Annu. Rev. Phys. Chem. 1998. 49:531–67

COMPUTER SIMULATIONS WITH EXPLICIT SOLVENT: Recent Progress in the Thermodynamic Decomposition of Free Energies and in Modeling Electrostatic Effects

Ronald M. Levy and Emilio Gallicchio
Department of Chemistry, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08855-0939; e-mail: ronlevy@lutece.rutgers.edu; emilio@lutece.rutgers.edu

KEY WORDS: entropy-enthalpy compensation, hydrophobic effect, linear response, protein electrostatics, thermodynamic perturbation

ABSTRACT

This review focuses on recent progress in two areas in which computer simulations with explicit solvent are being applied: the thermodynamic decomposition of free energies, and modeling electrostatic effects. The computationally intensive nature of these simulations has been an obstacle to the systematic study of many problems in solvation thermodynamics, such as the decomposition of solvation and ligand binding free energies into component enthalpies and entropies. With the revolution in computer power continuing, these problems are ripe for study but require the judicious choice of algorithms and approximations. We provide a critical evaluation of several numerical approaches to the thermodynamic decomposition of free energies and summarize applications in the current literature. Progress in computer simulations with explicit solvent of charge perturbations in biomolecules was slow in the early 1990s because of the widespread use of truncated Coulomb potentials in these simulations, among other factors. Development of the sophisticated technology described in this review to handle the long-range electrostatic interactions has increased the predictive power of these simulations to the point where comparisons between explicit and continuum

0066-426X/98/1001-0531$08.00

solvent models can reveal differences that have their true physical origin in the inherent molecularity of the surrounding medium.

INTRODUCTION

Computer simulations have provided the basis for much of our molecular understanding of solvation thermodynamics. At the most detailed level, the simulations include explicit molecular representations of the solvent. This approach is, in principle, the most realistic and accurate model for studying the physical chemistry of solvation, whether the solutes are simple organic molecules or complex biomolecules. Indeed, computer simulations of a large variety of solutes in solution have been performed in an effort to correlate molecular properties with macroscopic properties. With intermolecular potential functions and the machinery of statistical thermodynamics, all the excess thermodynamic quantities of interest can be calculated from Monte Carlo (MC) or molecular dynamic (MD) trajectories. In the middle 1980s, the development and application of free-energy simulations based on thermodynamic perturbation theory stimulated interest in the field; the primary applications have been in organic and biophysical chemistry. This methodology came to be called computational alchemy because the physical transformations corresponding to portions of a thermodynamic cycle were calculated via alchemical transformations represented by the remaining portions of the cycle. There have been several recent reviews (1–4). Ultimately, the predictive power of free-energy perturbation simulations depends on the accuracy of the intermolecular potential models and an ability to generate a statistical sample of the relevant configurations for the solute-solvent system. Many laboratories are engaged in efforts to improve the methodology. This review focuses on recent developments in two areas that are likely to have a significant impact on the future usefulness of explicit solvent simulations for studying thermodynamic transformations: (*a*) the decomposition of solvation free energies into component enthalpies and entropies, and (*b*) the accurate simulation of electrostatic effects in solution using explicit solvent models.

Computer simulations of explicit solvent models are computationally intensive. For many problems, the full atomic model has to be abandoned and empirical or mean field models that require less computational power must be adopted instead. It is often possible to parameterize empirical models to reproduce experimental results, but this is often accomplished at the expense of a clear molecular interpretation of the fitting parameters. The revolution in computer power, however, continues at a fast rate and with it the complexity of the problems we are able to study by explicit solvent models. In this review,

we focus on two problems—entropy-enthalpy decomposition of free energies and electrostatic effects—that require the use of explicit solvent models and the intelligent use of the available computer technology.

Although computer simulations have been applied to the calculation of solvation free energies for a wide range of molecules, the entropies and enthalpies of solvation have not received the same attention, partly because of the more stringent computer requirements. Accurate reproduction of entropies and enthalpies of solvation is essential if computer simulations are to become more useful tools for obtaining molecular insights into solvation and ligand binding phenomena. Furthermore, the enthalpies and entropies provide additional benchmarks to be used in developing force fields for condensed-phase simulations. In this review, we provide a critical evaluation of several different numerical approaches to the decomposition of free energies into component enthalpies and entropies. Enthalpy-entropy decomposition of free energies by explicit solvent simulations has been applied sporadically and only to small solutes. We summarize the current literature. The application of these methods to enthalpy-entropy decomposition for processes involving biological macromolecules represents a challenging computational problem, but such calculations could prove to be a source for developing deeper insights into the connection between biological structure and thermodynamics. We present a perspective on the role that computer simulations of component enthalpies and entropies may play in studies of protein folding and ligand binding.

Computer simulations of the response of complex systems to charge perturbations can be carried out using either detailed atomic models or reduced continuum solvent models. Continuum solvent models, which have attracted a considerable following in recent years, are computationally fast and arguably no less accurate than explicit solvent models when suitably parameterized. However, on a small enough length scale, the molecularity of the condensed phase will be manifest and yet it is precisely on this microscopic scale that continuum models are being applied. Computer simulations of electrostatic effects in solution with explicit solvent models have an important role to play, both in development of a more fundamental understanding of the physics of charge transfer processes in solution and in developing improvements in reduced models. In the second section of this review we consider recent progress in the modeling of electrostatic effects in solution from an explicit solvent viewpoint. We analyze the linear response approximation that is crucial for interpolating between all atom and continuum solvent models. Recent advances in the methods for modeling the long-range electrostatic interactions are discussed, as are applications to charging free energies of ions, pK_a shifts of organic acids, and electrostatic properties of proteins.

ENTHALPY-ENTROPY DECOMPOSITION

Entropy and enthalpy are an important complement to free energy because they provide additional information to help understand and interpret the connection between molecular perturbations and thermodynamic changes in condensed phases. For instance, two solutes may have similar hydration free energies even though their solubility may be determined by completely distinct phenomena at the molecular level. The solvation of one solute could be mostly driven by a favorable entropy, that is by an increase in the configurational space spanned by the solvent and the solute with little energy change. For the other solute, instead, the dominant effect could be a large and favorable interaction energy with the solvent that, however, causes a loss of entropy due to the system locking in a particular low-energy configuration. The hydrophobic effect, assumed to be entropically dominated, is a concrete example of how the entropy can be used to classify solvation processes.

Entropy and enthalpy are thermodynamic state functions and are independent of the thermodynamic path connecting two thermodynamic states of the system. The decomposition of the free energy into enthalpic and entropic components is, therefore, always thermodynamically meaningful. This is in contrast to decomposition schemes of the free energy in terms of components that depend on the thermodynamic path connecting the two states (for further discussion, see 5 and references therein) and other decomposition schemes that assume thermodynamic independence between different parts of the system (6). More importantly, enthalpies and entropies are experimentally measurable quantities. The comparison between calculated and experimental enthalpies and entropies may provide additional physical insights and may be used as benchmarks to optimize force fields for condensed-phase simulations.

Our ability to characterize solvation and binding phenomena by making general predictions based only on free energies is severely limited. The main obstacle is represented by the fact that the free energy is usually the result of a large cancelation between entropy and enthalpy and therefore presents small variations from one system to another that are difficult to rationalize. Entropy and enthalpy changes are more readily interpreted than are free-energy changes. The enthalpy change measures a change in the strength of the interactions between molecules while the entropy change measures a variation in the order of the system. Invariably, the analysis of free-energy changes involves the analysis of the relative importance of the corresponding enthalpy and entropy changes.

This section is organized as follows. First we summarize several available methodologies to compute relative enthalpies and entropies changes. The strengths and the weaknesses of each method are outlined. Issues related to ensemble dependence, solvation shell approximations, and comparison to

experimental quantities are discussed. We then review explicit solvent studies of enthalpies and entropies of hydration of noble gases and small organic molecules. Next, the phenomenon of entropy-enthalpy compensation and its theoretical interpretations are discussed. Finally, we present a perspective on future applications of entropy-enthalpy decomposition to the study of biological processes.

Methods

In this section, the accuracy and convergence properties of different methods are rated. The study of the convergence behavior of the different available numerical methods to perform the entropy-enthalpy decomposition of the free energy is of particular importance because the slow convergence is the main obstacle encountered in these calculations. The following discussion focuses on the calculation of differences of free energy and on entropy and enthalpy of solvation between two solutes. The methods are easily generalized to include a wide variety of processes in solution, such as ligand binding. Unless otherwise noted, we assume that the two solutes lack internal degrees of freedom and that the total potential energy is the sum of molecular pair interactions.

FREE ENERGY Two main methodologies are available to calculate the free-energy ΔF_{YX} of mutating solute X into solute Y in solution: the free-energy perturbation (FEP) method and the thermodynamic integration (TI) method (1, 7–10). A charging parameter λ is introduced in the potential function so that in going from $\lambda = 0$ to $\lambda = 1$, the system is removed from the initial state X and transformed into the final state Y. This is accomplished in several steps, or windows. In the i-th window, the system configurations are sampled by using an MD or MC algorithm having set $\lambda = \lambda_i$. In the FEP method, the Gibbs free-energy change in varying λ_i to λ_{i+1} is expressed as

$$\Delta F_{i+1,i} = -kT \ln\langle \exp\{-[U(\lambda_{i+1}) - U(\lambda_i)]/kT\}\rangle_{\lambda_i}, \qquad 1.$$

where $\langle \ldots \rangle_{\lambda_i}$ denotes an ensemble average for $\lambda = \lambda_i$ and $U(\lambda)$ is the λ-dependent total potential energy. The TI expression for the total free-energy change is

$$\Delta F = \int_0^1 d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda . \qquad 2.$$

Only the solute-solvent potential energy term couples with the charging parameter λ. The free-energy change, thus, is explicitly dependent only on the solute-solvent potential energy. The convergence rate of the free energy, therefore, is independent of the number N of solvent molecules. This is in contrast

to the system size-dependent fluctuations of some of the estimators of the entropy and energy changes discussed below and accounts for the relatively fast convergence of FEP or TI free-energy computations.

ENTHALPY AND ENTROPY Given that $\Delta G = \Delta H - T\Delta S$, and assuming that the Gibbs free-energy ΔG is available, to perform the entropy-enthalpy decomposition of the free energy, it is sufficient to compute either one of ΔH or ΔS. In the isobaric ensemble, the enthalpy change is the sum of the change in average potential energy $\Delta\bar{U}$ and the the pressure-volume term $P\Delta\bar{V}$, where $\Delta\bar{V}$ is the partial molar volume of the solute. Below, we outline some numerical methods that have been applied to calculate the average potential energy change and the entropy change.

The *direct method* is based on a simple end-point expression for the average potential energy change (11–13)

$$\Delta\bar{U} = \langle U(\lambda = 1)\rangle_{\lambda=1} - \langle U(\lambda = 0)\rangle_{\lambda=0}. \qquad 3.$$

Although easy to implement, the direct method requires two separate simulations, one at $\lambda = 0$ and another at $\lambda = 1$. Furthermore, a separate FEP (or TI) simulation has to be implemented to obtain the free-energy change and, consequently, the entropy change. The statistical uncertainty of the direct formula asymptotically grows with system size as $\sqrt{N}$ (14). This limits the applicability of the direct formula (and of the other methods, described below, whose uncertainty increases with system size) to small systems.

The *thermodynamic perturbation* (TP) formula for $\Delta\bar{U}$ (15–17) follows from the direct formula by calculating, employing a standard TP approach, the average at $\lambda = \lambda_{i+1}$ in the ensemble at $\lambda = \lambda_i$ [the TP formula can be also derived from the temperature derivative of the FEP formula (Equation 1)], namely

$$\Delta\bar{U}_{i,i+1} = \frac{\left\langle U(\lambda_{i+1})\, e^{-\Delta U_{uv}(i+1,i)/kT}\right\rangle_{\lambda_i}}{\left\langle e^{-\Delta U_{uv}(i+1,i)/kT}\right\rangle_{\lambda_i}} - \langle U(\lambda_i)\rangle_{\lambda_i}. \qquad 4.$$

Because in the TP formula all the averages are performed at the same λ, only one simulation per window is required. Moreover, the TP formula makes it possible to calculate the total energy change for the same thermodynamic path used for the free-energy change. Although the the statistical uncertainty of the TP formula grows as $\sqrt{N}$, in most cases the TP formula improves the convergence rate of the direct formula because the prefactor of the asymptotic scaling of the statistical uncertainty favors the TP formula over the direct formula whenever the two solutes *X* and *Y* have similar structure and similar interactions with the solvent (14).

