First-Shell Solvation of Ion Pairs: Correction of Systematic Errors in Implicit Solvent Models†

Zhiyun Yu,‡ Matthew P. Jacobson,§ Julia Josovitz,⊥ Chaya S. Rapp,⊥ and Richard A. Friesner*,‡

Department of Chemistry and Center for Biomolecular Simulation, Columbia University, New York, New York 10027, Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143, and Department of Chemistry, Stern College for Women, Yeshiva University, New York, New York 10016

Received: December 11, 2003; In Final Form: February 10, 2004

Salt bridges play an important role in protein stability, protein−protein interactions, and protein folding. The electrostatic solvation free energies of the analogues of charged amino acid side chains were calculated using both explicit solvent free energy perturbation methods and implicit solvation models such as Poisson Boltzmann and surface-generalized Born model. A systematic difference between explicit and implicit solvent results was observed, which we attribute to a specific first-shell solvation effect, which we refer to as bridging waters. We present a method for including a single explicit bridging water between the pairs that improves the implicit solvation models significantly.

I. Introduction

The development of accurate implicit models of aqueous solvation is crucial to the practical application of computational chemistry methods to modeling biological systems. The inclusion of explicit waters in simulations substantially increases the computational expense, and also requires extensive averaging over the positions of the water molecules to converge the solvation free energy of a macromolecular system. Furthermore, it is difficult to execute rapid, large-scale configurational changes of the solute, which are required for conformational search techniques such as docking of a ligand into a protein receptor, prediction of loops and side chains in protein structures, and ab initio protein folding. In an explicit solvation model, such changes would necessitate displacement of a large number of explicit waters and hence would have a low acceptance probability from an energetic point of view. In contrast, it is straightforward to make such large displacements using continuum solvent models, particularly if analytical gradients are available so that the structure can subsequently be optimized.

The construction of an accurate and computationally efficient continuum solvent model requires extensive parameter optimization and careful consideration of the appropriate functional form. The Poisson−Boltzmann (PB) equation provides a reasonable starting point for such developments; in particular, the treatment of water as a medium composed of infinitesimal dipoles, with dielectric constant 80, is rigorously correct at long distances. However, in the first solvation shell, this description is not accurate. Even from a purely electrostatic point of view, the finite size of actual water molecules becomes important at short range, and a purely electrostatic description of hydrogen bonding is often inadequate. A straightforward approach to this problem is to parametrize dielectric radii, and surface area corrections, for various chemical functional groups, fitting the parameters to experimental solvation free energies of small molecules. With an appropriate treatment of the nonpolar van der Waals and cavity terms, this approach is capable of achieving accuracies of ∼0.25−0.5 kcal/mol for a small molecule database of several hundred compounds.[1−3]

The key question, however, is how transferable these parameters are to larger, more complex structures in which the individual functional groups interact, e.g., salt bridges, which play an important role in protein stability and protein−protein interactions.[4−11] The solvation shell around a salt bridge, or for that matter any solvent exposed hydrogen bond, may differ qualitatively from that observed for the individual functional groups, and this in turn may lead to problems in transferring parameters obtained from small molecules to macromolecules. More generally, it is not yet clear how well continuum methods treat water molecules in confined spaces, as are observed routinely in the active sites of proteins and protein−ligand complexes. These questions are critical to protein structure prediction and structure based drug design applications, and go beyond arguments regarding the numerical accuracy of solving the PB equation (which is often taken, incorrectly, as a measure of the validity of a particular continuum treatment). If the errors inherent in the continuum methodology itself, as compared to experiment, are substantially larger than the difference between the various computational approximations to the PB equation, the use of precise (from a numerical standpoint) PB benchmarks to validate such approximations is not particularly meaningful.

There are two principal approaches to investigating the transferability of a continuum solvation parametrization for large, complex solutes. The first is direct comparison with experimental data for such systems. We have pursued this strategy extensively, via prediction of protein side chain and loop structures available in the protein data bank (PDB). Single side chain prediction (SSP) (leaving the remainder of the protein structure in its experimental conformation) is a particularly

† Part of the special issue "Hans C. Andersen Festschrift".
‡ Columbia University.
§ University of California, San Francisco.
⊥ Yeshiva University.

10.1021/jp037821l CCC: $27.50
Published on Web 05/05/2004

effective approach to uncovering major problems with a continuum model, because it is easy to perform accurate sampling for one side chain. We have used SSP to optimize our surface-generalized Born (SGB) continuum model, in particular developing empirical corrections for the treatment of surface salt bridges, which are overpredicted unless such correction terms are introduced. Loop prediction, a more challenging sampling problem, can be used to test the optimized continuum model. Using a novel buildup procedure, augmented by hierarchical screening and clustering approaches, we have recently carried out predictions for 600 loops of lengths 5−12, achieving substantial improvements in RMSD as compared with previous work in the literature (e.g., the median RMSD for 8 residue loops is 0.44 Å). While these results are encouraging, there remain a small but nontrivial number of cases where qualitative errors can be observed, many of which can be attributed to the solvation model. Furthermore, these tests measure the ability to predict structures as opposed to quantitating free energy differences between conformations; the latter is a more demanding task, but is required in many important biological applications (e.g., calculation of protein−ligand binding affinities).

The second approach is comparison with potentials of mean force (PMF)[12−14] obtained from explicit solvent simulations.[15−18] In principle, given a quantitatively accurate force field and explicit solvent simulation protocol, such comparisons would allow one to robustly interrogate all aspects of a continuum model, at a level of detail not possible in the experimental comparisons. In practice, it is far from clear how accurate a typical explicit solvent calculation actually is. The force field is imperfect, polarizability is rarely treated explicitly (possibly leading to qualitative as well as quantitative errors), sampling can be difficult to converge, and treatment of boundary conditions for systems containing charged groups requires careful consideration. Hence, at present, such comparisons must be used judiciously, in conjunction with experimental benchmarks, to investigate the functional form of the continuum model. In our view, detailed parametrization should be accomplished by direct fitting to experimental data.

Nevertheless, valuable insights can be obtained by comparing explicit and implicit solvent simulations, and the results can be a useful adjunct in parameter fitting. Levy and co-workers have carried out charging free energy perturbation (FEP) calculations for large systems (an octapeptide in free solution and bound to a receptor protein), and compared the results with SGB and PB calculations.[2] A simple uniform scaling of the dielectric radii relative to the molecular mechanics van der Waals values (fit to free energy calculations on small molecule analogues) was sufficient to yield excellent agreement between the SGB and FEP calculations for a substantial ensemble of such comparisons.[2] The Levy group then supplemented this parametrization of dielectric radii with a surface area and van der Waals parametrization in order to accurately fit experimental data.[3] This model, which we have used in all of our loop and side chain predictions, is clearly, as a result of the empirical parametrization, reasonably close to reproducing physically observed free energies.

In the present paper, we focus on a particular type of structural motif−explicit water molecules that simultaneously hydrogen bond to both charged groups in a salt bridge ("bridging waters")—which we will argue is not treated accurately by current continuum solvation models, regardless of the type of approximation to the PB equation that is employed. Previous work has demonstrated that inclusion of structural waters in SSP tests yields substantially greater accuracy in side chain RMSD, in many cases because of the inclusion of bridging waters of this type.[19] Here, we carry out explicit solvent FEP calculations on several small molecule dimers composed of charged amino acid side chain analogues, and investigate the dependence of the solvation free energy of these molecular pairs upon the intermolecular distance. The results are compared with several different continuum calculations of both the GB and PB variety. Despite differences in the details of the various continuum models, the qualitative comparison to the FEP results is remarkably similar. All of the methods make a serious error in relative free energy at the geometry in which a bridging water can hydrogen bond to both groups in the pair. The importance of the bridging water structure in the explicit solvent simulations is confirmed by examination of the occupation probability of the bridging position in these calculations.

Finally, based on these results, we present a heuristic attempt to improve continuum calculations via introduction of a *single* explicit water molecule, employing a well-defined protocol motivated by the explicit solvent simulation results and physical intuition concerning the behavior of the solvent at various intermolecular separations. Using a uniform protocol, and without employing any adjustable parameters, a dramatically improved prediction (as compared to the explicit solvent simulations) of the electrostatic solvation free energy for two different ion pairs is obtained. The approach needs to be tested on a much wider range of examples and will likely benefit from a formal statistical mechanical analysis (which we do not provide in the present paper); nevertheless, the results obtained are encouraging with regard to both accuracy and practicality in treating the larger and more complex systems that are of actual biological interest.

