"Development and Biological Applications of Quantum Mechanical Continuum Solvation Models," C. J. Cramer and D. G. Truhlar, in *Solute/Solvent Interactions (Theoretical and Computational Chemistry*, Vol. 1), edited by P. Politzer and J. S. Murray (Elsevier, Amsterdam, 1994), pp. 9–54.

DEVELOPMENT AND BIOLOGICAL APPLICATIONS OF QUANTUM MECHANICAL CONTINUUM SOLVATION MODELS

Christopher J. Cramer and Donald G. Truhlar

Department of Chemistry and Supercomputer Institute, University of Minnesota, Minneapolis, MN 55455.

1. INTRODUCTION

The development of computational methods for condensed phases that have predictive powers equivalent to those available for gas-phase systems has been a long-standing goal of theoretical chemists. The complementary nature of theory and experiment has become strikingly apparent for the study of molecular structure and dynamics in the gas-phase, and indeed we have reached the point where many gas-phase observables are more accurately and efficiently predicted by theoretical methods than they may be measured. However, the corresponding role of theory in *solution*-phase chemistry is nowhere near so well established. Since the bulk of preparative organic chemistry and all of biological chemistry occur in condensed phases, the importance of further progress in this area should be clear.

The transition from the gas phase to solution is by no means a small perturbation; often the effects of solvation on issues of structure and reactivity are extremely large [1]. For example, the Menschutkin reaction of ammonia and chloromethane is illustrated in Figure 1. The reaction involves nucleophilic displacement of chloride by ammonia and thereby converts the neutral reactants into a pair of charged products, methylammonium cation and chloride anion, a so-called Type II S_N2 reaction. This reaction is exothermic and proceeds readily in aqueous solution. Modeling efforts in the gas phase are unable to provide much information about this process, since the product ions are very high in energy in the absence of solvation—in fact the corresponding reaction path is entirely uphill in the gas phase. The reaction is made possible by virtue of aqueous solvation, which is roughly 150 kcal/mol more favorable for the products than for the reactants [2,3].

While the above example illustrates the effect of solvent on the equilibrium between reactants and products, another interesting case is the effect of solvent on the activation energy, i.e., the differential solvation of the transition state relative to the reactants. This effect has been well studied for S_N2 reactions of anions [4-11]. Figure 2 illustrates this effect for a more complex organic reaction, namely the Claisen rearrangement, an electrocyclic reaction which converts an allyl vinyl ether to a γ,δ-unsaturated carbonyl compound. The Claisen rearrangement has been demonstrated to be accelerated on the order of 1000-fold for the unsubstituted parent ether on going from the gas phase to aqueous

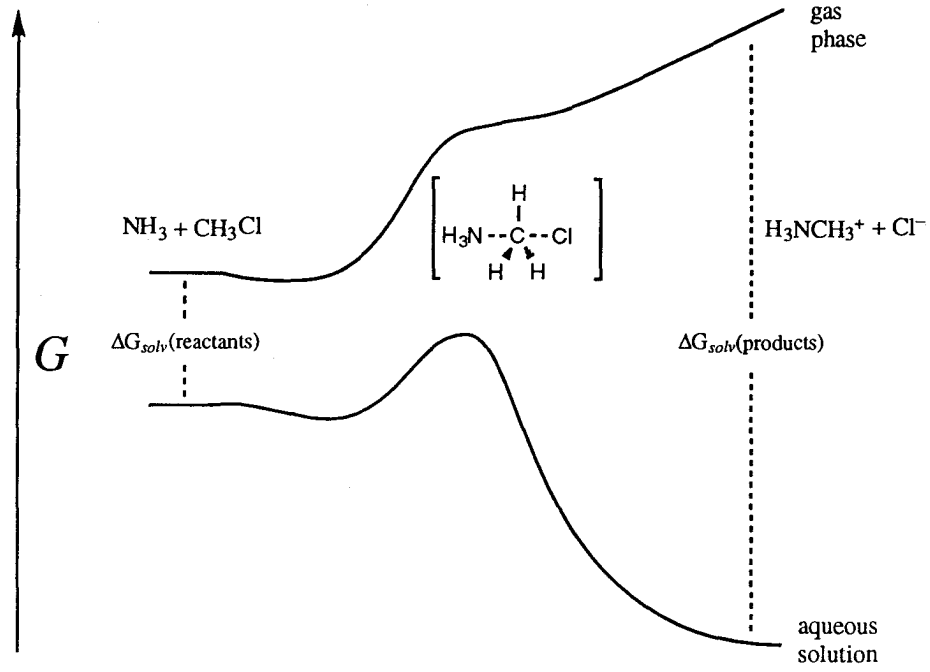


Figure 1. Schematic gas-phase and aqueous potentials of mean force for the Menschutkin reaction projected onto a single generalized reaction coordinate.

solution at room temperature [12-16]. The Claisen rearrangement is a particularly interesting case, insofar as there are two distinct transition state structures, a chair and a boat form, each of which leads to stereochemically distinct products when the reactive termini are appropriately substituted (illustrated with deuterium substitution in Figure 2). In addition to the differential solvation effect between reactants and transition states for this reaction, there are additional solvation differences between the two transition states, giving rise to the possibility that in specific instances one might tune the stereoselectivity of the reaction by judicious choice of solvent [14,17].

These kinds of effects also occur in a more biologically focused paradigm, which is illustrated in Figure 3. The top portion of the illustrated cycle corresponds to the unimolecular, gas-phase reaction of some molecule A. It is convenient to continue using the Claisen rearrangement as an example, since this appears to be the mechanism by which chorismate is converted into prephenate in vivo, an important step in the shikimic acid pathway [18,19] by which a family of sugars is transformed into aromatic amino acids. The middle portion of the cycle is similar to what we have discussed above, i.e., the effect of aqueous solvation has been included. The bottom portion of the cycle, on the other hand,

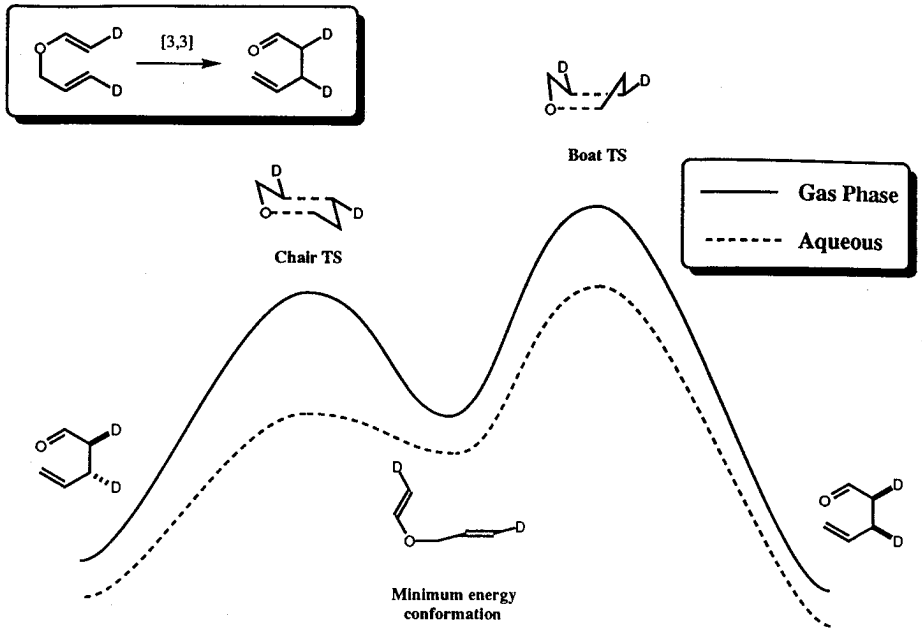


Figure 2. Schematic comparison of the differential effect of aqueous solvation on competing chair and boat transition states for the Claisen rearrangement.

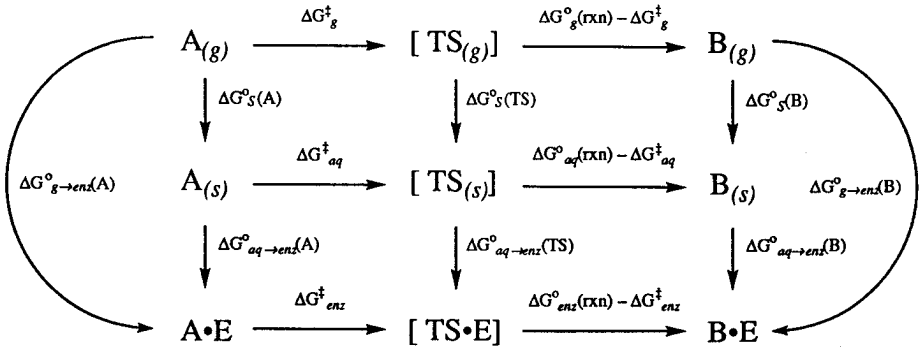


Figure 3. Thermodynamic cycles for a unimolecular reaction showing the relationship between the gas-phase, aqueous solution, and the enzyme-mediated reaction.

represents a reaction that is aqueous overall, but the reaction now proceeds within an enzyme active site (e.g., chorismate dismutase). The cycle illustrates that the manner in which an enzyme influences reaction rates is very similar to the manner in which solvent does [20-22]. Thus, enzyme catalysis involves the competition between solvation of the substrate by the bulk solvent and "solvation" by the enzyme (the latter more commonly associated with terms like "complexation," "binding," or "stabilization") [23].

Section 2 reviews theories of aqueous solvation based on combining quantum mechanics with continuum models, including both a continuum representation of the bulk solvent and also the continuum approach (based on atomic surface tensions) to first hydration shell effects. Section 3 reviews selected applications to biomolecules. Section 4 considers other solvents. Section 5 presents a brief comparison of the present approach to the method of theoretical linear solvation energy relationships, which also involves both quantum mechanics and a continuum treatment of the solvent.

2. MODELING AQUEOUS SOLVATION

The most obvious way to account for solvation in a theoretical calculation is to surround a substrate of interest with sufficient solvent molecules to mimic the effects of bulk solvation. Regrettably, the number of solvent molecules which is required to mimic bulk solvent is usually quite large. Moreover, the supersystem in general has a very large number of energetically accessible states (differing by the individual orientations of solvent molecules, for example) which must be statistically sampled to obtain thermodynamically averaged information [24-30]. This sampling may be performed using either probabilistic methods like Monte Carlo [21,25,27,28,30] or by following molecular dynamics trajectories in phase space [20-22,26-28,31,32], but the net result is that converged treatment of the explicit quantum mechanical representation of the entire system is effectively impossible. As a result, simulations along the lines described above are typically carried out with classical mechanical Boltzmann factors replacing the quantum mechanical density operators for atomic coordinates. Furthermore, the atomic potential energy function which, according to the Born-Oppenheimer separation of electronic and nuclear motions, is actually governed by the quantum mechanical adiabatic evolution of the electronic states, is often replaced by a set of pairwise interactions governed by a molecular mechanics force field. Even in such classical simulations, convergence with respect to long-range forces [33-35] and multidimensional sampling [36-42] remains problematic. In the latter respect, there is no general way to insure the adequate sampling of phase space during a simulation. Thus, in systems with hundreds to thousands of position and momentum coordinates, phase-space bottlenecks may prevent the system from exploring regions which are important for the accurate prediction of observed properties.

Classical simulations suffer from other drawbacks as well. Inevitably, force fields are derived from fitting functional forms for interaction energies to experimental or theoretical data [43,44], and the extension of such approaches to transition states is a severe difficulty because of the limited database available. Thus, the development of force fields capable of handling transition states is technically challenging [45].

In order to alleviate the problems associated with solvent sampling, it proves useful to simplify the microscopic representation of the solvent. For instance, Warshel and co-

workers pioneered replacement of individual water molecules with Langevin dipoles [46-50]. One critical advantage of eliminating some of the technical difficulties associated with carrying out the solvent sampling is that it permits a more accurate representation of the solute, e.g., it makes it much more practical to introduce a quantum mechanical treatment of the solute [10,46,51-75]. Another approach to simplifying the problem is to replace hydrogen bonding interactions in the first hydration shell by an effective potential method [76]. An older and more radical approach to simplifying the treatment of the solvent is to remove all microscopic coordinates and leave behind only a continuum characterized by one or more of its bulk properties. The simplest such solvation models retain no information about the solvent beyond its bulk dielectric constant. This level of solvent detail is all that is required for the cavity-based electrostatic approximations of Born [77], Kirkwood [78], and Onsager [79] which capture the essential physics of an ion or small dipolar molecule in a dipolar solvent. These classical treatments lead to simple analytic expressions for the solvation free energy when the solute can be assumed to be spherical, and these analytical expressions provide the essential scaling laws for qualitative understanding of how solvation energies vary with solute radii.

Three key elements have been involved in extending the usefulness of continuum solvation models to essentially arbitrarily complicated solutes: (1) By using computers, one can eliminate the restriction to spherical solutes. The modern approach is to represent the solute shape in terms of overlapping spheres centered at the atomic nuclei. Such space-filling models are now recognized to provide realistic models of the shape of essentially all molecules and molecular fragments, and even transition states. (2) By using molecular orbital theory, the charge distribution of any solute can be modeled reasonably well by a series of partial charges at the atomic centers; this is called the distributed monopole model. This allows the treatment of solutes that contain more than a single dipolar center. Of course, the more complicated models of solutes as arbitrarily large numbers of spheres, not necessarily arranged symmetrically, and each with an arbitrary partial charge at the center, means that analytic solutions are no longer available; however, the cost of computer solutions (with varying degrees of numerical complexity and auxiliary approximations) is a rapidly decreasing function of time. (3) Finally, modern extensions of the continuum solvent model have benefited from the realization that the main defect of the continuum treatment of the solvent is the breakdown of bulk properties in the first solvation shell. In early work this concern was expressed in attempts to define effective solute radii that encompass a portion of the solvent where "dielectric saturation" occurs. This saturation of the solvent orientational polarizability in the vicinity of a large electric field is accompanied by other first solvation shell effects, for example, "electrostriction", which refers to the loss of motional freedom experienced by dipolar solvent molecules in that same region. In modern work, one focuses more directly on the first hydration shell and explicitly accounts for its effects in addition to the bulk electrostatic effects.

This section will continue with more details of these modern extensions. Section 2.1.1 reviews the essentials of classical theory. Sections 2.1.2 and 2.1.3 review the classical analytical theories. Sections 2.1.4 and 2.1.5 review the elements of extensions (1) and (2). Section 2.2 reviews extension (3).

2.1. Electrostatic components of solvation free energy

We begin with a discussion of that portion of the solvation free energy which arises from solute-solvent, solvent-solvent, and internal solute electrostatic interactions.

2.1.1. The Poisson Equation

The key reason for choosing to characterize the solvent continuum by its dielectric constant is that it allows one to use the power of classical electrostatics. When the solute is represented explicitly, and the solvent is treated as a continuum, the Laplacian of the electrostatic potential, $\phi(\mathbf{r})$, is related to the free charge density (i.e., the charge density due exclusively to the solute), $\rho(\mathbf{r})$, by Poisson's equation [80-82],

$$\nabla^2\phi = -\frac{4\pi\rho(\mathbf{r})}{\varepsilon}, \tag{1}$$

where ε is the homogeneous dielectric constant (relative permittivity of the solvent, e.g., 78.3 for water at 298 K), $\mathbf{r}$ denotes the position in space, and the equation is written in gaussian units. If ε depends on $\mathbf{r}$, one can replace equation 1 by the slightly more complicated [81]

$$\nabla\cdot\varepsilon(\mathbf{r})\nabla\phi = -4\pi\rho(\mathbf{r}). \tag{2}$$

Assuming that thermal equilibrium is maintained by an external heat bath, the free energy of solvation G, which is the maximum work which may be extracted from the solvation process, is then obtained from [81,82]:

$$G = -\frac{1}{2}\int d^3\mathbf{r}\,\frac{1}{4\pi}\phi(\mathbf{r})\,\nabla\cdot\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r}) \tag{3}$$

or equivalently

$$G = \frac{1}{2}\int d^3\mathbf{r}\,\frac{1}{4\pi}\phi(\mathbf{r})\,\nabla\cdot\varepsilon(\mathbf{r})\,\mathbf{E}(\mathbf{r}) \tag{4}$$

where $\mathbf{E}(\mathbf{r})$ is the electric field given by the gradient of the electrostatic potential and is given by

$$\mathbf{E}(\mathbf{r}) = -\nabla\phi(\mathbf{r}). \tag{5}$$

In solution, the electric field contains contributions from both the intrinsic solute charges and from the polarization *induced* by the solute in the solvent. The latter contribution is called the reaction field. Equations 3 and 4 can also be written as

$$G = \frac{1}{2} \int d^3\mathbf{r}\, \mathbf{D}(\mathbf{r}) \cdot \mathbf{E}(\mathbf{r}), \tag{6}$$

where $\mathbf{D}(\mathbf{r})$ is the electric displacement due to the solute charges. This illustrates how the solute interacts with its own reaction field, which has a significant effect on the energy of the system. For conceptual purposes, equation 6 may be thought of as arising from an iterative process. That is, the gas-phase solute is placed into solution, inducing a reaction field determined from the gas-phase charge distribution. The interaction of the reaction field with the solute charge distribution in general induces a relaxation of the gas-phase nuclear and electronic structure in order to minimize the free energy of the whole system. Of course, as relaxation proceeds it changes the reaction field quantitatively, such that additional changes in the solute charge distribution may be favorable. The entire procedure reaches its terminus when the internal cost of additional change in the charge distribution of the solute and solvent begins to exceed the resultant gain in their interaction free energy.