The *thermodynamic integration* (TI) formula for the entropy (17–19),

$$-T\Delta S = \beta \int_0^1 d\lambda \left[\left\langle U(\lambda) \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda - \langle U(\lambda) \rangle_\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \right], \qquad 5.$$

is obtained by taking the temperature derivative of the TI formula (Equation 2) for the free-energy change. A statistical analysis (14) shows that, as for the direct and TP formulas for the potential energy change, the variance of the TI formula for the entropy grows with system size as $\sqrt{N}$. The asymptotic prefactor is similar to the one discussed above for the TP formula. It should also be pointed out that the slow-growth method (20) cannot be applied to the TI formula for the entropy change.

The entropy change between states *X* and *Y* at temperature *T* is given by the temperature derivative of the free energy. The *finite-difference* (FD) method approximates the derivative as (17, 21, 22)

$$\Delta S(T) \simeq -\frac{\Delta F(T + \Delta T) - F(T - \Delta T)}{2\Delta T}, \qquad 6.$$

where ΔT is a small temperature shift. The FD approximation is based on the assumption that the heat capacity is constant over a certain range of temperatures near the target temperature *T*. For aqueous solutions, this assumption normally holds near room temperature, with ΔT as large as 50 K (17, 23). The choice of ΔT is, however, important to minimize the statistical uncertainty of the FD formula that is inversely proportional to ΔT. As ΔT becomes smaller, significantly more expensive computations are required to achieve converged results from the FD formula. If ΔT is too large, on the other hand, the accuracy of the FD formula degrades because the assumption of linearity of the free-energy change over temperature does not hold any longer. The FD formula is an attractive method because—in contrast to the direct, TP, and TI formulas—it estimates entropy and enthalpy differences with a statistical uncertainty independent of system size. The limit of the FD formula may lay, however, in its accuracy. One source of systematic errors is the FD approximation itself. This is probably unimportant for the study of aqueous solvation at room temperature, but it may have important consequences in the study of ligand binding. Another source of systematic errors is in the sampling procedure adopted to estimate the free-energy change. Incorrect or insufficient equilibration at each perturbation window, potential energy cutoffs, failure to explore all relevant potential energy minima of the system, etc, all contribute to the loss of accuracy in the estimate of the free energy. The size of such errors, although small compared with ΔF itself, may be important for the FD estimate of $T\Delta S$, as the errors are amplified by the factor $T/\Delta T$. This sets a lower limit on the size of ΔT that can be safely selected. The direct, TP, and TI formulas do not suffer from this drawback.

A variation on the theme proposed for the FD formula has been developed (24) according to whether the averages necessary to compute $\Delta F(T + \Delta T)$ and $\Delta F(T - \Delta T)$ are all computed at temperature T employing a TP of the temperature (25, 26). For an accurate calculation of the entropy change with the temperature perturbation formula, it is necessary for the temperature perturbation ΔT to decrease with increasing system size. A temperature perturbation affects the entire system, and therefore, the larger the system the smaller the value of ΔT that must be chosen. This effectively makes the variance of the temperature perturbation formula grow as $\sqrt{N}$ (14), the same as for the direct, TP, and TI formulas. In addition, the temperature perturbation formula suffers from the accuracy problems caused by the amplifying factor $T/\Delta T$, as discussed previously for the FD formula. In conclusion, the temperature perturbation formula—like the direct, TP, and TI formulas—presents the same relatively poor convergence behavior as a function of system size without improving on the accuracy of the FD formula.

Dependence on insertion conditions In the thermodynamic limit, the chemical potential of a species in solution is independent of the boundary conditions of the system. In particular, the chemical potential at constant volume, defined as the change in Helmholtz free energy by inserting the solute in the solution at constant volume, is equal to the chemical potential at constant pressure, defined as the change in Gibbs free energy by inserting the solute at constant pressure, provided that the process is initiated at the same thermodynamic conditions (27). This property of the chemical potential makes it possible to compare free energies of solvation computed in a constant volume ensemble with experiments conducted at constant pressure. The partial molar energy and entropy, however, depend on insertion conditions (28). Simple formulas can be used (22, 27, 28) to convert excess energies and entropies of solvation calculated at constant volume to the corresponding quantities for the constant pressure process. These formulas depend only on the partial molar volume of the solute and thermodynamic properties of the pure solvent (isothermal compressibility and thermal expansion coefficient).

Solvation shell approximation The statistical uncertainty of the excess energy and entropy obtained from the direct, TP, and TI formulas grows as $\sqrt{N}$ because the computation of the change in solvent-solvent energy involves all the solvent molecules of the system. Including only those solvent molecules within a certain shell around the solute would, instead, yield a statistical uncertainty independent of the total number of solvent molecules included in the system.

Matubayasi et al (29) developed a solvation shell formula for the excess energy of solvation as a function of solvation shell radius. They found that the hydration shell formula yields the exact excess energy of hydration of methane

for a hydration shell radius of about 9 Å (29). They also showed that their solvation shell formula yields, in the thermodynamic limit, the excess energy at constant pressure regardless of the insertion conditions. Thus, the excess energy of solvation at constant pressure, unlike the excess energy at constant volume, can be studied by local models like the solvation shell model. This conclusion is at odds with a recent study by Cann & Patey (30), who, by studying a phenomenological model of solvation, concluded that nonlocal contributions are present even for constant pressure processes. This result, if valid, would invalidate much of the solvation shell analysis of experimental data conducted in the past (31, 32). However, the conclusion of Cann & Patey has been criticized (33). A different interpretation of their result is presented by Matubayasi et al (34). They conclude that the solvation shell formula of Matubayasi et al (29, 35) yields excess energies of solvation that can be directly compared with experimental data once the difference between the computational and experimental standard states is resolved. This issue is discussed below.

Experimental measurements and the Ben-Naim standard state convention Computer simulations do not usually take into account the ideal part of excess thermodynamic quantities of solvation. Effectively, they provide the excess thermodynamic quantities in the Ben-Naim (28) frozen ensemble, whereby the solute is inserted in the solvent in a fixed position. When comparing experimental measurements (that obviously contain such ideal contributions) with computer simulation results, the ideal contributions have to be separately analyzed. Fortunately, all of these ideal contributions either are trivial or are dependent only on the properties of the pure solvent that, therefore, are the same for the solvation of all the solutes in the same solvent (22, 28, 29, 35). These correction factors are sometimes called liberation terms (33). They can be, in fact, interpreted as the change in excess thermodynamic quantities from the frozen ensemble after releasing the constraint on the position of the solute.

Solvation of Small Organic Molecules

This section reviews the application of the methodologies presented in the previous section to the computational study of the thermodynamics of solvation of small solutes with explicit solvent. Particular effort has been devoted to the entropy-enthalpy decomposition of the free energy of hydration at ambient conditions of simple hydrophobic solutes. Only few instances of similar studies for complex hydrophobic solutes and polar organic molecules have been reported.

HYDROPHOBIC HYDRATION Most of the studies of the hydrophobic effect using explicit solvent models in computer simulations have focused on calculating the free energy of hydration of a hydrophobic solute and on elucidating the changes of solvent structure near a hydrophobic solute or surface (20, 36–40).

This section reviews those studies in which the the free energy of hydrophobic hydration has been decomposed into the enthalpic and entropic components. The enthalpy-entropy decomposition of the free energy is of particular importance in understanding the origins of the hydrophobic effect, as it has been recognized that hydrophobic hydration is characterized by a large decrease in entropy that has been interpreted either as a local ordering of water around the solute (41) or, alternatively, as a result of the small size of the water molecule (12, 42).

Guillot et al (15, 16) studied the temperature dependence of the chemical potential, excess energy, and excess entropy of gas hydrates and methane in water. They adopted the TP formula to calculate the excess free energy and excess energy of solvation through a particle insertion method (43) using neat water trajectories from constant-volume MD simulations. They were able to correlate the solubilities of the noble gases and methane to the average solute-water interaction energy E_μ and the solute-water entropy ΔS_μ (obtained from the difference between chemical potential and average solute-water energy). They observed that E_μ governs the hierarchy of the solubilities in water. In particular, the larger the solute, the more favorable the solute-solvent energy and the larger the solubility. The solute-water entropy, only marginally affected by the details of the solute-water interaction, governs, instead, the temperature dependence of the solubilities. At room temperature, $-T\Delta S_\mu$ has a maximum that coincides with a minimum in solubility. At higher temperatures, the higher occurrences of cavities in water causes a decrease of $-T\Delta S_\mu$ that results in larger solubilities. The decrease of $-T\Delta S_\mu$ at lower temperature is explained by the increase of rigidity of the hydrogen bonding network that favors naturally occurring cages of water molecules that can accept the solute with little entropic penalty. This view is confirmed by the information theory cavity distributions temperature study of Garde et al (44, 45) and Berne (46). The analysis of Guillot & Guissani (16) suggests that the solubility at high temperatures is governed by the distribution of cavities in the solvent whereas at lower temperatures it is governed by the caging effect first proposed by Frank & Evans (41). Guillot et al (15, 16) also estimated the solvent-solvent contribution to the total excess energy and entropy of hydration of the rare gases and methane. The solvent-solvent contribution was found to be small, so the analysis described above, based only on the solute-solvent components, is warranted.

Matubayasi et al (29, 35) analyzed the solvation shell model for the study of the excess energy, excess volume, and compressibility of a hydrophobic solute (methane). The accuracy of the model for the excess energy has been tested by a calculation of the exact excess energy from a test-particle insertion approach, as done by Guillot et al (15, 16). The chemical potential, excess energy, and excess entropy of solvation of krypton in water and in a Lennard-Jones fluid at

room temperature have been studied by Durell & Wallqvist (12), who employed the direct method. A standard deviation of 1 kcal/mol of the excess energy of hydration was obtained after a long MD sampling. They observe that most of the favorable excess energy is accounted for by the solute-solvent attractive interaction, the stabilization of the solvent-solvent energy around the solute being less important. Likewise, the loss of entropy is explained by the loss of translational freedom of the solvent more than by the loss of orientational freedom, as would be expected by a caging effect. Both observations contradict the iceberg model of Frank & Evans (41), bringing forward the idea that the hydrophobic effect is simply the consequence of the small size of the water molecule (42).

Durell & Wallqvist (12) and Matubayasi et al (29) investigated the binding energy of solvent molecules as a function of the distance from a hydrophobic solute. Durell & Wallqvist (12) found that the solvent-solvent energy in the first solvation shell is destabilized and that it decays monotonically to the bulk value. Matubayasi et al (29) found, instead, a stabilization of the solvent binding energy near the solute. This feature should be probably attributed to the somewhat overly structured ST2 water potential adopted by Matubayasi et al. They observed that for methane, most of the excess energy comes from a hydration shell around the solute with a radius of about 9 Å. They also show that it is necessary to calculate accurately the shift of solvent binding energy at a distance r far from the solute, because any error is amplified by a r^2 factor when summing over the water molecules at that distance r from the solute.

Zeng et al (13) analyzed the water structure around rare gases by using energy minimization and normal mode analysis. They showed that the harmonic approximation yields hydration entropies, in disagreement with the experiments. This suggests that the negative hydration entropy is due to a decrease of the number of the energy minima and to anharmonic contributions.

Lazaridis & Paulatis (47, 48) calculated the entropy of hydration of simple monoatomic solutes (methane, helium, argon, krypton, xenon, radon) as a function of temperature by evaluating a form of the Green-Wallace expansion of the entropy (49–51) up to the two-particle terms. In this approach, the solute-solvent distribution functions are obtained from explicit solvent computer simulations and are then processed to extract the solute-solvent component of the entropy of hydration. Three-particle and higher-order terms, including the entropy of solvent reorganization, are neglected. Lazaridis & Paulatis divided the solute-solvent entropy of hydration into a translational term and a rotational term, the latter including only the contributions from the water molecules in the first solvation shell. Their calculated entropies and heat capacities of hydration are in good agreement with the experiments and the explicit solvent calculations of the thermodynamics of solvation of monoatomic solutes summarized previously (12, 15, 16, 29). They found that the translational component of the

entropy of solvation in water is similar to the entropy of hydration in nonpolar solvents of similar size whereas the rotational part accounts for the peculiarities of hydrophobic hydration. They also found, contrary to other studies (12, 42), that the rotational component makes a substantial contribution to the total entropy change. Both the translational and rotational components depend more on the size of the solute than on the strength of the solute-water interaction. Furthermore, at room temperature, the loss of rotational entropy of the solvent near a hydrophobic solute was found to be mostly due to a reduction in the freedom of rotational motion of the water molecules around the solute; the strengthening of water-water interactions played a minor role. At lower temperatures this situation is reversed.

