are designed to be used with “rigid” gas-phase geometries. To denote these changes we add 2R to the names. We are also considering the preparation of a 5.42 non-rigid model based on AM1 and PM3 with CM2 changes. The SM5.42 models will be preferred to SM5.4 models for several reasons: (1) CM2 charges are more accurate than CM1 charges. (2) The SM5.42 models, coming later than the SM5.4 models, incorporate improvements in the training set, in some of the surface tension functionals, and in the empirical effective solvent radius.

2.1.1.2. SM5.2. The addition of “.2” to “SM5” indicates that the charge model employed for the electrostatic component of the free energy of solvation is a Class II charge model. For all SM5.2 calculations described in this paper, the charges employed are derived from Mulliken population analysis [39] of the zero-differential-overlap AM1 wave function. Although Class II charges are less realistic than Class IV charges, they are more efficient to calculate and do not require new parameterization, so they may

serve for instances where system complexity motivates the use of a faster method or for solutes involving atom types for which a Class IV charge model has not yet been parameterized. Models using Class II charges are, however, much less robust for applications to transition states or functional groups that were not included in the training set. The only currently available SM5.2 model is for water [5], and that model (called SM5.2PD/AM1) is only defined for the case where the dielectric descreening is calculated via the pairwise descreening approximation, as described more fully in Section 2.1.2.

A quartet of SM5.2R models with volume descreening are, however, currently being parameterized [8] for both water and organic solvents using the AM1, PM3, MNDO [44], and MNDO/d [45–47] electronic Hamiltonians and will probably be available by the time these proceedings are published. These models are called SM5.2R/AM1, SM5.2R/PM3, SM5.2R/MNDO, and SM5.2R/MNDO/d, where “R” again denotes that the solute geometry is rigid, i.e., not re-optimized in solution. The rigid models are parameterized using gas-phase geometrics. Thus geometry relaxation in solution is modeled only in an average implicit way through the parameters.

2.1.1.3. SM5.0R. The addition of “.0” to “SM5” indicates that *no* charge model is employed for the electrostatic component of the free energy of solvation, i.e., all charges are set to zero and consequently G_P is also zero (in fact, ΔG_{ENP} is zero) [8]. Thus, in the parameterization of an SM5.0R model, the model depends on the Hamiltonian only to the extent that different Hamiltonians predict different molecular geometries—there is no dependence on electronic density matrix elements through partial atomic charges. Furthermore, the total free energy of solvation is carried in the CDS term as described in Section 2.2 below, and this term is calculated without reoptimizing the gas-phase geometry in solution, although the empirical nature of the parameters make up for this computational restraint; we call the model SM5.0R to denote this “rigid” aspect. While the SM5.0R assumptions are not particularly physical, since they correspond to modeling long-range electrostatic effects entirely in terms of short-range functions involving surface areas, and since the geometry *does* change upon dissolution, the method can deliver high predictive accuracy for solutes with functional groups similar to those in the parameterization set, and calculation of the solvation free energy is extremely rapid given a previously obtained molecular geometry. All SM5.0R calculations described in this paper are for AM1 solute geometries and all solutes are uncharged. A modification of the SM5.0R model, SM5.05R, exists to handle certain kinds of charged solutes typically found in proteins, and interested readers are referred to the original paper for details [8].

2.1.2. Dielectric descreening. The polarization energy of a system comprised of two charged spheres in a homogeneous dielectric continuum can be solved analytically, but that ceases to be the case upon the in-

roduction of one or more additional spheres. If the solute is represented as a union of partially charged atomic spheres, accounting for the dielectric descreening that results from having additional solute atoms displace the dielectric medium must either be accomplished by solving the Poisson equation or by an approximation to this procedure [13]. The Generalized Born approach employs the latter method, calculating $\gamma_{kk'}$ in equation 3 by an empirical algorithm that accounts for descreening of solute atoms, i.e., displacement of a volume of the dielectric continuum by the remaining molecular volume of the solute [9, 17, 35] generalized Born model replaces the solvation of a differential equation (Poisson’s equation) by a quadrature. It involves determining effective atomic radii by a numerical integration of the free energy density surrounding each solute atom assuming that all other solute atoms displace the dielectric medium and assuming that the free energy density is given by Coulomb’s law for a single charge in the medium. It employs these radii in an empirical form for $\gamma_{kk'}$. Unless otherwise specified, it is this approach that is used for any SM5 model having non-zero ΔG_{ENP} .

An alternative method to calculate the dielectric descreening is to take advantage of the analytic form available for a system of two spheres, and to assume that the descreening of any solute atom can be calculated by summing all pairwise descreenings owing to other solute atoms [5, 18, 48]. Such an approach ignores overlap between the descreening spheres (i.e., some descreening will be counted multiple times because the descreened volume is shared by multiple spheres), but by scaling the atomic radii to compensate it is possible to account for that overlap in an average way. Such a descreening approach is referred to here as “pairwise descreening”, and is abbreviated PD. The SM5.2 models discussed below all employ the pairwise descreening approximation and the AM1 Hamiltonian, and are properly referred to as SM5.2PD/AM1 [5]. SM5.4PD models are also available [5].

2.2. Non-electrostatic components of solvation. As noted above, the non-electrostatic components of the solvation free energy are largely associated with interactions of the solute with its first solvation shell. In SM x models they are assumed to be proportional to the area of a surface through that shell. Thus, the SM5 first-solvation-shell term has the general form:

$$(4) \quad G_{CDS} = \sum_k \sigma_k A_k(R_S^{CD}) + \sigma^{CS} \sum_k A_k(R_S^{CS})$$

where k denotes an atom, $A_k(R_S^{CX})$ is the solvent-accessible surface area [17, 49] of atom k as calculated for a solvent effective radius R_S^{CX} , σ_k is the atomic surface tension (a surface tension is any quantity having units of energy per unit area), and σ^{CS} is a molecular surface tension calculated for a solvent effective radius R_S^{CS} . In the SM5.4 and SM5.2 models parameterized so far, $R_S^{CD} = R_S^{CS} = 1.7 \text{ \AA}$ in water, and $R_S^{CD} =$

$R_S^{CS}/2 = 1.7 \text{ \AA}$ in organic solvents. In the SM5.0R and SM5.05R models, we take $R_S^{CD} = R_S^{CS} = 0$. This reduces the solvent-accessible surface area to the van der Waals surface area. The atomic surface tensions depend upon the local geometry of the solute and one or more composite surface tension coefficients $\hat{\sigma}_Z, \hat{\sigma}_{ZZ'}$ and $\hat{\sigma}_{ZZ'}^{(2)}$ that are themselves dependent on the atomic numbers Z and Z' of the atoms and are parameters of the model. The set of functions describing these dependencies are part of what defines a model as being of the SM5 variety, and they have been described in detail [3–6]. The molecular surface tension σ^{CS} is a constant for a given solvent.