We will have cause to refer to this electrostatic portion of the free energy of solvation repeatedly in later discussion. We label it ΔG_{ENP} to indicate that it includes the work required to distort the electronic (E) and nuclear (N) structure (i.e., the molecular electronic wave function and the geometry, respectively) of the solute from their optimal gas-phase values, and this is driven by the gain in polarization (P) free energy, which is the net gain in solute-solvent interaction free energy minus the cost in solvent internal free energy.

2.1.2. The Born equation

In actual practice, analytical solutions to the Poisson equation exist only for rather simple cases. One example is a charge q on a conducting sphere of radius α. Since a charge on a metallic sphere is spread uniformly over its surface, but the effect of this outside the sphere is the same as for a point charge at the sphere center, this is a simple model for a monatomic ion. Solution of Poisson's equation in this instance for the gas-phase ($\varepsilon = 1$) and the dielectric medium ($\varepsilon > 1$) gives Born's formula [77,83,84] for the free energy of transfer of the charged sphere from a medium with a dielectric constant of unity (vacuum or sufficiently dilute gas phase) to a solvent of dielectric constant ε:

$$\Delta G_S^o = -\frac{1}{2}\left(1 - \frac{1}{\varepsilon}\right)\frac{q^2}{\alpha}. \tag{7}$$

Of course, the meaning of α is somewhat less clear for a monatomic ion than for the ideal case of the conducting sphere [84-88]. That is, the surrounding solvent dielectric is not homogeneous all the way up to the ionic "surface". Many models have been proposed for relating α to the electronic structure of the monatomic ion, but ΔG_S^o is very sensitive to the

precise value of α, so these models must be used with caution. We prefer the empirical route where α is determined from equation 7 using an experimental value for the free energy of solvation. Such an α is called an effective ionic radius, or a Born radius. Based on analysis of classical simulations employing explicit solvent representation, Jayaram et al. have concluded that dielectric saturation does not affect the quadratic dependence of ΔG_S^o on charge for spheres with a charge less than 1.1 electronic units [87]; in such instances, the ionic radii in equation 7 which reproduce the simulation free energies of solvation are generally in reasonable agreement with other standard measures of atomic radii, such as the van der Waals radii suggested by Bondi [89] (although, of course, they are not exactly the same). Hirata et al. [86] and Roux et al. [90] have separately discussed this result in terms of extended reference interaction site method (RISM) calculations [91,92], suggesting that sensible Born radii may be derived from analysis of the first peak in the solute-solvent radial distribution function, again in concert with a simple spherical model for the ion. Furthermore, it has been shown that the average distance between ions and nearest neighbor water molecules is about 1.4 Å larger than the ionic crystal radius for typical ions [93].

As mentioned at the end of Section 2.1.1, the electrostatic energies which we have so far discussed include not only the interaction of the solute with the solvent but also the change in solvent-solvent interactions when the solute is inserted. Under the mild assumptions of linear response theory, it can be shown that the latter increase in intrasolvent energy cancels half the favorable solute-solvent interactions, which is one way to think about the factors of $\frac{1}{2}$ appearing in equations 3, 4, 6, and 7 [63,67,69,72,86,87,90,94-100].

2.1.3. The reaction field approach

Clearly only monatomic ions may be unambiguously regarded as spherical. However, at large enough distance, the interaction of any multipole distribution with a surrounding field is dominated by its leading term. In other words, at long enough range the remaining interaction of an ion with a surrounding continuum is well approximated by the Born equation. Indeed, in molecular simulations of ions which *include* explicit solvent molecules, the calculation of electrostatic interactions is typically truncated at some maximal distance from the solute to maintain computational efficiency, and the remaining interactions out to infinite distance are often approximated using the Born equation [33-35].

The leading multipole moment for *uncharged* solutes is usually the dipole term. Once again, an analytic solution to Poisson's equation exists. For a point dipole of magnitude μ in a sphere of radius α, the result, derived in different ways by Kirkwood and Onsager, is [78,79]

$$\Delta G_S^o = -\frac{(\varepsilon - 1)\mu^2}{(2\varepsilon + 1)\alpha^3} \tag{8}$$

where ε is again the dielectric constant of the solvent modeled as a continuum. It should be emphasized that the spherical radius α appearing in equations 7 and 8 is only defined insofar as it represents the point at which a discontinuity in the dielectric constant occurs in the integral of equations 3 and 4. In other words, although it has clear mathematical utility in the solution of the Poisson equation, there is no prescription for its determination in molecular

cases. We note at this point that analytical solutions to the Poisson equation for a point dipole inside an ellipsoid are also available, and implementations of this approximation have also appeared [71].

As mentioned above, α may be assigned on the basis of fitting the Born equation to experimental data for ions, but for uncharged molecules it would be a much more severe and ambiguous approximation to determine α in the same way from the Kirkwood-Onsager equation because of the spherical cavity approximation; it is clearly impossible to construct an uncharged, spherically symmetric molecule which has a dipole moment. Furthermore, little work has appeared establishing whether or not there are any clear connections between optimal α values and known physical properties for atoms and molecules. In this regard, an obvious approach would be to choose α as the radius of a sphere whose volume matches the cavity enclosed by the solvent-excluding surface of the solute, also called the molecular surface [101-105]. This approach yields similar results to those obtained using the van der Waals surface to divide the regions of unequal dielectric constant [106]. Wong et al. [107] have used a quantum mechanical approach in which the van der Waals surface is replaced by an isodensity surface and empirical scale factors are employed. Since the choice of surface determines the value of α, and the electrostatic free energy of solvation depends on the third power of α in equation 8, the calculations are quite sensitive to the choice of surface, and some nonphysical results have been reported in the literature when insufficient care was taken in assigning a value to α. (Further discussion of alternative definitions of solute surface is provided in Section 2.2.)

One of the key differences between the Born and Kirkwood-Onsager equations is that the former involves the solute charge, which is unchanged by the polarization field induced in the dielectric (we neglect charge transfer to or from the solvent); the latter, however, involves the molecular dipole moment, which may readily increase by relaxation of the electronic structure as described in Section 2.1, thereby contributing more polarization free energy at the cost of reorganizational free energy. Since the polarization field now affects the solvated electronic structure, it should be treated self-consistently in any quantum mechanical calculations designed to incorporate the effects of continuum solvation.

Within the Kirkwood-Onsager approximation, the quantum mechanical Hamiltonian operator that includes reaction field effects for neutral solutes is

$$(H_0 - \lambda g \mu \cdot \langle \psi|\mu|\psi\rangle)\,|\psi\rangle = E\,|\psi\rangle \qquad (9)$$

where $\lambda = 0.5$, $g = 2\,(\varepsilon-1)/(2\,\varepsilon+1)\,\alpha^3$, α is the solute cavity radius, and H_0 is the gas-phase Hamiltonian constructed in the usual manner. The corresponding Hartree-Fock equations are then [57,66,71,107-123]

$$(F_0 - \lambda g \mu \cdot \langle \psi|\mu|\psi\rangle)\,|\phi_i\rangle = \varepsilon_i\,|\phi_i\rangle \qquad (10)$$

where F_0 is the usual gas-phase Fock operator [124,125], and the ε_i are the one-electron orbital energies associated with the molecular orbitals ϕ_i. Note that these equations must be solved self-consistently insofar as the Fock operator, the one-electron density matrix involved in the solution of the Hartree-Fock equations, and the molecular dipole moment are all mutually interdependent. It is easily seen that these equations capture the phenomenon of increased charge separation being favored in solvents of increasing dielectric constant. The electrostatic portion of the free energy of solvation, ΔG_{ENP}, is then simply the energy calculated from the orbitals obtained from equation 10 minus the gas-phase energy. Note that this will generally be less negative than the energy calculated from equation 8, which is the portion associated with ΔG_P, since the costs of distorting the electronic and molecular structure of the *solute* have been included in the calculations implicitly by the SCF formalism.

In addition, as discussed in Section 2.1.2, the cost of distorting the *solvent* structure is taken into account via the factor λ of 0.5 preceding the reaction field portion of the solution-phase Fock operator. This formalism thus treats all three effects—favorable solute-solvent interactions (twice ΔG_P), cost of solute reorganization (ΔE_{EN}), and cost of solvent reorganization (minus ΔG_P)—mutually self-consistently; changes in solute electronic and molecular structure cease when additional gain in ΔG_P fails to be larger than the costs of distortion, ΔE_{EN}. An alternative approach has been considered, however, which minimizes the solute energy plus solute-solvent interaction, rather than the system energy. Formally, this convention involves taking $\lambda = 1$ in equations 9 and 10; therefore, *solute* electronic and structural reorganization proceeds until that portion of ΔG_P associated *only with the solute* no longer exceeds any increase in ΔE_{EN}. Clearly, the solute distortion for the latter approach must be greater than that for the former. However, when optimization is carried out with $\lambda = 1$, the *total* energy of the system can only be obtained by adding back a factor of $0.5g\mu^2$, where μ is the relaxed dipole moment. In other words, if the cost of solvent reorganization is not accounted for self-consistently, it is included *ex post facto*.

Both methods, employing $\lambda = 0.5$ [71,107,109-113,116-123] and $\lambda = 1$ [57,66,108,114,115,122,126,127], have seen extensive use in the literature. Regrettably, it is not always clear in certain instances which method has been employed, and careful analysis of the equations in individual papers even reveals cases where the mathematical derivation switches haphazardly from one approach to the other! Not much work has been done to assess which, if either, of these two methods is to be preferred. Szafran et al. [122] compared the two approaches for the prediction of solvent effects on tautomeric equilibria in 2-hydroxypyridine$\rightleftharpoons$2-pyridone and 4-hydroxypyridine$\rightleftharpoons$4-pyridone and found little difference between them in comparison to experimental results. However, any comparison of the methods is somewhat ambiguous insofar as the spherical cavity radius α appearing in the coupling factor g has no obvious physical interpretation (vide supra). Treating the radius as a free parameter makes it fairly simple to obtain reasonable results with either value of λ. What is *unambiguous*, however, is that optimization of solute geometries is much more straightforward when one takes $\lambda = 0.5$. The alternative, $\lambda = 1$, requires solution of a coupled-perturbed Hartree-Fock equation each time analytic gradients are required, since the total energy of the system includes the term $0.5g\mu^2$ non-self consistently, i.e., the term depends on the dipole moment, and therefore the density matrix, but it is not accounted for in the Fock operator since it is added *ex post facto*. This latter consideration is significant enough to suggest that, from a purely practical standpoint, it would be better to adopt

$\lambda = 0.5$, regardless of the competing arguments over which is to be preferred based on fundamental principles.

In any case, the great simplicity of the Kirkwood-Onsager approach has prompted its incorporation into a number of quantum mechanical electronic structure programs, both at the ab initio and semiempirical levels [108,111,128-131]. Considerable caution should be exercised in its application, however. For instance, for charged solutes one should include an ionic Born term derived from equation 7, which will be a *considerably* larger contributor to the solvation free energy than the corresponding Kirkwood-Onsager term. However, for at least one commonly used electronic structure program this term is *not* included [130], perhaps because it does not require a self-consistent treatment in the Hartree-Fock equations and is thus easily added on post facto. Another consideration for ionic systems is that only the leading molecular multipole moment is independent of the origin of the chosen molecular coordinate system. Thus, Born-Kirkwood-Onsager calculations of such systems require a choice of where to evaluate the molecular dipole moment, e.g., at the molecular center of mass, center of charge, center of the encompassing sphere, etc. The final result will be correct only if a consistent choice of coordinates is used throughout the derivation and application. The two most critical considerations, however, are that (1) application of the model is justified only when higher-order multipole moments are negligible, and (2) application of the model is justified only for nearly spherical solutes. These issues are addressed in the next two sections.

2.1.4. Truncated single-center multipolar expansions

As discussed above, the Born-Kirkwood-Onsager approach includes only the solute's monopole and dipole interactions with the continuum. That is, the full classical multipolar expansion of the total solute charge distribution is truncated at the dipole term. Although these terms dominate at very large distances, one may imagine evaluating the electrostatic potential and the polarization contributions to the free energy of solvation while approaching more and more closely to the solute. Eventually, the contributions from higher-order moments cease to be negligible. The importance of such higher order moments is most obvious for neutral molecules whose dipole moments vanish as a result of symmetry. The Born-Kirkwood-Onsager model would require the electrostatic portion of the free energy of solvation for these molecules to be identically zero.

Generalization of the Born-Kirkwood-Onsager approach may be accomplished by solution of the Poisson equation for a single-center multipolar expansion to arbitrarily high order. This approach yields [60,78,98,132-135]

$$\Delta G_S^o = -\frac{1}{2} \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \sum_{l'=0}^{\infty} \sum_{m=-l'}^{+l'} M_l^m f_{ll'}^{mm'} M_{l'}^{m'} \qquad (11)$$

where each component $M_{l'}^{m'}$ of every multipole moment interacts with all of the reaction field multipole moments induced by the solute multipoles via a set of coupling factors $f_{ll'}^{mm'}$, called the reaction field factors. The assumption that the cavity is a sphere leads to the coupling factor being non-zero only for $l = l'$ and being independent of m and m'. The Born-Kirkwood-Onsager model is then seen as a special case involving only the net charge ($l = 0$)

and dipole moment ($l = 1$) terms. The assumption of a spherical or ellipsoidal solute cavity actually permits analytical determination of *all* of the reaction field factors in equation 11 [60,71,78,108]. When implemented into the SCF equations in a manner analogous to that described in the previous section with λ taken as 0.5, this simplification additionally permits efficient optimization of solvated geometries [71,132,133,135-137]. Even for more general cavities, e.g., cavities which more resemble a molecular van der Waals surface, it is possible to determine the reaction field factors numerically, since they appear in an overdetermined system of linear equations [133,134]. Both of these approaches have seen increasing use [134,138,139]. Although Tapia has discussed the competing derivations which yield $\lambda = 1$ (solvent as isothermal bath) and $\lambda = 0.5$ (work of solvent polarization included self-consistently) in the equations analogous to equations 9 and 10 for this more general multipole approach [66], it appears that implementations appearing in the recent literature have exclusively used $\lambda = 0.5$, probably to take advantage of the simplified geometry optimizations for this case [71,132-139].

A point of obvious interest is how fast the electrostatic portion of the solvation free energy converges with respect to multipole order. Interestingly, this convergence can be quite slow, even for fairly simple molecules. A specific example is Z-3-aminoacrylonitrile, which was studied by Pappalardo and co-workers [140]. Immersion of this solute in a continuum with a dielectric constant of 38.8 yields a total electrostatic polarization free energy of –13.2 kcal/mol. Decomposition of this energy finds 66% contained in the dipole term, 22% in the quadrupole, and the remaining 12% in the terms up through 2^6-pole, which was the highest multipole considered. Moreover, this slow convergence was still more pronounced for the transition state for rotation about the carbon-carbon double bond, where the polarization free energy for this nearly zwitterionic structure is –44.8 kcal/mol, now consisting of 64% dipole, 18% quadrupole, and 19% in the higher order terms. The convergence of the multipole expansion is also very dependent on the shape of the employed cavity [134,138].

Although typically the multipolar expansion of the electronic structure is performed at only a single point, e.g., the center of mass of the molecule, this is not a requirement. Instead, an arbitrary number of *distributed* multipoles may be placed at multiple points, e.g., the atomic coordinates or the atomic coordinates and bond midpoints [72,134,141-144]. Numerical fitting of the multipoles and reaction field factors proceeds equivalently. One of the simplest approaches is to use only atomic monopoles (i.e., partial charges); this may be made equivalent to a single-center expansion up through $l = N - 1$ where N is the number of atoms. Indeed, this is quite similar to the generalized Born approximate solution to Poisson's equation which is discussed in greater detail in Section 2.3.

As expected, this distributed approach is much more rapidly convergent in terms of the multipole order required at each center. In the modeling of formamide, for instance, a one-center expansion in a generalized cavity still has 1% fluctuations by the 2^6-pole term [134]. The distributed expansion at the atomic positions in the same cavity, on the other hand, has only a 1% contribution from the quadrupole, and is effectively converged after this point [134]. Moreover, it is generally more efficient computationally to describe the molecular electronic structure as a set of N distributed monopoles rather than a single multipolar expansion of order N. Of course, the precise method for determining the magnitude of the distributed monopoles (partial charges) remains controversial insofar as atomic partial charges are not physical observables; a large number of models and

algorithms is available [144-166]. Although including higher multipoles at *every* center clearly increases the flexibility of the approach, the cost is considerable in terms of computational effort and so far has limited its application to fairly simple systems, e.g., the NH_3-HCl complex [167].

