Accounting for Polarization Cost When Using Fixed Charge Force Fields. II. Method and Application for Computing Effect of Polarization Cost on Free Energy of Hydration

William C. Swope,* Hans W. Horn, and Julia E. Rice

IBM Almaden Research Center, 650 Harry Road, San Jose, California 95120

Received: December 10, 2009; Revised Manuscript Received: April 30, 2010

Polarization cost is the energy needed to distort the wave function of a molecule from one appropriate to the gas phase to one appropriate for some condensed phase. Although it is not currently standard practice, polarization cost should be considered when deriving improved fixed charge force fields based on fits to certain types of experimental data and when using such force fields to compute observables that involve changes in molecular polarization. Building on earlier work, we present mathematical expressions and a method to estimate the effect of polarization cost on *free* energy and enthalpy implied by a charge model meant to represent a solvated state. The charge model can be any combination of point charges, higher-order multipoles, or even distributed charge densities, as long as they do not change in response to environment. The method is illustrated by computing the effect of polarization cost on free energies of hydration for the neutral amino acid side chain analogues as predicted using two popular fixed charge force fields and one based on electron densities computed using quantum chemistry techniques that employ an implicit model to represent aqueous solvent. From comparison of the computed and experimental hydration free energies, we find that two commonly used force fields are too underpolarized in their description of the solute−water interaction. On the other hand, a charge model based on the charge density from a hybrid density functional calculation that used an implicit model for aqueous solvent performs well for hydration free energies of these molecules after the correction for dipole polarization is applied. As such, an improved description of the density (e.g., B3LYP, MP2) in conjunction with an implicit solvent (e.g., PCM) or explicit solvent (e.g., QM/MM) approach may offer promise as a starting point for the development of improved fixed charge models for force fields.

1. Introduction

Expressions and methods for the computation of the polarization cost of a molecule associated with a particular conformation and charge model were reported in a companion publication (DOI 10.1021/jp911699p).[1] However, the promise of (and primary motivation for) this method lies in its use to account for polarization in classical simulations that use fixed charge force fields to compute thermodynamic properties associated with processes that involve a change in polarization state. These include, for example, enthalpies of vaporization and solvation free energies. Since experimental observables of these sorts are often used for the development and testing of force fields, we feel that *improved* fixed charge models might be possible if proper account is made of polarization energetics. Such models could provide more accurate representations of molecular interactions for particular environments and obviate the need for polarizable models in some contexts.

As commented in the previous publication,[1] classical fixed charge models of water that take polarization into account (e.g., SPC/E,[2] TIP4P-Ew,[3] and TIP4P/2005[4]) are more polarized and seem to have more accurate properties for some observables for which they were not parametrized (such as the self-diffusion coefficient) than those that do not (e.g., TIP3P, TIP4P,[5] and SPC[6]). The water model efforts that took polarization cost into account all used a very simple approximation for the polarization cost, W_{pol}

$$W_{\mathrm{pol}} = \frac{(\Delta|\boldsymbol{\mu}|)^2}{2\bar{\alpha}} \tag{1}$$

where $\Delta|\boldsymbol{\mu}|$ is the change in the magnitude of the dipole moment between the gas phase value for a water molecule (1.85 D) and that implied by the charge model of the force field, and $\bar{\alpha}$ is the average of the diagonal elements of the dipole−dipole polarizability tensor (51.6633 Fm^2) for (gas phase) water.[7] This expression, although very effective for water model development, is rather crude, and it has been difficult to see how to generalize this approach to treat arbitrary molecules, particularly ones where the polarizability is not isotropic, where quadrupolar (or higher) polarization might be important, and where conformational averaging might need to be taken into account (all these water models being rigid, only a single structure needs to be considered). Equation 1 was derived, and its inherent assumptions and applicability were discussed in the companion publication.[1] Other workers[8−10] have discussed the potential and problems associated with the use of such an expression to account for polarization in hydration free energies computed from classical simulations that employ fixed charge models. In the end, however, they chose not to apply this type of correction.[11]

The method for computing polarization cost described in the companion paper[1] represents a generalization and improvement on this simple approximation. In this approach, there are two key relationships. The first one expresses induced multipole moments in terms of the polarizabilities and the electrostatic potential gradients, Hessians, and higher-order derivatives

* Author to whom correspondence should be addressed. E-mail: swope@almaden.ibm.com.