We note that in the parameterization of all SM5 models except the R models, solute geometries and electronic wave functions were allowed to fully relax in response to the surrounding solvent. All surface tension coefficients are different (separately optimized) for SM5.4/AM1 and SM5.4/PM3, but all other parameters are the same for the /AM1 and /PM3 cases. We note again that we will discuss only /AM1 models in this chapter.

In the special case of water, since the solvent radii R_S^{CD} and R_S^{CS} are both taken to be 1.7 \AA equation 4 simplifies to

$$(5) \quad G_{CDS} = \sum_k \sigma_k A_k(R_S).$$

As a consequence, σ^{CS} is not a separately optimized parameter in the SM5 water models, but is absorbed into the optimization of the atomic surface tension coefficients contributing to each σ_k . The parameterization procedure for water involves calculation of the electrostatic component of the free energy of solvation for a large number of solutes for which experimental data are available, subtracting this value from the experimental free energy of solvation, and optimizing, by linear regression, values for the atomic surface tensions. Details on the performance of the water models can be found in the next section.

In the case of organic solvents, it is uncommon for a single solvent to have sufficient data available to effect a parameterization along the lines of the water models. We addressed this challenge by making the atomic surface tensions themselves be parametric functions of certain macroscopic *solvent* properties; this ensures that the model can be employed for *all* organic solvents for which those properties are available or estimable. To accomplish this, the surface tensions appearing in equation 4 are calculated as

$$(6) \quad \sigma^{CS} = \hat{\sigma}^{CS,n} n + \hat{\sigma}^{CS,\gamma} \gamma$$

$$(7) \quad \bar{\sigma}_Z = \hat{\sigma}_Z^{(n)} n + \hat{\sigma}_Z^{(\alpha)} \alpha + \hat{\sigma}_Z^{(\beta)} \beta$$

$$(8) \quad \bar{\sigma}_{ZZ'} = \hat{\sigma}_{ZZ'}^{(n)} n + \hat{\sigma}_{ZZ'}^{(\alpha)} \alpha + \hat{\sigma}_{ZZ'}^{(\beta)} \beta$$

and

$$(9) \quad \tilde{\sigma}_{ZZ'}^{(2)} = \hat{\sigma}_{ZZ'}^{(2,n)} n$$

where n is the solvent index of refraction, γ is the macroscopic solvent surface tension, in units of $\text{cal mol}^{-1} \text{ \AA}^{-2}$, and α and β are Abraham's [50–52] indices of solvent hydrogen bonding acidity and basicity, respectively (more specifically, these are the $\sum \alpha_2^H$ and $\sum \beta_2^H$ values for a solvent molecule were it to be taken as a *solute* in the Abraham model). The solvent radii R_S^{CD} and R_S^{CS} are taken to be 1.7 and 3.4 \AA , respectively. Two radii were used in order to capture both short-range (cavitation-dispersion) and intermediate-range (cavitation-solvent-structure) effects. Equations 6–9 are used directly for all solvents in the SM5.42R, SM5.2R, and SM5.0R models, and in the SM5.4 models they are used for all solvents except chloroform, benzene, and toluene solvents. In the latter three cases, further adjustments were made to the parameters, and all calculations for these solvents include those adjustments.

2.3. Performance of the SM5 models. Table 1 provides information on the data sets used for parameterization of the various models, and Table 2 provides information on the accuracies of the various models broken down by chemical functionality of solute molecules. Although the SM5.0R, SM5.2PD/AM1, SM5.4PD/AM1, and SM5.4/AM1 models all offer similar accuracy by comparison to experiment, they differ in the degree to which the *ENP* and *CDS* components of the total solvation free energy can be regarded as physical. The level of interpretability is inversely correlated with computational speed, with SM5.4/AM1 being most physical but also most time-consuming, followed by SM5.4PD/AM1, which maintains the use of Class IV charges but approximates the descreening via the pairwise approach, followed by SM5.2PD/AM1 which further approximates the electrostatics by using Class II Mulliken charges, followed at last by SM5.0R, which does not make any attempt to separate electrostatic and non-electrostatic effects, but which is extremely rapid for the computation of solvation free energies, since an electronic structure calculation is not required; rather, for a given geometry, one simply requires a calculation of the solvent exposed surface area.

We now proceed to employ our most physical solvation model, SM5.4/AM1, for the calculation of solvent effects on conformational analysis (next section) and for the computation of partition coefficients (Section 4). In the latter instance, we illustrate how the model may be generalized to media not well described as homogeneous liquid solutions.

3. Solvent effects on conformational analysis. When a solute has multiple low-energy conformations, the population of those conformations at equilibrium is determined by their relative energies in a Boltzmann-

TABLE 1

Data employed in the parameterization of the SM5.4/AM1 model for aqueous and organic solvents and the SM5.4PD/AM1, and SM5.2PD/AM1 models for aqueous solvent.

Solute class	Water	Nonaqueous Solvents		
	number of data	number of solutes	number of solvent classes ^b	number of data
unbranched alkanes	8	9	15	76
branched alkanes	5	5	2	7
cycloalkanes	5	4	4	13
alkenes	9	8	3	18
alkynes	5	5	2	9
arenes	8	9	15	126
alcohols	16	17	15	369
ethers	9	9	15	71
aldehydes	6	7	5	32
ketones	12	12	14	191
carboxylic acids	5	5	11	119
esters	12	12	6	227
aliphatic amines	15	11	8	153
aromatic amines	10	11	10	71
nitriles	4	4	4	18
nitrohydrocarbons	6	6	6	32
amides	3 ^b	1	3	4
thiols	4	3	3	10
organic sulfides	4	4	3	10
disulfides	2	2	2	3
nonhalobifunctional	8	6	6	26
fluorohydrocarbons	3	5	3	13
chlorohydrocarbons	16	14	4	58
bromohydrocarbons	10	10	4	30
iodohydrocarbons	8	9	4	20
polyfunctional halocarbons	17	13	7	44
inorganic compounds	5	5	9	36
all solutes	215	206	5	1786

^a For the parameterization of the SM5.0R model for water several additional solute data were added: one urea and several halogenated compounds.

^b Solvent classes are defined as alkanes, arenes, alcohols, ketones, esters, ethers, amines, pyridines, nitriles, nitrohydrocarbons, tertiary amides, haloaliphatics, haloaromatics, miscellaneous acidic solvents, and miscellaneous non-acidic solvents (15 total solvent classes).

TABLE 2
Mean unsigned error (kcal/mol) for various SM5 models by class of solute^a.