Ashbaugh & Paulatis (52) extended the method of Lazaridis et al (47, 48, 53) to hydrophobic chains of arbitrary shape by recasting the solute-solvent entropy of solvation as a summation over correlation functions between the water molecules and atomic sites of the solute. Only the water molecules in the first hydration shell of each solute site are explicitly taken into account, and the solute orientation dependence of the water-site correlation function is neglected. Ashbaugh & Paulatis (52) found good agreement between their calculated entropies of hydration of methane, ethane, propane, and n-butane with the experimental measurements. In addition, they show that gauche conformation of n-butane is entropically favored but energetically unfavored with respect to the trans conformation. It should be noted that although the study of Ashbaugh & Paulatis (52) showed good agreement with the experiments, a full assessment of their approximations can be obtained only by comparison with exact calculations for the same potential model.

HYDROPHOBIC INTERACTION Hydrophobic interaction, the phenomenon of solvent-induced attraction between two hydrophobic solutes in water, has been the subject of numerous computational studies with explicit solvent models. Most of these studies focused on the free energy of hydrophobic association (54) and, more generally, on the calculation of the potential of mean force between hydrophobic solutes (55–59). In this section we review those studies in which the decomposition of the potential of mean force into its energetic and entropic components has been attempted.

Smith et al (17, 19) used the TI method to calculate the potential of mean force of two methane particles in water. They also used the direct, TI, and FD methods to decompose the potential of mean force into its entropic and energetic contributions. In a 106-water-molecule system, they obtained an uncertainty in the decomposition of the free energy of association of about 0.8 kcal/mol using the TI method. They also obtained a similar uncertainty by employing the direct method. They suggest that in a larger system the TI method should

perform better than the direct method. The calculations of Smith et al clearly show that the hydrophobic interaction is governed by the entropy of association. They obtained a value for the entropy of association of $T\Delta S = 1.5$ kcal/mol. This is consistent with the fact that a pair of hydrophobic solutes in contact, by displaying less solvent-exposed surface, order fewer water molecules than a solvent-separated pair. The calculations also showed that this entropic component was partially offset by the increase in solvent configurational energy.

In a number of studies, the entropy component of the hydrophobic interaction has been studied by analysis of the dependence of the propensity for hydrophobic association (60, 61) or of the potential of mean force (62–64) on temperature. Skipper (60) and Mancera & Buckingham (61) found that the propensity of hydrocarbon molecules to aggregate in water increases with increasing temperature, which suggests that hydrophobic association is driven by a favorable entropy component. The calculations of the potential of mean force between two methane molecules in water show that the occurrence of the contact pair is either unaffected (62) or increases with temperature (63, 64). These studies have shown that at room temperature the favorable entropy of association ($T\Delta S$) is between 1.5 (63) and 3 kcal/mol (64). A similar stabilization of the solvent-separated pair is not observed. Some of the discrepancies between different calculations may be ascribed to the different statistical ensembles adopted. The constant volume studies (17, 19, 62, 63) consistently yield a smaller entropy of association than do the constant pressure studies (64) (see section on Dependence on Insertion Conditions).

Tobias & Brooks (26) decomposed the torsional free-energy surface of n-butane in water and carbon tetrachloride into energetic and entropic components by employing a temperature perturbation method. The solvent effects on the conformational equilibrium are seen as an aspect of the hydrophobic interaction that should favor a more compact structure (gauche) over an extended structure (trans). As the analysis also involved a hydrophobic solvent, Tobias & Brooks were able to ascertain whether the hydrophobic interaction is unique to water or is a property of any solvent (solvophobic interaction). They found a significant increase in the population of the gauche conformer with respect to the gas phase in both water and CCl_4. The amount of solvophobic effect was found to be larger in water where the gauche conformer, in agreement with the accepted view of hydrophobic interaction, was entropically stabilized and energetically destabilized.

HYDRATION OF POLAR MOLECULES Fleischman & Brooks (25) applied a combination of the FEP method and the temperature perturbation method to determine the difference in excess free energy, excess entropy, and excess energy of solvation between a series of alkanes and the corresponding alcohol derivatives.

In particular, they converted methanol into ethane and ethanol into propane. Although their calculations were performed at constant volume without applying the needed ensemble correction factors (22), and the calculated values suffered from large statistical uncertainties, they achieved a satisfactory agreement with the experimental measurements. They observed that the free-energy differences were energy dominated. The large positive change of the excess energy in going from the alcohol to the alkane was related to the loss of hydrogen bonding and Coulomb interactions. On the other hand, the entropy change, in agreement with the experiments, was small because the decrease in entropy due to increasing hydrophobicity was offset by the increase in disorder caused by the loss of the hydrogen bond between the hydroxyl group and the water molecules.

Kubo et al (22) studied the free energies, enthalpies, and entropies of solvation of a series of simple amines, oxides, and sulfides. Starting from ammonia, water, and hydrogen sulfide, the molecules were fully methylated by addition of one methyl group at a time. The absolute solvation properties of the starting molecules were calculated by letting them disappear from solution. The free-energy changes were calculated by using the FEP method in a constant pressure MD simulation using the TIP3P model for water (65) and the Lennard-Jones AMBER (66) parameters for the solutes. Partial charges at each atom site were fitted to the molecular electrostatic potentials generated from CHELPG ab initio calculations (67). The entropy and enthalpy changes at 298 K were calculated by using the FD method from the free-energy changes at 268 K, 298 K, and 328 K. Similar results have also been obtained with the TP formula using a constant-pressure MC procedure (68). The small (often less than 1 kcal/mol) statistical uncertainty achieved for the ΔH and $T\Delta S$ of solvation allowed for a quantitative interpretation of the results. In agreement with the experiments, the free energies of solvation become less favorable with increasing degree of methylation; however, this effect was consistently overestimated by the calculations. Even more pronounced are the discrepancies between the calculated and experimental enthalpies and entropies of solvation. In general, the calculations predict a less favorable enthalpy of solvation and a more favorable entropy of solvation than did the experiments. Particularly noticeable is the case of the amines, where the qualitative trend of the calculated enthalpies and entropies of solvation with degree of methylation is opposite that of the experiments. Kubo et al (22) attributed the discrepancies between calculated and experimental free energies of solvation of the amines [observed also in several other studies (69–72)] to the failure of the model in reproducing even qualitatively enthalpies and entropies of solvation. The failure of the model is ascribed to the inability of the force field to correctly describe the changes of hydrogen bonding patterns in solution (73). The polarization of the alkyl groups induced by the solvent, not considered by Kubo et al (22), may also play a role in explaining the discrepancies between calculated and experimental enthalpies and entropies of solvation

(69). The replacement of the AMBER (66) Lennard-Jones parameters with the OPLS (74) Lennard-Jones parameters has been shown to yield enthalpies and entropies of solvation of the amines in better agreement with the experiments (68).

Theory of Enthalpy-Entropy Compensation

The phenomenon of entropy-enthalpy compensation—an empirical observation that, in most chemical processes, the change in enthalpy is partially compensated for by a corresponding change in entropy, resulting in a smaller net free-energy change—has been widely documented (75–78). The theoretical interpretations of entropy-enthalpy compensation given in the literature, however, lack unity and seem inadequate for the presentation of a clear explanation of the physical causes that produce the phenomenon. In this section, we give an overview of the numerous efforts in the field that are of relevance in the theory of solvation and, more generally, molecular association.

Entropy-enthalpy compensation and its analogues (for example, isokinetic relationship) are terms used to describe phenomena not necessarily related to each other. This has been one of the sources of confusion in the literature. The term entropy-enthalpy has been used, for example, to characterize solvation, ligand binding, and chemical reactions and their temperature and pressure dependence. The term entropy-enthalpy compensation has been also used to indicate the existence of a quantity that does not contribute to the free energy change because it appears with opposing signs in the expressions of ΔH and $-T\Delta S$.

Experimental evidence (77–82) shows that entropy-enthalpy compensation is likely in many chemical and biological processes. Recently, Dunitz (83) proposed a simple model that predicts the occurrence of entropy-enthalpy compensation at a certain temperature in any process characterized by weak molecular interactions. This is the case for biological systems in water where the hydrogen bond, which produces a stabilization of about 4 kcal/mol, determines entropy-enthalpy compensation at about physiological temperature. It has been argued (84, 85), however, that in some cases entropy-enthalpy compensation is an artifact of correlated errors in the measurements of entropy and enthalpies from Arrehenius or Van't Hoff plots. Calorimetrical measurements do not suffer from this difficulty.

In some cases, entropy-enthalpy compensation is a consequence of thermodynamic relations and therefore thermodynamically necessary. In other words, for these cases the phenomenon is independent of interaction potentials and therefore conveys no specific information about molecular interactions. For example (27) it can be shown that the processes of varying temperature or volume of a system are entropy-enthalpy compensating. It can be also shown that differences between the entropies and enthalpies calculated in different statistical ensembles are compensating because the free energy is an ensemble-independent

quantity whereas the entropy and the enthalpy are not. These cases, threfore, can be considered trivial occurrences of entropy-enthalpy compensation.

The phenomenon of entropy-enthalpy compensation in hydrophobic solvation has received particular attention. Most of the theoretical and computational studies focused on the role of the solvent reorganization. That the TP (76), TI (86), and Widom particle insertion formulae (16, 25) for the free energy do not explicitly involve the solvent-solvent potential energy term has led some researchers to conclude that the change in solvent-solvent energy (the solvent reorganization energy) compensates for a term in the entropy of solvation (the solvent reorganization entropy) and therefore does not contribute to the free energy of solvation. This definition of solvent reorganization entropy is not unique. Lee (87) proposed that the total entropy change should be regarded as the reorganization entropy, in which case the solvent reorganization entropy (the total entropy change in this case) makes a substantial contribution to the free energy of cavity formation. Ashbaugh & Paulatis (52) analyzed the expression for the entropy of solvation in terms of an infinite series involving multiparticle solvent correlation functions. They were able to show that the two-particle term is partially compensated for by the solvent reorganization energy, but they were unable to prove that the remaining terms, which are not compensated, are negligible at liquid-like densities. This is probably not the case (33). Furthermore, by a similar decomposition, Matubayasi et al (29) found that in addition to the solvent reorganization energy, a term of the entropy expansion is exactly canceled by the the solute-solvent energy. It would be useful to further develop these analyses by providing a physical interpretation and numerical evaluation of the compensating terms and to the remaining terms that, by definition, are noncompensating.