2.1.5. Generalized reaction fields from surface charge densities

If the multipolar expansion of the molecular charge density implicit in equation 11 were carried out to infinite order, the resulting equation would be a complete solution to the volume integral expression of Poisson's equation discussed so far. An alternative approach is application of Green's theorem to convert the volume integral of equation 6 to an integral over the molecular cavity surface *S*. In particular, the effect of the reaction field may be modeled by a continuous polarization charge density spread over that surface, where this virtual charge density, $\sigma(r)$, is in Gaussian units

$$\sigma(r) = \frac{1-\varepsilon}{4\pi\varepsilon}\frac{\partial}{\partial n}\left[\phi_\rho(r) + \phi_\sigma(r)\right]_{S_-} \tag{12}$$

with $\phi_\rho(\mathbf{r})$ being the electrostatic potential due to the solute charge distribution, and $\phi_\sigma(\mathbf{r})$ being the potential due to the virtual charges [65,74,168-171]. The derivative is with respect to an outward surface normal evaluated on the solute side (indicated by the S_- subscript) of the dielectric interface. The potential created by the surface virtual charge density is

$$\phi_\sigma(\mathbf{r}) = \int_S \frac{\sigma(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^2\mathbf{r}', \tag{13}$$

and it must be added to the potential due to the solute charge distribution to obtain the total electrostatic potential at r. The electrostatic portion of the free energy of solvation is then defined as

$$\Delta G_S^o = \langle \psi | H_0 + \tfrac{1}{2}\phi_\sigma | \psi \rangle - G_g^o. \tag{14}$$

In practice, the surface charge density is approximated by a discrete set of point charges which are distributed as uniformly as possible, and the appropriate integrals are then replaced by summations. This model is often referred to as the Polarized Continuum Model or PCM, and it saw most of its early development by Tomasi and co-workers [65,74,169-171]. More recently, it has been implemented in a variety of ab initio and semiempirical quantum chemistry programs [128,172-179]. Of all of the models discussed so far, PCM has seen the most effort spent upon the development of prescriptions for choosing the optimal cavity surface, to include as a function of basis set at the ab initio level [180].

It is probably worth emphasizing here, in case it isn't obvious, that the multipole methods of Section 2.1.4 and the surface-charge-density methods discussed in this section are physically identical, and they will yield the same result if (1) the same molecular surface

is used in both methods (e.g., the same solute atomic radii if the surface is constructed from overlapping spheres), (2) the expansion of equation 11 is well-converged, (3) the numerical representation of the surface charge density is well converged in the PCM, and (4) the factor λ used in the development of the SCF equations is identical. With respect to the last point, equation 14 explicitly takes $\lambda = 0.5$, as is most common in the multipolar reaction field schemes discussed previously. However, again a choice of $\lambda = 1$ may be made (i.e., assuming the solvent to be an isothermal bath) in which case, in exact analogy to the discussion in Section 2.1.3, the free energy of solvation must include the addition of a work of solvent polarization term *ex post facto*. While this latter approach has been pursued by Tomasi and co-workers [65,74,169-171], all other implementations of the PCM approach have used the formalism of equation 14 [128,172-179,181]. To our knowledge, no comparison between the two implementations has appeared.

The broadly general nature of the PCM technique makes it uniquely attractive, especially as it is somewhat more straightforward to implement than the analogous truncated multipolar expansion method taken to arbitrarily high order. Nevertheless, it remains extremely demanding in computational resources, primarily because of the time required to generate the surface virtual charges. Klamt and Schüürmann [178] have presented a particularly efficient algorithm [182] for accomplishing this at the Neglect of Diatomic Differential Overlap (NDDO) [183] semiempirical level of theory.

While the scope of this chapter is intended only to cover quantum mechanical continuum solvation models, we mention in passing that classical approaches to solving the Poisson equation also exist which are similar in spirit to the PCM model but involve representation of the solute charge density as a discrete, grid-mapped set of charges [184-198]. These models fail to allow for self-consistent relaxation of the molecular electronic structure, although obviously they are considerably faster than fully quantal models as a result. To make up for the lack of self-consistency, the dielectric constant is sometimes set equal to a value in the range of 2 to 4 in the non-self-consistent Poisson approaches, whereas it is properly set equal to 1 in the solute when solute polarization is included explicitly.

2.2. Non-electrostatic components of solvation free energy

So far in this chapter, we have concerned ourselves only with the bulk electric polarization of the *volume* surrounding the solute. However, there are other effects that are more specifically associated with the *surface* layer of solvent, i.e., the first solvation shell. One example is the free energy required to create a solute-sized vacuum within the solvent. This cavitation energy, which is approximately proportional to the surface area of the created cavity, is quite dependent on the particular solvent. Additional components at solvent-solute interfaces include the attractive dispersion forces between the solute and the nearby solvent molecules and local structural changes in the surrounding solvent as a result of the insertion of the solute. A key example of the latter effect is found in water, where the solute may induce solute-solvent hydrogen bonding and/or cause especially significant changes in solvent-solvent hydrogen bonding in the first solvation shells [199,200]. Although continuum solvation models arguably include the electrostatic component of hydrogen bonding to some degree in the dielectric polarization term, short-range components cannot be fully modeled by a uniform dielectric constant. Considering the other extreme, for solutes that do *not* hydrogen bond to an acceptor/donor solvent, the solvent structural change may be unfavorable due to the loss of solvent orientational entropy, i.e.,

the hydrophobic effect. We will refer to the sum of these effects as the CDS term, for Cavitation, Dispersion, and solvent-Structural rearrangement.

The difference between the electrostatic effect calculated using the bulk dielectric constant and that calculated taking account of local structural factors may be viewed as arising from an inhomogeneous dielectric constant. In the absence of a microscopic model for the solvent, one approach to incorporating the CDS term would be to allow the dielectric constant of the surrounding medium to take on different values at different locations. However, no clear prescription for accomplishing this in a physically meaningful way is available. An attractive alternative is to assume that the approximate proportionality of the cavitation term to the cavity surface area extends to the remaining terms as well. This seems intuitively reasonable for dispersion, which operates over so short a range that one expects it to be proportional to the number of molecules in the first solvation shell, which is clearly dependent on the solvent accessible surface area of the solute [201]. There is a key distinction to be made between cavitation and dispersion, however, and that is to note that cavitation is independent of the solute, while dispersion is expected to depend on the local polarizability of the solute in any given region. Thus, one might model the C and D terms by assigning a surface tension σ_i to each atom i in the solute and calculating:

$$\Delta G^{o}_{CD} = \sum_i \sigma_i^{CD} A_i \qquad (15)$$

where A_i is the solvent accessible surface area of atom i. The surface tension will contain a constant component which is independent of i, while the remainder will be associated with dispersion and will be dependent on the atomic polarizability of the individual atom. The solvent accessible surface area is most readily calculated following the definition of Lee and Richards [101,202], and it is calculated as the surface mapped out by the center of a solvent-sized ball rolling over the molecular van der Waals surface.

Finally, the sum of the solvent structural rearrangement free energy and the free energy due to specific electrostatic and hydrogen bonding effects in the first solvation shell (including the non-homogeneity of the dielectric constant) can also be assumed to be proportional to a cavity surface area, although it need not be the case that the CS cavity will be defined in the same way as the CD cavity. We include the C term in both cases because it is not rigorously clear how it might be separated from either dispersion or structural rearrangement when the CDS term is simply known in its entirety. Of course, if the two cavities are identical, then equation 15 may be used for the entire term ΔG^{o}_{CDS} with a single set of atom-specific surface tensions which would now be called σ_i^{CDS}, i.e., [203-209]

$$\Delta G^{o}_{CDS} = \sum_i \sigma_i^{CDS} A_i \,. \qquad (16)$$

It is worth noting that in cases where the CDS term dominates the entire free energy of solvation, it may be possible to calculate the full ΔG^{o}_{S} using the formalism of equation 16.

Such an approach works reasonably well for estimating the solvation free energies of hydrocarbons in water, where the ENP terms are very small [203-209]. Scheraga and co-workers had moderate success applying this idea to the twenty biologically important amino acids, although this is admittedly a crude approximation for these much more polar molecules [207].

Very few quantum mechanical solvation models have attempted to incorporate simultaneously both the ENP and some or all of the CDS components of solvation. One common approach is to model cavitation free energies using the scaled-particle theory of Pierotti [210], and some attempts to model dispersion have also appeared [58,67,135,177,211-220]. Rigorous quantum mechanical calculation of this term can be quite costly, and one must question whether the inherent accuracy of a continuum model, with its strong dependence on solute radii, warrants such an approach, as compared to a simplified approximation such as that offered by equation 15 or 16

One particularly interesting attempt to address both the ENP and hydrophobic CDS components of solvation has been described by Still and co-workers within the framework of molecular mechanics [221]. In their model the ENP term is arrived at non-self-consistently by using a generalized Born formalism [10,11,51-54,61-63,66,69,70,221,222] to approximately solve the Poisson equation, and the hydrophobic portion of the CDS term is calculated using an analog of equation 16 where all atoms have the same surface tension, which is a good start but is not flexible enough to model all CDS effects quantitatively. We will devote Section 2.3 to describing our own extension of these ideas to a quantum chemical implementation at the NDDO level, in which we attempt to include all of the important solvation terms in σ_i^{CDS} while simultaneously permitting self-consistent relaxation of molecular electronic structure as a function of solvation for one particular solvent, water.

Before closing this section, it is useful to comment in more detail on the precise interpretation of A_i in equations 15 and 16. Lee and Richards [101,202] and Pascual Ahuir, Silla, and Tuñon [104,105] have carefully distinguished three definitions of the surface area of a solute or its associated cavity. The three definitions will be given here for the case where the solute is modeled as a set of overlapping "atomic" spheres, one representing every atom i (or a group i consisting of a nonhydrogenic atom and its attached hydrogens), with radii R_i, and the solvent molecules are modeled as spheres of radius R_S. The van der Waals surface, also called WSURF [105], is composed of those surfaces of the atomic spheres that are not encompassed by any other sphere. The solvent-accessible surface [202], also called the cavity surface [203] or ASURF [105], is the surface traced out by the center of a solvent sphere rolling on the van der Waals surface. The third surface, originally called the Molecular Surface [101], but more clearly labeled the solvent-excluded surface [105] and also called ESURF [105], is the surface traced out by the points of contact of a solvent sphere rolling on the van der Waals surface in those regions where the solvent sphere touches only one solute sphere; in regions where the solvent sphere simultaneously contacts two or more solute spheres, it is defined to be the surface traced by the inward surface of the solvent sphere. Note that WSURF is independent of R_S, ASURF depends strongly on R_S, and ESURF has an intermediate dependence. We believe that the ASURF definition is most physical for equations 15 and 16. This is because the ASURF surface passes through the center of the first solvation shell and thus, when the solvent is modeled by a continuum, it is

proportional to the number of solvent molecules in the first solvation shell. Since dispersion, solvent-structural free-energy changes, and the effect of nonhomogeneous dielectric constant in the first solvation shell are all approximately proportional to the number of solvent molecules in that shell, this definition is ideal for equation 16.

2.3. The SMx models and absolute free energies of solvation

In this section we review our own SMx models, based on the generalized Born model [10,11,51-54,61-63,66,69,70,221,222] for electrostatic effects and a generalization of equation 16 for first-solvation-shell effects. Models AM1-SM1 [223], AM1-SM1a [223], AM1-SM2 [224], and PM3-SM3 [225] are named to indicate their extension of an underlying gas-phase semiempirical NDDO Hamiltonian, either Austin Model 1 (AM1) [226] or Parameterized Model 3 (PM3) [227,228], with a particular solvation model (SM) where the numbering of the solvation model is primarily chronological in nature. Two new models, AM1-CM1A-SM4A and PM3-CM1P-SM4P, are under development [229] and will be published soon. We have extensively reviewed the theory behind these models elsewhere [75,230], and we present here a less exhaustive description for purposes of completeness.

In the SMx models, the standard-state free energy of solvation, ΔG^{o}_{S}, is calculated from

$$\Delta G^{o}_{S} = G^{o}_{S} - E_{EN}(g), \tag{17}$$

where $E_{EN}(g)$ is the gas-phase electronic kinetic and electronic and nuclear coulombic energy, and G^{o}_{S} is that part of the aqueous free energy given by

$$G^{o}_{S} = E_{EN}(aq) + G_{P}(aq) + G^{o}_{CDS}(aq), \tag{18}$$

where $E_{EN}(aq)$ is equivalent to $E_{EN}(g)$ except now calculated in the presence of solvent, i.e., including distortion energy, $G_{P}(aq)$ is the electric polarization free energy in solution, and G^{o}_{CDS} is the cavitation-dispersion-structural free energy summarized in the previous section. Other contributions to the total free energy of the solute in solution (e.g., vibrational) are assumed (in work carried out so far) to make an identical contribution to the gas-phase and aqueous free energy, and thus not to affect ΔG^{o}_{S}. Clearly they would have to be added to G^{o}_{S} in order to obtain the total free energy in solution.

The polarization free energy is calculated from the generalized Born approximation [10,11,51-54,61-63,66,69,70,221,222] to the solution of the Poisson equation,

$$G_{P} = \frac{1}{2}\left(1 - \frac{1}{\varepsilon}\right)\sum_{k,k'} q_k q_{k'} \gamma_{kk'} \tag{19}$$

where q_k is the atomic partial charge on atom k, where k and k' run over all atoms ($k = 1,2,\ldots,k_{max}$), and where $\gamma_{kk'}$ is a coulomb integral. We have adopted the form for the coulomb integrals proposed by Still *et al.* [221]

$$\gamma_{kk'} = \{r_{kk'}^2 + \alpha_k \alpha_{k'} C_{kk'}(r_{kk'})\}^{-1/2}, \tag{20}$$

where α_k is the Born radius of atom k, $r_{kk'}$ is the interatomic distance between atoms k and k', and $C_{kk'}$ is in general given by

$$C_{kk'} = \exp(-r_{kk'}^2 / d^{(0)} \alpha_k \alpha_{k'}), \tag{21}$$

with $d^{(0)}$ being an empirically optimized constant equal to 4. Equation 20 is designed so that G_P behaves properly in three important limits for a pair of atoms k and k': infinite separation of atoms k and k', where it yields a sum of Born monatomic ion energies, coalescence of identical atoms, where it again yields the monatomic Born formula, and close approach of dissimilar atoms, where it approaches the Kirkwood-Onsager result.

For the monatomic case ($k = k'$ and $k_{max} = 1$), α_k is set equal to

$$\rho_k = \rho_k^{(0)} + \rho_k^{(1)} \left[-\frac{1}{\pi} \arctan \frac{q_k + q_k^{(0)}}{q_k^{(1)}} + \frac{1}{2} \right] \tag{22}$$

where $\rho_k^{(0)}$ and $\rho_k^{(1)}$ are empirically optimized parameters corresponding to positive and negative ionic radii, $q_k^{(0)}$ is the charge about which the switch is centered, q_k is the calculated partial charge, and $q_k^{(1)}$ has been fixed at 0.1 for all atoms. In the multicenter case, α_k is determined numerically so that the G_P which would be derived for the monatomic case (i.e., as if using equation 7) is equal to the G_P determined via a numerical integration [221,223,230,231] that corresponds to dividing the solute from solvent with a WSURF-type surface based on a set of spheres centered at nuclear location $\mathbf{x}_{k'}$ with radii $\rho_{k'}$, where $k' = 1, 2, \ldots, N$, and N is the number of atoms in the solute. In particular,

$$\frac{1}{\alpha_k} = \int_{\rho_k}^{\infty} \frac{dr}{r^2} a(r, \{\mathbf{x}_{k'}, \rho_{k'}\}_{k' \neq k}) \tag{23}$$

where $a(r, \{\mathbf{x}_{k'}, \rho_{k'}\}_{k' \neq k})$ is the fraction of the surface area of a sphere of radius r at the origin that is not contained in any of the $N - 1$ spheres specified by the set $\{\mathbf{x}_{k'}, \rho_{k'}\}_{k' \neq k}$ when the system is translated so that atom k is the sphere at the origin. The integration is

performed by precisely specified (and hence reproducible) quadrature rules in the SM1, SM1a, SM2, and SM3 models, and it is converged in SM4A and SM4P. The integral in equation 23 accounts for screening of the solute from the solvent by other parts of the solute; this solute screening effect is critical to solvation modeling [17,75,221,223,232] but is often neglected or—when included—often unappreciated.

At the valence-electron NDDO level, the electrostatic terms are calculated using the density matrix **P** of the aqueous-phase SCF calculation as

$$G_{ENP} = \frac{1}{2}\sum_{\mu\nu} P_{\mu\nu}\left(H_{\mu\nu} + F_{\mu\nu}\right) + \frac{1}{2}\sum_{k,k'\neq k} \frac{Z_k Z_{k'}}{r_{kk'}} - \frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_{k,k'} Z_k q_{k'} \gamma_{kk'} \qquad (24)$$

where **H** and **F** are, respectively, the one-electron and Fock matrices [233-235], μ and ν run over basis orbitals, Z_k is the effective (valence) nuclear charge of atom k. We point out here that both equation 5 of reference 225 and equation 19 of reference 230 (both being analogs of equation 24 above) are incorrect: the former is missing the last two terms of equation 24, and the latter is missing the final term and has "=" in the summation index instead of "≠".

The critical point in implementing the generalized Born model at this semiempirical level is that the Fock matrix is related to the energy functional of equation 18 as its partial derivative with respect to the density matrix. The partial charges which appear in the definition of G_P (equation 19) are themselves derived from the density matrix. In the SM1, SM1a, SM2, and SM3 models, this is accomplished by a simple Mulliken population analysis [147] under the assumptions of zero overlap

$$q_k = Z_k - \sum_{\mu\in k} P_{\mu\mu}. \qquad (25)$$

A more complicated dependence of atomic partial charges on the density matrix elements is used in SM4A and SM4P, which we call Charge Model 1A and 1P (CM1A or CM1P) [166]. In any case, using either formalism to define q_k, the partial differentiation of equation 18 is straightforward, and it delivers a solution-phase Fock matrix which *self-consistently* includes polarization effects. Thus, as for the other quantum models, self-consistency is required in these calculations: in particular, the Fock matrix, the density matrix, and the interacting solvent field are made self-consistent.