10.1021/jp911701h
Published on Web 06/11/2010

$$\begin{aligned} \mu_\alpha - \mu_\alpha^0 &= -\alpha_{\alpha\beta}V_\beta - \frac{1}{3}A_{\alpha,\beta\gamma}V_{\beta\gamma} + ... \\ \Theta_{\alpha\beta} - \Theta_{\alpha\beta}^0 &= -A_{\gamma,\alpha\beta}V_\gamma - C_{\alpha\beta,\gamma\delta}V_{\gamma\delta} + ... \end{aligned} \quad (2)$$

Here, μ_α and $\Theta_{\alpha\beta}$ are Cartesian components of the dipole and quadrupole moments implied by the charge model, meant to be appropriate for the solvated (polarized) molecule, and μ_α^0 and $\Theta_{\alpha\beta}^0$ are those for the gas phase (unpolarized) molecule, which for our purposes can be determined from a high enough level of quantum chemical calculation. (Einstein notation is used in these equations, so one is meant to sum over repeated indices.) The left side of eqs 2 represents the multipole moments that are *induced* by electrostatic fields presumed to have been generated in the solvent. These induced moments are expressed in terms of the dipole−dipole (α), dipole−quadrupole (**A**), and quadrupole−quadrupole (**C**) polarizabilities, all of which can also be evaluated from gas phase quantum chemical calculations, and the potential gradient, $V_\alpha = \nabla_\alpha V$, and Hessian, $V_{\alpha\beta} = \nabla_\alpha\nabla_\beta V$. Since the charge model of interest gives the solution phase multipole moments and quantum chemistry calculations give the gas phase multipole moments and polarizabilities, eq 2 is a set of linear equations for the unknown field gradient and Hessian components in terms of the known induced multipole moments and polarizabilities. Essentially, these, then, are the properties of an effective electrostatic potential that would polarize the molecule to the extent exhibited by the charge model and describe the *reaction field* assumed to be established by the solvent. (As explained in the companion article,[1] if one wishes to include terms beyond dipole−dipole polarizabilities, these equations need to be solved using real spherical multipoles, polarizabilities, gradients, and Hessians to avoid linear dependencies in the analogous Cartesian expressions.) Once one solves for these, they can be used with the polarizabilites to obtain the polarization cost as follows

$$W_{\mathrm{pol}} = \frac{1}{2}\alpha_{\alpha\beta}V_\alpha V_\beta + \frac{1}{3}A_{\alpha,\beta\gamma}V_\alpha V_{\beta\gamma} + \frac{1}{6}C_{\alpha\beta,\gamma\delta}V_{\alpha\beta}V_{\gamma\delta} + ... \quad (3)$$

Of special interest is the following expression that is appropriate when the effects of the quadrupolar and higher-order terms can be neglected

$$W_{\mathrm{pol}}^{\mathrm{D}} = \frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}^0)^{\mathrm{t}}(\boldsymbol{\alpha}^{-1})^{\mathrm{t}}(\boldsymbol{\mu} - \boldsymbol{\mu}^0) \quad (4)$$

In this paper, we extend the use of these expressions for computing polarization cost to address the effect of polarization cost on free energies and enthalpies. The derivation of this approach involves the construction and analysis of a thermodynamic cycle that relates an actual experimental process with what is performed in a classical simulation involving a fixed charge force field. This is applied to address the polarization cost exhibited for several small molecule amino acid side chain analogues by three different charge models. The polarization cost for these charge models is then included in the hydration free energies reported from classical simulations, and the results are discussed in light of experimental values. Section 2 derives the expressions for computing the effect of polarization cost on free energy and enthalpy in the context of solvation. Section 3 discusses the multipole moments, polarization costs, and polarization-corrected hydration free energies for the amino acid side chain analogues. Section 4 is a general discussion, and conclusions are presented in Section 5. Appendix A investigates an alternative thermodynamic cycle and develops other practical expressions for computing the effect of polarization cost on free energies. Appendix B provides justification for an important approximation used in this work by explicit computation of the effect of polarization cost on free energy using expressions derived in Section 2 and in Appendix A for four molecules. These evaluations involve conformational (Boltzmann) averages of functions of the polarization costs to produce polarization-adjusted free energies.

2. Expressions for Effect of Polarization Cost Free Energy and Enthalpy

In this section, we develop expressions for the effects of polarization cost on free energy and enthalpy. These effects should be considered in the computation of energy, enthalpy, and free energy changes at finite temperature when there is a change in the nature of the molecular environment, as in going from a condensed phase to a gas phase. Examples include the enthalpy of vaporization, free energies of solvation, chemical potentials, and transfer free energies.

Thermodynamic quantities such as these are temperature-dependent observables, and their evaluation from first principles involves the averaging of certain properties over many molecular conformations consistent with the temperatures of interest and a potential energy surface. The required expressions are derived using statistical mechanical approaches that employ state functions and partition functions. In fact, the quantities we wish to evaluate are actually energies that describe *changes* in state functions. Such changes are independent of the path taken between the end states, so one often constructs thermodynamic cycles, one side of which corresponds to some experimentally meaningful physical process of interest, but where the same end states are connected by an alternative path which may contain possibly nonphysical states and processes. Generally these nonphysical states and processes are used because they can be analytically or numerically characterized. In this section, we will develop expressions employing such a thermodynamic cycle for the effect of polarization cost on the free energy of solvation.