A recent paper by Masunov and Lazaridis also considered small molecule analogues of solvated salt bridges, and compared explicit and implicit solvent simulations.[20] This work differs both in technical aspects and motivation. Technical differences include the following: (1) we have employed fully explicit solvent simulations with periodic boundary conditions and Ewald summation, as opposed to the solvent boundary potential method; (2) we have used FEP, as opposed to umbrella sampling to calculate effective interactions between the charged groups in implicit solvent (i.e., we focus on the electrostatic aspects of solvation and specifically exclude nonpolar solvation free energy); (3) we have used the OPLS-AA force field, as opposed to CHARMM 19 in the Masunov and Lazaridis work. There are other minor differences in the systems considered, e.g., the geometries of approach of the ions. Our new contributions include comparison of the explicit solvent results with accurate solution of the Poisson−Boltzmann equation, using two different programs (PB methods were not considered in the work of Masunov and Lazaridis). We demonstrate that this important class of implicit solvent methodology also shows serious discrepancies from explicit solvent results. However, our most important new contributions are (1) the detailed analysis of the origin of the failures of the implicit solvent models and (2) the proposal of a heuristic, but practically useful, method of improving implicit solvent treatment of salt bridges by including *single* water molecules.

The paper is organized as follows. We begin in section II by discussing computational methods for explicit solvent, PB, and SGB calculations. We examine two rather different PB methods, which employ different algorithms, definitions of the dielectric boundary, and dielectric radii. Comparison of the explicit solvent results with the PB and SGB calculations is presented in section

III. In section IV, we introduce a hybrid approach, incorporating a single explicit water molecule into the continuum calculations, and compare with the explicit and implicit solvation results presented in section III. In section V, we discuss the implications of the results, along with various possible approaches to improving continuum methods (including the path of incorporating explicit waters that we have started to investigate here). The Discussion also presents an overview of bridging water structures located in the protein data bank (PDB), demonstrating the relevance of this type of structure to biomolecular modeling. Section VI, the Conclusion, summarizes our findings and briefly presents future directions.

II. Computational Methods

A. Implicit Solvation Models. A.1. Poisson-Boltzmann Methods. The Poisson−Boltzmann (PB) equation is one of the most commonly used implicit solvation models, and is based on a classical electrostatic formulation in which the aqueous solvent is treated as a dielectric continuum. In the PB approach, one solves eq 1 below to find the resulting electrostatic potential, $\phi(\vec{r})$, at point $(\vec{r})$ from the charge density, $\rho(\vec{r})$, and the dielectric constant, $\epsilon(\vec{r})$, of the system:

$$\nabla[\epsilon(\vec{r})\nabla\phi(\vec{r})] = -4\pi\rho(\vec{r}) \tag{1}$$

A PB model is specified by the charge distribution (arising from the molecular mechanics potential function of the solute) and the functional form of the dielectric of the system. Typically this is defined as a dielectric surface; all points in the system falling inside the surface (i.e., in the interior of the protein) are defined to have dielectric of unity (or whatever the internal dielectric is specified as−this itself is a subject of considerable controversy), whereas points falling outside of the surface are defined to have dielectric 80 (given that we are modeling aqueous solution).

To solve the PB equation, we use two different programs:

(1) The first is the DelPhi program of Honig and coworkers.[21−24] DelPhi uses a finite difference approach to solving the PB equation, with a number of specialized numerical techniques that ensure both accuracy and high efficiency. We have tested the grid resolution in Delphi for the present problem and obtained converged results using 0.5 Å/grid (that is, finer grids minimally change the results presented below). For maximum consistency, we have input the OPLS-AA charges and van der Waals radii into DelPhi.

(2) The second is the PBF program, which was developed in the Friesner group several years ago.[25,26] PBF employs numerical methodology based on tetrahedral finite-element meshes, rather than cubic ones, for solving the PB equations. For the same dielectric surface, assuming that the resolution of the numerical mesh is converged in both cases, this methodology should give the same answer as DelPhi. However, PBF in fact uses a Gaussian surface definition (primarily because this was required to implement an analytical gradient, a feature that is lacking in DelPhi) where as DelPhi employs the Connolly surface[27−30] as the dielectric surface. Thus, the results shown below provide an indication as to the importance of details of the definition of the dielectric surface for the ion pair potential of mean force.

A.2. Surface-Generalized Born Model. The SGB model is an approximation to the boundary element formulation of the PB equation, in which the polarization effects throughout the entire volume of the system can by exactly reproduced by an appropriate distribution of induced polarization charge at the surface of solute molecules. A derivation of this approximation has been presented previously[1] and was shown to be equivalent to the volume formulation of the GB method by Still and coworkers,[31] assuming equivalent dielectric surfaces.

The total electrostatic solvation free energy can be divided into two terms−a self-term and a pair term:

$$\Delta G_{\text{total}} = \Delta G_{\text{self}} + \Delta G_{\text{pair}} \tag{2}$$

The formulas to calculate the self-term and the pair term are given as

$$\Delta G_{\text{self}} = -\frac{q_k^2}{8\pi}(1 - 1/\epsilon)\int_s \frac{(\vec{R} - \vec{r}_k)\cdot\vec{n}(\vec{R})}{|\vec{R} - \vec{r}_k|^4}\,\mathrm{d}^2\vec{R}$$

$$= -\frac{1}{2}(1 - 1/\epsilon)\frac{q_k^2}{\alpha_k} \tag{3}$$

$$\Delta G_{\text{pair}} = -\frac{1}{2}(1 - 1/\epsilon)\frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_{ij}^2 \mathrm{e}^{-D}}} \tag{4}$$

where q_k is the charge on atom k, $\vec{r}_k$ is its coordinate, ϵ is the dielectric constant, $\vec{R}$ is the coordinate of a point on the dielectric boundary, $\vec{n}(\vec{R})$ is the normal at the point, α_k is the Born alpha radius for atom k, r_{ij} is the distance between a pair of atoms i and j

$$\alpha_{ij} = \sqrt{\alpha_i \alpha_j}$$

and

$$D = \frac{r_{ij}^2}{(2\alpha_{ij})^2}$$

The simplest dielectric surface (from the point of view of computational efficiency and programming effort) to employ in a GB methodology is the van der Waals surface. This is almost certainly a nonoptimal surface to employ, as it fails to exclude water molecules from small spaces into which they cannot fit, a flaw that is corrected by using the molecular surface, as in Delphi. The use of this surface certainly will yield poor comparisons of total energies as compared to PB calculations. However, it is far from clear what the quantitative consequences are in predicting relative energies of different structures as compared to experiment, which is the only meaningful benchmark, particularly if the model is extensively parametrized to experiment, as in our current SGB methodology.

Future work will report on the effects of using different surfaces from the point of view of reproducing experimental results. Here, we use the van der Waals surface, and compare the results with the explicit solvent simulations and PB calculations. If errors due to the SGB approximation and surface definition are qualitatively dominant over those due to the use of continuum dielectric methods, one would expect the SGB results to deviate substantially more from the explicit solvation results than either of the PB calculations. The results below examine this question and provide an answer, at least for the present problem of two interacting charged groups at various intermolecular separations.

B. Explicit Solvation Methods: Free Energy Perturbation. To obtain solvation free energies from explicit solvent simulations, we use the free energy perturbation technique.[15] The free

energy difference between two related systems/states is given by

$$\Delta G = -kT \ln\langle \exp(-[H_2 - H_1]/kT)\rangle_1 \quad (5)$$
$$= -kT \ln\langle \exp(-[V_2 - V_1]/kT)\rangle_1$$

where H_1 and H_2 are the Hamiltonians and V_1 and V_2 are the total potential energy of the two systems/states, respectively. Because the kinetic energy contributions to the two Hamiltonians are equal at each point in phase space, the difference of the Hamiltonians is replaced by the difference of the potential functions in the above equations. $\langle ... \rangle_1$ denotes an ensemble average corresponding to the Hamiltonian H_1. Equation 5 is exact, and called "perturbation" because, as a practical matter, it is directly useful for states 1 and 2 that are not too far apart. The difficulty arises because with finite computer time, a simulation of the ensemble corresponding to H_1 will predominantly sample microstates for which H_1 is small, which are not necessarily the same as those for which H_2 are small. Therefore, the evaluation of eq 5 is divided into multiple windows, each one involving a small enough interval characterized by parameter λ, to allow the free energy difference between the adjacent intermediate states to be calculated accurately. $\lambda = 0$ corresponds to state 1 and $\lambda = 1$ to state 2. ΔG is then evaluated as the sum of the free energy differences ΔG_i between the nearby windows corresponding to λ_{i-1} and λ_i:

$$\Delta G = \sum_{i=1}^{N} \Delta G_i$$
$$= \sum_{i=1}^{N} -kT \ln\langle \exp(-[V(\lambda_i) - V(\lambda_{i-1})]/kT)\rangle_{\lambda_{i-1}} \quad (6)$$

A useful trick in evaluating the free energy ΔG comes from noticing that either the state at λ_i or the state at λ_{i-1} may be used as the "initial" state in eq 6. If the states are reversed, the free energy change between them is just the negative of what it was, so

$$\Delta G_i = +kT \ln\langle \exp(-[V(\lambda_{i-1}) - V(\lambda_i)]/kT)\rangle_{\lambda i} \quad (7)$$

We also have

$$\Delta G_{i+1} = -kT \ln\langle \exp(-[V(\lambda_{i+1}) - V(\lambda_i)]/kT)\rangle_{\lambda i} \quad (8)$$

so we can use the ensemble at λ_i to evaluate the free energy differences between the state at λ_i and the states on either side of it. This "double-wide sampling" method allows increased accuracy at a given calculation speed, or increased speed at a given accuracy, since the free energy change from state $i - 1$ to state $i + 1$ is calculated in a single simulation using a state i that is closer to each of them, and thus better for sampling them, than they are to each other.