The last term in equation 18, which is required to accurately calculate absolute free energies of solvation, is calculated by a version of equation 16 that is modified in a way that depends on which of the SMx models is involved. For example, in SM2 and SM3, we use

$$G^{o}_{CDS} = \sum_{k'} \left\{\sigma^{(0)}_{k'} + \sigma^{(1)}_{k'}\left[f\left(B_{k'H}\right) + g\left(B_{k'H}\right)\right]\right\} A_{k'}\left(\beta_{k'}, \{\beta_k\}\right) \qquad (26)$$

where the $\sigma_{k'}$ are atomic surface tension parameters, and $A_{k'}(\beta_{k'},\{\beta_k\})$ is the solvent-accessible surface area for *non-hydrogen* atoms k'. The latter is defined as that portion of

ASURF associated with atom k' when the set of radii for all the atoms is $\{\beta_k\}$. We set R_S = 1.4 Å, and we set β_k = 0 for hydrogen atoms.

In the remainder of equation 26, $B_{k\mathrm{H}}$ is the sum of the bond orders, defined more specifically as covalent bond indices [236], of atom k to all hydrogen atoms in the solute, i.e.,

$$B_{k\mathrm{H}} = \sum_{\mu \in k, \nu \in \mathrm{H}} P_{\mu\nu}^2 \tag{27}$$

where μ runs over the atomic orbitals of atom k, and ν runs over all hydrogen orbitals. The hydrogen atom is specifically defined to have zero solvent-accessible surface area, and moreover not to block the accessible area of the atom(s) to which it is attached. Finally

$$f(B_{k\mathrm{H}}) = \tan^{-1}(\sqrt{3}\,B_{k\mathrm{H}}) \tag{28}$$

and

$$g(B_{k\mathrm{H}}) = \begin{cases} a_k \exp\left\{-b_k \Big/ 1 - \left[(B_{k\mathrm{H}} - c_k)/w_k\right]^2\right\}, & |B_{k\mathrm{H}} - c_k| < w_k \\ 0, & \text{otherwise.} \end{cases} \tag{29}$$

This more complicated expression of $G^{\mathrm{o}}_{\mathrm{CDS}}$ was found to be required in water because hydrogen atoms in that solvent interact with the first solvation shell differently depending on the heavy atom to which they are attached. For example, an alkane hydrogen is hydrophobic but an amine hydrogen is hydrophilic. To maintain flexibility and accuracy in the model, it was convenient to make the heavy atom surface tension be a function of the number of attached hydrogen atoms. We emphasize that the zero-radius treatment of hydrogen occurs only for the CDS term in SM2 and SM3, not for the other models, and never for the ENP term.

The various parameters (van der Waals radii, surface tensions, etc.) were fit to reproduce experimental aqueous solvation data. In practice, one begins the parameterization process by focusing on ions, for which most of the solvation free energy is found in the ENP term. This being the case, the ENP parameters are effectively "decoupled" from the CDS parameters, and indeed in the first pass through the ions, all surface tensions are set to zero. Once cavity radii have been defined for the ions, they are then used in neutral molecule calculations, with the difference between the calculated and experimental free energies being assigned to the CDS term. Surface tensions and radii are then fit using equation 26 in a multilinear regression to minimize residual error compared to experiment. These CDS terms are then used in another pass through the set of ions, and the ENP parameters are

allowed to relax accordingly to maximize agreement with experiment. This iterative procedure is continued until the parameters converge, usually by the second or third pass through the data. Throughout this process the parameters are monitored so as to ensure the location of a physically meaningful local minimum in parameter space.

As a consequence of this type of parameterization, the CDS parameters do more than account for cavitation, dispersion, and structural rearrangement. In particular, they correct for the impact on the ENP terms of errors in the NDDO wave function, in the representation of a continuous charge distribution by a set of partial charges on the atoms, and in the generalized Born approximation to solving the Poisson equation. In addition in models SM1, SM1a, SM2, and SM3, they also account for systematic numerical errors in the quadrature of equation 23.

We have reviewed the performance of the AM1-SM2, AM1-SM1a, and PM3-SM3 models elsewhere [75,230] and the SM4A and SM4P models will be described in a forthcoming paper [229]. We note that the parameterizations ultimately permit the prediction of absolute free energies of solvation for a large set of neutral molecules (> 150) with a mean unsigned error of less than 1 kcal/mol (where the data span a range from roughly +5 to –10 kcal/mol) and for about 30 ions with a mean unsigned error of about 3-4 kcal/mol. The latter number is well within experimental error since the measurements require the completion of thermodynamic cycles which include gas-phase deprotonation enthalpies. The models are all available in the semiempirical package AMSOL [237], and they have also been implemented in commercial software packages. All of the calculations discussed below were performed with various versions of AMSOL. The remainder of this contribution will focus on the application of these models to systems of biological interest.

2.4. Nonequilibrium solvation

The above considerations all apply to the solvent being equilibrated to the solute and vice versa. This is a reasonable assumption for free energies of solvation of solutes executing small-amplitude vibrations around a single equilibrium structure. For dynamics problems, though, one must sometimes consider non-equilibrium solvation. The theoretical treatment of nonequilibrium solvation involves a careful consideration of time scales [238], and it is much less well understood than equilibrium solvation. Nonequilibrium solvation effects can be included in dynamics calculations by treating solvent degrees of freedom explicitly, or they can be incorporated as corrections to transition state theory [239,240] The reader is directed elsewhere for recent literature on continuum models of nonequilibrium solvation [241-245].

Another area where nonequilibrium solvation is important is electronic spectroscopy. To a first approximation, excitation from an electronic ground state into an excited state occurs much more rapidly than reorganization of the structure of the surrounding solvent shells. As such, the solvent reaction field acting upon the excited state at the instant of excitation is just that field that was derived from the ground state. Over a longer time period, solvent relaxes to a new equilibrium reaction field, and if radiative emission occurs to create the ground state, again the instantaneous reaction field experienced by the ground state will be that for the excited state. This differential solvation can lead to a change in Franck-Condon factors and significantly shifted absorption maxima in solution compared to the gas phase. There is great interest in the theoretical prediction of the effects of solvation on

spectroscopy, and again we will simply refer the reader to recent discussions of these issues [126,177,179,246-252].

3. APPLICATIONS TO BIOLOGICALLY INTERESTING SYSTEMS

A major motivation for the development of aqueous solvation models is the modeling of biological and pharmacological structure and reactivity. For instance, it is clear that the development of the tertiary structure of non-membrane-associated proteins is driven in part by the energetically favorable tendency for them to bury hydrophobic amino acid residues in the interior of the protein while leaving hydrophilic residues exposed to the surrounding aqueous environment [104,209,253-257]. Moreover, the energetics of substrate-enzyme or substrate-receptor binding may be viewed as a differential solvation effect between the active site and the bulk solvent [23,258], i.e., electrostatic and binding stabilization of the substrate at the active site is in competition with solvation of the free substrate in aqueous solution. The references in this paragraph provide only a small sampling of the wide activity in this field.

Sections 3.1 and 3.2 are dedicated to specific examples of the effects of aqueous solvation on systems of biological import. Section 3.1 illustrates for the nucleic acids how the electronic structure changes as a result of solvent-induced polarization, and section 3.2 provides examples of how aqueous solvation influences the equilibrium population of conformers for flexible biomolecules.

3.1. Aqueous solvation effects on electronic structure — the nucleic acids

Quantum mechanical studies of the nucleic acids [113,129,259-262] and classical mechanical studies of their molecular dynamics, both in the gas phase and when surrounded by hundreds or thousands of water molecules [26,263-269], have done much to advance our understanding of these important molecules. As discussed in section 2, classical simulations involve molecular mechanics force fields which are parameterized by electronic structure calculations and/or semiempirical fitting to experimental data. The charges built into such force fields are not subject to solvent-induced polarization. That is, empirically optimized partial atomic charges on the individual atoms are taken to be constant and are not self-consistently determined with respect to either molecular conformation or environment. However, quantum mechanical calculations on solutes using the self-consistent models detailed in section 2 nearly always indicate electronic polarization to be a non-negligible component of the overall solvation free energy. In very polarizable systems, like heterocycles, the effect can in fact be quite large [259,270,271].

Such a finding raises interesting questions about the requirements for realistic force fields for molecular mechanics simulations of polarizable solutes in aqueous media. Potential functions optimized for liquid-phase simulations [267,269,272] must necessarily have larger partial charges and dipole moments than would be appropriate for the gas-phase molecules. One way to accomplish this is to parameterize the force field based on calculations on hydrogen bonded or ion-water complexes. A popular alternative has been to base the partial charges on ab initio Hartree-Fock wave functions derived using the 6-31G* basis set [273], since calculations with this basis set are known to overestimate molecular dipole moments by about 10% in many cases [125]—this appears to be roughly the

contribution of polarization to the overall electrostatic component of the free energy of solvation. It is clear, however, that such an approach will be of limited value in more polarizable systems or in systems where it is difficult to account for the bulk of the solvent-induced polarization by performing electronic structure calculations for the solute when it is complexed to only one or two water molecules. The nucleic acids appear to serve as a particularly important example of this phenomenon. We have reported AM1-SM2 and PM3-SM3 calculations for these solutes [259], and the two methods are in reasonable agreement for the absolute free energies of solvation. Both models predict that solvent induced polarization of the solute accounts for 23–34% of ΔG_S^o. Since experimental free energies of solvation remain to be measured for the nucleic acid bases, it is most instructive to compare diverse theoretical models in order to assess the relative importance of polarization to the aqueous solvation of these molecules.

One pertinent study is that of Gao and Xia [274], who estimated polarization effects on solute-solvent interaction energies by a combined quantum/classical mechanical approach which included 260 molecular mechanics water molecules in the AM1 Hamiltonian with the nucleic acid bases constrained to their gas-phase geometries. Although calculations of the relative free energies of hydration are expensive by this method and were therefore not carried out, canonical ensemble averages are less expensive, and average values of the molecular dipole moments were computed for each solute. Katritzky and Karelson [113] have also studied the nucleic acid bases, employing the Kirkwood-Onsager approximation described in section 2.1.3 as implemented into the AM1 Hamiltonian. Since both of these studies, as well as our own AM1-SM2 calculations, start from the AM1 gas-phase Hamiltonian, it is particularly interesting to compare them. Table 1 lists the dipole

Table 1.
Dipole moments of the nucleic acid bases in aqueous solution compared to the gas phase (Debyes)

		In solution			
		At gas-phase geometries			At optimized solution geometries
	In the gas phase	Katritzky and Karelson	Gao and Xia	AM1-SM2	AM1-SM2
cytosine	6.3	7.4	9.9	9.0	10.0
thymine	4.2	5.2	5.9	6.2	6.6
uracil	4.3	5.0	5.9	6.4	6.9
adenine	2.2	2.9	3.8	3.1	3.1
guanine	5.9	6.6	8.5	8.5	9.3

Table 2.
Absolute free energies of solvation for the methylated nucleic acid bases (kcal/mol)

	AMBER [275]	AMBER [276]	OPLS [277]	OPLS [278]	6-31G* [279]	AM1-SM2 frozen	AM1-SM2 relaxed
1-Methylthymine	−7.7	−9.4	−10.4	−9.9	−8.6	−10.8	−13.3
9-Methyladenine	−12.8	−14.9	−10.8	−9.4	−6.5	−16.7	−20.9
1-Methylcytosine	−12.9		−16.8	−17.9	−13.0	−14.4	−18.7
9-Methylguanine	−19.8		−19.7	−19.5	−16.0	−18.2	−24.3

moments found for the nucleic acid bases in the gas phase, and with each of the solvation models, for the gas-phase geometries [i.e., only the electronic structure (not the geometry) has been permitted to relax in the presence of solvent]. We have also optimized the solutes at the AM1-SM2 level in order to study the additional effects of geometric relaxation.

Three points are especially worthy of note. First, the increase in the molecular dipole moment is smallest for the Kirkwood-Onsager model, as expected given the severely limiting assumptions of that model (as discussed in section 2.1.3). Second, the results of Gao and Xia are remarkably consistent with our own, especially considering the different representations of the solvent either as explicit and classical or as a continuum dielectric. Finally, the additional effects of geometric relaxation are not trivial for several of these solutes, illustrating the importance of reoptimizing geometries in solution.

To illustrate the importance of the polarization contribution to the absolute free energy of solvation, Table 2 presents the calculated free energies of solvation for the methylated nucleic acid bases of DNA as calculated by several different methods [259,275-279]. Elcock and Richards [278] calculated only relative free energies of solvation, and for comparison purposes their results have been arbitrarily normalized to sum to the same value as those of Mohan et al. [277] who used the same charge model. The results of Elcock and Richards and of Bash et al. [275] are from classical simulations employing the Optimized Parameters for Liquid Simulations (OPLS) [267,272] and AMBER [269] charge sets, respectively, the results of Mohan et al. [277] are from a numerical solution of the classical Poisson equation (i.e., non-self-consistent charges are employed) using the OPLS charge set, and the results of Young and Hillier [279] are based on a multipole expansion through l=7, with the 6-31G* basis set. The results of Ferguson et al. [276] are based on repeating the calculations of Bash et al. [275] with new algorithms. AM1-SM2 results are provided as calculated both for the frozen gas-phase molecules (i.e., no relaxation of either electronic structure or molecular geometries) and for self-consistently optimized solutes. Although it will require experimental data to establish which model is most consistently accurate for these molecules, it is clear that consideration of polarization dramatically increases the

AM1-SM2 calculated absolute free energies of solvation, which are otherwise in reasonable agreement with all of the other non-self-consistent models, with the exception of 9-methyladenine. The wide variation between even the different classical models illustrates the need for continued study of the solvation free energies of nucleic-acid bases.

3.2. Aqueous solvation effects on molecular conformation

The nucleic acids are examples of systems where solvation induces large changes in electronic structure. In addition to the change in *electronic* structure, there is an additional small change in molecular geometry, as evidenced by comparing the last two columns of Table 1. Further inspection, however, reveals that really there have been only very small changes in bond lengths and bond angles, the net result of which has been to permit additional electronic relaxation with minimal geometric distortion. Put another way, bond stretching potentials and bond angle bending potentials are very steep for the nucleic acids compared to the additional solvation free energy which may be gained by distortion from the gas-phase equilibrium point for any given degree of freedom. In order to observe a significant effect of solvation on molecular geometry, one of two situations must hold: either the gain in solvation free energy to be had by geometric distortion must be quite large, so that it competes with such steep potentials as those mentioned above or alternatively the geometric potential itself must be rather shallow or involve more than one minimum, in which case solvation may be instrumental in determining the shape of the potential. A good example of the latter situation is a torsional coordinate in a flexible molecule. Torsions often exhibit multiple minima and are characterized by fairly low-energy barriers separating the minima. In section 3.2.1, we discuss this situation in more detail for the neurotransmitter dopamine. In section 3.2.2, we expand the level of complexity to consider a number of conformational issues relevant to polyalcohols, to include ethylene glycol and the two anomers of D-glucopyranose.

3.2.1. Dopamine

At physiological pH, the neurotransmitter dopamine exists predominantly in its protonated form [280]. As a consequence, dielectric shielding may play a significant role in the stabilization/destabilization of particular conformers [281]. We focus in particular on the torsional isomerism about the carbon-carbon single bond which is defined by dihedral angle ϕ_1 in Figure 4. In general, torsion about the other carbon-carbon single bond, denoted by dihedral angle ϕ_2, is not particularly affected by solvation, and we restrict our discussion to cases where that torsion has been fixed to its near optimal value of 90°. Figure 4 illustrates that in the absence of solvation, the global minimum occurs for $\phi_1 = 60°$, i.e., when the ammonium group is *gauche* to the aromatic ring (the catechol hydroxyl groups introduce a slight asymmetry into the system, so that $\phi_1 = -60°$ is not quite isoenergetic). This is consistent with numerous theoretical [282,283] and experimental [284,285] studies which illustrate that, in the absence of solvation, aromatic pi clouds interact very favorably with ammonium cations, the most notable examples being interactions in enzyme binding sites [282,285].

When aqueous solvation was included by using the original AM1-SM1 method, the situation changed markedly [281]. The *trans* isomer ($\phi_1 = 180°$) is considerably stabilized relative to the *gauche* rotamers. As a result, the *trans* rotamer became the global minimum

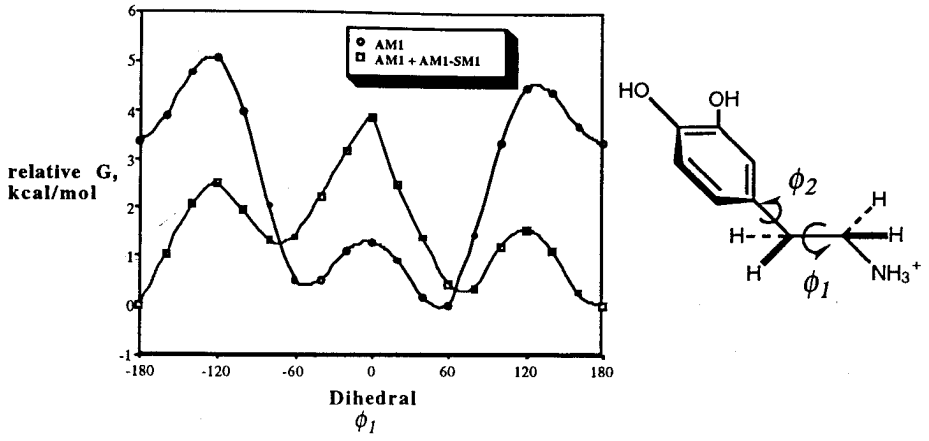


Figure 4. Electronic-nuclear energy of dopamine in the gas phase and free energy of dopamine in aqueous solution as a function of the torsion about the sp^3-sp^3 carbon-carbon single bond. For all points on the torsional coordinate, the dihedral angle ϕ_2 is fixed at 90°, which is the value illustrated in the molecular structure.

on the rotational coordinate because of the considerably smaller dielectric shielding experienced by the positively charged amino group when it is more distant from the bulky aromatic ring. Expressedly differently, the hydrophilic ammonium group is more accessible to solvent in the extended *trans* conformation.