Solute class	Aqueous			Nonaqueous	
	SM5.0R	SM5.2PD /AM1	SM5.4PD /AM1	SM5.4 /AM1	SM5.4 /AM1
unbranched alkanes	0.5	0.2	0.5	0.6	0.4
branched alkanes	0.3	0.2	0.3	0.7	0.4
cycloalkanes	0.5	0.2	0.2	0.2	1.0
alkenes	0.2	0.4	0.6	0.5	0.5
alkynes	0.1	0.2	0.2	0.3	0.3
arenes	0.5	0.2	0.2	0.1	0.3
alcohols	0.3	0.4	0.4	0.5	0.4
ethers	0.7	0.5	0.5	0.8	0.4
aldehydes	0.3	0.3	0.3	0.3	0.5
ketones	0.4	0.4	0.2	0.4	0.4
carboxylic acids	0.5	0.7	1.0	0.8	0.6
esters	0.3	0.6	0.7	0.4	0.5
aliphatic amines	0.5	0.5	0.7	0.8	0.4
aromatic amines	1.0	0.6	0.7	0.7	0.4
nitriles	0.6	0.3	0.4	0.5	0.4
nitrohydrocarbons	0.4	0.4	0.5	0.5	0.5
amides	2.8	1.2	1.5	1.8	1.1
thiols	0.3	0.6	0.4	0.3	0.4
organic sulfides	0.3	0.7	0.4	0.4	0.4
disulfides	0.1	0.4	0.3	0.2	0.0
nonhalobifunctional	0.8	0.7	0.5	0.4	0.9
fluorohydrocarbons	1.1	0.4	0.4	0.5	0.4
chlorohydrocarbons	0.5	0.5	0.4	0.4	0.3
bromohydrocarbons	0.4	0.5	0.2	0.2	0.3
iodohydrocarbons	0.9	0.4	0.1	0.2	0.2
polyfunctional halocarbons	0.9	0.7	0.5	0.5	0.6
inorganic compounds	0.8	1.7	1.7	1.8	1.4
all solutes	0.6	0.5	0.5	0.5	0.5

^a All calculations for AM1 Hamiltonian. These errors do include the additional data mentioned in footnote a of Table 1.

weighted sense. The fraction of the molecules found in conformation *A* is

$$(10) \quad F(A) = \frac{e^{-G_A^0/RT}}{\sum_i e^{-G_i^0/RT}}$$

where G_A^0 is the free energy of conformer *A*, and *R* and *T* are the universal gas constant and the temperature, respectively.

Since different conformers may have different interactions with the surrounding solvent (because of different charge distributions or different

degrees of exposure of solvophobic or solvophilic surface area, for example) the equilibrium distribution of conformers is typically modified by solvation (i.e., the relative G_A^0 values of various conformers may be different in the gas phase than in solution). The SM x solvation models have previously been used to investigate aspects of the aqueous conformational equilibria of heterocycles (where tautomeric equilibria have also been extensively studied) [13, 53–55], sugars [19, 56], cocaine [57], HIV protease inhibitors [58], neurotransmitters [59], nitroaromatic radical anions [60], 1,3-dimethylthiourea [61], 1,2-ethanediol [62], diazene [63], and allyl vinyl ether [64]. The next two sections discuss aspects of the anomeric equilibrium of D-glucose in water and the *gauche/anti* equilibrium of 1,2-dichloroethane in multiple solvents, both being modeled using the SM5.4 method.

3.1. Conformational isomerism of D-glucose. Portions of the conformational potential energy surface for D-glucose have been studied by many researchers [19, 56, 65–76]. The significant conformational issues addressed in these studies revolve around one or more of the following questions: (1) What is the preferred configuration at the anomeric center, and how much higher in energy is the other anomer? (2) What are the preferred orientations of the exocyclic hydroxyl and hydroxymethyl substituents? (3) What is the energetic separation between the two possible chair forms? (4) What are the energetic consequences that arise from the coupling of different conformational issues, e.g., to what extent can one consider the anomeric equilibrium as separate from the hydroxymethyl conformation? (5) How does a surrounding solvent change the intrinsic (gas-phase) conformational preferences? In particular, for hydrogen-bond donor and/or acceptor solvents, to what extent do intermolecular hydrogen bonds between glucose and the solvent compete with intramolecular hydrogen bonds between different glucose hydroxyl groups?

Figure 1 presents the relative free energies for six glucose conformers (three hydroxymethyl rotamers of each anomer) in the gas phase and in aqueous solution. The gas phase results are derived from composite calculations including examination of basis set effects up to valence quadruple zeta, electron correlation effects through coupled cluster with single and double substitutions, geometry optimization using second-order perturbation theory, and thermochemical corrections using the standard rigid-rotor, harmonic oscillator approach [77]. This composite approach is described in greater detail elsewhere [19, 78]. The aqueous relative energies are derived by adding SM5.4/AM1 aqueous solvation free energies to the gas-phase values at the best gas-phase geometries. (Thus we make the approximation of neglecting geometry-relaxation effects in solution.) In both the gas phase and solution, the energy of the αG conformer is taken as 0, and other free energies are relative to this. The bottom rows of the table correspond to the average over the 3 conformations, calculated using equation 10, and to the experimental result; both these rows are relative to the α case, which is arbitrarily set to 0.

As is evident from Figure 1, the experimentally known [79] free energy of anomerization, 0.3 kcal/mol for the $\alpha \rightarrow \beta$ process, is well predicted by the SM5.4/AM1 calculations, which agree with experiment within 0.2 kcal/mol, and the effect of solvation is predicted to stabilize the beta anomer by 0.8 kcal/mol relative to the alpha anomer.

Although this example is averaged over only six conformers, more complete surveys of conformational space using classical mechanical models are entirely consistent with an anomeric solvation effect of this magnitude [73]. The SM5.4/AM1 model has also been used to examine the effect of aqueous solvation on other aspects of glucose conformational analysis [80].

The calculations in Figure 1 are particularly satisfying because they are based on combining state-of-the-art gas-phase calculations with a carefully parameterized solvation model. Previous attempts to understand this issue often suffered from deficiencies in the gas-phase relative free energy as well as from difficulties in the solvation calculation. Another point worth stressing in this example is that SM5.4 models do not suffer from having to make *ad hoc* assumptions about the dependence of partial atomic charges on torsional angles or on the presence of solvent; in SM5.4 models these charges are predicted naturally as an intrinsic part of the self-consistent-field calculation.

3.2. Internal rotational isomerism of 1,2-dichloroethane. As a disubstituted ethane, 1,2-dichloroethane can adopt two equilibrium conformers that differ as rotamers about the C-C bond. One rotamer places the two chlorines *trans* to each other (and by symmetry this conformer has a dipole moment of zero), and the other places the two chlorines *gauche* to each other. One would expect the electrostatic interactions of the *gauche* conformer with the dielectric medium to be greater in magnitude than those of the *trans* conformer, because of the finite dipole moment for the former. Of course, other moments play a role as well, and the generalized Born approach accounts for these (unlike, say, the Kirkwood-Onsager model [13, 81, 82], which, by truncating the electronic distribution at the dipole moment, would predict an electrostatic free energy of solvation for the *trans* conformer of zero in all solvents). Non-electrostatic effects will also be different for the two conformers. Table 3 compares the predicted solvent-induced shifts in the *gauche/trans* equilibrium to experiment [83] for eight organic solvents. Agreement is within 0.07 kcal/mol for all solvents except benzene and diethyl ether, where the error is 0.22 and 0.14 kcal/mol, respectively.