In our FEP calculation, the whole process is divided into 10 windows. In the initial state, all atoms of the solute are uncharged, and in the final state all atoms of the solute are fully charged. At each window they are charged to one tenth more of their full charges, respectively, and inserted in a box of dimensions $24.862 \times 24.862 \times 24.862$ Å^3 containing 525 SPC water molecules. Then the system is equilibrated to 298.15 K in several stages: 1 ps each at 50, 100, 150, 200, and 250 K, and finally 3.4 ps equilibration at 298.15 K. Data collection over 2.4 ps follows equilibration. Electrostatic intermolecular interactions are evaluated using the Ewald formula, and Lennard-Jones (LJ) intermolecular dispersion interactions are truncated at 8.0 Å, according to the atomic positions of the solute and according to the water center of mass for the solvent. The calculations at each window are performed in parallel on several processors. This reduces the required wall clock time to one-tenth of what it would be.

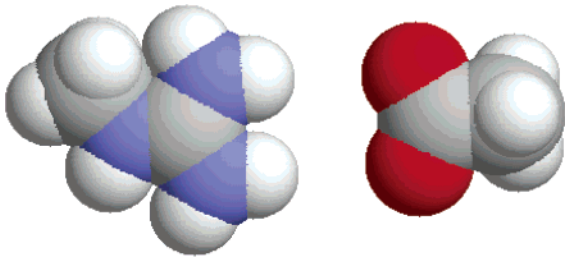

Figure 1. Acetate−guanidinium ion pair.

C. Force Field and Continuum Solvent Parameters. We use the OPLS-AA force field for the explicit solvent FEP simulations. Parameters for the various species considered below are listed in the Supporting Information. Our belief is that the key qualitative features of the results are not critically dependent upon the details of the parametrization (assuming that this is reasonable from a physical point of view), however this remains to be examined in a subsequent publication. We use the SPC water model; again, it is unlikely that the results would be qualitatively affected by a different choice, but this should be checked in a detailed calculation.

The charges for the SGB, PBF, and DelPhi calculations are identical to the OPLS-AA charges used in the explicit solvent simulations. The DelPhi PB calculations use the molecular surface constructed from the OPLS-AA van der Waals radii to define the dielectric boundary. In contrast, the SGB model uses modified dielectric radii for each atom (typically a simple 10% scaling of the OPLS-AA van der Waals radii), and these are given in the Supporting Information. These parameters have been optimized by Levy and co-workers to reproduce electrostatic charging free energies obtained from explicit solvent simulations for small molecules.[3]

Thus, the charges are the same in all three calculations, but the surface definitions are significantly different for the two PB methods and for the SGB calculations, as are various algorithmic approximations employed. As discussed above, the results below will provide some calibration as to how important these differences are as compared to the basic fact of using a continuum approach.

III. Results for Explicit Solvent, PB, and SGB Simulations

The first system we study is the acetate ion−guanidinium pair, where the two carbon atoms of acetate and the carbon and nitrogen atoms of guanidinium are aligned in straight line, as shown in Figure 1. This system is a simple model for the interaction of arginine with aspartic or glutamic acid residues in a protein, which can form a surface-exposed salt bridge in this fashion (although the geometrical constraints can be much more complicated, of course).

Using the PB, SGB, and FEP methods described above, we have calculated the solvation free energy of this pair of solutes as a function of their intermolecular separation, in terms of N−O distance, enforcing the angular geometry described above for each distance. Table 1 presents the values of the electrostatic charging free energy for each method at the various distances. The absolute value of the charging free energies at large

TABLE 1: Electrostatic Solvation Free Energies of the Acetate−Guanidinium Ion Pair at Various N−O Distances, As Calculated by DELPHI, PBF, SGB, and Explicit Solvent FEP

N-O dist (Å)	DELPHI (kcal/mol)	PBF (kcal/mol)	SGB (kcal/mol)	FEP (kcal/mol)
3.159	−60.59	−61.03	−53.74	−56.1
3.559	−69	−69.95	−61.42	−63.05
3.959	−78.93	−78.46	−69.54	−72.85
4.359	−86.46	−84.11	−76.82	−81.58
4.759	−92.72	−92.51	−83.06	−89.99
5.159	−98.76	−99.12	−88.36	−97.13
5.559	−103.9	−102.72	−92.86	−99.38
5.959	−107.79	−107.22	−96.69	−100.66
6.359	−111.43	−110.21	−100	−102.52
6.759	−114.44	−112.69	−102.91	−107.2
7.159	−116.06	−114.86	−105.47	−109.19
7.559	−118.34	−117.27	−107.74	−112.11

separation (which in essence corresponds to solvation of the individual small molecules) differs across the various methods. These differences are highly dependent upon parametrization, and it is not necessarily clear that the explicit solvent values are correct as compared to experiment. Therefore, we set the charging free energy at the largest separation to zero for all calculations, and plot the results with these shifts incorporated in Figure 2. Finally, we subtract the FEP charging free energy from that of the remaining calculations (using the shifted values in all cases) and plot these differences in Figure 3.

The results for the three implicit solvation models display a striking qualitative similarity. The charging free energy is more positive (i.e., less favorable) for conformations where e.g., in the SGB model, the N−O distance lies in the range 4.3−5.9 Å (the corresponding H−O distance between the charged H atom of guanidinium and the charged O atom of acetate ion ranges from 3.3 to 4.9 Å). The divergence from the explicit solvent results peaks at the N−O distance of about 5.2 Å (the corresponding H−O distance is 4.2 Å), which corresponds to the solvent-separated minimum, and the magnitude of the peak divergence is approximately 4.5 kcal/mol. These results agree qualitatively with recent results from Lazaridis,[20] where the divergence of the potential of mean force of implicit solvation models from explicit solvation model peaks at an N−O distance of 5.2 Å with a magnitude of about 2.3 kcal/mol. These positive divergences indicate an underestimation of the solvation free energy by the implicit models at the corresponding N−O distances. As we argue below, this is precisely the distance at

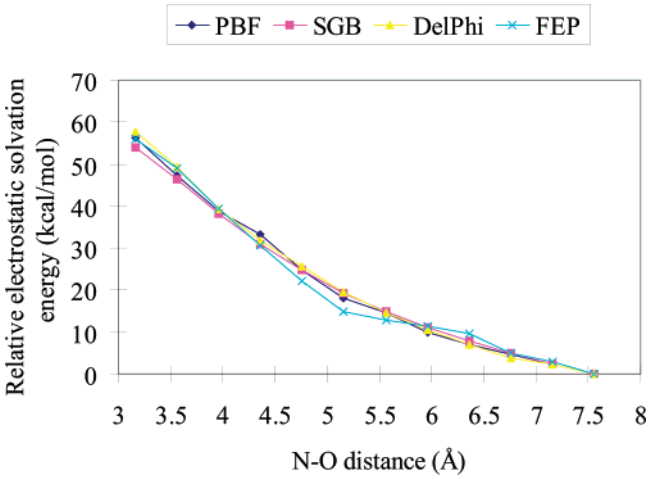

Figure 2. Electrostatic solvation free energies of the acetate−guanidinium ion pair at various N−O distances after setting the one at the largest separation to zero by DELPHI, PBF, SGB, and FEP.