In order to assess the quantitative accuracy of the model, it is useful to compare to rotameric populations determined from aqueous nuclear magnetic resonance (NMR) studies [286]. In particular, the gauche:anti ratio about ϕ_1 has been observed by NMR to be 58:42. Of the two theoretical models, AM1 (gas phase) calculations predict a >99:1 ratio and AM1 + AM1-SM1 (implying a standard-state free energy in aqueous solution arrived at by adding the AM1 gas-phase energy to the AM1-SM1 free energy of solvation) predicts 37:63. The latter represents an error of only 0.5 kcal/mol. We have here repeated these calculations using the more recent AM1-SM2 model, and the results are similar: the predicted aqueous gauche:anti ratio is 12:88 which corresponds to an error of 1.4 kcal/mol. It is worth noting that the calculated ratios rely in part on the relative accuracy of the AM1 gas-phase surface to which the solvation free energies are being added in order to arrive at Boltzmann-averaged equilibrium populations. Thus, it is impossible to judge which model, SM1 or SM2, is actually predicting the differential free energies of solvation more accurately. The fact that they are within one kcal/mol of each other reflects the similar way in which the two models treat the ENP portion of the solvation free energy, which is the dominant term for the dopamine cation. We will not devote extensive discussion to ϕ_2,

however we do note that AM1-SM1 predicts aqueous solvation to lower the rotational barrier about the indicated bond [281], and this result is consistent with the rapid rotation observed experimentally by NMR [286].

3.2.2. Ethylene glycol and glucose

Accurately modeling the aqueous solvation of 1,2-ethanediol (ethylene glycol) provides a very challenging test for continuum solvation models insofar as much of the favorable solvation of this molecule results from numerous hydrogen bonding interactions with the water molecules found in the first solvation shell. In the SMx models, such interactions are accounted for by parameterization in the surface tension terms, and it is of considerable interest to analyze the robustness of this scheme. Ethylene glycol is additionally challenging because its equilibrium population at room temperature is composed of numerous isomers. Figure 5 illustrates the ten possible symmetry-unique isomers for this system. The nomenclature refers to the torsion angles about the leftmost C-O bond (lower case), the central C-C bond (upper case), and the rightmost C-O bond (lower case). A "g" (or "G") indicates a gauche torsion angle (either positive or negative, as marked) while a "t" (or "T") indicates the torsion to be trans. Although the presence of three torsions, each of which is characterized by three rotational minima, would normally give rise to $3^3 = 27$ separate minima, symmetry in this system causes a number of these cases to be degenerate. The numbers in parentheses beneath the molecular structures in figure 5 indicate that degeneracy, and, as required, these sum to 27.

Experimentally, it has been established by a number of methods that isomer tG^+g^- predominates, there is a smaller fraction of $g^+G^+g^-$, and there is a barely detectable fraction of $g^-G^+g^-$ [287-295]. High level ab initio calculations which employ very large basis sets

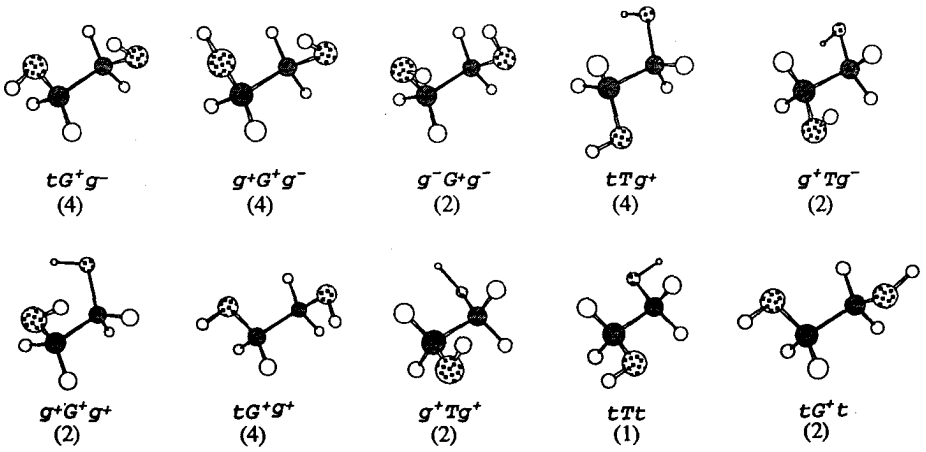


Figure 5. The ten unique structures of 1,2-ethanediol. Geometries were optimized at the MP2/cc-pVDZ level of theory.

(the correlation-consistent polarized valence triple-ζ basis of Dunning [296]) and account for correlation at sophisticated levels (coupled cluster analysis including all single and double excitations with a perturbative treatment of triples [297]) reproduce these trends nicely [298]. Results are summarized in Table 3. It is noteworthy that lower levels of theory, especially semiempirical models, do very poorly at reproducing these results.

Table 3 also presents the equilibrium population predicted by addition of PM3-SM3 solvation free energies for structures fully optimized in aqueous solution to the ab initio gas-phase energies. Comparison may again be made to experiment, in this case referring to aqueous NMR measurements which estimated the percentage of conformers *gauche* about the C–C bond to be 88±3% [299]. This is in good agreement with the predicted value of 92%. Results based on adding AM1-SM2 solvation free energies to the ab initio solute energies were found to be quite similar, although in that instance the gas-phase geometries were employed since AM1 geometries are qualitatively incorrect [298].

Table 3.
Equilibrium populations of ethylene glycol conformers in the gas phase and in aqueous solution (%)

Isomer	Triple-ζ[a] $G^o_{298}(gas)$	Triple-ζ[a] + PM3-SM3 $G^o_{298}(aqueous)$
tG^+g^-	55.9	45.8
$g^+G^+g^-$	27.4	31.0
$g^-G^+g^-$	13.4	14.0
tTg^+	1.3	4.2
g^+Tg^-	0.6	2.5
$g^+G^+g^+$	0.4	0.4
tG^+g^+	0.3	0.6
g^+Tg^+	0.2	0.9
tTt	0.2	0.6
tG^+t	0.2	not stationary
Total C–C gauche	97.6	91.8
Total C–C trans	2.4	8.2
Internal H-bond	83.4	76.8
No internal H-bond	16.7	23.2

[a] Ab initio Hartree-Fock calculations.

Finally, the population-averaged absolute free energy of solvation may be calculated from [300]

$$\exp\left[-\Delta G_S^0/RT\right] = \sum_C P_C \exp\left[-\Delta G_S^0(C)/RT\right] \tag{30}$$

where P_C is the equilibrium mole fraction of conformation C in the gas phase. Following this procedure, we find the PM3-SM3 value for ΔG_S^0 to be –9.4 kcal/mol, in outstanding agreement with a very recent experimental measurement of –9.6 kcal/mol [301].

The solvation of glucose, a monosaccharide, is expected to be dominated by the same effects as those which are important for ethylene glycol. Of course, the degree of complexity present in the sugar molecule is considerably greater than that observed for the simple diol! Since thorough discussions of the energetics of glucose are available [302-306], especially as regards the influence of aqueous solvation, we focus here on the contributions of the SMx models to further illustrate the applicability of these continuum solvation models [307].

Several noteworthy conformational equilibria have been experimentally determined. NMR experiments have been interpreted to indicate that only the $\bar{G}$ and G

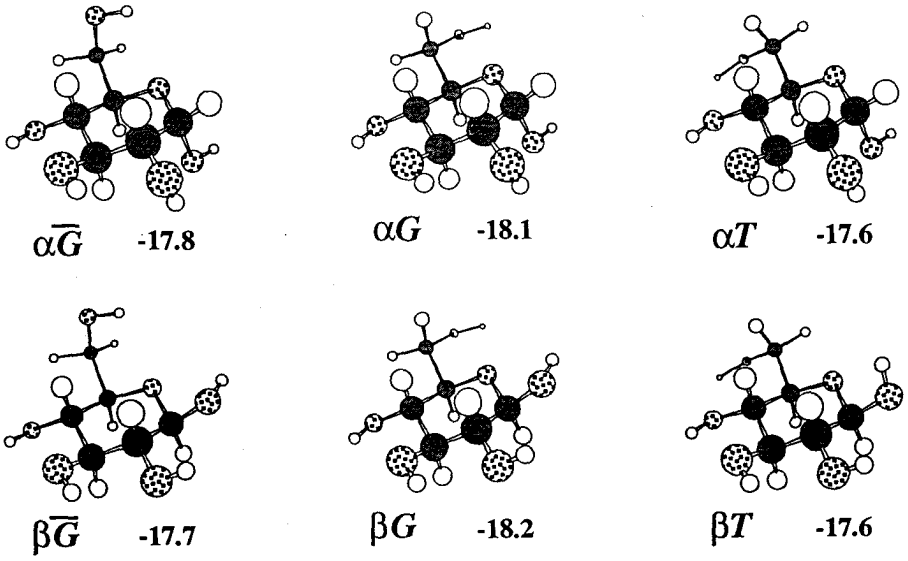


Figure 6. The important conformers for D-glucose. Under each structure is the nomenclature designating the anomer (α or β 2-hydroxyl group), the hydroxymethyl conformer ($\bar{G}$, G, or T), and the AM1-SM2 value predicted for ΔG_S^0 in kcal/mol.

hydroxymethyl conformers are present in aqueous solution (Figure 6) [308]. Consistent with this observation, we find the alternative T conformer to be the highest in energy in aqueous solution. This is due in part to a differential solvation effect: the T conformer is 0.6 kcal/mol less well solvated than the lowest energy G conformer for both anomers. This preferential solvation of the G conformer is in good agreement with results from explicit-solvent molecular simulations [304,305], once again illustrating the ability of the SMx models to accurately reflect first-shell solvation effects. This agreement extends to consideration of the effect of solvation on the anomeric equilibrium, which is known to be 36:64 α:β in aqueous solution [309]. The effect of solvation on the anomeric equilibrium, as calculated from the Boltzmann-averaged equilibrium populations using equation 34, is predicted to be zero. This may be compared to the results of Ha et al. [305] and van Eijck et al. [302] who used explicit-solvent aqueous simulations to predict a differential solvation effect for the β anomer relative to the α of only 0.6 ± 0.5 kcal/mol and -0.5 ± 0.2 kcal/mol, respectively. The SMx prediction is thus halfway between the two simulation results. Moreover, when the AM1-SM2 solvation free energies are added to Hartree-Fock gas-phase relative free energies [306], the predicted anomeric equilibrium is within 0.8 kcal/mol of the experimental result. While more work clearly remains to be done with respect to converging the gas-phase relative free energies and examining in more detail the individual components of solvation, this is heartening agreement.

4. OTHER SOLVENTS

Section 4.1 discusses the development of solvation models for solvents other than water, and Section 4.2 briefly reviews empirical relationships for the prediction of solvent-dependent properties.

4.1. Modeling solvation in non-aqueous solvents

Although water is clearly the single most important solvent in which chemistry of biological relevance takes place, there are many solvents which find extensive use in other areas, e.g., organic synthesis, or which find use in the modeling of biological processes, e.g., octanol and hexadecane are sometimes used to model lipid membranes. For such solvents, all of the quantum mechanical continuum models discussed in Section 2 may be used for calculation of the ENP portion of the standard-state gas to condensed phase solvation free energy (i.e., free energy of transfer). One simply uses the dielectric constant appropriate for the solvent being modeled. The CDS portion of the solvation free energy, on the other hand, is not so straightforwardly addressed, and unfortunately very little work has been carried out incorporating the CDS effects into quantum mechanical continuum solvation models.

Within the SMx models, the CDS terms must be parameterized separately against available experimental data for each new solvent, and it is not clear, since they account for several physical effects as well as for errors in the ENP terms, that they will be related to any particular bulk property of the solvent in question. This presents the most significant hurdle to the development of new solvent parameter sets. Presently, the only published SMx models are specific for water. However, a *n*-hexadecane parameter set is in the final stages of completion, an octanol parameter set will be available shortly, and other ethereal, hydroxylic, and other solvents will be parameterized as well. Improved methods for arriving at the atomic partial charges (e.g., CM1A, CM1P) and/or solving the Poisson equation will

reduce these "error-correcting" contributions, and more clear correlations between the remaining surface tensions and such bulk properties as macroscopic surface tension, viscosity, cohesive energy density, etc., may be discernible.

In the parameterization of the *n*-hexadecane model [310], we have noted one refinement which appears to be required for larger solvent molecules: significant dispersion interactions involve only that part of the solvent within a distance R_{CD} of the solute, where $R_{CD} < R_S$. (Recall that R_S is the solvent radius.) In such instances, it is useful to generalize the concept of solvent-accessible surface area. In particular, we separate the CDS effects, very approximately, into a CD part and a CS part. Then, for the former, we use the ASURF definition taking this smaller radius, R_{CD}, for the rolling probe—this might be called DSURF. For water, which is quite small, we have implicitly taken R_{CD} equal to R_S, which is 1.4 Å. This is a reasonable value for the "radius" of a water molecule in liquid water [93]. Since R_S for water is so small, it is reasonable to assume that dispersive interactions will operate over the entire molecule. For a molecule like *n*-hexadecane, on the other hand, an R_{CD} on the order of 1.7 Å works best [310]. This radius mimics the size of one of the methyl or methylene groups forming a portion of the *n*-hexadecane solvent molecule. This is quite a bit smaller than the R_S of about 5 Å which may be calculated either from the bulk density of liquid *n*-hexadecane by assuming a spherical molecule, or from taking the calculated volume of the molecule using overlapping spheres with van der Waals radii and choosing R_S to be the spherical radius providing the same volume. The former method is less ambiguous for a flexible molecule like *n*-hexadecane, although studies on both the aqueous ASURF of *n*-heptane [311] and the WSURF of *n*-decane [312] have found that the average molecular surface areas for the Boltzmann-weighted population of conformers at 298 K are only 3% and 4% smaller, respectively, than the surface area for the fully extended chain conformation—obviously the effect on the volume will be similarly slight. Although it is clearly possible that the remaining cavitation and structural rearrangement terms might also each require a different radius, i.e., an R_{CS} that is not simply equal to R_S, we did not observe that to be required for *n*-hexadecane.

We have illustrated in Section 3 a number of instances where molecular structures and energetics change significantly on going from the gas phase to aqueous solution. Of course, experimental observations in the gas phase are not generally possible for medium- to large-sized organic and biological molecules. However, there is considerable data available detailing the effects of *changing* solvents from one to another. Indeed, certain partition coefficients, which describe the equilibrium concentration of a solute in two different liquids, prove to be quite useful in drug design [313,314]. The most notable is the octanol-water partition coefficient [313-320]; the octanol-water system is considered to be a good model for biological membranes which are composed of amphiphiles like octanol.

The development of new semiempirical quantum chemical continuum solvation models will permit a comparison to more empirical models which exist to explain the effect of different solvents on various chemical properties. Those latter models are the subject of the next section.

4.2 Quantitative Structure-Activity Relationships and Linear Solvation Energy Relationships

Although the calculation of molecular properties from first principles is an esthetically appealing proposition, it is often impractical. Difficulties may arise for relatively

simple reasons, e.g., the system of interest is simply too large to be tractable for trustworthy levels of theory. A more fundamental problem may be that no theoretical model exists to accurately predict the chemical property which is of interest. As a concrete example, consider that the antitumor activity of a series of structurally related organic molecules is known, and the objective is to predict the related activity of an as yet unsynthesized set of additional congeners. In such instances, it often proves useful to pursue an empirical approach which seeks to relate a variety of "simpler" molecular properties to the biological activity. Such an approach is embodied in the techniques of quantitative structure-activity relationships (QSAR) [321,322]. In essence, a QSAR analysis involves a regression equation that correlates microscopic features of a set of chemical compounds with some macroscopic property. This approach has demonstrated its utility repeatedly and it finds widespread use in medicinal and pharmaceutical chemistry.