This example illustrates the quantitative applicability of the SM5.4/AM1 model to conformational equilibria both in a single solvent and also across multiple solvents. We anticipate that these models will prove useful in examining conformational equilibria for drug molecules, both for understanding their binding and also for modeling their partitioning from one phase to another, as described more fully in the following section.

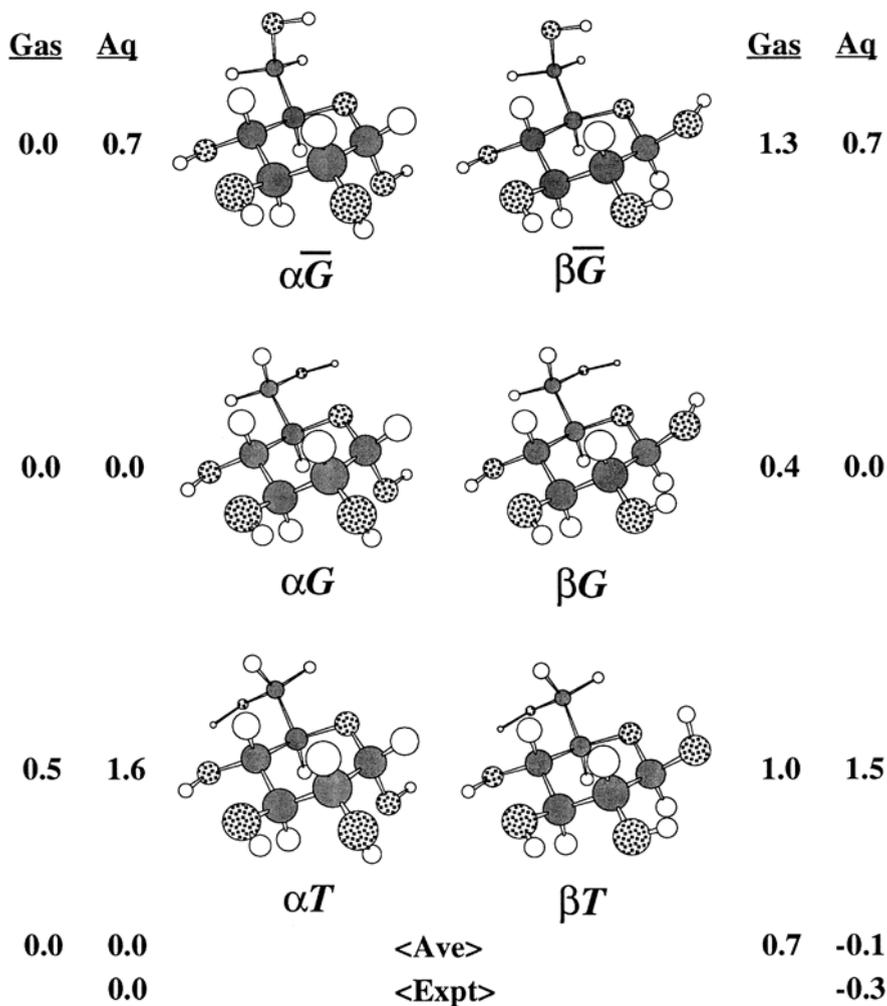
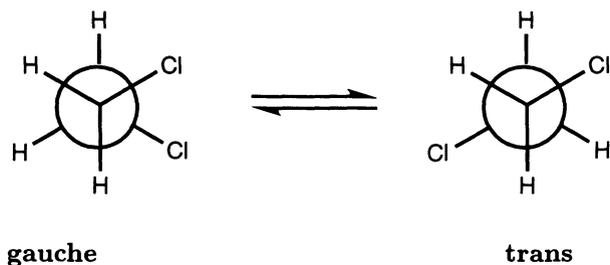


FIG. 1. Relative free energies (kcal/mol) in the gas phase and in aqueous solution for six glucose conformers (identified as α or β and T , G , or \bar{G} hydroxymethyl rotamers). Gas-phase free energies were calculated from a composite scheme using basis sets up to $cc\text{-}pVQZ$, accounting for electron correlation at the $CCSD$ level, and vibrational frequencies from the $HF/cc\text{-}pVDZ$ level. Aqueous free energies were derived by adding $SM5.4/AM1$ solvation free energies, calculated without geometry relaxation, to the gas-phase values.

TABLE 3
 Magnitude of the solvent-induced shift (kcal/mol) in the gauche/trans equilibrium of 1,2-dichloroethane relative to the gas phase.



Solvent	$\Delta\Delta G_s^0$	
	SM5.4/AM1	Experimental ^a
Tetrachloroethylene	0.38	0.31
Carbon disulfide	0.44	0.37
Diethyl ether	0.65	0.51
Benzene	0.38	0.60
Ethyl acetate	0.76	0.78
1,2-Dichloroethane (neat liquid)	0.87	0.89
Acetone	0.96	1.02
Acetonitrile	1.00	1.05

^a Reference [83].

4. Solute partitioning. Solute molecules will partition between multiple phases so as to equalize the concentration-dependent chemical potential of the solute in all phases [84–86]. This behavior affects the bioactivity of the solute. For example, a drug may have very high specific activity against a target enzyme, but if its concentration in aqueous biophases or fatty tissues is extremely low it may not be suitable for oral delivery (since it must be carried by the bloodstream, and it must pass through a lipid bilayer membrane to be absorbed in the small intestine). This illustrates the practical importance of being able to predict concentrations of solutes in different media.

The following two sections focus on the partitioning of solutes between two homogeneous liquid solutions (chloroform and water) and between water and a phospholipid bilayer. Prior studies using the SM x models to correlate solute partitioning between two phases have also appeared [6, 87–90].