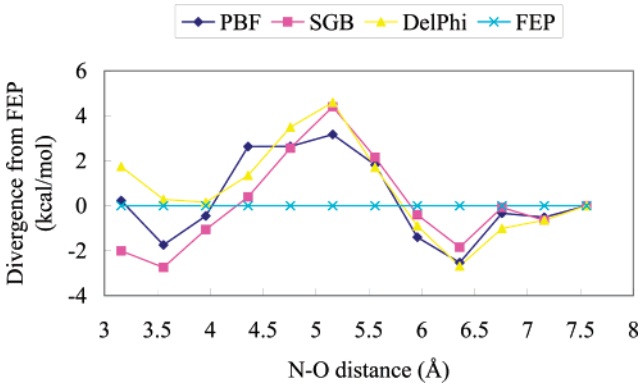

Figure 3. Comparison of the difference of the electrostatic solvation free energies of acetate−guanidinium ion pair at various N−O distances to that at the most extended distance, by DELPHI, PBF, SGB, and FEP. The distance 3.159 Å corresponds to contact ion pair while 5.159 Å, where the peak divergence appears, corresponds to solvent-separated pair.

TABLE 2: MD Evidence for Bridging Water in an Acetate−Guanidinium Ion Pair[a]

	no. of snapshots				
N−O dist (Å)	with no bridging water	with 1 bridging water	with 2 bridging waters	with 3 bridging waters	av no. of bridging waters
4.359	782	381	36	1	0.38
4.759	245	566	389	0	1.12
5.159	33	267	900	0	1.73
5.559	366	630	204	0	0.865

[a] Each snapshot corresponds to 0.25 ps. We define a bridging water as a water that hydrogen bonds simultaneously to both charged groups (the criterion for hydrogen bond length is defined as 2.0 Å or less).

which a structure with a single bridging water between the two ions is most likely to be formed. Interestingly, all three implicit solvent calculations display reasonable qualitative, although not quantitatively accurate, agreement with the FEP results as the ions come together to form a salt bridge.

We have examined the MD simulations for the configurations that demonstrate the greatest divergence between the implicit and explicit solvent results. These data clearly demonstrate that, most of the time, one or more water molecules is "trapped" near the pair because of strong hydrogen bonds formed with both ions simultaneously, as shown in Table 2, and the animation provided in the Supporting Information. We emphasize that the water is not physically trapped, but rather strongly restrained in this "bridging" conformation by highly favorable electrostatic forces. This water molecule, which we refer to as a "bridging" water, should be treated explicitly since it has extremely strong interactions with the salt bridge and thus differentiates itself from other water molecules in the bulk water. However, in implicit solvation models, this bridging water is unrecognized and treated uniformly with all other water molecules. As discussed further in section IV, it is this deficiency that leads to the underestimation of the electrostatic solvation free energies of the pair at H−O distance ranging from 3.3−4.9 Å.

The second system we study is acetate paired with methylammonium. Once again, the two carbon atoms of acetate ion and the nitrogen and carbon atoms of methylammonium are aligned in straight line, as shown in Figure 4. In this system, we replace the small molecule analogue of arginine with that for lysine. The results, in the same format as discussed above, are shown in Table 3 and Figures 5 and 6. In all implicit solvation models, a positive divergence from the explicit solvent

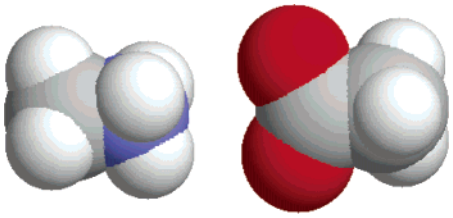

Figure 4. Acetate−methylammonium ion pair.

TABLE 3: Electrostatic Solvation Free Energies of the Acetate−Methylammonium Ion Pair at Various N−O Distances, As Calculated by DELPHI, PBF, SGB, and Explicit Solvent FEP

N−O dist (Å)	DELPHI (kcal/mol)	PBF (kcal/mol)	SGB (kcal/mol)	FEP (kcal/mol)
3.018	−59.14	−60.25	−48.63	−52.65
3.394	−69.33	−71.29	−56.91	−64.02
3.775	−78.78	−79.6	−64.96	−73.92
4.159	−87.83	−86.62	−74.3	−82.98
4.546	−95.06	−96.04	−81.92	−90.5
4.935	−102.3	−101.98	−88.25	−95.63
5.326	−108.12	−108.33	−93.55	−99.89
5.914	−114.97	−115.35	−100.09	−106.19
6.308	−118.37	−118	−103.77	−109.68
6.702	−121.21	−121.48	−107.11	−112.67
7.098	−124.32	−124.28	−109.75	−115.27
7.492	−126.41	−126.56	−112.24	−118.34

FEP results appears when the N−O distance ranges, e.g., in the SGB model, from 3.3 to 5.3 Å, with the peak divergence of 2.8 kcal/mol occurring at a distance of 3.8 Å corresponding to the solvent-separated minimum. In this case, the region over which strong deviations are observed, in terms of N−O distance, is broader than it is for the previous example (4.3−5.9 Å). We argue that this is because there are six possible H−O pairs, where H is one of three protons on the ammonium group, and O is one of the two oxygens on acetate. That is, in the acetate−ammonium system, the H−O pair forming hydrogen bonds with the bridging water could come from either of the three positive charged H atoms of the ammonium plus either of the two negative charged O atoms of the acetate ion. All six of the possible H−O atom pairs can form hydrogen bonds with a bridging water molecule, at various internuclear distances. In contrast, the geometry of the guanidinium-acetate system dictates that there are only two (identical) H−O pairs, with the consequence that the "bridging water" effect is more narrowly peaked. Lazaridis also observed a smaller peak divergence and a broader bridging−water region in their similar system.[20] Again, for the acetate−ammonium system, the MD demonstrates the existence of bridging waters, as shown in Table 4.

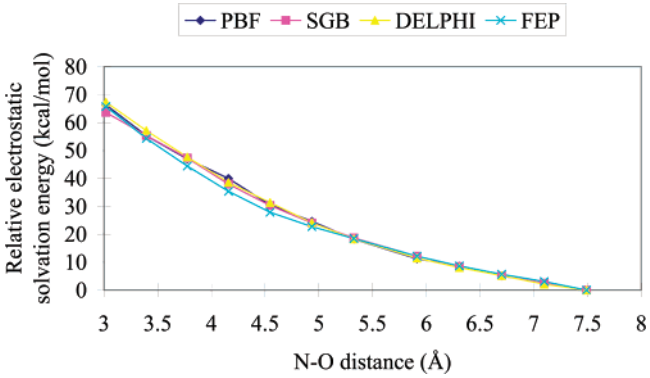

Figure 5. Electrostatic solvation free energies of the acetate−methylammonium ion pair at various N−O distances after setting the one at the largest separation to zero by DelPhi, PBF, SGB, and FEP.

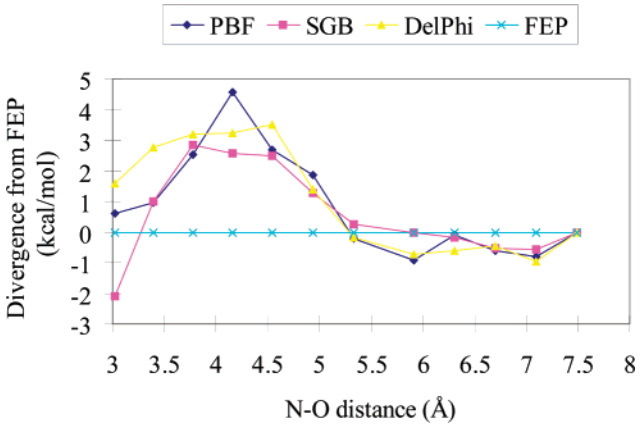

Figure 6. Comparison of the difference of the electrostatic solvation free energies of acetate−methylammonium ion pair at various N−O distances to that at the most extended distance, by DelPhi, PBF, SGB, and FEP. The distance 3.018 Å corresponds to contact ion pair while 3.775 or 4.159 Å, where the peak divergence appears in different models, corresponds to solvent-separated pair.

TABLE 4: MD Evidence for Bridging Waters in an Acetate−Ammonium Ion Pair[a]

	no. of snapshots				
N−O dist (Å)	with no bridging water	with 1 bridging water	with 2 bridging waters	with 3 bridging waters	av no. of bridging waters
3.394	188	477	450	85	1.36
3.775	140	473	470	117	1.47
4.159	254	556	355	35	1.14
4.546	207	599	376	18	1.17
4.935	618	491	91	0	0.56

[a] Each snapshot corresponds to 0.25 ps. We define a bridging water as a water that hydrogen bonds simultaneously to both charged groups (the criterion for hydrogen bond length is defined as 2.0 Å or less).