A different application of regression analysis, which is nevertheless similar in spirit, is to be found in the so-called "linear solvation energy relationship" (LSER) formalism developed by Kamlet, Taft, Abraham, and co-workers [323-326]. Seeking to explain the effects of different solvents on chemical properties and reactivities, these researchers proposed an equation of the form

$$\log \gamma = c_0 + c_1 \alpha + c_2 \beta + c_3 \pi^* \tag{31}$$

where γ is the solvent-influenced property of interest for a particular solute (typically free-energy based, and hence the logarithm), and α, β, and π^* are solvent-specific constants which describe the solvent's hydrogen bond donating, hydrogen bond accepting, and dipolarity/polarizability properties, respectively. These constants are determined by measurement of some reference process in the solvent, e.g., the shift in the absorption maximum for a particular dye whose absorbance is sensitive to hydrogen bond interactions. As a result of choosing this kind of a reference process, these parameters have come to be known as "solvatochromic parameters". In practice, this equation finds use in the following fashion. Following a series of experimental measurements of $\log(\gamma)$ in a number of solvents for which the solvatochromic parameters are known, one obtains optimal constants c_i by regression analysis of the data. In principle, one may now predict $\log(\gamma)$ for *any* solvent for which the solvatochromic parameters have been tabulated. There are clearly dangers involved in extrapolating from a limited data set if the solvents selected for experimental measurement were not representative of those for which predictions are being made; nevertheless the method is quite powerful and has demonstrated itself to be reasonably robust.

It will be particularly interesting to explore the interplay of ENP and CDS effects in the parameters α, β, and π^* when continuum solvation models are available for comparison to LSER approaches.

5. PROPERTIES OF SOLVATED MOLECULES AS PREDICTORS FOR STRUCTURE-ACTIVITY RELATIONSHIPS

This section discusses how calculated molecular properties which take account of solvation may be used in empirical structure-activity relationships to provide guidance in molecular design. Although the formalism discussed in the preceding section was originally developed to model the effect of solvents on some specific solute property, the regression analysis has been generalized to allow prediction of some solute property based on *solute* solvatochromic parameters [325,327]. That is,

$$\log \Gamma = c_0 + c_1 \alpha_2 + c_2 \beta_2 + c_3 \pi_2^* \tag{32}$$

where Γ is upper case to indicate a general property exhibited by a large number of solutes (as opposed to γ, which may be unique to a single solute within the formalism of equation 31) and the subscript "2" indicates the solvatochromic parameters to be associated with the solute, not the solvent (which is kept constant in equation 32). Solute-specific solvatochromic parameters continue to be arrived at by measurement of some reference process. An important point to note is that this is analogous to a QSAR analysis, i.e., a chemical property is being correlated with some set of other, measurable properties, albeit the latter approach developed more from a background of linear free-energy relationships [328-330] than from a purely empirical impetus. Equations 31 and 32 can be combined in a very general form, viz.,

$$\log \Xi = \sum_i C_i^{\text{solvent}} d_i^{\text{solute}} \tag{33}$$

where the parameters C_i and d_i may be solvatochromic parameters, or other general parameters [324,331-340].

An interesting variation on this theme, which has been pursued extensively by Famini and Wilson [341-344], is to maintain the formalism of equation 32, but to replace the solute parameters α_2, β_2, and π_2^*, which must be experimentally determined, with other constants derived from theoretical calculations. The "theoretical" linear solvation free energy relationship (TLSER) ansatz is expressed as

$$\log \Gamma = c_0 + c_1 V_{mc} + c_2 \pi^* + c_3 \varepsilon_\alpha + c_4 \varepsilon_\beta + c_5 q^+ + c_6 q^- \tag{34}$$

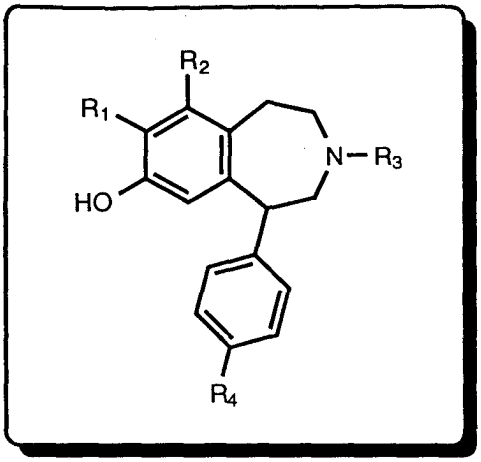
where V_{mc} is the molecular van der Waals volume, and π^* is a polarizability term derived from the calculated polarization volume (the latter should not be confused with the solvatochromic descriptors π^* and π_2^*, which are experimentally measured quantities). Just as with LSER, hydrogen bonding is separated into donor and acceptor components. Since

all intermolecular interactions can be considered to have varying degrees of covalent and electrostatic components, separate descriptors have been chosen to address each of these two limiting paradigms. The covalent contribution to Lewis basicity, ε_β, is taken as the difference in energy between the lowest unoccupied molecular orbital (E_{LUMO}) of the solute and the highest occupied molecular orbital (E_{HOMO}) of water, i.e., smaller values of ε_β denote a greater covalent basicity. The electrostatic basicity contribution is denoted by q^-, the magnitude of the most negative atomic partial charge in the molecule. Analogously, the hydrogen bonding acidity is divided into two components: ε_α is the energy difference between E_{HOMO} of the solute and E_{LUMO} of water, and q^+ is the magnitude of the most positively charged hydrogen atom in the molecule. Calculation of these descriptors may be performed at any level of theory, of course, but to date the emphasis has been on using the NDDO semiempirical level of theory (as mentioned in Section 2.3, this is the level used in the AM1 and PM3 models) so as to take advantage of the relative efficiency of this method for larger molecules.

Although equation 37 has been quite successful at predicting a number of interesting chemical and biological properties [341-346], it is by no means the only possible way to relate calculated molecular properties with activities/toxicities. Lewis has recently provided an extensive review detailing other descriptors which have found use in TLSER-like regression analyses [347]. Politzer and Murray and co-workers have also provided alternative formulations of this approach [348-350]. One additional point which should be mentioned is that it is perfectly logical to consider developing regression equations which consider *both* experimentally determined and calculated descriptors [351,352]. Of course, part of the attraction of the purely theoretical methodology is that it may be considerably simpler and more economical to perform the calculations rather than synthesize the solute in order to measure some parameter if it is not already known.

An interesting question is how solvation may influence the descriptors present in the regression models [353]. As has been discussed at length above, both electronic and geometric structure may change significantly for a solute in an aqueous environment (e.g., in vivo) and it seems clear that any TLSER-like regression equation being used to predict properties in such a situation should accurately take account of that. In particular, the free energy of solvation *itself* may be a particularly important descriptor. Activity in a biomolecule typically requires its interaction with a receptor and/or, as mentioned in Section 4, its crossing of a hydrophobic cell membrane; since these processes usually require desolvation which may or may not be balanced by specific interactions within a receptor site, the cost of desolvation can influence the overall activity.

One interesting example illustrating this point has been provided by Alkorta et al. [354], who have analyzed the affinity (expressed as $1/K_i$, where K_i is the dissociation constant of the enzyme-substrate complex) of a set of 3-benzazepine cations for the dopamine D_1 receptor. Figure 7 illustrates the gross structure of these substrates. In this case, the descriptors chosen were the AM1-SM1 calculated free energies of solvation (ΔG^0_S), the dot product of the solute dipole moments with the unit vector parallel to the activity-weighted sum of *all* of the solute dipole moments (μ^*, i.e., this descriptor measures both the magnitude of the molecular dipole moment and the degree to which it is aligned with the dipole moments of the compounds observed to be most active), and the molecular polarizability (α). Using this approach they obtained a regression equation of:



R_1 = H, F, Cl, Br, Me, OH, OMe, CN

R_2 = H, F, Cl

R_3 = H, Me

R_4 = H, Cl

Figure 7. 3-Benzazepines examined by Alkorta et al. Not all possible permutations of the listed R groups were explored.

$$\log (1/K_i) = -3.46 - 1.67\ \mu^* + 0.16\ \Delta G_S^o + 0.07\ \alpha \qquad (35)$$

with a multiple correlation coefficient R of 0.925 for a data set of 13 compounds. An analysis of the statistical significance of α suggests that it is the least useful parameter in the regression equation, which is perhaps not surprising since it represents a gas-phase property; its removal leads to only a slight drop in R (0.853). The most interesting aspects of this analysis, however, were the observations of Alkorta et al. [354] that (1) not even qualitative activities could be predicted from regression analyses performed for the neutral (i.e., non-protonated) 3-benzazepines, consonant with the expected protonation of these compounds in vivo, and (2) methylation of the azepine nitrogen leads to reduced solvation free energies as a result of dielectric shielding and loss of hydrogen bonding opportunities; this *increases* the affinity of the methylated compounds since it decreases the cost of their desolvation.

We anticipate that the continued development of rapid and accurate continuum solvation models will give rise to increasing use of such analyses in structure-activity predictions. The cost-effectiveness of the methodology for the pre-screening of potential synthetic targets gives it considerable practical importance.

6. CONCLUDING REMARKS

Quantum mechanical continuum solvation models span a wide range of complexity and utility in their ability to calculate electrostatic contributions to free energies of solvation.

When supplemented with models that also include those portions of the solvation free energy associated with the first surrounding shell of solvent, they become particularly powerful tools for studying complicated, condensed-phase systems. In particular, they can be used to provide insight into solutes that play important roles in biological systems. This information includes quantitative estimates of free energies of solvation, predictions of the detailed changes in electronic and geometric structure of biomolecules in solution, and calculation of micro- and macroscopic properties for use in structure-activity prediction. Continued comparison to explicit-solvent simulations and experimental results will be instrumental in improving the models.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under grant no. CHE89-22048 and by the ILIR program of the United States Army.

REFERENCES

1. C. Reichardt, Solvents and Solvent Effects in Organic Chemistry, VCH, New York, 1990.
2. J.-L.M. Abboud, R. Notario, J. Bertrán, and M. Solà, Prog. Phys. Org. Chem., 19 (1993) 1.
3. J.L. Gao and X.F. Xia, J. Am. Chem. Soc., 115 (1993) 9667.
4. C.-C. Han, J.A. Dodd, and J.I. Brauman, J. Phys. Chem., 90 (1986) 471.
5. J. Chandrasekhar, S.F. Smith, and W.L. Jorgensen, 107 (1985) 154.
6. J. Chandrasekhar and W.L. Jorgensen, J. Am. Chem. Soc., 107 (1985) 2974.
7. P.A. Bash, M.J. Field, and M. Karplus, J. Am. Chem. Soc., 109 (1987) 8094.
8. J.-K. Hwang, G. King, S. Creighton, and A. Warshel, J. Am. Chem. Soc., 110 (1988) 5297.
9. S.E. Huston, P.J. Rossky, and D.A. Zichi, J. Am. Chem. Soc., 111 (1989) 5680.
10. T. Kozaki, M. Morihashi, and O. Kikuchi, J. Am. Chem. Soc., 111 (1989) 1547.
11. S.C. Tucker and D.G. Truhlar, Chem. Phys. Lett., 157 (1989) 164.
12. J.J. Gajewski, J. Jurayj, D.R. Kimbrough, M.E. Gande, B. Ganem, and B.K. Carpenter, J. Am. Chem. Soc., 109 (1987) 1170.
13. E. Brandes, P.A. Grieco, and J.J. Gajewski, J. Org. Chem., 54 (1989) 515.
14. P.A. Grieco, Aldrichim. Acta, 24 (1991) 59.
15. D.L. Severance and W.L. Jorgensen, J. Am. Chem. Soc., 114 (1992) 10966.
16. J. Gao, submitted for publication.
17. C.J. Cramer and D.G. Truhlar, J. Am. Chem. Soc., 114 (1992) 8794.
18. S. Sogo, T.S. Widlanski, J.H. Hoare, C.E. Grimshaw, G.A. Berchtold, and J.R. Knowles, J. Am. Chem. Soc., 106 (1984) 2701.
19. H. Dugas, Bioorganic Chemistry, 2 ed., Springer-Verlag, New York, 1989.
20. C.L. Brooks, M. Karplus, and B.M. Pettit, Adv. Chem. Phys., 71 (1989) 1.
21. P.A. Kollman and K.M. Merz, Acc. Chem. Res., 23 (1990) 246.

22. A. Warshel, Computer Modeling of Chemical Reactions in Enzymes and Solutions, Wiley-Interscience, New York, 1991.
23. A. Warshel, J. Åqvist, and S. Creighton, Proc. Natl. Acad. Sci., USA, 86 (1989) 5820.
24. T.L. Hill, An Introduction to Statistical Thermodynamics, Addison-Wesley, Reading, MA, 1960, 119.
25. M.P. Allen and D.J. Tildesley, Computer Simulations of Liquids, Oxford University Press, London, 1987.
26. J.A. McCammon and S.C. Harvey, Dynamics of Proteins and Nucleic Acids, Cambridge University Press, Cambridge, 1987.
27. W.L. Jorgensen, Acc. Chem. Res., 22 (1989) 184.
28. D.W. Heermann, Computer Simulation Methods in Theoretical Physics, 2nd ed., Springer-Verlag, Berlin, 1990.
29. P.M. King, C.A. Reynolds, J.W. Essex, G.A. Worth, and W.G. Richards, Mol. Sim., 5 (1990) 262.
30. K. Binder and D.W. Heermann, Monte Carlo Simulation in Statistical Physics, 2nd corrected ed., Springer-Verlag, Berlin, 1992.
31. D.L. Beveridge and F.M. DiCapua, Annu. Rev. Biophys. Biophys. Chem., 18 (1989) 431.
32. J.M. Haile, Molecular Dynamics Simulation, Wiley-Interscience, New York, 1992.
33. F.S. Lee and A. Warshel, J. Chem. Phys., 97 (1992) 3100.
34. K. Tasaki, S. McDonald, and J.W. Brady, J. Comp. Chem., 14 (1993) 278.
35. J. Guenot and P.A. Kollman, J. Comp. Chem., 14 (1993) 295.
36. C.B. Post, C.M. Dobson, and M.K. Karplus, Proteins, 5 (1989) 337.
37. D.A. Pearlman and P.A. Kollman, J. Chem. Phys., 94 (1991) 4532.
38. M. Mazor and B.M. Pettit, Mol. Sim., 6 (1991) 1.
39. M.J. Mitchell and J.A. McCammon, J. Comp. Chem., 12 (1991) 271.
40. C.A. Reynolds, J.W. Essex, and W.G. Richards, Chem. Phys. Lett., 199 (1992) 257.
41. J.E. Straub and D. Thirumalai, Proteins, 15 (1993) 360.
42. J.J. Urban and G.R. Famini, J. Comp. Chem., 14 (1993) 353.
43. J.P. Bowen and N.L. Allinger, in: Reviews in Computational Chemistry, Vol. 2, K.B. Lipkowitz and D.B. Boyd (eds.), VCH, New York, 1991, 81.
44. D.M. Ferguson, I.R. Gould, W.A. Glauser, S. Schroeder, and P.A. Kollman, J. Comp. Chem., 13 (1992) 525.
45. J.E. Eksterowicz and K.N. Houk, Chem. Rev., 93 (1993) 2439.
46. A. Warshel, J. Phys. Chem., 83 (1979) 1640.
47. V. Luzhkov and A. Warshel, J. Comp. Chem., 13 (1992) 199.
48. Y.W. Xu, C.X. Wang, and Y.Y. Shi, J. Comp. Chem., 13 (1992) 1109.
49. J. Åqvist and A. Warshel, Chem. Rev., 93 (1993) 2418.
50. F.S. Lee, Z.T. Chu, and A. Warshel, J. Comp. Chem., 14 (1993) 161.
51. F. Peradejordi, Cahiers Phys., 17 (1963) 343.
52. O. Chalvet and I. Jano, Compt. Rend. Acad. Sci. Paris, 259 (1964) 1867.
53. I. Jano, Compt. Rend. Acad. Sci. Paris, 261 (1965) 103.
54. R. Daudel, Adv. Quantum Chem., 3 (1967) 121.
55. S. Yomosa and M. Hasegawa, J. Phys. Soc. Japan, 29 (1970) 1329.
56. M. Newton, J. Phys. Chem., 79 (1975) 2795.