TABLE 4
 log $P_{c/w}$ predicted by two different methods vs. experiment.

	log $P_{c/w}$		
	SM5.4/AM1 ^a	Reynolds ^b	Experiment ^c
benzene	2.7	4.3	2.8
toluene	3.3	5.0	3.1
ethylbenzene	3.9	5.5	3.7
methanol	-2.0	-1.5	-1.3
ethanol	-0.9	-0.5	-0.9
2-methylpropan-1-ol	0.7	1.0	0.3
1-pentanol	1.1	1.1	1.1
phenol	0.4	1.0	0.4
1-hexanol	1.6	1.6	1.7
1-naphthol	2.2 ^d	2.3	1.2
ethanal	-0.7	-0.6	0.1
propanone	0.1	1.6	0.4
methyl phenyl ketone	2.3	3.5	2.4
ethanoic acid	-1.3	-1.4	-1.4
butanoic acid	0.2	-0.1	-0.3
hexanoic acid	1.6	1.0	1.0
<i>p</i> -toluic acid	1.8	1.4	1.4
methylamine	-1.5	-0.6	-1.0
diethylamine	1.4	2.8	0.9
pyridine	0.5	2.1	1.3
benzotrile	3.0	3.1	2.3
2-propen-1-ol	-0.4	0.1	-0.5
ethanamide	-2.2	-3.6	-2.0
chlorobenzene	3.1	4.4	3.2
bromobenzene	3.2	4.3	3.4
<i>p</i> -bromophenol	1.0	0.8	1.1
<i>p</i> -hexylpyridine	4.4	5.4	5.0
propanamide	-1.2	-2.7	-1.4
<i>m</i> -chlorophenol	0.9 ^d	0.9	1.0
<i>p</i> -chlorophenol	0.9	0.9	1.0
Mean signed error	0.00	0.44	
Mean unsigned error	0.32	0.73	
Root-mean-square error	0.42	0.95	

^a Reference [6]. ^b Reference [91]. ^c Reference [92]. ^d Statistically averaged for two hydroxyl rotamers.

4.1. Partitioning between two homogeneous liquids. A reasonable number of data are available for the partitioning of an organic solute between chloroform and water. Reynolds [91] has emphasized the utility of water/chloroform partition coefficients for correlating membrane permeability and bioactivity that depends on this permeability. Reynolds [91] has developed this correlation based on calculations using the GB/SA model of Still *et al.*, [35] which is a nonpolarizable molecular mechanics model with fixed charges and a single molecular surface tension. Table 4 compares the performance of our SM5.4/AM1 models [6] to that of Reynolds' calculations [91] for the training set he used to parameterize a subsequent quantitative structure-activity relationship for lipophilicity. The polarizable SM5.4/AM1 models outperform the classical model for the 30 solutes considered by Reynolds. Of these 30 molecules, the first 23 in Table 4 were included in the SM5.4/AM1 Nonaqueous Solvent Parameterization Set (described in Table 1 above), but the remaining 7 molecules were not included in the parameterization set. The results in Table 4 show that the latter molecules are treated with equivalent accuracy. In particular, the mean unsigned error for the first 23 molecules is 0.31 kcal/mol, whereas the mean unsigned error for the last 7 is 0.37 kcal/mol.

4.2. Partitioning between water and phosphatidyl choline bilayers. Although the SM5.4 model for organic solvents was parameterized using data for 90 commonly available solvents, there is every reason to expect that the model will be applicable to *any* condensed phase that has liquid-like properties, e.g., a lipid membrane above its melting temperature. In such an instance, of course, it is not obvious how to go about choosing "correct" values of ϵ , n , γ , α and β , as are required to define an organic SM5.4 model. However, the solvation free energy as calculated at the SM5.4 level is multilinear in the latter four parameters. A practical approach to dealing with such ill-defined systems is to treat these latter four parameters as fitting parameters—but because the solvation free energy is non-linear in ϵ , it is more practical to fix this value at some best estimate.

An example of this approach is illustrated for the case of water/phosphatidyl choline unilamellar bilayer partitioning. Experimental partitioning data in this system are available [93] for the solutes illustrated in Figure 2. In the case of phosphatidyl choline, the dielectric constant is unknown; it is probably not unreasonable to assume an effective value that is intermediate between methyl nonadecanoate ($\epsilon = 3.0$) [94] and tricresyl phosphate ($\epsilon = 6.7$) [94], and so we used a value of 5. We further assumed that the α value is zero (there are no hydrogen bond donor groups in this molecule), and we regressed the experimental partition coefficients on the remaining three solvent parameters, namely n , γ , and β , employing the AM1 electronic Hamiltonian for the solutes. The regression provides values of 1.40, 27, and 1.15, respectively for these parameters. Such parameter values are very much in keeping with what one might expect based on

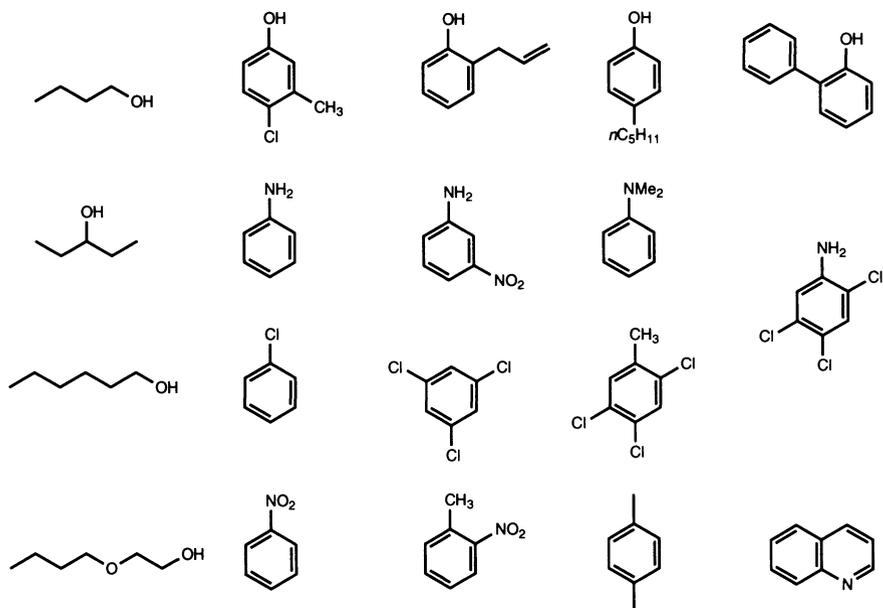


FIG. 2. Organic solutes for which water/phosphatidyl choline bilayer partitioning data are available.

analogous molecules as solvents. The regression was also allowed to have a constant term, which is calculated to be 0.59 log units. The quality of the regression, which has an R^2 correlation coefficient of 0.80, is illustrated in Figure 3.

The finding of a constant term in the regression is probably attributable to the amphiphilic nature of several of the solutes. The free energy of solvation calculated by our model for phosphatidyl choline corresponds to complete immersion of the solute in a homogeneous phosphatidyl choline medium. However, many of the experimental solutes presumably maintain hydrophilic contact with the aqueous phase, but bury hydrophobic surface in the lipid phase. This makes the solutes undetectable to the experimental [93] analytic method (which involves microsyringe sampling of an aqueous phase in which the phosphatidyl choline bilayers are suspended), and this makes it appear that the solutes are more soluble in the bilayer than they really are. A more refined model of this process would be one in which a planar boundary separates two homogeneous dielectric media, and this would permit an examination of solute behavior near or at the interface. However, the model presented here is expected to have general utility for the examination of solute partitioning between phases that would otherwise not easily be characterized as solvents, per se.