Entropy-enthalpy compensation in computational alchemy has also been addressed (22, 27, 88). The thermodynamic process involves a mutation of a molecular species (solute or ligand) into another by varying the force field parameters (charges, atomic radii, etc) to extract the change in free energy, enthalpy, and entropy. A similar process can be studied experimentally by analyzing the differences between experimentally determined solubilities and affinities. Qian & Hopfield (27) derived equations for the infinitesimal changes in the entropy and enthalpy by varying the strength of the interaction potential and noticed that in summing these two contributions to obtain the free energy change, a term related to the fluctuations of the potential energy is canceled and thus does not contribute to the free-energy change. The sign and magnitude of this term, however, are not specified and cannot be used, in general, to prove entropy-enthalpy compensation. In their explicit calculations of free energies, entropies, and enthalpies of methylation of simple organic molecules, Kubo et al (22) observe that entropy-enthalpy compensation is observed for all the mutations they attempted. Gallicchio et al (88), however, presented

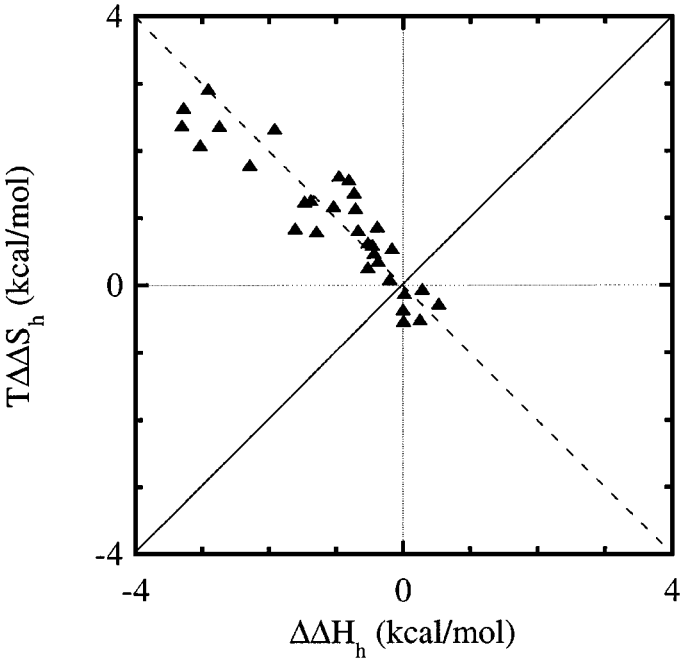

Figure 1 Experimental relative entropy of hydration ($T\Delta\Delta S_h$) versus relative enthalpy of hydration ($\Delta\Delta H_h$) for a series of entropy-enthalpy noncompensating solute pairs. (*solid line*) Perfect entropy-enthalpy compensation; (*dashed line*) perfect entropy-enthalpy reinforcement. (See Table 1 for the list of pairs and corresponding relative thermodynamic data.)

an analysis of experimental relative enthalpies and entropies of hydration and ligand binding, and several instances are reported in which entropy and enthalpy changes do not compensate each other; rather, they add constructively to produce an even greater free-energy change. The experimentally derived hydration data reported by Gallicchio et al is shown in Figure 1 and Table 1. The data indicates that entropy-enthalpy reinforcement is more likely to occur in cases where, as postulated by Lee (87), the solute volume is approximately unchanged. Gallicchio et al (88) observe that the process of designing molecules that optimize properties, like solubility or free energy of ligand binding, is often impeded by the fact that changes directed to strengthen the association of the solute with the solvent, or the ligand to the host, are accompanied by a compensating reduction in entropy, resulting in a small change in free energy (which sometimes is not even in the expected direction). They propose that the study of the exceptions to entropy-enthalpy compensation may be helpful in the process of rational molecular design.

Table 1 Experimental relative free energy of hydration ($\Delta\Delta G_h$), relative entropy of hydration ($T\Delta\Delta S_h$), and relative enthalpy of hydration ($\Delta\Delta H_h$) at 25°C for entropy-enthalpy noncompensating solute pairs

Solute pair[a]		$\Delta\Delta G_h$[b]	$\Delta\Delta H_h$[b]	$T\Delta\Delta S_h$[b]
Methane	→ Ethene[c]	−0.73	−0.37	0.36
	→ Fluoromethane[c]	−2.22	−1.04	1.18
Ethane	→ Chloromethane[c]	−2.39	−0.81	1.58
	→ Chloroethane[c]	−2.46	−1.61	0.85
	→ Bromomethane[c]	−2.65	−1.38	1.27
	→ Iodomethane[c]	−2.72	−1.47	1.25
Ethene	→ Fluoromethane[c]	−1.49	−0.67	0.82
Ethyne	→ 1-Propyne[c]	−0.29	−0.20	0.09
	→ Cyclopropane[c]	0.76	0.26	−0.51
Propane	→ 2-Methylpropane[c]	0.36	0.00	0.36
	→ 2-Methyl-1-propene[c]	−0.79	−0.52	0.27
	→ Chloroethane[c]	−2.58	−0.96	1.63
	→ Acetonitrile[c]	−5.84	−2.91	2.93
1-Propene	→ Butane[c]	0.81	0.54	−0.27
	→ 2-Propanone[c]	−5.12	−3.03	2.09
	→ Propanenitrile[c]	−5.11	−2.74	2.37
1-Propyne	→ 1-Butyne[c]	0.14	0.03	−0.12
	→ 1-Buten-3-yne[c]	0.35	0.29	−0.05
Cyclopropane	→ 1-Propyne[c]	−1.05	−0.45	0.60
	→ 1-Butyne[c]	−0.91	−0.43	0.48
	→ 1-Buten-3-yne[c]	−0.71	−0.16	0.54
Butane	→ Propanenitrile[c]	−5.92	−3.28	2.65
2-Methylpropane	→ 2-Methyl-1-propene[c]	−1.15	−0.52	0.63
1,3-Butadiene	→ Tetrahydrofuran[c]	−4.08	−2.29	1.79
	→ 2-Butanone[c]	−4.25	−1.92	2.33
Hexane	→ Cyclohexane[c]	−1.26	−0.38	0.87
Cyclohexane	→ Methylbenzene[c]	−2.11	−0.73	1.38
Heptane	→ 4-Methyl-2-pentanone[c]	−5.68	−3.30	2.38
1,3-Dimethylbenzene	→ (1-Methylethyl) benzene[c]	0.54	0.01	−0.52
Propylbenzene	→ Naphthalene[c]	−1.86	−0.71	1.15
(1-Methylethyl) benzene	→ Naphthalene[c]	−2.09	−1.29	0.80

[a] The numerical values reported correspond to the changes of thermodynamic quantities in the direction indicated.
[b] In kilocalories per mole.
[c] From Reference 79.

Applications to Biological Macromolecules

Enthalpy-entropy decomposition of the free energies of solvation from explicit solvent models has been applied sporadically and only to small solutes. The application of the methodologies described previously to processes involving biological macromolecules would undoubtedly represent not only a challenging computational problem, but also an extremely useful source of information for the understanding of the thermodynamics behind many biological processes. In this section, we review some issues regarding protein folding and ligand binding and present a perspective on the role that the calculations of enthalpies and entropies may play in these areas of research.

The entropy and enthalpy of hydration of small organic molecules are important elements used to analyze the thermodynamics of protein folding (89, 90). Computer simulations directed to the study of hydrophobic interactions should be able to determine to what extent the dehydration of nonpolar residues upon folding contributes to the increase in the entropy of folding (91–93). The thermodynamics of protein folding is partially determined by a delicate balance between hydrophilic and hydrophobic interactions (94). The contribution of polar residues to the thermodynamics of protein folding (95) could be elucidated by using computer models. In the study of the thermodynamics of protein folding it is also critical to be able to understand the temperature dependence of all the thermodynamic determinants involved in the process (96). The existence of the isoenthalpic and isoentropic temperatures, for instance, at which the enthalpy and entropy of hydrophobic hydration is the same for large classes of molecules has been a source of speculation (23, 44, 97–99). The work of Lazaridis et al (100) represents a first, large-scale attempt to validate, through atom-based computer modeling, the phenomenological models developed to describe the thermodynamic stability of the protein native state. By observing that some aspects of the phenomenological models are qualitatively correct whereas others need to be reexamined, Lazaridis et al (100) show that computer simulations of explicit models of biomolecules in solution can substantially increase our understanding of protein thermodynamics.

Component analysis is a tool widely used to interpret the thermodynamics of biological processes. This approach assumes that different portions or characteristics of the system contribute independently and additively to the thermodynamic quantities. Although generally successful in studying the thermodynamics of reactions involving the breaking/forming of covalent bonds, the application of additivity principles to biological processes that involve weak molecular interaction has been criticized (6). Computer simulations using explicit atomic models yield formally exact thermodynamic quantities and can be applied to analyze and test the validity of additivity principles applied to the study of the thermodynamics of protein folding (100) and ligand binding.

The understanding of the thermodynamics of binding to biological macromolecules could also benefit from the decomposition of the free energy of binding into its entropic and enthalpic components. The FEP method has been applied to the calculation of the free energy of ligand binding (3, 4, 7, 10, 101), but this information alone is rarely sufficient to determine quantitatively which are the most important determinants for the process. Our ability to make general predictions is consequently severely limited. The main obstacle is the fact that the free energy of binding is usually the result of a large cancelation between entropy and enthalpy and therefore presents small variation from one molecule to another that are difficult to interpret and are often not in the expected direction (77, 78, 80, 82, 102–106). Several optimization criteria—such as hydrogen bonding matching, maximization of hydrophobic contacts, minimization of configurational strain, rigidification, removal of formal charge, and covalent bond formation—are usually considered in the process of optimizing drug-receptor affinity. Often, however, it is observed (102, 104) that the effects of these changes on the binding affinity are difficult to predict. They can be, instead, successfully characterized by using additional experimental data. Entropy and enthalpy of binding data from computer simulations could be particularly valuable in validating such characterizations.

ELECTROSTATIC EFFECTS IN SOLUTION

The correct and computationally efficient treatment of the long-range interactions in simulations of aqueous solutions, and particularly simulations of biomolecules in solution, continues to be an area of active research. The problem involves the modeling of the response to charge perturbations and is intimately related to the difficulty encountered extracting thermodynamic properties from simulations of finite-size systems. There has been significant progress in recent years in our understanding of the mathematical and physical concepts involved and in the development of algorithms for efficiently simulating transformations involving charging processes in solution using explicit solvent.

Two classes of models are used to simulate charging processes in solution. Continuum models (107–110) replace the solvent, which accounts for the bulk of the system simulated, with a dielectric continuum. Explicit solvent models (1–3) consider the solute surrounded by a large number of solvent molecules treated in atomic detail. There are also hybrid models that treat part of the solvent explicitly while attempting to deal with the effects of the bulk solvent by a mean field approximation (111–113). An advantage, in principle, of the explicit solvent models is that the dielectric response of the solute and solvent are obtained from the simulation itself, rather than given as adjustable input

parameters. Also, continuum models introduce an unphysically sharp dielectric boundary between the solute and the solvent. Evidence based on explicit solvent simulations shows that the dielectric response function is not charge independent close to the solute (114–117); to some extent this can be accounted for by adjusting the dielectric boundary appropriately in continuum models. Continuum models assume a particularly simple form for the dielectric response of the medium and are therefore computationally much faster. Also, as discussed below, significant numerical problems are encountered when simulating the dielectric response of the solution using detailed atomistic models. There are relatively few systems for which the results of continuum solvent and explicit solvent calculations have been directly compared (118). A better understanding of the complementary strengths and weaknesses of the two models is likely to be obtained from more focused studies of this kind. A key approximation in the continuum models is that the reaction field generated by a charge in an inhomogeneous dielectric is linearly proportional to that charge. There exists a fully molecular analogue to the linear response approximation employed in continuum models (119). Therefore, when comparing continuum and explicit solvent models, it is possible to separate nonlinear effects from those that arise from the molecularity of the medium. This is discussed further below. In this section we review the state of the art concerning modeling of electrostatic effects in organic and biochemical systems using explicit solvent models.

Models for Long-Range Interactions

The treatment of electrostatic properties using fully atomistic simulations of charged species in polar solvents is difficult because of the long-range character of the Coulomb potential. One always simulates a finite system, which furthermore is of microscopic dimensions. The thermodynamic limit is introduced via the use of different models for the boundary of the simulated system. Periodic boundary conditions are commonly employed in simulations of liquids and solutions, and the long-range interactions are simply truncated at a distance of less than half the box length (120). Many papers have been devoted to the effects of truncation (120–126). There is a growing awareness that for ionic systems, truncating the Coulomb interactions results in large errors in the estimates of thermodynamic parameters. Most importantly, for the analysis of electrostatic properties, quantities that strongly depend on the solvent orientational structure, such as the dielectric constant, are not accurately reproduced.

The truncation of the Coulomb interactions beyond a cutoff distance has a long history in computer simulations of liquids (127, 128). For simulations of aqueous solutions and biomolecules, a cutoff radius of between 8 Å and 16 Å is typically employed. The motivation for the use of a cutoff is both to reduce

the computational cost and to remove artifacts arising from the imposition of periodic boundary conditions. A variety of methods for truncating the coulomb interactions using switching functions and/or group-based cutoffs has been proposed (121, 122). When a spherical cutoff is applied to the coulomb forces in liquid simulations, there is little effect on the radial distribution functions (127, 129–132), and the widespread use of cutoffs in simulations of aqueous solutions is probably due to this observation. However, it has also been known for some time that the orientational correlations in the liquid are strongly perturbed by the use of a cutoff (127, 130–132). Because the polarization of a liquid on the introduction of a charge into a cavity is mainly determined by the reorientation of the polar solvent molecules, the sensitivity of the orientational correlation functions in polar liquids to cutoffs is an indicator of severe problems for modeling charging processes when cutoffs are employed.