57. O. Tapia and O. Goscinski, Mol. Phys., 29 (1975) 1653.
58. J. Hylton-McCreery, R.E. Christofferson, and G.G. Hall, J. Am. Chem. Soc., 98 (1976) 7191.
59. J.L. Burch, K.S. Raghuveer, and R.E. Christofferson, in: Environmental Effects on Molecular Structure and Properties, B. Pullman (ed.), Reidel, Dordrecht, 1976, 17.
60. J.-L. Rivail and D. Rinaldi, Chem. Phys., 18 (1976) 233.
61. I. Fischer-Hjalmars, I. Hendricksson-Entlo, and C. Hermann, Chem. Phys., 24 (1977) 167.
62. R. Constanciel and O. Tapia, Theor. Chim. Acta, 48 (1978) 75.
63. O. Tapia and B. Silvi, J. Phys. Chem., 84 (1980) 2646.
64. G. Klopman and P. Andreozzi, Theor. Chim. Acta, 55 (1980) 77.
65. S. Miertus, E. Scrocco, and J. Tomasi, Chem. Phys., 55 (1981) 117.
66. O. Tapia, in: Quantum Theory of Chemical Reactions, Vol. 2, R. Daudel, A. Pullman, L. Salem, and A. Viellard (eds.), Reidel, Dordrecht, 1980, 25.
67. P. Claverie, in: Quantum Theory of Chemical Reactions, Vol. 3, R. Daudel, A. Pullman, L. Salem, and A. Veillard (eds.), Reidel, Dordrecht, 1982, 151.
68. O. Tapia, in: Molecular Interactions, Vol. 3, H. Rajaczak and W.J. Orville-Thomas (eds.), John Wiley & Sons, London, 1982, 47.
69. R. Constanciel and R. Contreras, Theor. Chim. Acta, 65 (1984) 1.
70. R. Contreras and A. Aizman, Int. J. Quant. Chem., 27 (1985) 293.
71. J.L. Rivail, B. Terryn, and M.F. Ruiz-Lopez, J. Mol. Struct. (Theochem), 120 (1985) 387.
72. O. Tapia, F. Colonna, and J.G. Angyan, J. Chem. Phys., 87 (1990) 875.
73. O. Tapia, J. Mol. Struct. (Theochem), 226 (1991) 59.
74. J. Tomasi, R. Bonaccorsi, R. Cammi, and F.J. Olivares del Valle, J. Mol. Struct. (Theochem), 234 (1991) 401.
75. C.J. Cramer and D.G. Truhlar, in: Reviews in Computational Chemistry, Vol. 6, K.B. Lipkowitz and D.B. Boyd (eds.), VCH, New York, 1994, in press.
76. J.H. Jensen, P.N. Day, M.S. Gordon, H. Basch, D. Cohen, D.R. Garmer, M. Krauss, and W.J. Stevens, ACS Monograph Series on Hydrogen Bonding, in press.
77. M. Born, Z. Physik, 1 (1920) 45.
78. J.G. Kirkwood, J. Chem. Phys., 2 (1934) 351.
79. L. Onsager, J. Am. Chem. Soc., 58 (1936) 1486.
80. N.H. Frank and W. Tobocman, in: Fundamental Formulas of Physics, Vol. 1, D.H. Menzel (ed.), Dover, New York, 1960, 307.
81. R.K. Wangsness, Electromagnetic Fields, John Wiley & Sons, New York, 1979.
82. D.R. Corson and P. Lorrain, Introduction to Electromagnetic Fields and Waves, W.H. Freeman, San Francisco, 1962.
83. N. Hush, Aust. J. Sci. Res., Ser. A, 1 (1948) 480.
84. A.A. Rashin and B. Honig, J. Phys. Chem., 89 (1985) 5588.
85. A.A. Rashin and K. Namboodiri, J. Phys. Chem., 91 (1987) 6003.
86. F. Hirata, P. Redfern, and R.M. Levy, Int. J. Quant. Chem. Symp., 15 (1988) 179.
87. B. Jayaram, R. Fine, K. Sharp, and B. Honig, J. Phys. Chem., 93 (1989) 4320.
88. H.-S. Kim and J.-J. Chung, Bull. Korean Chem. Soc., 14 (1993) 220.
89. A. Bondi, J. Phys. Chem., 68 (1964) 441.
90. B. Roux, H.-A. Yu, and M. Karplus, J. Phys. Chem., 94 (1990) 4683.

91. D. Chandler and H.C. Andersen, J. Chem. Phys., 57 (1972) 1930.
92. F. Hirata, B.M. Pettit, and P.J. Rossky, J. Chem. Phys., 77 (1982) 509.
93. Y. Marcus, J. Solution Chem., 12 (1983) 271.
94. C.J.F. Bottcher, O.C. van Belle, P. Bordewicijk, and A. Rip, Theory of Electric Polarization, 2nd ed., Elsevier, Amsterdam, 1973, 94.
95. A. Warshel and S.T. Russell, Quart. Rev. Biophys., 17 (1984) 283.
96. H.L. Friedman, Mol. Phys., 29 (1975) 29.
97. H.-A. Yu and M. Karplus, J. Chem. Phys., 89 (1988) 2366.
98. J.L. Rivail, in: New Theoretical Concepts for Understanding Organic Reactions, J. Bertrán and I.G. Czismadia (eds.), Kluwer, Dordrecht, 1989, 219.
99. F.S. Lee, Z.-T. Chu, M.B. Bolger, and A. Warshel, Protein Eng., 5 (1992) 215.
100. G. King and R.A. Barford, J. Phys. Chem., 97 (1993) 8798.
101. F.M. Richards, Annu. Rev. Biophys. Bioeng., 6 (1977) 151.
102. M.L. Connolly, Science, 221 (1983) 709.
103. J.L. Pascual-Ahuir and E. Silla, J. Comp. Chem., 11 (1990) 1047.
104. I. Tuñón, E. Silla, and J.L. Pascual-Ahuir, Protein Eng., 5 (1992) 715.
105. J.L. Pascual-Ahuir, E. Silla, and I. Tuñon, J. Comp. Chem., submitted for publication.
106. T. Furuki, A. Umeda, M. Sakurai, Y. Inoue, and R. Chûjô, J. Comp. Chem., 15 (1994) 90.
107. M.W. Wong, K.B. Wiberg, and M.J. Frisch, J. Am. Chem. Soc., 114 (1992) 1645.
108. M.M. Karelson, A.R. Katrizky, and M.C. Zerner, Int. J. Quant. Chem. Symp., 20 (1986) 521.
109. K.V. Mikkelson, H. Agren, H.J.A. Jensen, and T. Helgaker, J. Phys. Chem., 89 (1988) 3086.
110. M.M. Karelson, A.R. Katritzky, M. Szafran, and M.C. Zerner, J. Org. Chem., 54 (1989) 6030.
111. M.M. Karelson, T. Tamm, A.R. Katritzky, S.J. Cato, and M.C. Zerner, Tetrahedron Comput. Methodol., 2 (1989) 295.
112. M.M. Karelson, A.R. Katritzky, M. Szafran, and M.C. Zerner, J. Chem. Soc., Perkin Trans. 2, (1990) 195.
113. A.R. Katritzky and M. Karelson, J. Am. Chem. Soc., 113 (1991) 1561.
114. H.S. Rzepa, M.Y. Yi, M.M. Karelson, and M.C. Zerner, J. Chem. Soc., Perkin Trans. 2, (1991) 635.
115. H.S. Rzepa and M.Y. Yi, J. Chem. Soc., Perkin Trans. 2, (1991) 531.
116. M.W. Wong, M.J. Frisch, and K.B. Wiberg, J. Am. Chem. Soc., 113 (1991) 4776.
117. M.W. Wong, K.B. Wiberg, and M.J. Frisch, J. Chem. Phys., 95 (1991) 8991.
118. O.G. Parchment, I.H. Hillier, and D.V.S. Green, J. Chem. Soc., Perkin Trans. 2, (1991) 799.
119. M.W. Wong, K.B. Wiberg, and M.J. Frisch, J. Am. Chem. Soc., 114 (1992) 523.
120. L.C.G. Freitas, R.L. Longo, and A.M. Simas, J. Chem. Soc., Faraday Trans., 88 (1992) 189.
121. M.W. Wong, R. Leung-Toung, and C. Wentrup, J. Am. Chem. Soc., 115 (1993) 2465.
122. M. Szafran, M.M. Karelson, A.R. Katritzky, J. Koput, and M.C. Zerner, J. Comp. Chem., 14 (1993) 371.

123. P. Young, D.V.S. Green, I.H. Hillier, and N.A. Burton, Mol. Phys., 80 (1993) 503.
124. J. Avery, The Quantum Theory of Atoms, Molecules, and Photons, McGraw-Hill, London, 1972.
125. W.J. Hehre, L. Radom, P.v.R. Schleyer, and J.A. Pople, Ab Initio Molecular Orbital Theory, Wiley, New York, 1986.
126. M.M. Karelson and M.C. Zerner, J. Phys. Chem., 96 (1992) 6949.
127. A. Dega-Szafran, M. Gdaniec, M. Grunwald-Wyspianska, Z. Kosturkiewicz, J. Koput, P. Krzyzanowski, and M. Szafran, J. Mol. Struct., 270 (1992) 99.
128. M.F. Guest and J. Kendrick, GAMESS-UK,Daresbury Laboratory, 1986.
129. R.L. Longo and L.C.G. Freitas, Int. J. Quant. Chem., Quant. Biol. Symp., 17 (1990) 35.
130. M.J. Frisch, G.W. Trucks, M. Head-Gordon, P.M.W. Gill, M.W. Wong, J.B. Foresman, B.G. Johnson, H.B. Schlegel, M.A. Robb, E.S. Replogle, R. Gomperts, J.L. Andres, K. Raghavachari, J.S. Binkley, C. Gonzalez, R.L. Martin, D.J. Fox, D.J. Defrees, J. Baker, J.J. Stewart, and J.A. Pople, Gaussian 92, Revision B, Gaussian, Inc., Pittsburgh, PA, 1992.
131. D. Rinaldi, P.E. Hogan, and A. Cartier, QCPE Bull., in press.
132. J.L. Rivail and B. Terryn, J. Chem. Phys., 79 (1982) 1.
133. D. Rinaldi, J.-L. Rivail, and N. Rguini, J. Comp. Chem., 13 (1992) 675.
134. V. Dillet, D. Rinaldi, J.G. Angyán, and J.-L. Rivail, Chem. Phys. Lett., 202 (1993) 18.
135. J.-L. Rivail, in: Theoretical and Computational Models for Organic Chemistry, S.J. Formosinho, I.G. Czismadia, and L.G. Arnaut (eds.), Kluwer, Dordrecht, 1991, 79.
136. D. Rinaldi, M.F. Ruiz-Lopez, and J.-L. Rivail, J. Chem. Phys., 78 (1983) 834.
137. G.P. Ford and B. Wang, J. Comp. Chem., 13 (1992) 229.
138. M.J. Frisch, in: Abstracts of Papers, 205th National Meeting of the American Chemical Society, American Chemical Society, Washington, DC, 1993.
139. D. Rinaldi and R.R. Pappalardo, program to be submitted to the Quantum Chemistry Program Exchange.
140. R.R. Pappalardo, E.S.Marcos, M.F. Ruiz-Lopez, D. Rinaldi, and J.-L. Rivail, J. Am. Chem. Soc., 115 (1993) 3722.
141. A.J. Stone, Chem. Phys. Lett., 83 (1981) 233.
142. A.D. Buckingham and P.W. Fowler, J. Chem. Phys., 79 (1983) 6426.
143. A.J. Stone and M. Alderton, Mol. Phys., 56 (1985) 1047.
144. F. Colonna, E. Evleth, and J.G. Angyán, J. Comp. Chem., 13 (1992) 1234.
145. R.S. Mulliken, J. Chem. Phys., 3 (1935) 564.
146. C.A. Coulson and H.C. Longuet-Higgins, Proc. R. Soc. London, 191 (1947) 39.
147. R.S. Mulliken, J. Chem. Phys., 23 (1955) 1833.
148. F.L. Hirshfeld, Theor. Chim. Acta, 44 (1977) 129.
149. I. Mayer, Chem. Phys. Lett., 97 (1983) 270.
150. I. Mayer, Chem. Phys. Lett., 110 (1984) 440.
151. R.F. Bader, Acc. Chem. Res., 18 (1985) 9.
152. A.E. Reed, R.B. Weinstock, and F. Weinhold, J. Chem. Phys., 83 (1985) 735.
153. L.E. Chirlian and M.M. Francl, J. Comp. Chem., 8 (1987) 894.
154. A.E. Reed, L.A. Curtiss, and F. Weinhold, Chem. Rev., 88 (1988) 899.
155. J. Cioslowski, J. Am. Chem. Soc., 111 (1989) 8333.

156. B.H. Besler, J. K.M. Merz, and P.A. Kollman, J. Comp. Chem., 11 (1990) 431.
157. C.M. Breneman and K.B. Wiberg, J. Comp. Chem., 11 (1990) 361.
158. E.R. Davidson and S. Chakravorty, Theor. Chim. Acta, 83 (1992) 319.
159. K.M. Merz, J. Comp. Chem., 13 (1992) 749.
160. M.N. Ramos and B.d.B. Neto, Chem. Phys. Lett., 199 (1992) 482.
161. M.A. Spackman, Chem. Rev., 92 (1992) 1769.
162. T.K. Ghanty and S.K. Ghosh, J. Mol. Struct. (Theochem), 276 (1992) 83.
163. T.R. Stouch and D.E. Williams, J. Comp. Chem., 14 (1993) 858.
164. S.M. Bachrach, in: Reviews in Computational Chemistry, Vol. 5, K.B. Lipkowitz and D.B. Boyd (eds.), VCH, New York, 1993, 171.
165. G. Rauhut and T. Clark, J. Comp. Chem., 14 (1993) 503.
166. J.W. Storer, D.J. Giesen, C.J. Cramer, and D.G. Truhlar, J. Comp. Chem., submitted for publication.
167. C. Chipot, D. Rinaldi, and J.-L. Rivail, Chem. Phys. Lett., 191 (1992) 287.
168. J.L. Pascual-Ahuir, E. Silla, J. Tomasi, and R. Bonaccorsi, J. Comp. Chem., 8 (1987) 778.
169. R. Bonaccorsi, P. Palla, and J. Tomasi, J. Comp. Chem., 4 (1983) 567.
170. R. Bianco, S. Miertius, M. Persico, and J. Tomasi, Chem. Phys., 168 (1992) 281.
171. R. Bonaccorsi, R. Cammi, and J. Tomasi, J. Comp. Chem., 12 (1991) 301.
172. S. Woodcock, D.V.S. Green, M.A. Vincent, I.H. Hillier, M.F. Guest, and P. Sherwood, J. Chem. Soc., Perkin Trans. 2, (1992) 2151.
173. M. Peterson and R. Poirer, MONSTERGAUSS, (Department of Chemistry, University of Toronto, Toronto, Ontario, Canada,
174. M. Negre, M. Orozco, and F.J. Luque, Chem. Phys. Lett., 196 (1992) 27.
175. B. Wang and G.P. Ford, J. Chem. Phys., 97 (1992) 4162.
176. B. Wang and G.P. Ford, J. Am. Chem. Soc., 114 (1992) 10563.
177. G. Rauhut, T. Clark, and T. Steinke, J. Am. Chem. Soc., 115 (1993) 9174.
178. A. Klamt and G. Schüürmann, J. Chem. Soc., Perkin Trans. 2, (1993) 799.
179. T. Fox, N. Rösch, and R.J. Zaubar, J. Comp. Chem., 14 (1993) 253.
180. M.A. Aguilar and F.J. Olivares del Valle, Chem. Phys., 129 (1989) 439.
181. M. Orozco, W.L. Jorgensen, and F.J. Luque, J. Comp. Chem., 14 (1993) 1498.
182. H. Hoshi, M. Sakurai, Y. Inouye, and R. Chûjô, J. Chem. Phys., 87 (1987) 1107.
183. J.A. Pople and D.A. Beveridge, Approximate Molecular Orbital Theory, McGraw-Hill, New York, 1970.
184. R. Fowler and E.A. Guggenheim, Statistical Thermodynamics, Cambridge University Press, London, 1956.
185. M.K. Gilson and B. Honig, Proteins, 4 (1988) 7.
186. H. Nakamura, J. Phys. Soc. Japan, 57 (1988) 3702.
187. S.C. Harvey, Proteins, 5 (1989) 78.
188. M.E. Davis and J.A. McCammon, J. Comp. Chem., 10 (1989) 386.
189. M.E. Davis and J.A. McCammon, Chem. Rev., 90 (1990) 509.
190. D.A. Sharp and B. Honig, Annu. Rev. Biophys. Biophys. Chem., 19 (1990) 301.
191. B.J. Yoon and A.M. Lenhoff, J. Comp. Chem., 11 (1990) 1080.
192. A. Nicholls and B. Honig, J. Comp. Chem., 12 (1991) 435.
193. A. Jean-Charles, A. Nicholls, K. Sharp, B. Honig, A. Tempczyk, T.F. Hendrickson, and W.C. Still, J. Am. Chem. Soc., 113 (1991) 1454.