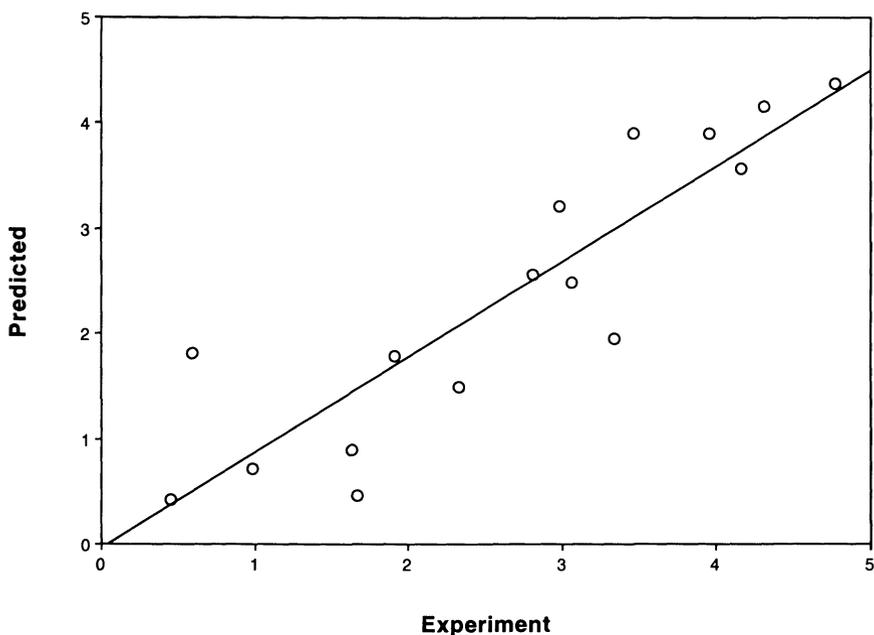


FIG. 3. Predicted phosphatidyl choline bilayer/water partition coefficients (defined as [molar concentration in phosphatidyl choline bilayer]/ [molar concentration in water] in \log_{10} units) vs. experiment for the organic solutes in Figure 2.

This example illustrates a very general approach to using the SM x models for drug design, an approach that we have labeled as the SRP approach to denote specific reaction parameters or specific range parameters. In this approach we adopt the general framework of the model and most of its parameters from a general parameterization (e.g., in the present case the Coulomb radii, van der Waals radii, and surface tensions were taken from the general SM5.4/AM1 model, but a few parameters (in this case solvent parameters) are adjusted to a specific small data set on a limited range of molecules, e.g., a lead molecule and selected analogs. Having done this, the model could be used for computational screening of other lead analogs as a guide to which ones are most likely to be worthwhile for synthesis. The SRP approach has also been used in other contexts [16, 19, 95–97] and we anticipate it could be extremely powerful for rational drug design.

5. Concluding remarks. SM5 solvation models are designed to offer flexibility to the user in terms of choosing a model based on a balance of speed and interpretability. For neutral organic solutes for which experimental free energies of transfer from the gas phase to a solvent are available, the models typically have mean unsigned errors on the order of 0.5 kcal/mol. When examining solvent effects in individual systems, these errors can cancel out for similar solutes (e.g., conformational isomers) per-

mitting very accurate calculations of solvent effects on conformational equilibria. Furthermore, by being able to calculate free energies of solvation in multiple solvents, it is possible to calculate free energies of partitioning between solvents in a highly accurate fashion. Finally, the SM5.4 model for organic solvents lends itself to specific parameterization for more complicated condensed phases, like lipid bilayers. We anticipate that several of these latter features will make the models useful tools in drug discovery and development.

Acknowledgments. We are grateful for high-performance vector and parallel computing resources made available by the Minnesota Supercomputer Institute and the University of Minnesota-IBM Shared Research Project, respectively. This work was supported by the National Science Foundation, the U.S. Army Research Office, and the Alfred P. Sloan Foundation.

REFERENCES

- [1] HANSCH, C.; FUJITA, T., *J. Am. Chem. Soc.*, 1964, **86**, 1616.
- [2] SILVERMAN, R.B., *The Organic Chemistry of Drug Design and Drug Action*, Academic: San Diego, 1992, 26-34.
- [3] CHAMBERS, C.C.; HAWKINS, G.D.; CRAMER, C.J.; TRUHLAR, D.G., *J. Phys. Chem.*, 1996, **100**, 16385.
- [4] GIESEN, D.J.; GU, M.Z.; CRAMER, C.J.; TRUHLAR, D.G., *J. Org. Chem.*, 1996, **61**, 8720.
- [5] HAWKINS, G.D.; CRAMER, C.J.; TRUHLAR, D.G., *J. Phys. Chem.*, 1996, **100**, 19824.
- [6] GIESEN, D.J.; CHAMBERS, C.C.; CRAMER, C.J.; TRUHLAR, D.G., *J. Phys. Chem.*, 1997, **101**, 2061.
- [7] GIESEN, D.J.; HAWKINS, G.D.; LIOTARD, D.A.; CRAMER, C.J.; TRUHLAR, D.G., *Theor. Chem. Acc.*, in press.
- [8] HAWKINS, G.D.; CRAMER, C.J.; TRUHLAR, D.G., to be published.
- [9] CRAMER, C.J.; TRUHLAR, D.G., *J. Am. Chem. Soc.*, 1991, **113**, 8305.
- [10] CRAMER, C.J.; TRUHLAR, D.G., *J. Comput.-Aid. Mol. Des.*, 1992, **6**, 629.
- [11] CRAMER, C.J.; TRUHLAR, D.G., In *Quantitative Treatments of Solute/Solvent Interactions*; Politzer, P. and Murray, J.S., Eds.; Elsevier: Amsterdam, 1994, **1**, 9.
- [12] CRAMER, C.J.; TRUHLAR, D.G., In *Reviews in Computational Chemistry*; Lipkowitz, K.B. and Boyd, D.B., Eds.; VCH: New York, 1995, **6**, 1.
- [13] CRAMER, C.J.; TRUHLAR, D.G., In *Solvent Effects and Chemical Reactivity*; Tapia, O. and Bertrán, J., Eds.; Kluwer: Dordrecht, 1996, 1.
- [14] CRAMER, C.J.; TRUHLAR, D.G., *Science*, 1992, **256**, 213.
- [15] CRAMER, C.J.; TRUHLAR, D.G., *J. Comp. Chem.*, 1992, **13**, 1089.
- [16] STORER, J.W.; GIESEN, D.J.; HAWKINS, G.D.; LYNCH, G.C.; CRAMER, C.J.; TRUHLAR, D.G.; LIOTARD, D.A., In *Structure and Reactivity in Aqueous Solution*; Cramer, C.J. and Truhlar, D.G., Eds.; American Chemical Society: Washington, DC, 1994, 24.
- [17] LIOTARD, D.A.; HAWKINS, G.D.; LYNCH, G.C.; CRAMER, C.J.; TRUHLAR, D.G., *J. Comp. Chem.*, 1995, **16**, 422.
- [18] HAWKINS, G.D.; CRAMER, C.J.; TRUHLAR, D.G., *Chem. Phys. Lett.*, 1995, **246**, 122.