Renewed attention has been focused on the effects of different truncation schemes on simulations of inhomogeneous systems. There have been several studies reported on the influence of Coulomb truncation on ion pairing. Bader & Chandler (133) found that a smooth spherical truncation of the electrostatic interaction in a system of ferrous-ferric ions in aqueous solution introduced an unphysical attraction between the ions at relatively large separations. When the long-range electrostatic interactions were calculated with the Ewald method, this artifact was removed; for the latter case, the calculated potential of mean force approached the expected dielectric continuum limit at large ion separations. Earlier studies of sodium dimethyl phosphate ion pairing in solution showed repulsion of the oppositely charged ions at long distances; this behavior was also attributed to the use of spherical truncation (134). Conflicting results regarding the first minimum of the potential of mean force have been found for studies of the Cl^- ion pair in solution (135–138). Two studies employing spherical truncation (135, 136) predicted the existence of deep minimum wells at short but different $Cl^- - Cl^-$ separations. Dang & Pettitt (136) observed a stable contact ion pair, whereas Buckner & Jorgensen (135) predicted a stable solvent-separated ion pair.

These results disagree with those reported by Guardia et al (137) and Hummer et al (138) in which no evidence for anion pairing was found, although they did report a weak solvent-separated minimum. Both groups used the Ewald potential. In addition, Hummer et al (138) performed analogous studies using two different water models as well as reaction field techniques; all their results were consistent with those obtained using the Ewald potential.

Del Buono et al (126) studied the dielectric response of two oppositely charged ions in aqueous solution in two ways: by calculating charging free energies and by analyzing the slope of the ion-ion potential of mean force in the asymptotic region. The correct high dielectric shielding by the ions was

obtained by either of the two methods. The Ewald potential was used to model the long-range interactions. In contrast, simply truncating the Coulomb interactions lead to unphysical behavior.

THE PERIODIC COULOMB POTENTIAL The correct dielectric behavior of ionic solutions, obtained from simulations employing the periodic Coulomb potential (PCP) (also called Wigner potential), has motivated further studies of the properties of this model and the development of methods for applying the model to larger systems. The solution to the periodic Poisson equation is a potential of the form:

$$\Phi'_{\mathrm{Ew}}(\mathrm{x}) = \sum_{\mathrm{n}} \frac{\operatorname{erfc}(\alpha \|\mathrm{x} - L\mathrm{n}\|)}{\|\mathrm{x} - L\mathrm{n}\|} + \frac{1}{\pi} \sum_{\mathrm{k}\neq 0} (L\,\|\mathrm{k}\|^2)^{-1} \exp\left(-\pi^2 \frac{\|\mathrm{k}\|^2}{\alpha^2 L^2} + \frac{2\pi i}{L}\mathrm{x}\cdot\mathrm{k}\right) + \frac{\lambda}{V}\mathrm{x}\cdot\mathrm{d}, \qquad 7.$$

where L is the box length, α is a convergence parameter that controls the relative weights of the real and k-space series, and λ is a parameter that depends on the dielectric constant $\varepsilon_{\mathrm{rf}}$.

De Leeuw et al (139) showed that the last term, which depends on the dipole moment of the unit cell, represents the interaction of a large sphere of replicas of the unit cell with a surrounding medium of dielectric $\varepsilon_{\mathrm{rf}}$. For tinfoil or conducting boundary conditions, $\varepsilon_{\mathrm{rf}} = \infty$ and the dipolar term is absent. The potential is then periodic. Figueirido et al (140) explained the origin of the strong finite-size effect that arises when using this potential to simulate nonperiodic systems. They showed that when calculating the interaction energy between a pair of charges in a periodic system, the $\mathbf{k} = 0$ Fourier mode of the potential must be removed:

$$\Phi_{\mathrm{Ew}}(\mathrm{x}) = \Phi'_{\mathrm{Ew}}(\mathrm{x}) - \frac{1}{L}\frac{\pi}{(\alpha L)^2}, \qquad 8.$$

so that

$$\frac{1}{V}\int_V d^3x\, \Phi_{\mathrm{Ew}}(\mathrm{x}) = 0. \qquad 9.$$

The average of $\Phi_{\mathrm{Ew}}(\mathbf{x})$ over the unit cell is zero. Figure 2 shows a comparison between the PCP calculated by using the Ewald summation method and the Coulomb potentials for a unit charge placed at the origin in a box with sides of length 32 Å. Figueirido et al (140) showed that to a first approximation, the PCP potential can be obtained by shifting the Coulomb potential by its volume average (the $\mathbf{k} = 0$ Fourier mode). Physically, this corresponds to the

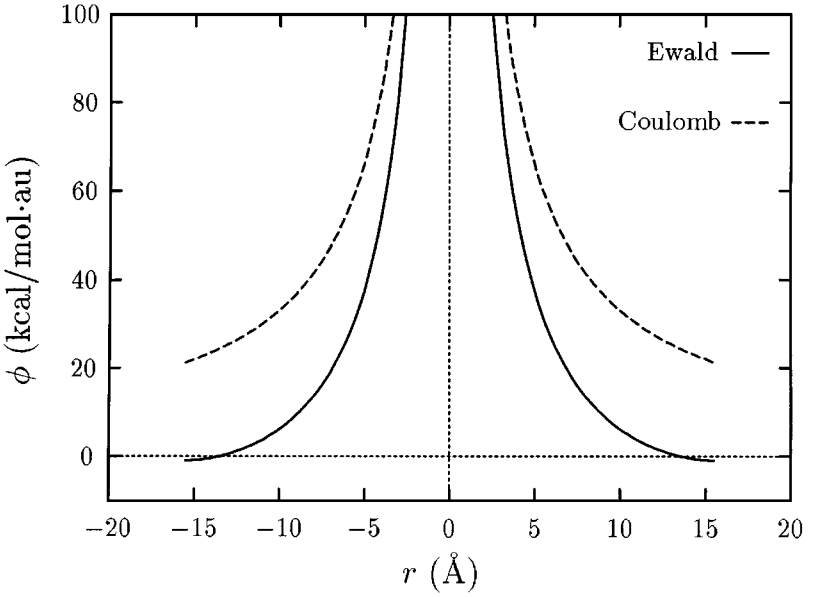

Figure 2 Comparison between Ewald (*full curve*) and Coulomb (*dashed curve*) potentials for a unit charge at the origin in a box of side length 32 Å.

self energy of a charge interacting with the neutralizing background plus the charges in all the image cells,

$$\zeta_{\mathrm{Ew}} = \lim_{\|\mathrm{x}\|\to 0}\left[\Phi_{\mathrm{Ew}}(\mathrm{x}) - \frac{1}{\|\mathrm{x}\|}\right]. \qquad 10.$$

For PCP in a cubic lattice, the self-term is $\zeta_{\mathrm{Ew}} = -2.837297/L$, where L is the length of the cube (141). Thus, the difference between the PCP and Coulomb potentials is closely related to the self-energy of an ion in a periodic lattice. Self-energies are not typically included in molecular simulations. However, Hummer et al (142) made the interesting observation that, when self-energies are included in simulations of ion charging using the PCP, the computed free energy is largely independent of the size of the system for systems containing as few as eight water molecules. Figuerido et al (143) presented a simple theoretical argument that explains why the solvent polarization free energy contains a contribution that largely cancels the lattice self-energy. For a large-enough cell size L, the behavior of the periodic system containing the ion in molecular solvent should approach the continuum limit, for which the potential decays like $\Phi_{\mathrm{Ew}}/\varepsilon_{\mathrm{rf}}$. In this limit, the self-energy of the ion is almost entirely screened by the solvent polarization. These authors (143) proposed a charging free-energy

formula that goes to the continuum limit for a point charge in a periodic lattice filled with a uniform dielctric,

$$\Delta F_{\text{corr}}(q_0 \to q_1) = \Delta F_{\text{total}}(q_0 \to q_1) - \frac{c_0}{L}$$

$$= \Delta F_{\text{sim}}(q_0 \to q_1) + \frac{1}{2}\zeta_{\text{Ew}}\left(1 - \frac{1}{\epsilon}\right)\left(q_0^2 - q_1^2\right). \qquad 11.$$

The Born charging free-energy formula for an ion in an unbounded medium results from a corresponding cancelation of the self-energy of the ion with the polarization free energy of the solvent. Hummer et al (144) have recently shown how to correct Equation 11 for the finite size of the ion and how to extract the ionic radius from a fit to the ion charging free energies in simulations using the PCP with cells of varying size. It is remarkable and not really understood why systems containing as few as eight water molecules in the unit cell are "large" enough to obtain a good approximation to the thermodynamic limit for ion charging free energies. Bogusz et al (145) also proposed a Born-like term to correct for finite-size effects in simulations using the PCP. Their proposed functional form, however, does not in the limit of large cell size correspond to the solution of the periodic Poisson equation in a continuum dielectric.

FAST COMPUTATIONS WITH THE PERIODIC COULOMB POTENTIAL With a better understanding emerging of the deficiencies of cutoff schemes, especially for simulating processes in which the charge state changes, and with a better understanding of how to correct for finite-size effects when using the periodic Coulomb potential, there is increased interest in using this long-range potential model to simulate macromolecular systems in solution. Attention has now turned to developing algorithms for speeding up the calculations for these very large systems. Two approaches are being pursued: the particle mesh Ewald method (PME) and the periodic fast multipole method (FMM). The PME method and a smooth variant developed by Darden et al (146) are based on Hockney & Eastwood's idea (147) of assigning charges to a mesh according to their real space positions: The CPU time saving comes from applying the fast Fourier transform to the particle mesh to accelerate the reciprocal space calculations of the Ewald sum. The algorithms are found to be on the order of $O(N \log N)$. The FMM of Greengard & Rokhlin (148), which scales like $O(N)$, is based on two approximations. First, the potential produced by a collection of charges can be approximated by a sum of potentials produced by the corresponding multipoles of the charge distribution. Second, the multipolar expansions of the potentials produced by several clusters can be lumped together into a local expansion. Board et al (149) described one of the first implementations of FMM in the context of MD. An implementation of the FMM for

systems with periodic boundary conditions was described by Schmidt & Lee (150). Reports of the performance of this method have appeared from several laboratories (151–153).

Recently, Figueirido et al (154) described a new algorithm in which the periodic FMM is combined with a multiple time–step algorithm, the reversible reference system propagator (r-RESPA) method. A recent review of multiple time–step methods applied to biomolecules has recently appeared (155). In r-RESPA, developed by Tuckerman et al (156), the system is decomposed into two or more subsystems through a Trotter factorization of the Liouville propagator. The basic idea is to update the rapidly varying forces more frequently while recomputing the more slowly varying forces less often. For accomplishing this, r-RESPA provides a systematic algorithm that conserves energy to third order in time. Figueirido et al (154) showed that the best results for the combined RESPA periodic FMM algorithm (RFMM) were obtained by dividing the nonbonded interactions into three ranges: (*a*) short range—those arising from within the smallest cluster used to construct the multipole moments and their first neighbors; (*b*) medium range—those arising from the second neighbors; and (*c*) long range—the contribution from the local expansion of the distant multipoles. With this decomposition, if the energy drift per step is required to be less than $\log\Delta E < -2.5$, which gives acceptable numerical accuracy, then the fastest varying forces can be updated with a 0.5-fs time step whereas the most slowly varying forces can be updated with a 12-fs time step. In this work, six multipole moments were required for energy conservation. Bishop et al (157) find that for a simulation of a highly charged system, a protein-DNA complex in water, additional multipole terms are needed to conserve energy.

The combined r-RESPA/FMM is much faster then the standard Verlet/Ewald method (154). For protein plus water systems containing between 8,000 and 44,000 atoms, the CPU time saving varied between 3 and 23. For systems containing ~20,000 atoms, r-RESPA/FMM and PME exhibit comparable performance. They are both about twice as fast as a simple spherical cutoff ($r_c = 10$ Å). The smoothed particle mesh Ewald (SPME) benefits most from the use of a small, real space part (as small as $r_c = 6$ Å) so that fast Fourier transforms, which speed up the *k*-space part, can be used to maximum advantage. In contrast, to gain the most from r-RESPA, it is desirable to use a large cutoff for the real space sum. Thus, there is a trade off between the SPME and r-RESPA methods. Although the two methods appear to be competitive for ~20,000 atoms, RESPA/FMM is expected to be faster than RESPA/SPME for even larger systems because of their respective computational complexities.