Using this approach one conceives of the physical process of solvation as occurring along a nonphysical pathway in multiple nonphysical steps, some of which involve the transfer of molecules from one environment to another *while holding their charge distribution fixed*, and others which involve a subsequent change in charge distribution in response to the change of environment. The energy, enthalpy, or free energy change associated with the first kind of process can be evaluated for fixed charge force fields using relatively standard techniques such as thermodynamic integration and thermodynamic perturbation theory. Methods for evaluating the changes in state functions associated with the second kind of process, which involve changes in polarization, will be described in this section and are based on calculations of polarization cost described in the earlier work.[1]

Processes occurring under conditions of controlled temperature and pressure are described using the Gibbs free energy, $G(N, p, T)$, a state function of the temperature, pressure and number (composition) of molecules of each type considered to be the system of interest. There is a corresponding isothermal−isobaric partition function related to the Gibbs free energy by $G = -kT \ln \Xi(N, p, T)$ where the partition function

$$\Xi = \int \mathrm{d}V\mathrm{d}r\mathrm{d}p\mathrm{e}^{-\beta(H(r,p)+pV)}$$

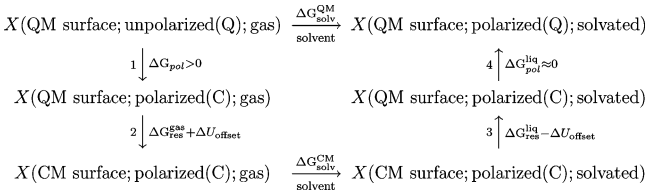

Figure 1. Thermodynamic Cycle 1.

involves an integral over volume and all particle positions and momenta, generically represented by the variables r and p. $H(r, p)$ is the Hamiltonian. β is $1/kT$, where k is the Boltzmann constant.

Taking solvation free energy as the property for which one wishes to obtain a measure of the effect of solute polarization by solvent, one considers the following thermodynamic cycle (cycle 1, Figure 1).

In this cycle, one is *primarily* interested in obtaining the free energy difference ΔG^{QM}_{solv} described on the top line for the process of solvating some molecule, X, where the electron distribution is able to respond to the change in environment in going from the gas phase (unpolarized) to the solution phase (polarized) and to have the energetics of this process properly accounted for in the computed free energy. This, of course, is the experimental situation with which we ultimately want to compare; however, in general such an accurate representation of the polarization process is too expensive, so one instead considers an alternative pathway broken here into five process steps between the same two end states.

$$\Delta G^{QM}_{solv} = \Delta G^{gas}_{res} + \Delta G_{pol} + \Delta G^{CM}_{solv} + \Delta G^{liq}_{pol} + \Delta G^{liq}_{res}$$

One of these steps (bottom) is the computation of the solvation free energy using a fixed charge force field just as is usually done. However, the other steps, which treat polarization issues, have been neglected by most users of fixed charge models. The main point of this section of this paper is to derive expressions for these missing components so that they may be evaluated or approximated and applied as corrections to solvation free energies computed using fixed charge models. We will discuss the top line of this thermodynamic cycle and the five steps of the alternative process in greater detail in what follows. The final result that will be demonstrated is that under appropriate circumstances $\Delta G^{QM}_{solv} \approx \Delta G^{CM}_{solv} + W_{pol}$, and so the polarization cost of a single molecular conformation may be added as an adjustment to the free energy computed using a fixed charge force field to produce values that may be compared directly with experimental solvation free energies. A demonstration of the degree to which this approximation is valid is described in Appendix B.

2.1. Desired Solvation Process, ΔG^{QM}_{solv}. The top line of the cycle represents the main process of interest, with a free energy difference represented by ΔG^{QM}_{solv}. The use of QM in the notation indicates that we will assume some level of quantum chemistry and, although prohibitively expensive, would in principle be able to provide an adequately accurate representation of the experimental situation and the polarization effects; however, one may also in the future be able to substitute for this a suitably accurate polarizable force field. The notation also includes a designation for the environment of the molecule (gas or solvent) to describe the end states as well as its polarization state and the nature of the potential energy surface that determines the conformational dynamics and thermodynamics of the solute molecule. The polarization state can be unpolarized, as in the gas phase, or polarized in a way that is consistent with the solution phase. So, as indicated by the notation, the desired process takes the molecule from the gas to the solvated state, using a quantum chemical potential surface, and since it can respond to the environment, the charge model changes from one appropriate to the gas phase to one appropriate to the solution phase.

In anticipation of the need to connect the desired process with ones that will use a fixed charge model, which tend to be polarized to some degree, we extend the notation for the charge model to indicate whether it was obtained from some particular fixed charge force field (designated with C, for classical charge model) or from some representation (designated with Q, for quantum chemical charge model) that is capable of responding to the environment. Examples of the latter include the polarizable continuum model (PCM) and related techniques discussed in the companion article,[1] hybrid quantum mechanics/molecular mechanics (QM/MM) approaches, or ab initio molecular dynamics (AIMD) simulation approaches, as well as polarizable force fields. Whatever the charge model, we are primarily concerned with the multipole moments of the charge distribution, in particular the dipole and quadrupole moments, and these may be obtained from either a continuous electron density (e.g., from a wave function) or from a set of point charges and multipoles as determined by the force field, polarizable model, or a fit to the electrostatic potential around the molecule (e.g., a restrained electrostatic potential fit (RESP)[12]).

We also provide for the notation to describe the potential energy surface that determines the dynamical motion of the system, which can be either from a classical force field (designated CM surface) or produced using a quantum chemical or other accurate method (designated QM surface).