- [19] BARROWS, S.E.; DULLES, F.J.; CRAMER, C.J.; TRUHLAR, D.G.; FRENCH, A.D., *Carbohydr. Res.*, 1995, **276**, 219.
- [20] GIESEN, D.J.; CRAMER, C.J.; TRUHLAR, D.G., *J. Phys. Chem.*, 1995, **99**, 7137.
- [21] GIESEN, D.J.; STORER, J.W.; CRAMER, C.J.; TRUHLAR, D.G., *J. Am. Chem. Soc.*, 1995, **117**, 1057.
- [22] DEWAR, M.J.S.; ZOBISCH, E.G.; HEALY, E.F.; STEWART, J.J.P., *J. Am. Chem. Soc.*, 1985, **107**, 3902.
- [23] DEWAR, M.J.S.; ZOBISCH, E.G., *J. Mol. Struct. (Theochem)*, 1988, **180**, 1.
- [24] DEWAR, M.J.S.; YATE-CHING, Y., *Inorg. Chem.*, 1990, **29**, 3881.
- [25] STEWART, J.J.P., *J. Comp. Chem.*, 1989, **10**, 221.
- [26] HOIJTINK, G.J.; DE BOER, E.; VAN DER MEIJ, P.H.; WEIJLAND, W.P., *Recl. Trav. Chim. Pays-Bas*, 1956, **75**, 487.
- [27] PERADEJORDI, F., *Cahiers Phys.*, 1963, **17**, 343.
- [28] JANO, I., *Compt. Rend. Acad. Sci. Paris*, 1965, **261**, 103.
- [29] KLOPMAN, G., *Chem. Phys. Lett.*, 1967, **1**, 200.
- [30] TAPIA, O., In *Quantum Theory of Chemical Reactions*; Daudel, R., Pullman, A., Salem, L. and Viellard, A., Eds.; Reidel: Dordrecht, 1980, **2**, 25.
- [31] CONSTANCIEL, R.; CONTRERAS, R., *Theor. Chim. Acta*, 1984, **65**, 1.
- [32] CONTRERAS, R.; AIZMAN, A., *Int. J. Quant. Chem.*, 1985, **27**, 293.
- [33] KOZAKI, T.; MORIHASHI, M.; KIKUCHI, O., *J. Am. Chem. Soc.*, 1989, **111**, 1547.
- [34] TUCKER, S.C.; TRUHLAR, D.G., *Chem. Phys. Lett.*, 1989, **157**, 164.
- [35] STILL, W.C.; TEMPczyk, A.; HAWLEY, R.C.; HENDRICKSON, T., *J. Am. Chem. Soc.*, 1990, **112**, 6127.
- [36] KING, P.M., In *Computer Simulation of Biomolecular Systems*; van Gunsteren, W.F., Weinger, P.K. and Wilkinson, A.J., Eds.; ESCOM: Leiden, 1993, **2**, 267.
- [37] KIKUCHI, O.; MATSUOKA, T.; SAWAHATA, H.; TAKAHASHI, O., *J. Mol. Struct. (Theochem)*, 1994, **305**, 79.
- [38] STORER, J.W.; GIESEN, D.J.; CRAMER, C.J.; TRUHLAR, D.G., *J. Comput.-Aid. Mol. Des.*, 1995, **9**, 87.
- [39] MULLIKEN, R.S., *J. Chem. Phys.*, 1955, **23**, 1833.
- [40] SINGH, U.C.; KOLLMAN, P.A., *J. Comp. Chem.*, 1984, **5**, 129.
- [41] BESLER, B.H.; MERZ, K.M.; KOLLMAN, P.A., *J. Comp. Chem.*, 1990, **11**, 431.
- [42] MERZ, K.M., *J. Comp. Chem.*, 1992, **13**, 749.
- [43] CORNELL, W.D.; CIEPLAK, P.; BAYLY, C.I.; KOLLMAN, P.A., *J. Am. Chem. Soc.*, 1993, **115**, 9620.
- [44] DEWAR, M.J.S.; THIEL, W., *J. Am. Chem. Soc.*, 1977, **99**, 4899.
- [45] THIEL, W.; VOITYUK, A.A., *Int. J. Quant. Chem.*, 1992, **44**, 807.
- [46] THIEL, W.; VOITYUK, A.A., *Theor. Chim. Acta*, 1992, **81**, 391.
- [47] KOLB, M.; THIEL, W., *J. Comp. Chem.*, 1993, **14**, 775.
- [48] SCHAEFER, M.; FROEMMEL, C., *J. Mol. Biol.*, 1990, **216**, 1045.
- [49] LEE, B.; RICHARDS, F.M., *J. Mol. Biol.*, 1971, **55**, 379.
- [50] ABRAHAM, M.H.; GRELLIER, P.L.; PRIOR, D.V.; DUCE, P.P.; MORRIS, J.J.; TAYLOR, P., *J. Chem. Soc., Perkin Trans.*, 1989, **2**, 699.
- [51] ABRAHAM, M.H., *Chem. Soc. Rev.*, 1993, 73.
- [52] ABRAHAM, M.H., In *Quantitative Treatments of Solute/Solvent Interactions*; Politzer, P. and Murray, J.S., Eds.; Elsevier: Amsterdam, 1994, 83.
- [53] CRAMER, C.J.; TRUHLAR, D.G., *J. Am. Chem. Soc.*, 1991, **113**, 8552 and 9901(E).
- [54] CRAMER, C.J.; TRUHLAR, D.G., *J. Am. Chem. Soc.*, 1993, **115**, 8810.
- [55] CONTRERAS, J.G.; ALDERETE, J.B., *J. Mol. Struct. (Theochem)*, 1994, **309**, 137.
- [56] CRAMER, C.J.; TRUHLAR, D.G., *J. Am. Chem. Soc.*, 1993, **115**, 5745.
- [57] YANG, B.; WRIGHT, J.; ELDEFRAWI, M.E.; POU, S.; MACKERELL, A.D., *J. Am. Chem. Soc.*, 1994, **116**, 8722.
- [58] RAO, B.G.; KIM, E.E.; MURCKO, M.A., *J. Comput.-Aid. Mol. Des.*, 1996, **10**, 23.
- [59] URBAN, J.J.; CRAMER, C.J.; FAMINI, G.R., *J. Am. Chem. Soc.*, 1992, **114**, 8226.
- [60] BARROWS, S.E.; CRAMER, C.J.; TRUHLAR, D.G.; WEBER, E.J.; ELOVITZ, M.S., *Environ. Sci. Technol.*, 1996, **30**, 3028.