194. A.H. Juffer, E.F.F. Botta, B.A.M. van Keulen, A. van der Ploeg, and H.J.C. Berendsen, J. Comp. Phys., 97 (1991) 144.
195. B.A. Luty, M.E. Davis, and J.A. McCammon, J. Comp. Chem., 13 (1992) 768.
196. B.A. Luty, M.E. Davis, and J.A. McCammon, J. Comp. Chem., 13 (1992) 1114.
197. B. Honig, K. Sharp, and A.-S. Yang, J. Phys. Chem., 97 (1993) 1101.
198. A.A. Rashin, Prog. Biophys. Molec. Biol., 60 (1993) 73.
199. A.J. Hopfinger, Conformational Properties of Macromolecules, Academic Press, New York, 1973.
200. E.M. Huque, J. Chem. Ed., 66 (1989) 581.
201. G. Neméthy, W.J. Peer, and H.A. Scheraga, Annu. Rev. Biophys. Bioeng., 10 (1981) 459.
202. B. Lee and F.M. Richards, J. Mol. Biol., 55 (1971) 379.
203. R.B. Hermann, J. Phys. Chem., 76 (1972) 2754.
204. G.L. Amidon, S.H. Yalkowsky, S.T. Anik, and S.C. Valvani, J. Phys. Chem., 79 (1975) 2239.
205. S.C. Valvani, S.H. Yalkowsky, and G.L. Amidon, J. Phys. Chem., 80 (1976) 829.
206. G.D. Rose, A.R. Geselowitz, G.J. Lesser, R.H. Lee, and M.H. Zehfus, Science, 229 (1985) 834.
207. T. Ooi, M. Oobatake, G. Nemethy, and H.A. Scheraga, Proc. Natl. Acad. Sci., USA, 84 (1987) 3086.
208. D. Eisenberg and A.D. McLachlan, Nature, 319 (1986) 199.
209. B.v. Freyberg and W. Braun, J. Comp. Chem., 14 (1993) 510.
210. R.A. Pierotti, J. Phys. Chem., 67 (1963) 1840.
211. M.J. Huron and P. Claverie, J. Phys. Chem., 76 (1972) 2123.
212. P. Claverie, J.P. Daudey, J. Langlet, B. Pullman, D. Piazzola, and M.J. Huron, J. Phys. Chem., 82 (1978) 405.
213. B.T. Thole and P.T. van Duijnen, Theor. Chim. Acta, 55 (1980) 307.
214. D. Rinaldi, B.J. Costa Cabral, and J.-L. Rivail, Chem. Phys. Lett., 125 (1986) 495.
215. F. Floris and J. Tomasi, J. Comp. Chem., 10 (1989) 616.
216. J.G. Ángyán and G. Jensen, Chem. Phys. Lett., 175 (1990) 313.
217. J.A.C. Rullman and P.T. van Duijnen, Reports Molec. Theory, 1 (1990) 1.
218. F.M. Floris, J. Tomasi, and J.L. Pascual-Ahuir, J. Comp. Chem., 39 (1991) 784.
219. F.J. Olivares del Valle and M.A. Aguilar, J. Mol. Struct. (Theochem), 280 (1993) 25.
220. F.M. Floris, A. Tani, and J. Tomasi, Chem. Phys., 169 (1993) 11.
221. W.C. Still, A. Tempczyk, R.C. Hawley, and T. Hendrickson, J. Am. Chem. Soc., 112 (1990) 6127.
222. G.J. Hoijtink, E. de Boer, P.H. Van der Meij, and W.P. Weijland, Recl. Trav. Chim. Pays-Bas, 75 (1956) 487.
223. C.J. Cramer and D.G. Truhlar, J. Am. Chem. Soc., 113 (1991) 8305.
224. C.J. Cramer and D.G. Truhlar, Science, 256 (1992) 213.
225. C.J. Cramer and D.G. Truhlar, J. Comp. Chem., 13 (1992) 1089.
226. M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, and J.J.P. Stewart, J. Am. Chem. Soc., 107 (1985) 3902.
227. J.J.P. Stewart, J. Comp. Chem., 10 (1989) 209.
228. J.J.P. Stewart, J. Comp. Chem., 10 (1989) 221.
229. J.W. Storer, D.J. Giesen, C.J. Cramer, and D.G. Truhlar, to be published.

230. C.J. Cramer and D.G. Truhlar, J. Comput.-Aid. Mol. Des., 6 (1992) 629.
231. D.A. Liotard, G.D. Hawkins, G.C. Lynch, C.J. Cramer, and D.G. Truhlar, to be published.
232. C.J. Cramer, G.D. Hawkins, and D.G. Truhlar, J. Chem. Soc., Faraday Trans., in press.
233. J.J.P. Stewart, in: Reviews in Computational Chemistry, Vol. 1, K.B. Lipkowitz and D.B. Boyd (eds.), VCH, New York, 1989, 45.
234. W. Thiel, Tetrahedron, 44 (1988) 7393.
235. M.C. Zerner, in: Reviews in Computational Chemistry, Vol. 2, K.B. Lipkowitz and D.B. Boyd (eds.), VCH, New York, 1990, 313.
236. D.R. Armstrong, P.G. Perkins, and J.J.P. Stewart, J. Chem. Soc., Dalton Trans., (1973) 838.
237. C.J. Cramer, G.C. Lynch, G.D. Hawkins, and D.G. Truhlar, QCPE Bull., 13 (1993) 78.
238. R.A. Marcus, J. Chem. Phys., 24 (1956) 979.
239. M.M. Kreevoy and D.G. Truhlar, in: Investigation of Rates and Mechanisms of Reactions, Part I, 4th ed., C.F. Bernasconi (ed.), Wiley, New York, 1986, 13.
240. B.C. Garrett and G.K. Schenter, Int. Rev. Phys. Chem., in press.
241. S. Lee and J.T. Hynes, J. Chem. Phys., 88 (1988) 6863.
242. J. Gehlen, D. Chandler, H.J. Kim, and J.T. Hynes, J. Phys. Chem., 96 (1992) 1748.
243. H.J. Kim and J.T. Hynes, J. Chem. Phys., 96 (1992) 5088.
244. D.G. Truhlar, G.K. Schenter, and B.C. Garrett, J. Chem. Phys., 98 (1993) 5756.
245. M.V. Basilevsky, G.E. Chudinov, and M.D. Newton, preprint.
246. J.S. Kwiatkowski and A. Tempczyk, Chem. Phys., 85 (1984) 397.
247. A. Razynska, A. Tempczyk, and Z. Grzonka, J. Chem. Soc., Faraday Trans. 2, 81 (1985) 1555.
248. A. Tempczyk, Z. Gryczynski, A. Kawski, and Z. Grzonka, Z. Naturforsch., 43a (1988) 363.
249. M. Karelson and M.C. Zerner, J. Am. Chem. Soc., 112 (1990) 9405.
250. T. Fox and N. Rösch, Chem. Phys. Lett., 191 (1992) 33.
251. T. Fox and N. Rosch, J. Mol. Struct. (Theochem), 276 (1992) 279.
252. M.A. Aguilar, F.J. Olivares del Valle, and J. Tomasi, J. Chem. Phys., 98 (1993) 7375.
253. G. Nemethy and H.A. Scheraga, J. Chem. Phys., 36 (1962) 3401.
254. T.J. Richmond, J. Mol. Biol., 178 (1984) 63.
255. J.T. Kellis, K. Nyberg, D. Sali, and A.R. Fersht, Nature, 788 (1988)
256. K.A. Dill, Biochemistry, 29 (1990) 7133.
257. K. Sharp, A. Nicholls, R. Friedman, and B. Honig, Biochemistry, 30 (1991) 9686.
258. J.W. Grate, R.A. McGill, and D. Hilvert, J. Am. Chem. Soc., 115 (1993) 8577.
259. C.J. Cramer and D.G. Truhlar, Chem. Phys. Lett., 198 (1992) 74 and 202 (1993) 567 (Erratum).
260. R. Rein, in: Intermolecular Interactions: From Diatomics to Biopolymers, B. Pullman (ed.), John Wiley & Sons, Chichester, 1978, 307.
261. R. Bonaccorsi, E. Scrocco, and J. Tomasi, Proc. Intl. Symp. Biomol. Struct. Interact., Suppl. J. Biosci., 8 (1985) 627.
262. R. Bonaccorsi, E. Scrocco, and J. Tomasi, Int. J. Quant. Chem., 29 (1986) 717.

263. M. Mezei, Mol. Sim., 10 (1993) 225.
264. A. Laaksonen, L. Nilson, B. Jönsson, and O. Teleman, Chem. Phys., 4 (1989) 81.
265. L. Nilsson and M. Karplus, J. Comp. Chem., 7 (1986) 59.
266. A. Pohorille, W.S. Ross, and J. I. Tinoco, Intl. J. Supercomput. Appl., 4 (1990) 81.
267. J. Pranta, S.G. Wierschke, and W.L. Jorgensen, J. Am. Chem. Soc., 113 (1991) 2810.
268. G.L. Seibel, U.C. Singh, and P.A. Kollmann, Proc. Natl. Acad. Sci., USA, 82 (1985) 6537.
269. S.J. Weiner, P.A. Kollman, D.T. Nguyen, and D.A. Case, J. Comp. Chem., 7 (1986) 230.
270. C.J. Cramer and D.G. Truhlar, J. Am. Chem. Soc., 113 (1991) 8552.
271. C.J. Cramer and D.G. Truhlar, J. Am. Chem. Soc., 115 (1993) 8810.
272. W.L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc., 110 (1988) 1657.
273. W.J. Hehre, R. Ditchfield, and J.A. Pople, J. Chem. Phys., 56 (1972) 2257.
274. J. Gao and X. Xia, Biophys. Chem., submitted for publication.
275. P. Bash, U.C. Singh, R. Langridge, and P.A. Kollman, Science, 236 (1987) 564.
276. D.M. Ferguson, D.A. Pearlman, W.C. Swope, and P.A. Kollman, J. Comp. Chem., 13 (1992) 362.
277. V. Mohan, M.E. Davis, J.A. McCammon, and B.M. Pettitt, J. Phys. Chem., 96 (1992) 6428.
278. A.H. Elcock and W.G. Richards, J. Am. Chem. Soc., 115 (1993) 7930.
279. P.E. Young and I.H. Hillier, Chem. Phys. Lett., 215 (1993) 405.
280. P. Seeman, Pharmacol. Rev., 32 (1987) 229.
281. J.J. Urban, C.J. Cramer, and G.R. Famini, J. Am. Chem. Soc., 114 (1992) 8226.
282. D.A. Daugherty and D.A. Stauffer, Science, 250 (1990) 1558.
283. J. Gao, L.W. Chou, and A. Auerbach, Biophys. J., 65 (1993) 43.
284. M. Meot-Ner and C.A. Deakyne, J. Am. Chem. Soc., 107 (1985) 469.
285. M.F. Perutz, G. Fermi, D.J. Abraham, C. Poyart, and E. Bursaux, J. Am. Chem. Soc., 108 (1986) 1064.
286. P. Solmajer, D. Kocjan, and T. Solmajer, Z. Naturforsch., 38c (1983) 758.
287. O. Bastiansen, Acta Chem. Scand., 3 (1949) 415.
288. P. Buckley and P.A. Giguère, Can. J. Chem., 45 (1967) 397.
289. W. Caminati and G. Corbelli, J. Mol. Spectrosc., 90 (1981) 572.
290. H. Frei, T.K. Ha, R. Meyer, and H.H. Günthard, Chem. Phys., 25 (1977) 271.
291. T.K. Ha, H. Frei, R. Meyer, and H.H. Günthard, Theor. Chim. Acta, 34 (1974) 277.
292. L.P. Kuhn, J. Am. Chem. Soc., 74 (1952) 2492.
293. C.G. Park and M. Tasumi, J. Phys. Chem., 95 (1991) 2757.
294. H. Takeuchi and M. Tasumi, Chem. Phys., 77 (1983) 21.
295. E. Walder, A. Bander, and H.H. Günthard, Chem. Phys., 51 (1980) 223.
296. D.E. Woon and T.H. Dunning, J. Chem. Phys., 98 (1993) 1358.
297. K. Raghavachari, G.W. Trucks, J.A. Pople, and M. Head-Gordon, Chem. Phys. Lett., 157 (1989) 479.
298. C.J. Cramer and D.G. Truhlar, J. Am. Chem. Soc., submitted for publication.
299. K.G.R. Pachler and P.L. Wessels, J. Mol. Struct., 6 (1970) 471.
300. A. Ben-Naim, Solvation Thermodynamics, Plenum, New York, 1987.
301. D. Suleiman and C.A. Eckert, J. Chem. Eng. Data, submitted for publication.

302. B.P. van Eijck, R.W.W. Hooft, and J. Kroon, J. Phys. Chem., 97 (1993) 12093.
303. B.P. van Eijck, L.M.J. Kroon-Batenburg, and J. Kroon, J. Mol. Struct., 237 (1990) 315.
304. L.M.J. Kroon-Batenburg and J. Kroon, Biopolymers, 29 (1990) 1243.
305. S. Ha, J. Gao, B. Tidor, J.W. Brady, and M. Karplus, J. Am. Chem. Soc., 113 (1991) 1553.
306. P.L. Polavarapu and C.S. Ewig, J. Comp. Chem., 13 (1992) 1255.
307. C.J. Cramer and D.G. Truhlar, J. Am. Chem. Soc., 115 (1993) 5745.
308. Y. Nishida, H. Ohrui, and H. Meguro, Tetrahedron Lett., 25 (1993) 1575.
309. S.J. Angyal, Angew. Chem., Int. Ed. Engl., 8 (1969) 157.
310. D.J. Giesen, C.J. Cramer, and D.G. Truhlar, to be published.
311. R.B. Hermann, Proc. Natl. Acad. Sci., USA, 74 (1977) 4144.
312. I. Tuñón, E. Silla, and J.L. Pascual-Ahuir, Chem. Phys. Lett., 203 (1993) 289.
313. A. Leo, in: Environmental Health Chemistry, J.D. McKinney (ed.), Ann Arbor Science, Ann Arbor, MI, 1981, 323.
314. C. Hansch, J.P. Bjorkroth, and A. Leo, J. Pharm. Sci., 76 (1987) 663.
315. A. Leo, C. Hansch, and D. Elkins, Chem. Rev., 71 (1971) 525.
316. R. Balducci, A. Roda, and R.S. Pearlman, J. Solution Chem., 18 (1989) 355.
317. S.C. DeVito and R.S. Pearlman, Med. Chem. Res., 1 (1991) 461.
318. A.K. Debnath, R.L.d. Compadre, A.J. Shusterman, and C. Hansch, Envrion. Mol. Mutagen., 19 (1992) 53.
319. J. Grogan, S.C. DeVito, R.S. Pearlman, and K.R. Korsekwa, Chem. Res. Toxicol., 5 (1992) 548.
320. A.J. Leo, Chem. Rev., 93 (1993) 1281.
321. D. Hadzi and B. Jerman-Blazic (eds.), QSAR in Drug Design and Toxicology, Elsevier, Amsterdam, 1987 269.
322. C. Silipo and A. Vittoria (eds.), New Perspectives in QSAR, Elsevier, Amsterdam, 1991.
323. J.-L.M. Abboud, M.J. Kamlet, and R.W. Taft, J. Am. Chem. Soc., 99 (1977) 8325.
324. M.J. Kamlet and R.W. Taft, Prog. Org. Chem., 48 (1983) 2877.
325. M.J. Kamlet, M.H. Abraham, R.M. Doherty, and R.W. Taft, J. Am. Chem. Soc., 106 (1984) 464.
326. M.J. Kamlet, Prog. Phys. Org. Chem., 19 (1993) 295.
327. M.J. Kamlet and R.W. Taft, Acta Chem. Scand. B, 40 (1986) 619.
328. O. Exner, Prog. Phys. Org. Chem., 18 (1990) 129.
329. L.P. Hammett, Chem. Rev., 17 (1935) 125.
330. H.H. Jaffé, Chem. Rev., 53 (1953) 191.
331. F.W. Fowler, A.R. Katritzky, and R.J.D. Rutherford, J. Chem. Soc. B, (1971) 460.
332. I.A. Koppel and V.A. Palm, in: Advances in Linear Free Energy Relationships, N.B. Chapman and J. Shorter (eds.), Plenum Press, London, 1972, 203.
333. T.M. Krygowski and W.R. Fawcett, J. Am. Chem. Soc., 97 (1975) 2143.
334. R.C. Dougherty, Tetrahedron Lett., (1975) 385.
335. M.J. Kamlet, J.L. Abboud, and R.W. Taft, J. Am. Chem. Soc., 99 (1977) 6027.
336. M.J. Kamlet, J.L. Abboud, and R.W. Taft, J. Am. Chem. Soc., 99 (1977) 8325.
337. U. Mayer, Monatsh. Chem., 109 (1978) 421, 775.

338. C.G. Swain, M.S. Swain, A.L. Powell, and S. Alunni, J. Am. Chem. Soc., 105 (1983) 502.
339. R.W. Taft, J.-L.M. Abboud, M.J. Kamlet, and M.H. Abraham, J. Solution Chem., 14 (1985) 153.
340. R.S. Drago and A.P. Dadmun, J. Am. Chem. Soc., 115 (1993) 8592.
341. G.R. Famini, R.J. Kassel, J.W. King, and L.Y. Wilson, Quant. Struct.-Act. Relat., 10 (1991) 344.
342. L.Y. Wilson and G.R. Famini, J. Med. Chem., 34 (1991) 1668.
343. G.R. Famini, C.A. Penski, and L.Y. Wilson, J. Phys. Org. Chem., 5 (1992) 395.
344. G.R. Famini, W.P. Ashman, A.P. Mickiewicz, and L.Y. Wilson, Quant. Struct.-Act. Relat., 11 (1992) 162.
345. G.R. Famini and S.C. DeVito, in: Biomarkers of Human Exposure, J. Blancato and M. Saleh (eds.), American Chemical Society, Washington, DC, in press.
346. G.R. Famini, B.C. Marquez, and L.Y. Wilson, J. Chem. Soc., Perkin Trans. 2, (1993) 773.
347. D.F.V. Lewis, in: Reviews in Computational Chemistry, Vol. 3, K.B. Lipkowitz and D.B. Boyd (eds.), VCH, New York, 1992, 173.
348. J.S. Murray, P. Lane, T. Brinck, and P. Politzer, J. Phys. Chem., 97 (1993) 5144.
349. P. Politzer, J.S. Murray, P. Lane, and T. Brinck, J. Phys. Chem., 97 (1993) 729.
350. P. Politzer, J.S. Murray, M.C. Concha, and T. Brinck, J. Mol. Struct. (Theochem), 100 (1993) 107.
351. L.H. Hall and L.B. Kier, in: Reviews in Computational Chemistry, Vol. 2, K.B. Lipkowitz and D.B. Boyd (eds.), VCH, New York, 1991, 367.
352. I.B. Bersuker and A.S. Dimoglo, in: Reviews in Computational Chemistry, Vol. 2, K.B. Lipkowitz and D.B. Boyd (eds.), VCH, New York, 1991, 423.
353. C.J. Cramer, G.R. Famini, and A.H. Lowrey, Acc. Chem. Res., 26 (1993) 599.
354. I. Alkorta, H.O. Villar, and J.J. Perez, J. Comp. Chem., 14 (1993) 620.