2.2. Fixed Charge Solvation Process, ΔG^{CM}_{solv}. Although one would like to include polarization effects directly in the free energy calculation, since this is often impractical, one instead performs solvation free energy calculations using simulations that employ fixed charge force fields in the context of thermodynamic integration and/or free energy perturbation. With these methods the free energy, ΔG^{CM}_{solv}, of introducing potential energy terms that describe the solute−solvent interaction energy is computed. This is indicated in the bottom line in the thermodynamic cycle shown above, where it is also indicated that these calculations employ a fixed charge model designed to represent some degree of solvent polarization (polarized(C)), and perform thermodynamic averaging using a potential surface described by a fixed charge force field (CM surface).

2.3. Effect of Polarization Cost on Free Energy, ΔG_{pol}. As suggested in Figure 1, one can obtain the desired solvation free energy (top) from the one computed using fixed charges (bottom) by including appropriate contributions (sides) to account for polarization and also for the fact that the classical fixed charge representation of the potential energy surface that was designed to be used in the solvated state is an approximation that may be inappropriate in the gas phase. The effect of polization cost on free energy (ΔG_{pol}) indicated in step 1 on the left accounts for the process of deforming the charge distribution from one appropriate for the gas phase (unpolarized (Q)) to one appropriate for the solution phase as described by the polarized but fixed charge model (polarized (C)) to be used for the solvated molecule. Note that this free energy change is measured in the gas phase, and it employs quantum chemically derived potential energy surfaces but in an interesting way that exploits the method described earlier[1] for computing polarization energy. If we denote the Hamiltonian that describes Born−Oppenheimer motion of the nuclei on a quantum chemically derived gas phase potential energy surface as $H^{QM}_{gas}(r, p)$, one is using the fact that for each configuration of nuclear coordinates given by r there is a well-defined electron density and

electronic energy obtained from solving the quantum chemical problem. This Hamiltonian describes the starting state for process step 1.

For the end state, we wish to employ a Hamiltonian that differs from $H_{\text{gas}}^{\text{QM}}(r, p)$ by our estimate of the polarization cost, $W_{\text{pol}}(r)$, evaluated at the nuclear coordinates r. W_{pol} depends on nuclear coordinates both because the computed gas phase polarizabilities ($\boldsymbol{\alpha}$, $\mathbf{A}$, and $\mathbf{C}$) depend on them *and* because the induced dipole and quadrupole moments depend on them through the two charge models being compared. ($\boldsymbol{\mu}$ and $\boldsymbol{\Theta}$ are determined by the fixed charge model (polarized (C)) and nuclear coordinates, and $\boldsymbol{\mu}^0$ and $\boldsymbol{\Theta}^0$ are determined by the unpolarized charge model (unpolarized (Q)) and nuclear coordinates.) Note that since W_{pol} is a function of coordinates it contributes to the forces and affects the dynamics of the nuclei in the end state of process step 1. Although we wish to use $H_{\text{gas}}^{\text{QM}} + W_{\text{pol}}$ to characterize the end state of the process, it should be understood that the corresponding potential energy surface does not come directly from the solution of a quantum chemical problem. The free energy difference for process step 1 is given below

$$\begin{aligned}\Delta G_{\text{pol}} &= G(p, T; \text{QM surface; polarized (C); gas}) \\ &\quad - G(p, T; \text{QM surface; unpolarized (Q); gas}) \\ &= -kT \ln \frac{\Xi(p, T; \text{QM surface; polarized (C); gas})}{\Xi(p, T; \text{QM surface; unpolarized (Q); gas})} \\ &= -kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p)+W_{\text{pol}}(r)+pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p)+pV]}} \\ &= -kT \ln\langle e^{-\beta W_{\text{pol}}(r)}\rangle_{NpT;\ \text{QM; unpolarized (Q)}}\end{aligned}$$

This expression indicates that the thermodynamic average is to be taken using an NpT ensemble and using the Hamiltonian corresponding to the unpolarized electron density and the gas phase quantum chemical potential energy surface. (An alternative expression for the effect of polarization cost on free energy that is based on an ensemble average that uses the fixed charge potential surface is derived in Appendix A.)

If W_{pol} is constant, or, at least, approximately so, over the range of structures that are thermodynamically accessible to the gas phase solute molecule, then it is useful to factor from the numerator in these expressions the quantity $e^{-\beta W_{\text{pol}}}$, with W_{pol} evaluated at the equilibrium gas phase structure of the molecule, which we denote $r_{\text{g,min}}^{\text{QM}}$.

$$\begin{aligned}\Delta G_{\text{pol}} &= W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}}) \\ &\quad - kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p)+(W_{\text{pol}}(r)-W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}}))+pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p)+pV]}} \\ &= W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}}) - kT \ln\langle e^{-\beta[W_{\text{pol}}(r)-W(r_{\text{g,min}}^{\text{QM}})]}\rangle_{NpT;\text{QM;unpolarized (Q)}}\end{aligned} \tag{5}$$

If it is approximately constant, $W_{\text{pol}}(r) \approx W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}})$, the second term is small, and $\Delta G_{\text{pol}} \approx W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}})$. So, a reasonable approximation, but one which needs to be justified on a case-by-case basis, is to use for ΔG_{pol} the value of W_{pol} evaluated at the equilibrium structure on the gas phase quantum chemically generated potential surface. (This approximation is discussed further in Appendix B, where it is shown to be justified for the set of molecules studied later in this paper.)