- [61] CHAMBERS, C.C.; ARCHIBONG, E.F.; JABALAMELI, A.; SULLIVAN, R.H.; GIESEN, D.J.; CRAMER, C.J.; TRUHLAR, D.G., *J. Mol. Struct. (Theochem)*, in press.
- [62] CRAMER, C.J.; TRUHLAR, D.G., *J. Am. Chem. Soc.*, 1994, **116**, 3892.
- [63] MCKEE, M.L., *J. Phys. Chem.*, 1993, **97**, 13608.
- [64] SEVERANCE, D.L.; JORGENSEN, W.L., In *Structure and Reactivity in Aqueous Solution*; Cramer, C.J. and Truhlar, D.G., Eds.; American Chemical Society: Washington, DC, 1994, **568**, 243.
- [65] ANGYAL, S.J., *Aust. J. Chem.*, 1968, **21**, 2737.
- [66] HA, S.N.; GIAMMONA, A.; FIELD, M.; BRADY, J.W., *Carbohydr. Res.*, 1988, **180**, 207.
- [67] BRADY, J.W., *J. Am. Chem. Soc.*, 1989, **111**, 5155.
- [68] KROON-BATENBURG, L.M.J.; KROON, J., *Biopolymers*, 1990, **29**, 1243.
- [69] VAN EIJCK, B.P.; KROON-BATENBURG, L.M. J.; KROON, J., *J. Mol. Struct.*, 1990, **237**, 315.
- [70] HA, S.; GAO, J.; TIDOR, B.; BRADY, J.W.; KARPLUS, M., *J. Am. Chem. Soc.*, 1991, **113**, 1553.
- [71] ZHBANKOV, R.G., *J. Mol. Struct.*, 1992, **275**, 645.
- [72] POLAVARAPU, P.L.; EWIG, C.S., *J. Comp. Chem.*, 1992, **13**, 1255.
- [73] VAN EIJCK, B.P.; HOOFT, R.W.W.; KROON, J., *J. Phys. Chem.*, 1993, **97**, 12093.
- [74] GLENNON, T.M.; ZHENG, Y.J.; LE GRAND, S.M.; SHUTZBERG, B.A.; MERZ, K.M., JR., *J. Comp. Chem.*, 1994, **15**, 1019.
- [75] DOWD, M.K.; FRENCH, A.D.; REILLY, P.J., *Carbohydr. Res.*, 1994, **264**, 1.
- [76] ZUCARELLO, F.; BUEMI, G., *Carbohydr. Res.*, 1995, **273**, 129.
- [77] HEHRE, W.J.; RADOM, L.; SCHLEYER, P.V.R.; POPLE, J.A., *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.
- [78] BARROWS, S.E.; STORER, J.W.; CRAMER, C.J.; TRUHLAR, D.G.; FRENCH, A.D., *J. Comp. Chem.*, in press.
- [79] ANGYAL, S.J., *Angew. Chem., Int. Ed. Engl.*, 1969, **8**, 157.
- [80] CRAMER, C.J.; TRUHLAR, D.G.; FRENCH, A.D., *Carbohydr. Res.*, 1997, **298**, 1.
- [81] KIRKWOOD, J.G., *J. Chem. Phys.*, 1934, **2**, 351.
- [82] ONSAGER, L., *J. Am. Chem. Soc.*, 1936, **58**, 1486.
- [83] MÍCOVIĆ, L.D.; ŽIGMAN, V., *J. Chem. Soc. Perkin Trans. II*, 1985, 625.
- [84] DENBIGH, K., *The Principles of Chemical Equilibrium*, 4th ed., Cambridge University Press: London, 1981, 256–257.
- [85] BEN-NAIM, A., *Statistical Thermodynamics for Chemists and Biochemists*, Plenum: New York, 1992, 421.
- [86] 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. and Frurip, D.J., Eds.; American Chemical Society: Washington, DC, in press.
- [87] LOMBARDO, F.; BLAKE, J.F.; CURATOLO, W.J., *J. Med. Chem.*, 1996, **39**, 4750.
- [88] TROHALAKI, S.; PACHTER, R., In *Abstracts of the 211th National Meeting of the American Chemical Society*; American Chemical Society: Washington, DC, 1996, COMP 60.
- [89] RIBADENEIRA, M.D.; AUNGST, B.J.; EYERMANN, C.J.; HUANG, S.-M., *Pharm. Res.*, 1996, **13**, 227.
- [90] GIESEN, D.J.; CHAMBERS, C.C.; CRAMER, C.J.; TRUHLAR, D.G., *J. Phys. Chem.*, in press.
- [91] REYNOLDS, C.H., *J. Chem. Inf. Comput. Sci.*, 1995, **35**, 738.
- [92] LEO, A.J., Masterfile database (1994) from MedChem Software.
- [93] VAES, W.H.J.; URRESTARAZU RAMOS, E.; HAMWIJK, C.; VAN HOLSTEIJN, I.; BLAAUBOER, B.J.; SEINEN, W.; VERHAAR, H.J.M.; HERMENS, J.L.M., *Chem. Res. Toxicol.*, 1997, **10**, 1067.
- [94] WOHLFARTH, C., In *CRC Handbook of Chemistry and Physics*, 76th ed.; Lide, D.R., Ed.; CRC Press: Boca Raton, FL, 1995, 6–159ff.
- [95] GONZALEZ-LAFONT, A.; TRUONG, T.N.; TRUHLAR, D.G., *J. Phys. Chem.*, 1991, **95**, 4618.

- [96] LIU, Y.-P.; LU, D.-H.; GONZALEZ-LAFONT, A.; TRUHLAR, D.G.; GARRETT, B.C., *J. Am. Chem. Soc.*, 1993, **115**, 7806.
- [97] BASH, P.; HO, L.L.; MACKERELL, A.D.; LEVINE, D.; HALLSTROM, P., *Proc. Natl. Acad. Sci., USA*, 1996, **93**, 3698.

Chambers C.C., Giesen D.J., Hawkins G.D., Cramer C.J., Truhlar D.G., Vaes W.H.J. (1999) Modeling The Effect of Solvation on Structure, Reactivity, and Partitioning of Organic Solutes: Utility in Drug Design. In: Truhlar D.G., Howe W.J., Hopfinger A.J., Blaney J., Dammkoehler R.A. (eds) Rational Drug Design. The IMA Volumes in Mathematics and its Applications, vol 108. Springer, New York, NY, 1999; pp. 51-72.

doi.org/10.1007/978-1-4612-1480-9_5