REACTION FIELD METHODS Reaction field models have been proposed as another way to avoid the introduction of cutoffs when calculating the long-range

electrostatic interactions (120). The conventional reaction field model was developed to correct for the perturbations resulting from the use of a coulomb cutoff in simulations of dipolar liquids (158). In this model, each dipolar molecule in the liquid is considered in turn to be the center of a sphere of radius R_{cut}; the sphere is imagined to be surrounded by a continuum solvent with dielectric ε_{rf}. The effect of the continuum solvent on the central molecule, which replaces the interactions neglected when the cutoff is introduced, is to induce a reaction field that is proportional to the total dipole moment within the sphere. Alper & Levy (159) developed a generalized reaction field model (GRF) for use in MD simulations of a finite system constrained to a spherical geometry based on Kirkwood's solution of the electrostatic equations for the reaction field acting on an arbitrary charge distribution in a spherical cavity (160). The reaction field acting on a charge at any point in the cavity depends on the multipole moments of the charge distribution in the cavity. As a charge is displaced from the center of the cavity, an increasing number of multipoles must be included in the reaction field. Alper & Levy (159) studied ion charging free energies using the GRF model. They reported that although the hydration free energy of the ion agrees with the Born hydration free energy when the ion is close to the center of the sphere, there are significant deviations when the ion is moved toward the boundary, even when the reaction field includes up to 20 multipoles. Beglov & Roux (161) combined a GRF-type model for the electrostatic interactions with a solvent boundary potential for the van der Waals terms. Lee & Warshel (153) proposed a local reaction field model that retains the first four terms of the multipolar expansion, but they did not analyze the sensitivity of the model to the distance between the charges and the boundary. Wallqvist (162) and Wang & Hermans (163) studied an image charge approximation first proposed by Friedman (164) that is similar to the GRF model. Wang & Hermans also reported that ion charging free energies strongly depended on the position of the ion in the spherical cavity. Relatively little effort has been devoted to exploring improvements in reaction field models for use in computer simulations of biomolecules. There are attractive features of these models that warrant further efforts.

FREE ENERGIES OF IONS AND LINEAR RESPONSE Considerable effort has been devoted during the past decade to the use of computer simulations as a tool to study the response of complex systems to charge perturbations, such as occurs in ionization equilibria, in electron transfer, and in molecular association. The linear response approximation provides the theoretical foundation for most of the models used to interpret these phenomena. Computer simulations with explicit solvent provide a powerful tool to study the range of validity of this approximation. The paradigm problem is the charging free energy of ions. The

linear response approximation is easily and informatively derived starting from a perturbation expansion of the free energy. Consider a set of solute atoms of charge (λq_i) immersed in a solute. The free energy for charging the ion(s) from state λ_0 to state λ can be written as a Taylor series expansion of the free energy in the charges

$$\begin{aligned}\Delta F(\lambda - \lambda_0) &= -k_B T \ln\langle \exp[-\beta(\lambda - \lambda_0)V]\rangle_{\lambda_0} \\ &= \sum_{n=1}^{\infty} \frac{(\lambda - \lambda_0)^n}{n!} F^{(n)}(\lambda_0),\end{aligned} \quad 12.$$

where $F^{(n)}$ is the *n*-th derivative of the free energy that is related to the *n*-th cumulant C_n of the probability distribution of *V* by

$$F^{(n)}(\lambda) = (-\beta)^{n-1} C_n(\lambda), \quad n \geq 1, \quad 13.$$

where to order four

$$\begin{aligned} C_1(\lambda) &= \langle V\rangle_\lambda \\ C_2(\lambda) &= \left\langle (V - \langle V\rangle_\lambda)^2 \right\rangle_\lambda = \sigma_\lambda^2 \\ C_3(\lambda) &= \left\langle (V - \langle V\rangle_\lambda)^3 \right\rangle_\lambda = \sigma_\lambda^3 \mathcal{S}(V) \\ C_4(\lambda) &= \left\langle (V - \langle V\rangle_\lambda)^4 \right\rangle_\lambda - 3\sigma_\lambda^4 = \sigma_\lambda^4 \mathcal{K}(V), \end{aligned} \quad 14.$$

where $\mathcal{S}(V)$ and $\mathcal{K}(V)$ are, respectively, the skewness and kurtosis of the probability distribution of *V*.

The first two cumulants are the average and variance of the fluctuations of the electrostatic potential at the solute partial charge sites due to the solvent. For the special case that the distribution of fluctuations in the electrostatic potential is Gaussian, all the terms in Equation 12 beyond the second are zero, and the charging free energy of the system has a particularly simple form:

$$\Delta F = \sum_i \bar{V}_i \Delta q_i - \frac{\beta}{2} \sum_{i,j} \langle \Delta V_i \Delta V_j \rangle \Delta q_i \Delta q_j. \quad 15.$$

Levy et al (119) proposed using this expression for estimating free energies in transformations involving charge perturbations, which they called the Gaussian fluctuation formula for electrostatic free-energy changes in solution. It has a simple physical interpretation. The first term corresponds to the free-energy change due to the difference between the initial- and final-state charge distributions of the solute interacting with the electrostatic potential of the reference system. The second term contains the induction effects related to the change in the polarization of the solvent by the change in the solute charge distribution. In the Gaussian fluctuation model, the free energy is quadratic in the charge

perturbation and the electrostatic potential is linear in the perturbation—hence the name linear response. Several studies of the linear response model for ion charging by computer simulations with explicit solvent have confirmed its qualitative validity (116, 119, 142, 143, 165) over a range of solute shapes and charges. The origins of the quantitative deviations from linearity are of interest. For example, the slope of the ion free-energy change exhibits charge asymmetry; the slope is steeper for negative ions than for positive ions (117, 142, 166). This is attributed to the molecular structure of the ion solvation shell; for negative ions, the water hydrogens point toward the ion and are able to solvate it more strongly than for positive ions of equal size. Larger deviations from linear response have been observed for ions with a charge greater than ± 1 (116, 117, 165, 167). Levy et al (119) reported a 20% deviation from linearity when charging an ion pair between 0.8 and 1.0 unit charge but subsequently found better agreement with linear response in this charge range (126). The discrepancy may be due to the use of different long-range potential models in the two calculations: Coulomb cutoffs were used in the earlier simulations, whereas the Wigner potential was employed in the more recent study. Recently, Ichiye and co-workers developed a framework for determining an apparent local dielectric response from simulations of ions in water which they used to quantify deviations from continuum behavior. They found that, while the spatial variation of the dielectric response reflects specific molecular interactions close to the ion, the integrated response was satisfactorily fit by a linear Born charging model (167a,b).

The Gaussian fluctuation formula Equation 15 is a useful approximation that has a number of applications to problems involving charge transfer in solution. Although the linear response equations are nothing more than second-order perturbation theory of the free energy, these equations become really useful when the physics of the bath fluctuations are indeed close to Gaussian so that the higher-order terms in the perturbation expansion can be reasonably ignored. The Gaussian model provides an estimate, based on simulations of a reference state, of the electrostatic free-energy change due to the creation or redistribution of charge on the solute atoms. As discussed in the following section, it is closely related to continuum electrostatic models and provides the physical basis for these models. Levy et al (119) suggested how the formula could be used in TP simulations of both initial and final states to estimate free-energy changes when the linear response approximation was not accurate enough to connect the two states. "Hummer & Szabo (169) showed how to extract the best estimate of δF to any order in $\delta\lambda$ from such simulations." The use of higher-order cumulants to estimate free-energy changes via computer simulations with explicit solvent has been studied by Liu et al (168) and Hummer & Szabo (169). The drawback to these perturbation approaches to estimating free-energy differences is the difficulty in obtaining sufficient precision in the higher-order cumulants so that it

is difficult to study the convergence properties of the expansion. An empirically adjusted linear response formula has been proposed by Aqvist et al (170) and calibrated by Jorgensen for solvation free energies (171). It is interesting to note that the Gaussian nature of the solvent bath has also provided new insights into the molecular basis for the hydrophobic effect (172, 173).

Electrostatic properties of proteins by computer simulations The study of electrostatic properties of proteins in solution is a subject of major interest for both theoreticians and experimentalists. Many protein functions are related to the use of of charge-charge interactions for a variety of purposes—energy storage, signal transduction, catalytic activity. Furthermore, the stability of proteins depends on a balance between electrostatic and other physical forces that are of approximately equal magnitudes. Several computational approaches have been used to describe the electrostatics and dynamics of solvated proteins. These approaches have ranged from all atom microscopic models (174–180), to semi-microscopic approaches (111, 181, 182), to continuum models where the solvent is represented by a dielectric continuum (109, 110, 113, 183, 184, 186). These studies, using different electrostatic models, have increased our understanding of the dielectric behavior of proteins in solution; the differences among the models have provided useful insights. Nevertheless, both qualitative and quantitative improvements are needed in the models currently used to predict the electrostatic properties that govern protein stability and activity.

In the early 1990s, great progress was made in the development of continuum solvent models for protein electrostatics, as the focus shifted away from explicit solvent simulations. This shift in interest toward continuum solvent models occurred for two reasons: The explicit solvent simulations were simply too slow to permit many computer experiments to be carried out, and they suffered from numerical instabilities related to the approximations used to treat the long-range interactions. An important benchmark for electrostatic models of proteins has been the prediction of pK_a shifts. Groups that use continuum methods have generally obtained reasonably accurate pK_a shifts (187, 188), although the dielectric constant used for the protein interior has been a source of controversy. It has been argued on physical grounds that the interior dielectric of a protein is about 2–4, and most electrostatic calculations with implicit solvent models use values in this range. However, in a provocative paper appearing in 1995, Antosiewicz et al (189) reported that an interior dielectric constant of 20 gave the best pK_a shift results. It was also observed that even the best results did not give much better agreement with experiment than the null hypothesis (i.e. assuming zero pK_a shift), which, of course, carries no information about the dielectric properties of the protein. Demchuk & Wade (190) showed how optimization of the dielectric constant individually for different groups within a protein could lead to significant improvements in the accuracy of pK_a predictions. These

studies raise questions about the microscopic meaning of the parameters used in continuum electrostatic models for proteins and the predictive accuracy that can be obtained with them. The superficially surprising lack of sensitivity of predicted pK_a shifts to the choice of protein dielectric constant simply reflects the surface accessibility of many ionizable groups. For these groups, the pK_a shift is dominated by their proximity to the high dielectric solvent. It should be possible to obtain more useful information about the dielectric behavior of proteins by focusing on residues that have large shifts relative to model compounds and on residues that have functionally significant shifts. More systematic studies of these cases are needed.

Molecular dynamics simulations of proteins with explicit solvent provide a laboratory for studying the approximations that are implicit in continuum electrostatic models and for developing improvements that retain the computational speed and simplicity of these reduced models. The issues one would like to address include the following: (*a*) How can the dielectric properties of boundary layer solvent that behaves differently from the bulk be most accurately captured in continuum models; (*b*) what is the "best" value to use for the effective dielectric constant of a protein in continuum models; (*c*) what is the spatial variation of the dielectric response in a protein and is there a simple way to capture that variation within the framework of continuum models; and (*d*) how can protein conformational averaging (as distinct from small amplitude relaxation) be incorporated into reduced electrostatic models? More generally, for what problems must the full nonlinearity of the protein response be modeled explicitly? These are difficult questions, the answers to which are only beginning to emerge.

One focus area using explicit solvent simulations has been the analysis of the role of nonlinear solvent and protein dynamics in electron transfer reactions (178, 179, 191–193). Marcus theory, a linear response model, has been shown to work remarkably well for many different systems (194). It is interesting to note, as Zhou & Szabo (195) have recently shown, that Marcus theory is a good approximation in the normal regime even when the solvent response is significantly nonlinear. This is an example of a more general phenomenon; it is often possible to fit experimental data concerning protein electrostatics and dynamics with linear models even when the underlying dynamics is not linear.