A key conclusion arising from these results is that the differences among the various implicit solvent models are much smaller than the difference of any of the models as compared to the explicit solvent simulations, despite the use of a different dielectric surface in all three cases, and the different mathematical representation of the reaction field employed in the SGB, as opposed to PB, methodology. There is no evidence that the use of the full PB equation, or the more elaborate Connolly or Gaussian dielectric surface, provides systematically better agreement with explicit solvent results. There are significant differences in behavior in the salt bridge (short distance) limit, but these appear, at least with all three of the present models, to randomly differ by roughly similar amounts from the explicit solvent results. Of course, these conclusions may not hold for more complex structures such as those found at the surface of a protein; however, the charging free energy results of Levy and co-workers, which compared SGB and PBF results with explicit solvent simulations for various peptides and peptide−protein complexes, yielded similar conclusions.[2]

The second key finding is that all of the implicit solvation models differ, systematically and reproducibly, from the explicit solvent simulations, in what appears to be a physically understandable fashion, which we interpret qualitatively based on the presence of quasi-bound (bridging) water molecules. The 3−4 kcal/mol discrepancies demonstrated here even for high resolution, converged PB calculations are highly nontrivial, e.g., in the context of protein structure prediction or calculations of protein−ligand binding affinities, and addressing them would appear to be a first order of business whether one chooses to use PB or GB methods for continuum electrostatics. Errors of

this type may be even larger, and more prevalent, in the confined spaces that characterize the great majority of protein active sites. Below, we present an initial attempt to tackle this problem in a computationally cost-effective fashion, based on inclusion of discrete bridging water molecules in a continuum methodology. In the present paper, we include at most one discrete water; more complex systems will require a generalization of this idea, briefly considered in the conclusion.

IV. Incorporation of a Single Explicit Water into a Continuum Solvation Calculation

At distances where significant deviations are observed between explicit and implicit solvent free energies, ion pairs manifest, on average, roughly one bridging water in the explicit solvent simulations, as shown in Tables 2 and 4. This observation suggests that the continuum solvent calculations are likely to be improved by incorporation of a very small number of explicit water molecules. The question of how to incorporate explicit waters in a continuum solvent model in a rigorous fashion is a fundamental problem in formal statistical mechanical theory; some progress in this direction has been made, for example by Pratt and co-workers via their quasichemical approximation.[32,33] However, the solutes they have examined to date are considerably less complicated than those that we have studied here (primarily single ions). Application of quasichemical theory to the present system would be of great interest. Here we adopt a considerably simpler approach. Although heuristic in nature, it does have the advantage of being straightforward to extend to much more complicated systems. The fact that encouraging results are obtained for both ion pair systems, despite possible deficiencies in the formal justification of the proposed protocol, will hopefully encourage further, more sophisticated efforts along these lines.

Consider a system consisting of the two solute molecules which compose the ion pair, plus a single water molecule. We are interested not in the absolute solvation free energy of the ion pair, but rather in the relative probability of observing various configurations of the ion pair in solution (i.e., the potential of mean force). If the dielectric continuum calculations gave exact results for the solvation free energy at any configuration of the solute molecules, the straightforward continuum protocol would yield the correct potential of mean force. Furthermore, if the error in the continuum model was identical for each solute configuration of interest, the errors would cancel and the potential of mean force would again be correct. A problem arises when the continuum approximation exhibits differential errors for different solvent configurations. Here, our argument is that the errors are largest, for specific reasons discussed further below, in the region where a bridging water can form an optimal hydrogen bond between the charged atoms of the ion pair.

We hypothesize that, if a water is added in the bridging position, the continuum model provides an accurate treatment of the resulting solute−water complex. This idea is similar to the motivation behind the quasi-chemical approximation; only in the "inner sphere" are explicit waters required. For our purposes, a reasonable diagnostic for an "inner-sphere" water is one which is observed to be localized in a specific position and orientation a high fraction of the time during explicit solvent simulations (the question of whether one can identify such locations without carrying out a simulation is one that we will briefly consider below, but not pursue in any detail in the present paper).

The potential of mean force for any particular single water-solute complex is given by the electrostatic solvation free energy of that complex plus the Coulombic and LJ interaction energy of the components of the complex. If, for a given solute configuration (in this case, ion pair separation), a water molecule is *always* observed in a single configuration (which may be a good approximation for many bound waters in proteins), then the probability of observing the solute configuration itself is the same as the probability of observing the solute-single water configuration. If on the other hand the center of mass position and orientation of a water molecule in the volume of interest is observed to vary considerably for a particular solute configuration, one has to average over the relevant orientations and positions. One way to approach this averaging problem is to place a point at the center of the volume occupied by the original water molecule, and then define the average as over the water molecule in an arbitrary simulation configuration which is closest to the specified point.

The approach that we take in the present paper is to define, for each solute configuration, a single "bound" water molecule located at one position and orientation. The positions and orientations are generated by a set of simple rules, developed primarily from examination of the explicit solvent simulations, as follows:

(1) We assume that, at the conformation corresponding to peak divergence from explicit solvent results, the bridging water is trapped the most tightly (that is, the specific bridging geometry, in which hydrogen bonds are formed with both ions according to reasonable geometric criteria, has the highest occupation probability over the course of a simulation) and that it becomes less tight at other conformations where less divergence is observed. Therefore, we first insert the bridging water at the peak conformation. The MD simulation statistics are used as a reference to determine the length of the two hydrogen bonds the bridging water forms with the pair. The geometry of the water molecule is such that it is coplanar with the collinear axis of the system and the H−O pair of atoms from the ions that traps the water molecule. The H−O bond of the water is in parallel with the line connecting the H−O atoms in the ion pair, as shown in Figures 7 and 8. This geometry also applies to the explicit water molecule added in all other conformations. We note that for other types of ion pairs, the bridging geometry described above might not be optimal (i.e., have a particularly high occupation probability); the present results are specific to the two systems that we have chosen to study here.

(2) Then, we keep the bridging water fixed, and evenly pull the ion pair apart along the collinear axis until the various specified intermolecular distances are obtained. The bridging water is now less tightly trapped, forming longer hydrogen bonds with the pair. When the pair is pushed closer together from the peak distance, the water molecule is pushed farther away from the ion pair to avoid steric clash to the distance as if they are pulled apart from the peak distance by the same distance. The process is shown in Figures 7 and 8.

(3) Finally, for those conformations that are outside the bridging−water region, we also added an explicit water, which is placed at a location such that the distances between the water O and a positively charged H on the acceptor group, and the water H and a carboxylate oxygen, are both 3 Å, and the geometry is otherwise similar to that in the bridging region, as specified above in point 1. The value of 3 Å is estimated from statistics obtained from the molecular dynamics simulations in the large separation regions—on average, this is the distance at which the water in the relevant region is observed. The orientation of the water is identical to that in bridging region,