2.4. Polarization Adjustment Energy, $\Delta G_{\text{pol}}^{\text{liq}}$. The effect of polarization cost on the free energy change considered in step 4 would account for the possibility that the charge model of the fixed charge force field might be over or under polarized relative to that for a molecule that can appropriately adjust its charge distribution to respond to the solvent. The fixed charge model might be over (or under) polarized if it had been designed to be used with water as a solvent but is being used with a different solvent with a smaller (or larger) dielectric constant. In either case, as defined by the direction of the arrow in the diagram, $\Delta G_{\text{pol}}^{\text{liq}} < 0$. However, one hopes the fixed charge model is appropriately polarized, or nearly so, so that $\Delta G_{\text{pol}}^{\text{liq}} \approx 0$. A key difference between the free energy change associated with steps 1 and 4 is that the second one involves a solvent, and a change in polarization of the solute will change not only the energy of the solute, just as in step 1, but also the structure of the solvent and the solute−solvent interaction energy. The computation of this free energy could be carried out in a number of ways and could even be approximated using a reaction field approach. For now, we will assume that the charge model has been designed to provide an accurate description for the solute polarization in the solvent under consideration. This amounts to assuming this contribution is zero.

2.5. Restructuring Free Energy, ΔG_{res}. The *restructuring* free energies $\Delta G_{\text{res}}^{\text{gas}}$ and $\Delta G_{\text{res}}^{\text{liq}}$ in steps 2 and 3 address differences in conformational preferences between the classical and quantum potential energy surfaces *after* the difference in polarization has been taken into account. The potential energy offset, ΔU_{offset}, addresses differences in the reference values for the energy used for the classical and quantum surfaces. Potential surfaces generated by quantum chemical methods are usually energies measured with respect to infinitely separated electrons and nuclei. However, most force field based potential surfaces have an arbitrary zero of energy that is usually not relevant because such surfaces are only used to generate trajectories for kinetic studies or for thermodynamic sampling. This difference in reference energies must be accounted for in each of steps 2 and 3, and formally this can be done by factoring appropriate values from the partition function expressions so that energies are measured relative to the minima on their corresponding surfaces. One can represent by $U_{\text{gas}}^{\text{QM}} + W_{\text{pol}}$ the potential energy surface for the beginning state of process step 2 and by $U_{\text{gas}}^{\text{CM}}$ that for the ending state. The solute conformations of minimum energy on these surfaces can be represented by $\tilde{r}_{\text{g,min}}^{\text{QM}}$ and $r_{\text{g,min}}^{\text{CM}}$. The appropriate energies can be factored from partition function expressions as in the following

$$\begin{aligned}\Delta G_2 &= -kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p)+pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p)+W_{\text{pol}}(r)+pV]}} \\ &= \{U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}}) - [U_{\text{gas}}^{\text{QM}}(\tilde{r}_{\text{g,min}}^{\text{QM}}) + W_{\text{pol}}(\tilde{r}_{\text{g,min}}^{\text{QM}})]\} \\ &\quad - kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p)-U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}})+pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p)+W_{\text{pol}}(r)-U_{\text{gas}}^{\text{QM}}(\tilde{r}_{\text{g,min}}^{\text{QM}})-W_{\text{pol}}(\tilde{r}_{\text{g,min}}^{\text{QM}})+pV]}} \\ &= \Delta\tilde{U}_{\text{offset}} + \Delta G_{\text{res}}^{\text{gas}}\end{aligned}$$

The energy difference between these two conformations constitutes one choice for the potential energy offset, $\Delta\tilde{U}_{\text{offset}}$. One then uses exactly the same energy offset for step 3 but with the sign reversed since the process corresponds to going from the fixed charge classical surface to the quantum surface. Therefore, when combining process steps 2 and 3, the offset values cancel. However, different choices for the offset values affect what amount of ΔG_2 is considered to be restructuring free energy.

Different approaches may be taken to evaluate $\Delta\tilde{G}_{\text{res}}^{\text{gas}}$ for process step 2. One commonly used technique, which is actually only appropriate for relatively rigid molecules that have a single thermodynamically accessible conformation and small amplitude vibrations, is to factor the gas phase partition function into translational, rotational, and vibrational components and then treat the vibrational components as uncoupled harmonic modes. The translational and volume-dependent factors cancel to leave the following expression

$$\begin{aligned}\Delta G_{\text{res}}^{\text{gas}} &= -kT\ln\frac{Q_{\text{rot}}(r_{\text{g,min}}^{\text{CM}})Q_{\text{vib}}(U_{\text{gas}}^{\text{CM}})}{Q_{\text{rot}}(\tilde{r}_{\text{g,min}}^{\text{QM}})Q_{\text{vib}}(U_{\text{gas}}^{\text{QM}}+W_{\text{pol}})}\\ &= -kT\ln\left[\prod_{j=1}^{3}\left(\frac{I_j(r_{\text{g,min}}^{\text{CM}})}{I_j(\tilde{r}_{\text{g,min}}^{\text{QM}})}\right)^{1/2}\prod_{\text{modes},k}\left(\frac{\omega_k^{\text{QM}}}{\omega_k^{\text{CM}}}\right)\right]\end{aligned} \tag{6}$$