The effective dielectric constant for several proteins have been estimated from MD simulations by calculating the fluctuations in the protein dipole moment and using the Kirkwood-Frohlich relation (196) to relate the fluctuations to the dielectric constant (174, 175, 197). Questions have been raised about the self-consistency of these calculations, e.g. how to correctly apply the Kirkwood-Frohlich relations to a macromolecule with net charge (198). In any case, the protein dielectric constants calculated in this way depend strongly on how much of the protein is included in the dipole moment fluctuations. When only the core of the protein is included, a low dielectric constant of $\sim$3 is obtained;

the effective dielectric increases to ~25 when the entire protein is included in the calculation of the dipole moment fluctuations. Warshel and coworkers have tried to develop a scheme for interpolating between microscopic and macroscopic dielectric models for proteins based on an analysis of simulations of proteins in a solvent represented as Langevin dipoles (199). They suggest that the effective dielectric constant that should be used to reproduce the effect of protein relaxation on charge-charge interactions in a reduced model is not equal to the effective dielectric that reproduces the corresponding effect upon formation of individual charges. The results of all these different microscopic studies bring out the inhomogeneous nature of the protein dielectric response and are not easily reconciled with the simplest versions of continuum electrostatic models for proteins.

The appropriate theoretical framework for bridging the gap between explicit and continuum solvent models is the linear response approximation, as discussed in the previous section. Conceptually, one can imagine reducing the exact nonlinear response of a protein to a charge perturbation, as obtained from completely atomistic simulations, to a continuum solvent picture by making the reduction in two steps. First, the "exact" results can be computed from FEP simulations and compared with the molecular version of the linear response predictions (Equation 15)—this comparison provides a measure of the nonlinearity inherent in the problem. Second, the parameters extracted by linearizing the results of atomistic simulations can be fit to an even more reduced set of continuum model parameters. There have been few studies of either of the two steps in this reduction of the problem, and a complete analysis of this kind has yet to be reported. An important issue is the choice of reference system for linearizing the molecular results. Del Buono et al computed the intrinsic pK_as of 14 ionizable groups in hen egg white lysozyme, using a single 140-ps simulation of the completely uncharged protein as a reference system to extract the linear response parameters in Equation 15. Overall agreement between the pK_a's calculated from the linearized results of the fully molecular simulations and the results of FD Poisson-Boltzmann calculations were obtained (180). It was found that the contribution of the first solvation shell of the protein to the charging free energy (as approximated by the linear response formulas) was large and distinct from the bulk solvent behavior of more distant solvent molecules. Unfortunately, because of the errors introduced by the use of Coulomb cutoffs in these simulations, only qualitative information could be extracted from the analysis. Although the explicit solvent simulations of charging processes in complex systems are more expensive in terms of computer time, the development of sophisticated technology to handle the long-range interactions has increased the predictive power of these simulations to the point where comparisons between explicit and continuum solvent models can reveal differences that

have their true physical origin in the inherent molecularity of the surrounding medium (118).

CONCLUDING REMARKS

This review has focused on recent progress in two areas to which computer simulations with explicit solvent are being applied: the thermodynamic decomposition of free energies, and modeling electrostatic effects. The computationally intensive nature of these simulations has been an obstacle to the systematic study of interesting problems in solvation thermodynamics, such as the molecular basis for and limitations of the group additivity principle in solvation, or the decomposition of macromolecular binding affinities into enthalpy and entropy components. With the revolution in computer power continuing, these problems are ripe for study but will require an intelligent choice of algorithms and, for the thermodynamic decomposition of binding affinities, additional judicious approximations. Progress in computer simulations of electrostatic effects involving charge perturbations in solution was slow in the early 1990s because of the widespread use of truncated Coulomb potentials in these simulations, among other factors. More recently, there has been considerable progress in the development of more accurate and faster methods for treating the long-range interactions. The goal of creating a consistent molecular framework for modeling the electrostatic properties of biomolecules, within which it will be possible to interpolate between detailed explicit solvent models on the one hand and continuum models on the other, is within reach.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge many discussions about the subject of this review with Gabriela Del Buono, Francisco Figuerido, Masahito Kubo, and Anders Wallqvist. The work has been supported by grants from the National Institute of Health (GM-30580 and GM-52210).

Literature Cited

1. Beveridge DL, Di Capua FM. 1989. *Annu. Rev. Biophys. Chem.* 18:431
2. Kollman PA. 1996. *Acc. Chem. Res.* 29: 461–69
3. Lamb ML, Jorgensen WL. 1997. *Curr. Opin. Chem. Biol.* 1:449–57
4. Gilson MK, Given JA, Bush BL, McCammon JA. 1997. *Biophys. J.* 72:1047–69
5. Brady GP, Szabo A, Sharp KA. 1996. *J. Mol. Biol.* 263:123–25
6. Dill KA. 1997. *J. Biol. Chem.* 272:701–4
7. Tembe BL, McCammon JA. 1984. *Comput. Chem.* 8(4):281

8. Jorgensen WL. 1989. *Acc. Chem. Res.* 22:184
9. Warshel A, Sussman F, King G. 1986. *Biochemistry* 25:8368
10. Kollman P. 1993. *Chem. Rev.* 93:2395
11. Jorgensen WL, Gao J, Ravimohan C. 1985. *J. Phys. Chem.* 89:3470
12. Durell SR, Wallqvist A. 1996. *Biophys. J.* 71:1695
13. Zeng J, Hush NS, Reimers JR. 1993. *Chem. Phys. Lett.* 206:318
14. Gallicchio E, Levy RM. Manuscript in preparation
15. Guillot B, Guissani Y, Bratos S. 1991. *J. Chem. Phys.* 95(5):3643
16. Guillot B, Guissani Y. 1993. *J. Chem. Phys.* 99(10):8075
17. Smith DE, Haymet ADJ. 1993. *J. Chem. Phys.* 98(8):6445
18. Chialvo AA. 1990. *J. Chem. Phys.* 92(1):673
19. Smith DE, Zhang L, Haymet ADJ. 1992. *J. Am. Chem. Soc.* 114:5876–78
20. Pearlman DA, Kollman PA. 1989. *J. Chem. Phys.* 90:2460
21. Ding Y. 1996. *Computational studies of medium effects*. PhD thesis. New Brunswick, NJ: Rutgers Univ.
22. Kubo MF, Gallicchio E, Levy RM. 1997. *J. Phys. Chem.* 101(49):10527–34
23. Privalov PL, Gill SJ. 1988. *Adv. Protein Chem.* 39:191
24. Brooks CL. 1986. *J. Phys. Chem.* 90: 6680–84
25. Fleischman SH, Brooks CL. 1987. *J. Chem. Phys.* 87:3029
26. Tobias DJ, Brooks CL. 1990. *J. Chem. Phys.* 92(4):2582
27. Qian H, Hopfield JJ. 1996. *J. Chem. Phys.* 105(20):9292
28. Ben-Naim A, Marcus Y. 1984. *J. Chem. Phys.* 81(4):2016
29. Matubayasi N, Reed LH, Levy RM. 1994. *J. Phys. Chem.* 98:10640
30. Cann NM, Patey GN. 1997. *J. Chem. Phys.* 106:8165
31. Franks F, ed. 1972–1982. *Water, A Comprehensive Treatise*. New York: Plenum. Vols. 1–7
32. Nemethy G, Sheraga HA. 1962. *J. Chem. Phys.* 36:3401
33. Lazaridis T. 1998. *J. Phys. Chem.* In press
34. Matubayasi N, Gallicchio E, Levy RM. Submitted
35. Matubayasi N, Levy RM. 1996. *J. Phys. Chem.* 100:2681
36. Straatsma TP, Berendsen JC, Postma JPN. 1986. *J. Chem. Phys.* 85:6720
37. Jorgensen WL, Blake JF, Buckner JK. 1989. *Chem. Phys.* 129:193
38. Kaminski G, Duffy EM, Matsui T, Jorgensen WL. 1994. *J. Phys. Chem.* 98: 13077–82
39. Mancera RL, Buckingham AD. 1995. *J. Phys. Chem.* 99:14632–40
40. Mountain RD, Thirumalai D. 1998. *J. Chem. Phys.* In press
41. Frank HS, Evans MW. 1945. *J. Chem. Phys.* 13(11):507–32
42. Madan B, Lee B. 1994. *Biophys. Chem.* 51:279–89
43. Widom B. 1963. *J. Chem. Phys.* 39:2808
44. Garde S, Hummer S, Pratt LR. 1996. *Phys. Rev. Lett.* 77:4966
45. Hummer G, Garde S, García AE, Pohorille A, Pratt LR. 1996. *Proc. Natl. Acad. Sci. USA* 93:8951–55
46. Berne BJ. 1996. *Proc. Natl. Acad. Sci. USA* 93:8800–3
47. Lazaridis T, Paulatis ME. 1992. *J. Phys. Chem.* 96:3847–55
48. Lazaridis T, Paulatis ME. 1994. *J. Phys. Chem.* 98:635–42
49. Green HS. 1952. *The Molecular Theory of Fluids*. Amsterdam: North-Holland
50. Wallace DC. 1987. *J. Chem. Phys.* 87: 2282
51. Wallace DC. 1989. *Phys. Rev. A* 39:4843
52. Ashbaugh HS, Paulatis ME. 1996. *J. Phys. Chem.* 100:1900–13
53. Lazaridis T, Karplus M. 1996. *J. Chem. Phys.* 105(10):4294
54. Jorgensen WL, Buckner JK, Boudon S, Tirado-Rives J. 1988. *J. Chem. Phys.* 89(6):3742
55. Watanabe K, Anderesen HC. 1986. *J. Phys. Chem.* 90:795–802
56. Wallqvist A, Berne BJ. 1988. *Chem. Phys. Lett.* 145:26
57. Van Belle D, Wodak SJ. 1993. *J. Am. Chem. Soc.* 115:647
58. New MH, Berne BJ. 1995. *J. Am. Chem. Soc.* 117:7172–79
59. Payne VA, Matubayasi N, Reed Murphy L, Levy RM. 1997. *J. Phys. Chem. B* 101:2054–60
60. Skipper NT. 1993. *Chem. Phys. Lett.* 207:424
61. Mancera RL, Buckingham AD. 1995. *Chem. Phys. Lett.* 234:296–303
62. Dang LX. 1994. *J. Chem. Phys.* 100: 9032
63. Lüdemann S, Schreiberg H, Abseher R, Steinhauser O. 1996. *J. Chem. Phys.* 104:286
64. Rick SW, Berne BJ. 1997. *J. Phys. Chem. B* 101:10488
65. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. *J. Chem. Phys.* 79(2):926
66. Cornell WD, Cieplak P, Bayly CI, Gould