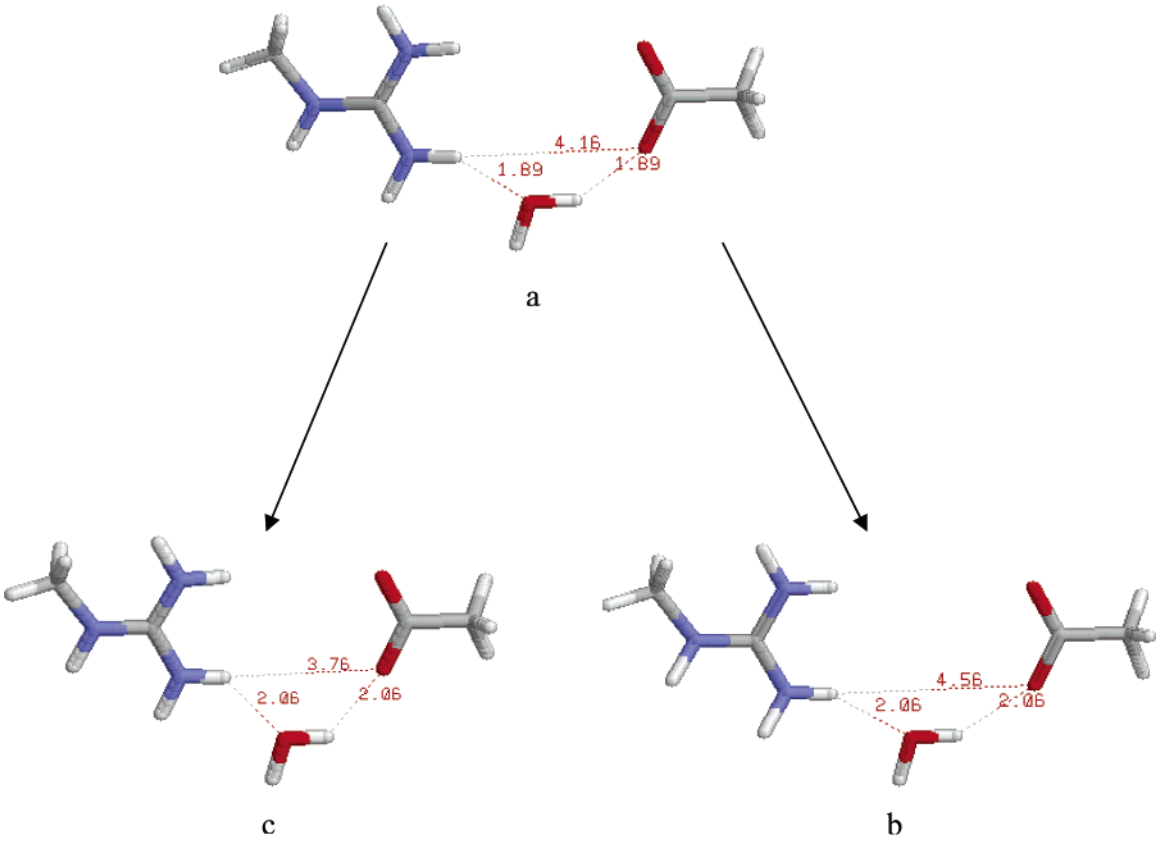

Figure 7. Geometry and orientation of the included explicit water molecule in the acetate−guanidinium ion system. The water molecule is coplanar with the collinear axis of the system and the charged H−O pair that traps the water molecule. The H−O bond of the water is in parallel with the line connecting the charged H−O pair. At O−H distances (i.e., between the ions) smaller than 4.16 Å, the water molecule must be "pushed back" from the ion pair to avoid steric clashes. At distances larger than 4.16 Å, the water molecule is held fixed as the ion pairs are symmetrically separated. The numbers in the figure denote the distances in Å between the two atoms connected by the broken lines. Note that the H-bond length between the water and the ion pair in c is the same as in b; this symmetry arises because in both cases the O−H distances between the ions differ from the reference state (4.16 Å) by 0.4 Å.

only the distance is increased, by ∼0.6 Å. If the distance is kept the same as that in the bridging region, the results shown below degrade by ∼1−2 kcal/mol in terms of agreement with the explicit solvent FEP calculations.

This protocol clearly contains substantial heuristic elements. First, a more rigorous formulation presumably would average over the zero, one, and two bridging water states; our approach will be effective if the correction terms are approximately additive. Second, in the less tightly bound bridging regions, there presumably will be more fluctuation in the water position, an effect that we neglect by employing one predominant conformation. Third, in the longer distance regime, the single water we have incorporated presumably has even greater conformational freedom (within the restrictions specified above); the treatment of this region (as compared to the bound water region) has to be considered as highly approximate. However, it is also possible that the effects we are neglecting are significantly smaller than the intrinsic errors in the continuum model, in which case one would expect some improvement in the results as compared with the fully continuum alternative; this is at present an empirical question, a formal analysis of the expected errors being difficult to formulate. As the protocol is identical for both ion pairs, there is at least some sort of consistency check available by examining the data for both systems.

The results are shown in Table 5 and Figure 9, for the guanidinium−carboxylate system, and in Table 6 and Figure 10 for the ammonium−carboxylate system. As with the previous data, the absolute potential of mean force of the single water−ion pair system in continuum solvent is set to zero at the longest distance considered. The difference in the potential of mean force at each separation with that obtained at the longest distance—taken to be proportional to the relative probability of observing the specified ion pair separation—is calculated and compared with same differential from the explicit solvent simulations. Under the assumptions discussed above, this represents a fair comparison.

The improvement in agreement with the explicit solvent data, as compared to the fully continuum calculations, is remarkable for both ion pair systems. Qualitatively better results are obtained not only in the optimal bridging water region but also in the contact region, where the explicit water presumably provides a more reliable description of the first-shell solvation of the ion pair. The errors are uniformly reduced to less than ∼1 kcal/mol, likely within the accuracy of the potential function itself. However, it should be emphasized that, due to the heuristic nature of the approximations as outlined above, a great deal more work, both in the theory and in computational tests for a wider range of systems, will be required before a robust, accurate protocol can be said to have been constructed. Despite these caveats, the results do suggest that it is worth putting substantial effort into developing and testing methods that combine continuum and explicit solvent approaches, even in the context of quite complex solutes, and that relatively simple approximations may work better than one might expect.

V. Discussion

It is clear from the above results that the differences between explicit solvent simulations and various continuun models are

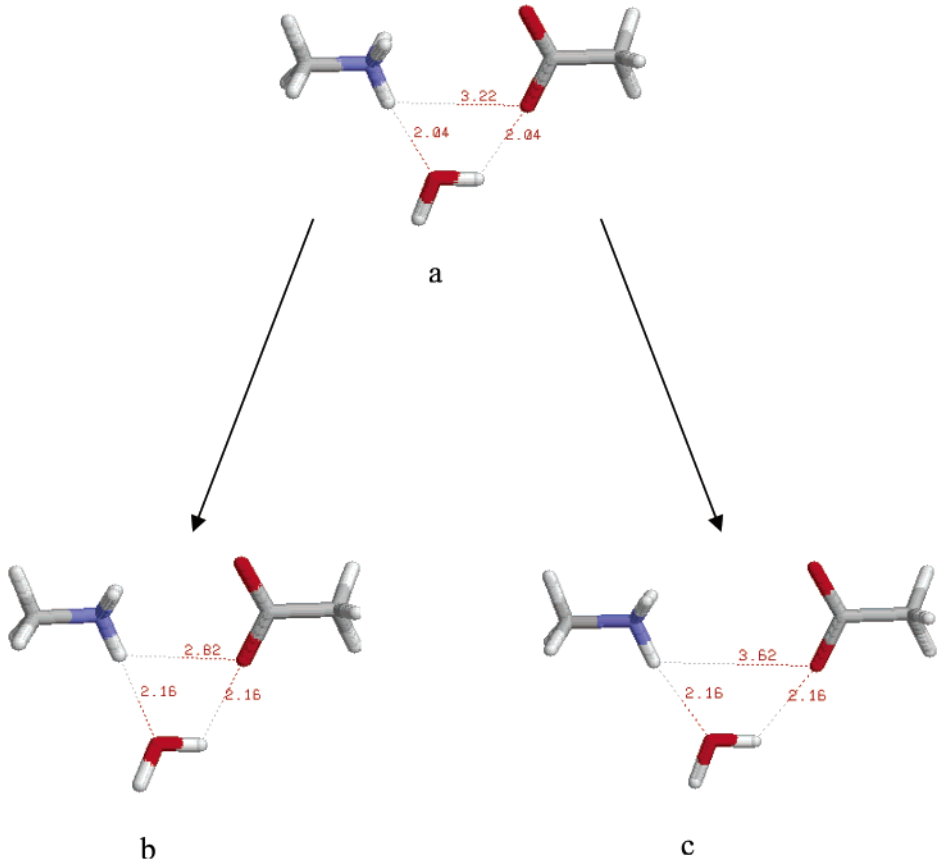

Figure 8. Geometry and orientation of the included explicit water molecule in the acetate−methylammonium ion system. The water molecule is coplanar with the collinear axis of the system and the H−O pair of atoms from the ions that traps the water molecule. The H−O bond of the water is in parallel with the line connecting the charged H−O pair. The numbers in the figure denote the distances in Å between the two atoms connected by the broken lines. For other details, see Figure 7.