Here, the notation emphasizes that the rotational partition functions are to use minimum energy conformations on the appropriate potential energy surfaces and that the vibrational partition functions[13] are to use Hessians computed using appropriate conformations and surfaces. In the second line, I_j represents the moment of intertia along the jth principal axis computed using a minimum energy molecular structure, and a classical form for the vibrational partition function results in a product of ratios of the normal-mode frequencies, ω_k. This expression has limited utility, but shows clearly how a change in molecular structure and vibrational frequencies between the two surfaces affects the restructuring free energy. For example, if the classical surface has *softer* modes, $\omega_k^{\text{QM}} > \omega_k^{\text{CM}}$, more phase space is thermally accessible on the classical surface and the free energy decreases in moving from the $U_{\text{gas}}^{\text{QM}} + W_{\text{pol}}$ to the $U_{\text{gas}}^{\text{CM}}$ surface. This effect is additive in the vibrational modes.

Equation 6 is not generally useful because many molecules of practical interest exhibit rotational−vibrational coupling, as well as low frequency, large amplitude, nonharmonic motions (e.g., methyl rotation), anharmonic potential surfaces, and significant coupling between modes. Another technique that is capable of addressing some of these issues in a classical context is to express the restructuring free energy as a thermodynamic average of some quantity over a Boltzmann distribution of conformations that can be generated through molecular dynamics or Metropolis Monte Carlo techniques. This is done as follows

$$\begin{aligned}\Delta G_2 &= -kT\ln\frac{\int \mathrm{d}V\,\mathrm{d}r\,\mathrm{d}p\mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p)+pV]}}{\int \mathrm{d}V\,\mathrm{d}r\,\mathrm{d}p\mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p)+W_{\text{pol}}(r)+pV]}}\\ &= -kT\ln\frac{\int \mathrm{d}V\,\mathrm{d}r\,\mathrm{d}p\mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p)+pV]}\mathrm{e}^{-\beta[U_{\text{gas}}^{\text{QM}}(r)+W_{\text{pol}}(r)-U_{\text{gas}}^{\text{CM}}(r)]}}{\int \mathrm{d}V\,\mathrm{d}r\,\mathrm{d}p\mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p)+pV]}}\\ &= [U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}}) - (U_{\text{gas}}^{\text{QM}}(r_{\text{g,min}}^{\text{QM}}) + W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}}))]\\ &\quad + kT\ln\left\langle\exp\{-\beta[(U_{\text{gas}}^{\text{QM}}(r) + W_{\text{pol}}(r)) - (U_{\text{gas}}^{\text{QM}}(r_{\text{g,min}}^{\text{QM}})\right.\\ &\quad \left. + W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}})) - (U_{\text{gas}}^{\text{CM}}(r) - U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}}))]\}\right\rangle_{NpT;\,\text{CM;\,polarized (C)}}\\ &= [U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}}) - (U_{\text{gas}}^{\text{QM}}(r_{\text{g,min}}^{\text{QM}}) + W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}}))]\\ &\quad + kT\ln\langle\mathrm{e}^{-\beta[\Delta U_{\text{gas}}^{\text{QM}}(r)-\Delta U_{\text{gas}}^{\text{CM}}(r)+W_{\text{pol}}(r)-W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}})]}\rangle_{NpT;\,\text{CM;\,polarized (C)}}\\ &= \Delta U_{\text{offset}} + \Delta G_{\text{res}}^{\text{gas}}\end{aligned} \tag{7}$$

This expression is useful if one has a set of Boltzmann distributed conformations generated using the classical (polarized (C)) potential surface, which is the easiest of the surfaces under consideration to sample extensively. The thermodynamic average involves potential energies measured relative to minima on the classical, $\Delta U_{\text{gas}}^{\text{CM}}(r) = U_{\text{gas}}^{\text{CM}}(r) - U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}})$, and quantum chemical, $\Delta U_{\text{gas}}^{\text{QM}}(r) = U_{\text{gas}}^{\text{QM}}(r) - U_{\text{gas}}^{\text{QM}}(r_{\text{g,min}}^{\text{QM}})$, potential surfaces with corresponding structures denoted $r_{\text{g,min}}^{\text{CM}}$ and $r_{\text{g,min}}^{\text{QM}}$. ($\tilde{r}_{\text{g,min}}^{\text{QM}}$ of the earlier discussion represented a minimum energy structure on the $U_{\text{gas}}^{\text{QM}} + W_{\text{pol}}$ surface rather than the $U_{\text{gas}}^{\text{QM}}$ surface, but in our experience these have been nearly the same. However, the energy offset is defined differently than in the previous expression for ΔG_2.) One can see from this expression that if the $U_{\text{gas}}^{\text{QM}} + W_{\text{pol}}$ and $U_{\text{gas}}^{\text{CM}}$ surfaces are parallel the restructuring energy vanishes.