IR, Merz KM, et al. 1995. *J. Am. Chem. Soc.* 117:5179
67. Breneman CM, Wiberg KB. 1990. *J. Comput. Chem.* 5:129
68. Gallicchio E, Kubo MM, Levy RM. Unpublished results
69. Ding Y, Bernardo DN, Krogh-Jespersen K, Levy RM. 1995. *J. Phys. Chem.* 99: 11575
70. Morgantini P-Y, Kollman PA. 1995. *J. Am. Chem. Soc.* 117:6057
71. Rao B, Singh U. 1989. *J. Am. Chem. Soc.* 111(9):3125
72. Orozco M, Jorgensen WL, Luque FJ. 1993. *J. Comput. Chem.* 14(12):1498
73. Marten B, Kim K, Cortis C, Friesner RA, Murphy RB, et al. 1996. *J. Phys. Chem.* 100:11775
74. Jorgensen WL, Maxwell DS, Tirado-Rives J. 1996. *J. Am. Chem. Soc.* 118: 11225
75. Lumry R, Rajender S. 1970. *Biopolymers* 9:1125–227
76. Ben-Naim A. 1975. *Biopolymers* 14: 1337–55
77. Kuroki R, Nitta K, Yutani K. 1992. *J. Biol. Chem.* 267(34):24297–301
78. Breslauer KJ, Remeta DP, Chou W-Y, Ferrante R, Curry J, et al. 1987. *Proc. Natl. Acad. Sci. USA* 84:8922–26
79. Cavani S, Gianni P, Mollica V, Lepori L. 1981. *J. Solut. Chem.* 10(8):563
80. Danil de Namor AF, Ritt M-C, Schwing-Weill M-J, Arnáud-Neu F, Lewis DVF. 1991. *J. Chem. Soc. Faraday Trans.* 87: 3231–39
81. Gilli P, Ferretti V, Gilli G, Borea PA. 1994. *J. Phys. Chem.* 98:1515–18
82. Peterson BR, Walliman P, Carcanague DR, Diederich F. 1995. *Tetrahedron* 51: 401–21
83. Dunitz JD. 1995. *Chem. Biol.* 2:709–12
84. Linert W, Jameson RF. 1989. *Chem. Soc. Rev.* 18:477
85. Krug RR, Hunter WG, Geiger RA. 1976. *J. Phys. Chem.* 80:2335
86. Yu H-A, Karplus M. 1988. *J. Chem. Phys.* 89(4):2366
87. Lee B. 1994. *Biophys. Chem.* 51:271
88. Gallicchio E, Kubo MF, Levy RM. 1998. *J. Am. Chem. Soc.* In press
89. Makhatadze GI, Privalov PL. 1993. *J. Mol. Biol.* 232:639–59
90. Privalov PL, Makhatadze GI. 1993. *J. Mol. Biol.* 232:660–79
91. Dill KA. 1985. *Biochemistry* 24:1501–9
92. Dill KA. 1990. *Biochemistry* 29:7133–55
93. Spolar RS, Livingstone JR, Record MT Jr. 1992. *Biochemistry* 31:3947–55
94. Yang A-S, Sharp KA, Honig B. 1992. *J. Mol. Biol.* 227:889–900
95. Makhatadze GI, Privalov PL. 1996. *Protein Sci.* 5:507–10
96. Sturtevant JM. 1977. *Proc. Natl. Acad. Sci. USA* 74:2236–40
97. Baldwin RL. 1986. *Proc. Natl. Acad. Sci. USA* 83:8069–72
98. Lee B. 1991. *Proc. Natl. Acad. Sci. USA* 88:5154–58
99. Privalov PL, Makhatadze GI. 1992. *J. Mol. Biol.* 224:715–23
100. Lazaridis T, Archontis G, Karplus M. 1995. *Adv. Protein Chem.* 47:231–306
101. Hermans J, Wang L. 1997. *J. Am. Chem. Soc.* 119:2707–14
102. Connelly PR, Aldape RA, Bruzzese FJ, Chambers SP, Fitzgibbon MJ, et al. 1994. *Proc. Natl. Acad. Sci. USA* 91:1964–68
103. Kato M, Makino R, Iizuka T. 1995. *Biochim. Biophys. Acta* 1246:178–84
104. Charifson PS, Shewchuk LM, Rocque W, Hummel CW, Jordan SR, et al. 1997. *Biochemistry* 36:6283–93
105. Searle MS, Williams DH, Gerhard U. 1992. *J. Am. Chem. Soc.* 114:10697–704
106. Gilli P, Ferretti V, Gilli G, Borea PA. 1994. *J. Phys. Chem.* 98:1515–18
107. Hawkins GD, Cramer CJ, Truhlar DG. 1996. *J. Phys. Chem.* 100:19824–34
108. Still A, Tempczyk WC, Hawley RC, Hendrikson R. 1990. *J. Am. Chem. Soc.* 112:6127–29
109. Gilson MK. 1995. *Curr. Opin. Struct. Biol.* 5:216–23
110. Sharp K, Nicholls A, Fine R, Honig B. 1991. *Science* 252:106–9
111. Warshel A, Aqvist J. 1991. *Annu. Rev. Biophys. Biomol. Struct.* 20:267–98
112. Resat H, McCammon JA. 1996. *J. Chem. Phys.* 104:7645–51
113. Simonson T, Archontis G, Karplus M. 1997. *J. Phys. Chem. B* 101:8349–62
114. Alper H, Levy RM. 1990. *J. Phys. Chem.* 94:8402
115. Papazyan A, Warshel A. 1997. *J. Chem. Phys.* 107:7975
116. Pratt LR, Hummer G, Garcia AE. 1994. *Biophys. Chem.* 51:147–60
117. Hirata F, Redfern P, Levy RM. 1988. *J. Quantum Chem.* 15:179–88
118. Figueirido F, Del Buono G, Levy RM. 1996. *J. Phys. Chem.* 100:6389–92
119. Levy RM, Belhadj M, Kitchen DB. 1991. *J. Chem. Phys.* 95:3627–33
120. Allen MP, Tildesley DJ. 1993. *Computer Simulation of Liquids*. New York: Oxford Univ. Press
121. Loncharich RJ, Brooks B. 1989. *Proteins Struct. Funct. Genet.* 6:32–35

122. Brooks CL, Pettitt BM, Karplus M. 1985. *J. Chem. Phys.* 83:5897–908
123. Linse P, Andersen HC. 1986. *J. Chem. Phys.* 85:3027–41
124. Madura JD, Pettitt BM. 1988. *Chem. Phys. Lett.* 150:105–8
125. York DM, Darden TA, Pedersen LG. 1993. *J. Chem. Phys.* 99:8345–51
126. Del Buono G, Figueirido F, Levy RM. 1996. *Chem. Phys. Lett.* 263:521–29
127. Stillinger F, Rahman A. 1974. *J. Chem. Phys.* 60:1545–52
128. Pangali C, Rao M, Berne BJ. 1980. *Mol. Phys.* 40:661–70
129. Brooks CL. 1987. *J. Chem. Phys.* 86: 5156–62
130. Andrea T, Swope A, Andersen HC. 1983. *J. Chem. Phys.* 79:4576–84
131. Alper HE, Levy RM. 1989. *J. Chem. Phys.* 91:1242–51
132. Neumann M. 1985. *J. Chem. Phys.* 82: 5663–70
133. Bader JS, Chandler D. 1992. *J. Phys. Chem.* 96:6423–29
134. Houstin SE, Rossky PJ. 1989. *J. Phys. Chem.* 93:7888–93
135. Buckner JK, Jorgensen WL. 1989. *J. Am. Chem. Soc.* 111:2507–10
136. Dang LX, Pettitt BM. 1990. *J. Phys. Chem.* 94:4303–11
137. Guardia E, Rey R, Padro JA. 1991. *J. Chem. Phys.* 95:2823–28
138. Hummer G, Soumpasis DM, Neumann M. 1993. *Mol. Phys.* 81:1155–62
139. De Leeuw SW, Perram JW, Smith ER. 1980. *Proc. R. Soc. London Ser. A* 373: 27–26
140. Figueirido F, Del Buono G, Levy RM. 1995. *J. Chem. Phys.* 103:6133–42
141. Cichocki B, Felderhof B, Hinsen K. 1989. *Phys. Rev. A* 39:5350–56
142. Hummer G, Pratt LR, Garcia AE. 1996. *J. Phys. Chem.* 100:1206–15
143. Figueirido F, Del Buono G, Levy RM. 1997. *J. Phys. Chem. B* 101:5622–23
144. Hummer G, Pratt LR, Garcia AE. 1997. *J. Chem. Phys.* 107:9275–77
145. Bogusz S, Cheatham TE, Brooks BR. 1998. *J. Chem. Phys.* In press
146. Darden TA, York DM, Pedersen LG. 1993. *J. Chem. Phys.* 98:10089–92
147. Hockney RW, Eastwood JW. 1989. *Computer Simulations Using Particles*. Bristol, UK: Adam Hilger
148. Greengard L, Rokhlin V. 1989. *Phys. A* 29:139–52
149. Board JA, Causey JW, Leathrum JF, Windemuth A, Schulten K. 1992. *Chem. Phys. Lett.* 198:89–97
150. Schmidt KA, Lee MA. 1991. *J. Stat. Phys.* 63:1223–31
151. Ding HQ, Karasawa N, Goddard WA. 1992. *J. Chem. Phys.* 97:4309–18
152. Saito M. 1992. *Mol. Simul.* 8:321–30
153. Lee FS, Warshel A. 1992. *J. Chem. Phys.* 97:3100–9
154. Figueirido F, Levy RM, Zhou R, Berne BJ. 1997. *J. Chem. Phys.* 106:9835–49
155. Schlick T, Poarth E, Mandziuk M. 1996. *Annu. Rev. Biophys. Biomol. Struct.* 26: 181–22
156. Tuckerman AM, Berne BJ, Martyna GJ. 1992. *J. Chem. Phys.* 97:1990–98
157. Bishop TC, Skeel RD, Schulten K. 1997. *J. Comp. Chem.* 18:1785–91
158. Barker JA, Watts RO. 1973. *Mol. Phys.* 26:789–96
159. Alper H, Levy RM. 1993. *J. Chem. Phys.* 99:9847–52
160. Tanford C, Kirkwood JG. 1957. *J. Am. Chem. Soc.* 79:5333–37
161. Beglov D, Roux B. 1994. *J. Chem. Phys.* 100:9050–62
162. Wallqvist A. 1993. *Mol. Simul.* 10:13–15
163. Wang L, Hermans J. 1995. *J. Phys. Chem.* 99:12001–7
164. Friedman HL. 1975. *Mol. Phys.* 29: 1533–38
165. Jayram B, Fine R, Sharp K, Honig B. 1989. *J. Phys. Chem.* 93:4320–26
166. Figueirido F, Del Buono G, Levy RM. 1994. *Biophys. Chem.* 51:235–41
167. Smith PE, van Gunsteren WF. 1994. *J. Chem. Phys.* 100:577–85
167a. Hyun JK, Babu C, Ichiye T. 1995. *J. Phys. Chem.* 99:5187–95
167b. Hyun JK, Ichiye T. 1995. *J. Phys. Chem.* 101:3596–604
168. Liu H, Mark AE, van Gunsteren WF. 1996. *J. Phys. Chem.* 100:9485–94
169. Hummer G, Szabo A. 1996. *J. Chem. Phys.* 105:2004–10
170. Aqvist J, Medina C, Samuelsson JE. 1994. *Protein Eng.* 7:385–91
171. Jorgensen WL. 1995. *J. Phys. Chem.* 99: 10667–73
172. Chandler D. 1993. *Phys. Rev. E* 48: 2898–905
173. Hummer G, Garde S, Garcia AE, Pohorille A, Pratt LR. 1996. *Proc. Natl. Acad. Sci. USA* 93:8951–55
174. Simonsen T, Perahia D. 1995. *J. Am. Chem. Soc.* 117:7987–8000
175. Simonsen T, Brooks CL. 1996. *J. Am. Chem. Soc.* 118:8452–58
176. Yang L, Valdeavella C, Pettitt BM. 1996. *Biophys. J.* 71:3022–29
177. Merz KM, Banci L. 1996. *J. Phys. Chem.* 100:17414–20
178. Treutlein H, Schulten K, Brünger AT, Karplus M, Deisenhofer J, Michel H.

1992. *Proc. Natl. Acad. Sci. USA* 89:75–79
179. Gehlen JN, Marchi M, Chandler D. 1994. *Science* 263:499
180. Del Buono G, Figueirido F, Levy RM. 1994. *Proteins Struct. Funct. Genet.* 20: 85–87
181. Sham Y, Chu ZT, Warshel A. 1997. *J. Phys. Chem. B* 101:4458–72
182. Muegge I, Qi AJ, Wand P, Chu Z, Warshel A. 1997. *J. Phys. Chem.* 101: 825–36
183. Bashford D, Karplus M. 1990. *Biochemistry* 29:10219–25
184. Froloff N, Windemuth A, Honig B. 1997. *Protein Sci.* 6:1293–301
185. Deleted in proof
186. You TJ, Bashford D. 1995. *Biophys. J.* 69:1721–33
187. Yang AS, Gunner M, Sampogna R, Sharp K, Honig B. 1993. *Proteins* 15: 252–65
188. Bashford D, Gerwert K. 1992. *J. Mol. Biol.* 224:473–86
189. Antosiewicz J, McCammon JA, Gilson MK. 1996. *Biochemistry* 35:7819–33
190. Demchuk E, Wade R. 1996. *J. Phys. Chem.* 100:17373–87
191. King G, Warshel A. 1990. *J. Chem. Phys.* 93:8682–92
192. Carter EA, Hynes JT. 1989. *J. Phys. Chem.* 93:2184–87
193. Yelle RB, Ichiye T. 1997. *J. Phys. Chem. B* 101:4127–35
194. Marcus RA. 1960. *Discuss. Faraday Soc.* 29:21–56
195. Zhou HX, Szabo A. 1995. *J. Chem. Phys.* 103:3481–94
196. Bottcher CJF. 1973. *Theory of Electric Polarization*. Amsterdam: Elsevier
197. Smith PE, Brunne R, Mark AE, van Gunsteren WF. 1993. *J. Phys. Chem.* 97:2009–14
198. Loftler G, Schreiber H, Steinhauser O. 1997. *J. Mol. Biol.* 270:520–34
199. Sham Y, Muegge I, Warshel A. 1998. *Biophys. J.* In press