TABLE 5: Electrostatic Solvation Free Energies of the Acetate−Guanidinium Ion Pair at Various N−O Distances as Calculated by SGB with an Explicit Water Molecule Included

N−O dist (Å)	SGB−water (kcal/mol)
3.159	−58.22
3.559	−65.97
3.959	−73.9
4.359	−82.59
4.759	−90.54
5.159	−98.6
5.559	−102.29
5.959	−102.23
6.359	−105.79
6.759	−108.93
7.159	−111.69
7.559	−114.06

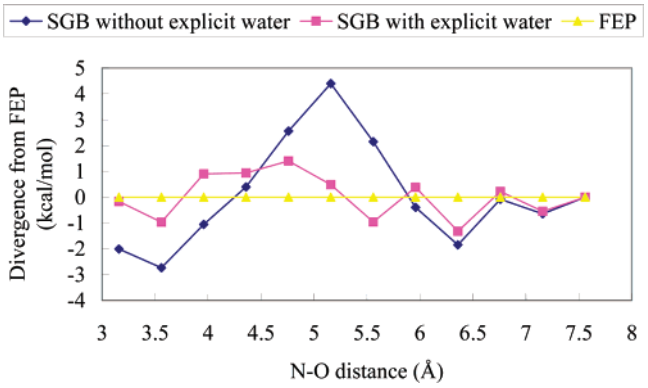

Figure 9. Comparison of the difference of the electrostatic solvation free energies of acetate−guanidinium ion pair at various N−O distances to that at the most extended distance, by SGB with or without the inclusion of explicit water molecule.

systematic, are relatively independent of details of the continuum model, and afford a straightforward physical interpretation. The continuum models treat the solvent as an infinitesimal dipole, as opposed to a species with finite dimensions. At the separation where a single water molecule can bridge the ion pair in an energetically favorable manner, the continuum model incorporates extensive dielectric screening of the ion pair which simply is not present in the actual physical system. These results are in accord with earlier work, in which bridging waters were observed to stabilize ion pairs in even simpler systems such as sodium chloride, sodium dimethyl phosphate and tetramethylammonium chloride, etc. in aqueous solution.[34−41] The single bridging water efficiently captures the first shell energy of both ions without creating dipole layers around each ion. It is true that the water oxygen that is hydrogen bonded to the positive species interacts unfavorably with the negative ion in the pair, and that the water hydrogen similarly interacts unfavorably with the positive ion. However, the geometrical structure of these interactions is considerably different from what is assumed in the continuum model, and apparently much less unfavorable. Thus, this appears to be a fundamental problem with continuum models which is relatively independent of the details of the types of continuum approximations made or the parametrization of those approximations.

The explicit solvent model that has been used here is not a perfectly faithful representation of physical reality; for example, it does not explicitly include polarization, no doubt is imperfectly parametrized, etc. However, it is difficult to see why any of

TABLE 6: Electrostatic Solvation Free Energies of the Acetate−Methyl Ammonium Ion Pair at Various N−O Distances as Calculated by SGB with an Explicit Water Molecule Included

N−O dist (Å)	SGB−water (kcal/mol)
3.018	−53.42
3.394	−64.27
3.775	−73.84
4.159	−83.12
4.546	−90.1
4.935	−95.57
5.326	−99.3
5.914	−106.06
6.308	−109.82
6.702	−113.29
7.098	−115.94
7.492	−118.2

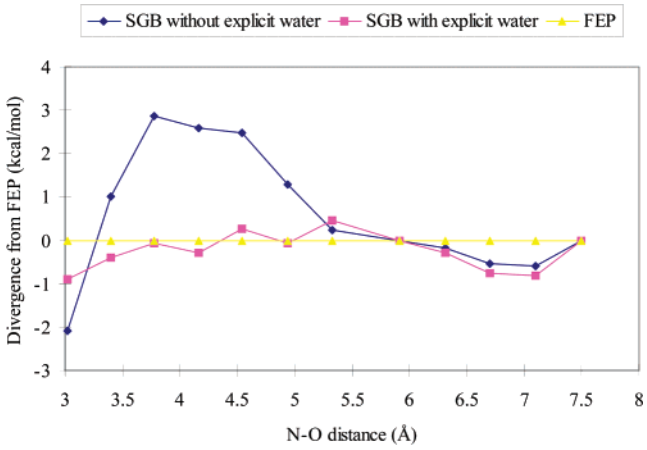

Figure 10. Comparison of the difference of the electrostatic solvation free energies of acetate−methylammonium ion pair at various N−O distances to that at the most extended distance, by SGB with or without the inclusion of explicit water molecule.

these flaws would invalidate the qualitative conclusions drawn above. Explicit inclusion of polarization would, if anything, further stabilize the bridging water structure, as the hydrogen bond of the water hydrogen to one charged species should enhance the favorable electrostatics of the water oxygen to the second, oppositely charged species. Simulations with a water model containing explicit polarization would be desirable, however, to confirm this assertion, and we plan to pursue such calculations in the near future. A number of polarizable simulations of ions in solution have been successfully carried out over the past few years,[42−44] so investigations of this type will be tractable.

This picture furthermore is in agreement with our attempts to predict side chain structures using a continuum solvation model. Our side chain prediction results indicate that the SGB continuum model systematically underpredicts such structures, and that the inclusion of crystallographic waters[45−47] results in a substantial improvement in prediction accuracy, even for a small, relatively simple side chain such as serine.[48,49]

Finally, this picture agrees with our studies of bound waters in the PDB, which demonstrates that the bridging water motif is commonplace across the PDB. Our results suggest that bridging waters may be a generic feature of the first shell solvation of ion pairs. However, the complex protein environment is qualitatively different than the small molecule results considered thus far. As an initial attempt to address the biological significance of our results, we have searched a small but diverse set of protein structures from the PDB for possible bridging waters. The rationale is that bridging waters, because

Figure 11. T-cell signal transduction protein (PDBid 1d4t). Residues ARG55 and GLU67, and the possible bridging water HOH 35, are shown in a space-filling representation.

of their strong interactions with the ion pair, are much less mobile than bulk water, and thus should constitute a subset of the observed crystallographic waters.

Specifically, we have searched a total of 35 proteins, with resolution better than 2.0 Å, *R* factor better than 0.2, and at most 30% sequence identity to each other.[1] Possible bridging waters were identified based on reported waters in close proximity to salt bridges. Specifically, the water oxygen must be within 4.0 Å of *both* electronegative atoms involved in the salt bridge. We have considered only salt bridges between Arg/Lys and Asp/Glu, to be consistent with the small molecule results. Thus, the water oxygen must be within 4.0 Å of both a nitrogen atom from the Arg/Lys charged groups, and one of the 2 oxygen atoms from the carboxylate groups. A total of 243 possible bridging waters were identified in the selected protein structures according to this criterion. An example is shown in Figure 11, and a detailed table appears in the Supporting Information. In several cases, more than one possible bridging water is identified for a particular salt bridge. Although our molecular dynamics simulations indicate that it is possible for more than one bridging water to exist, there is no guarantee that this is the case in the protein structures.

The salt bridges associated with the potential bridging waters ranged from very strong (<3 Å O−N distance between the ions) to rather weak (>6 Å O−N distance between the ions). The average O−N distance between the ions is 4.5 Å. Figure 12 displays a histogram of the number of bridging waters found as a function of the distance between the ion pairs. These plots are at least qualitatively consistent with the observed "bridging water regime" discussed above, i.e., bridging waters are most commonly found between ion pairs separated by about 4−5 Å.

We wish to emphasize that we cannot state with any certainty whether the waters we have identified are in fact "bridging" waters according to our definition, which requires that the water simultaneously hydrogen bond to both ions. The problem, of course, is that only the oxygen coordinates are reported for crystallographic waters, even for high-resolution crystal structures. Close proximity of a water to a salt bridge, in the complex protein environment, does not necessarily indicate that it hydrogen bonds to both ions. Moreover, the crystal structures are determined at various temperatures, crystal packing may contribute to the observed positions of crystallographic waters, and the identification of bridging waters relies on careful refinement of the structures. Nonetheless, the large number of

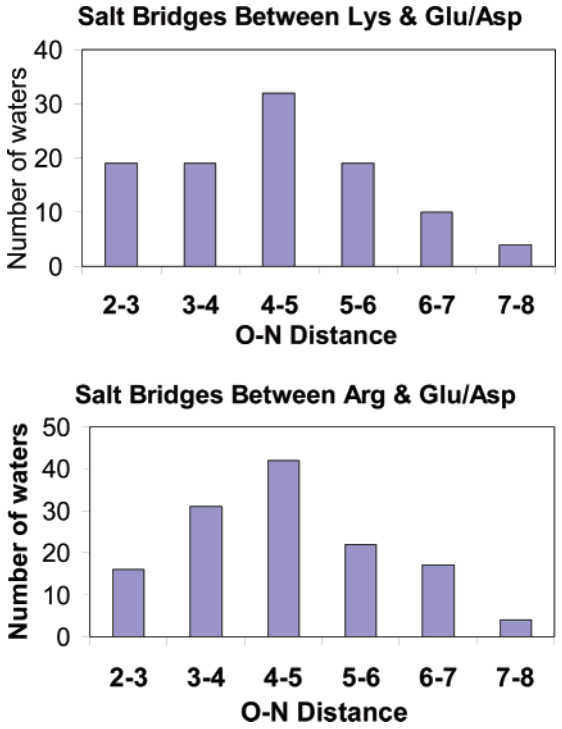

Figure 12. Histograms of the distances between ion pairs, when there exists one or more potential "bridging waters" in their vicinity. The salt bridges were identified in the 35 protein data set discussed in the test, and then grouped according to the closest O−N distance (Å) between the charged groups. The number of bridging waters founds for each category is shown.

possible bridging waters identified by our crude procedure suggests that they are likely to be important contributors to the detailed structure and stability of proteins. The total number of salt bridges (between Lys/Arg and Asp/Glu) in the protein structures chosen is 836, using a generous cutoff of 6 Å for the O−N distance.