When the thermodynamically dominant structures associated with the $U_{\text{gas}}^{\text{QM}} + W_{\text{pol}}$ and $U_{\text{gas}}^{\text{CM}}$ surfaces are similar and their vibrational frequencies are similar, one can see from eq 6 that $\Delta\tilde{G}_{\text{res}}^{\text{gas}}$ can be very small and may be neglected. However, accounting for this restructuring free energy allows for the possibility that the fixed charge force field ($U_{\text{gas}}^{\text{CM}}$), polarized and designed to be used in the context of a solvated state, produces in the gas phase a different distribution of conformational states and vibrational frequencies than the gas phase surface, even with the polarization surface (W_{pol}) included to produce a resultant surface ($U_{\text{gas}}^{\text{QM}} + W_{\text{pol}}$) that represents a polarized state for the molecule. For process step 3, the argument for neglect of the restructuring free energy is different. Here, it involves a possible conformational restructuring between the classical and quantum surfaces in the liquid phase. However, since it was designed to do so, we might expect an ideal fixed charge force field to produce a correct distribution of conformations (and frequencies) in the liquid phase, and so $\Delta G_{\text{res}}^{\text{liq}}$ can be neglected.

There may also be cases where there is significant restructuring free energy, but due only to the fact that the classical surface is *not* a good representation of the true (e.g., a high quality quantum chemical) surface in either the gas or liquid phase. In this situation, one might be able to argue that for some properties the effect of an inaccurate structure and frequency spectrum results in a large restructuring free energy in both the gas (thermodynamic cycle step 2) and the liquid phase (step 3) but that these contributions approximately cancel. This justification for the neglect of the restructuring free energy may be warranted for properties such as solvation free energy, which is more sensitive to the nature of the solvent−solute interactions than on the precise structure and vibrational motion of the solute, but these issues need further investigation.

2.6. Final Approximation. The net result is that unless there is a real and significant conformational change in going from the gas to the liquid the main difference between the solvation free energies obtained using the fixed charge model and experiment comes from the effect of the polarization cost.

$$\begin{aligned}\Delta G_{\text{solv}}^{\text{QM}} &= \Delta G_{\text{res}}^{\text{gas}} + \Delta G_{\text{pol}} + \Delta G_{\text{solv}}^{\text{CM}} + \Delta G_{\text{pol}}^{\text{liq}} + \Delta G_{\text{res}}^{\text{liq}}\\ &\approx \Delta G_{\text{solv}}^{\text{CM}} + \Delta G_{\text{pol}}\end{aligned}$$

Furthermore, the free energy associated with the polarization cost can be approximated directly by the polarization cost of the gas phase (i.e., unpolarized) equilibrium structure when the cost is relatively constant over the range of thermally accessible conformations. This is demonstrated in Appendix B for a set of molecules to be studied later in this paper, so that one is justified in using the following

$$\begin{aligned} \Delta G_{\mathrm{pol}} &\approx W_{\mathrm{pol}}(r_{\mathrm{g,min}}^{\mathrm{QM}}) \\ \Delta G_{\mathrm{solv}}^{\mathrm{QM}} &\approx \Delta G_{\mathrm{solv}}^{\mathrm{CM}} + W_{\mathrm{pol}}(r_{\mathrm{g,min}}^{\mathrm{QM}}) \end{aligned} \tag{8}$$

and one can obtain an expression for the effect of polarization cost on enthalpy by differentiating the above expression with respect to β to obtain the following

$$\begin{aligned} \Delta H_{\mathrm{pol}} &= \frac{\partial(\beta \Delta G_{\mathrm{pol}})}{\partial \beta} \\ &\approx W_{\mathrm{pol}}(r_{\mathrm{g,min}}^{\mathrm{QM}}) \end{aligned}$$

where we have assumed that the temperature dependence of all approximations is negligible.

It should also be noted that alternative thermodynamic cycles can be drawn that use different intermediate states. One such possibility is discussed in Appendix A, and analysis of this cycle results in expressions, some of which are more computationally tractable.

3. Polarization Cost For Hydration Free Energy of Fixed Charge Force Fields, Amino Acid Side Chain Analogues

The method for computing the cost of polarization on free energy resulting in eq 8 was applied to a small set of molecules described with three different force fields so that a polarization correction could be applied to the corresponding hydration free energies computed using free energy simulations. The goal of this was to understand how consideration of the polarization cost component of hydration free energy might affect the assessment of fixed charge force fields.

A set of small molecule amino acid side chain analogues (and *N*-methylacetamide, NMA) were selected for this study. The side chain set consisted of analogues for 15 of the 20 commonly occurring amino acids. (Glycine is omitted because the side chain analogue would be molecular hydrogen (H_2), and four others, Arginine, Lysine, Aspartic, and Glutamic acid were omitted because they either protonate or deprotonate at normal pH ranges and are, therefore, ionic in solution.) This set has been the subject of many previous studies and is important because force field parameters for these molecules are often the starting point for the development of the protein force fields. Table 1 shows the correspondence between the amino acids and their side chain analogues.