The conclusions to be drawn from these results is that some method must be found to improve the continuum solvation potential of mean force at bridging distances, particularly for ion pairs. The size of the relative energy error, 4−6 kcal/mol, is substantial, and it is material for a wide range of problems in protein structure prediction and protein−ligand binding calculations. In fact, in protein active sites, water molecules are intercalated into considerably more complex and confined geometries than we consider here, and the errors of a continuum treatment in such cases may be even larger.

We have presented in section IV one way in which a single explicit water can be incorporated into the continuum calculations; the results obtained from this method exhibited considerable improvement over the fully continuum calculalations. Other, more sophisticated methods for including explicit waters are likely to provide even better results. A second alternative is to empirically modify the effective screened interactions between solute atoms in a continuum description to take into account the phenomenon we have elucidated above. This could be accomplished in the context of a GB model by fitting the GB pair functions to reproduce the explicit solvent potentials of mean force.

VI. Conclusion

We have investigated the free energy profiles of two ion pairs in water as a function of separation for both implicit and explicit solvation models, and obtained a systematic difference between them at the distance where a single bridging water molecule can be inserted. We have shown that the occupation probability in the explicit solvent simulations at this position is very high, confirming our hypothesis as to the physical origin of the discrepancy of the explicit and implicit models.

The present results also demonstrate that our parametrization of the SGB model is, at least for the present case, more or less equivalent to numerically accurate PB solutions. More specifically, the deviations as compared to explicit solvent simulations are far larger than the differences between the various implicit solvent calculations. To some extent this attests to the careful parametrization of the SGB model against PB data, but it also serves as a reminder that the ultimate goal is accurate reproduction of experiment and not of a particular implicit solvation model whose ability to reproduce experimental results when short distance interactions are involved (as opposed to long distances, where there are numerous elegant validations of the PB approach) is far from clear.

The key to future improvements is extensive comparison with experiment in cases where structural waters are present. The structural data in the PDB should enable this type of comparison to be made rigorously; the trick is to come up with approaches that are computationally tractable for large data sets, for without statistical assessment of performance, it is difficult to draw reliable conclusions. This will require new approaches, along the lines discussed above.

Acknowledgment. We would like to thank Barry Honig for useful discussions. M.P.J. wishes to acknowledge support from an NSF Postdoctoral Fellowship in Biological Informattics. This work was supported in part by a grant to R.A.F. from the NIH (GM-52018).

Supporting Information Available: A table containing the full list of atomic charges and radii for the two systems, a table of possible "bridging waters" in PDB structures, and an mpeg file showing an animation of the water trapped between the pairs. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Ghosh, A.; Rapp, C. S.; Friesner, R. A. *J. Phys. Chem. B* **1998**, *102*, 10983.
(2) Zhang, L. Y.; Gallicchio, E.; Friesner, R. A.; Levy, R. M. *J. Comput. Chem.* **2000**, *22*, 591.
(3) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. *J. Comput. Chem.* **2002**, *23*, 517.
(4) Luo, R.; David, L. Hung, H.; Devaney, J.; Gilson, M. K. *J. Phys. Chem. B* **1999**, *103*, 727.
(5) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3*, 211.
(6) Sindelar, C. V.; Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1998**, *7*, 1898.
(7) Xiao, L.; Honig, B. *J. Mol. Biol.* **1999**, *289*, 1435.
(8) Strop, P.; Mayo, S. L. *Biochemistry* **2000**, *39*, 1251.
(9) Takano, K.; Tsuchimori, K.; Yamagata, Y.; Yutani, K. *Biochemistry* **2000**, *39*, 12375.
(10) Ngola, S. M.; Kearney, P. C.; Mecozzi, S.; Russell, K.; Dougherty, D. A. *J. Am. Chem. Soc.* **1999**, *121*, 1192.
(11) Searle, M. S.; Gfiffiths-Jones, S. R.; Skinner-Smith, H. *J. Am. Chem. Soc.* **1999**, *121*, 11615.
(12) Rashin, A. A. *J. Phys. Chem.* **1989**, *93*, 4664.
(13) Soetens, J. C.; Millot, C.; Chipot, C.; Jansen, G.; Angyan, J. G.; Maigret, B. *J. Phys. Chem. B* **1997**, *101*, 10910.
(14) Rozanska, X.; Chipot, C. *J. Chem. Phys.* **2000**, *112*, 9691.
(15) Beveridge, D. L.; Di Capua, F. M. *Annu. Rev. Biophys. Chem.* **1989**, *18*, 431.
(16) Kollman, P. A. *Acc. Chem. Res.* **1996**, *29*, 461.
(17) Lamb, M. L.; Jorgensen, W. L. *Curr. Opin. Chem. Biol.* **1997**, *1*, 449.
(18) Gilson, M. K.; Given, J. A.; Bush, G. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047.

(19) Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. *J. Mol. Biol.* **2002**, *320*, 597.
(20) Masunov, A.; Lazaridis, T. *J. Am. Chem. Soc.* **2003**, *125*, 1722.
(21) Gilson, M. K.; Sharp, K.; Honig, B. *J. Comput. Chem.* **1987**, *9*, 327.
(22) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435.
(23) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144.
(24) Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. *J. Comput. Chem.* **2002**, *23*, 128.
(25) Cortis, C. M.; Friesner, R. A. *J. Comput. Chem.* **1997**, *18*, 1570.
(26) Cortis, C. M.; Friesner, R. A. *J. Comput. Chem.* **1997**, *18*, 1591.
(27) Connolly, M. L. *Science* **1983**, *221*, 709.
(28) Connolly, M. L. *J. Appl. Crystallogr.* **1983**, *16*, 548.
(29) Connolly, M. L. *J. Appl. Crystallogr.* **1985**, *18*, 499.
(30) Connolly, M. L. *J. Mol. Graphics* **1993**, *11*, 139.
(31) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
(32) Pratt, L. R.; LaViolette, R. A.; Gomez, M. A.; Gentile, M. E. *J. Phys. Chem. B* **2001**, *105*, 11662.
(33) Asthagiri, D.; Pratt, L. R.; Ashbaugh, H. S. *J. Chem. Phys.* **2003**, *119*, 2702.
(34) Hirata, F.; Rossky, P. J.; Pettitt, B. M. *J. Chem. Phys.* **1982**, *78*, 4133.
(35) Pettitt, B. M.; Rossky, P. J. *J. Chem. Phys.* **1986**, *84*, 5836.
(36) Dang, L. X.; Pettitt, B. M. *J. Am. Chem. Soc.* **1987**, *109*, 5531.
(37) Jorgensen, W. L.; Buckner, J. K.; Huston, S. E.; Rossky, P. J. *J. Am. Chem. Soc.* **1987**, *109*, 1891.
(38) Buckner, J. K.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1989**, *111*, 2507.
(39) Patey, G. N.; Carnie, S. L. *J. Chem. Phys.* **1983**, *78*, 5183.
(40) Fries, P. H.; Patey, G. N. *J. Chem. Phys.* **1984**, *80*, 6253.
(41) Huston, S. E.; Rossky, P. J. *J. Phys. Chem.* **1989**, *93*, 7888.
(42) Stuart, S. J.; Berne, B. J. *J. Phys. Chem.* **1996**, *100*, 11934.
(43) Grossfield, A.; Ren, P.; Ponder, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 15671.
(44) Rick, S. W.; Stuart, S. J. *Rev. Comput. Chem.* **2002**, *18*, 89.
(45) Makarov, V. A.; Andrews, B. K.; Smith, P. E.; Pettitt, B. M. *Biophys. J.* **2000**, *79*, 2966.
(46) Brunne, R. M.; Liepinsh, E.; Otting, G.; Wüthrich, K.; van Gunsteren, W. F. *J. Mol. Biol.* 1**993,** *231*, 1040.
(47) Thanki, N.; Thornton, J. M.; Goodfellow, J. M. *J. Mol. Biol.* 1**988,** *202*, 637.
(48) Andrec, M.; Harano, Y.; Jacobson, M. P.; Friesner, R. A.; Levy, R. M. *J. Struct, Funct. Genomics* **2002,** *2*, 103.
(49) Jacobson, M. P.; Kaminski, G. A.; Friesner, R. A.; Rapp, C. S. *J. Phys. Chem. B* **2002,** *106*, 11673.