Because of the importance of this set of molecules, there are very precise calculated hydration free energies available for a number of force fields. We chose three force fields, representing somewhat different approaches to the formulation of the charge model. Atomic charges for these molecules that were based on an AM1-BCC procedure were taken from Mobley[15] and co-workers. Their procedure used the ANTECHAMBER[16] software package following a protocol often used with the AMBER simulation package.[17] Charges designed for use with the OPLS-AA[18−20] force field for these molecules are from Shirts[21] and co-workers and represent a set tuned to reproduce a selection of experimental bulk properties. The third set of atomic site charges we examined were taken from earlier work of Mobley[9] and co-workers. These were derived as the charges that best fit the electrostatic potential around these molecules produced by quantum chemistry calculations that used a hybrid density functional method, B3LYP,[22] with a polarizable continuum model (c-PCM)[23−25] and employing a cc-pVTZ basis set[26−28] (labeled in their publication as "B3LYP-TZ SCRF"). (We will hereafter refer to this as the B3LYP/PCM charge model.) There

TABLE 1: Selected Amino Acids and Their Small Molecule Side Chain Analogues[a]

amino acid	codes	side chain analogue
Alanine	Ala, A	methane
Asparagine	Asn, N	acetamide
Cysteine	Cys, C	methanethiol
Glutamine	Gln, Q	propionamide
Histidine	His, H	4-methylimidazole, 5-methylimidazole
Isoleucine	Ile, I	butane
Leucine	Leu, L	isobutane
Methionine	Met, M	methylethylsulfide
Phenylalanine	Phe, F	toluene
Serine	Ser, S	methanol
Threonine	Thr, T	ethanol
Tryptophan	Trp, W	3-methylindole
Tyrosine	Tyr, Y	paracresol
Valine	Val, V	propane

[a] Note that the analogue for Histidine has two isomeric forms which differ by the placement of hydrogen.

are several charge models discussed in the Mobley publication,[9] but the calculations that produced these represent the highest level of theory reported, since the values reported in this work as MP2 were actually SCF results that were mislabeled.[10] Precise hydration free energies computed using force fields that employ these three charge models in conjunction with a choice of water model were available. For the AM1-BCC charge model, used with the Generalized AMBER Force Field (GAFF) and the TIP3P water model, we used hydration free energies reported by Mobley.[15] For the OPLS-AA charge model and force field and the TIP4P-Ew water model, we used the hydration free energies reported by Hess[8] and co-workers. Finally, for the B3LYP/PCM charge model we used the hydration free energy values produced using GAFF and TIP3P reported by Mobley.[9] (We note that all hydration free energies reported in this work[9] using the TIP4P-Ew water model are in error.[10])

Gas and Liquid Phase Multipole Moments. For each molecule, the structure was first optimized for the gas phase using a B3LYP level of theory and a cc-pV(T+d)Z basis set.[26−29] Then, using these structures, MP2 calculations with an aug-cc-pV(T+d)Z basis set were performed to produce gas phase dipole and quadrupole moments and polarizabilities.[30] This protocol, level of theory, and basis set have been shown[31,32] to produce moments and polarizabilities of high accuracy, rivaling experimentally obtained values, and sufficient for the needs of this study. The GAMESS/US Quantum Chemistry package[33] was used for all of these quantum chemistry calculations.

For each of the three choices of point charge model, the multipole moments were computed using the same vacuum phase B3LYP-optimized structure as was used to obtain the gas phase multipoles and polarizabilities. The resulting dipole and quadrupole moments are shown in Tables 2 and 3. The quadrupole magnitude is evaluated as

$$\begin{aligned} |\Theta|^2 &= \sum_m |Q_{2m}|^2 = Q_{20}^2 + Q_{21c}^2 + Q_{21s}^2 + Q_{22c}^2 + Q_{22s}^2 \\ &= \Theta_{zz}^2 + \frac{4}{3}(\Theta_{xy}^2 + \Theta_{xz}^2 + \Theta_{yz}^2) + \frac{1}{3}(\Theta_{xx} - \Theta_{yy})^2 \end{aligned}$$

It is apparent from the tables that the vacuum phase MP2 and B3LYP calculations yield consistent dipole and quadrupole moments, although the quadrupoles are not in such good agreement as the dipoles. A striking feature in these two tables is that in many cases, indicated by asterisks, the charge models

TABLE 2: Dipole Moment Magnitudes for Selected Amino Acid Side Chain Analogues (and *N*-Methylacetamide, NMA)[a]

molecule	MP2 $\lvert\mu_g\rvert$	B3LYP $\lvert\mu_g\rvert$	B3LYP/PCM $\lvert\mu_l\rvert$	OPLS-AA $\lvert\mu_l\rvert$	AM1-BCC $\lvert\mu_l\rvert$
butane	0.10	0.10	0.39	0.02*	0.02*
isobutane	0.14	0.13	0.25	0.01*	0.00*
methane	0.00	0.00	0.00	0.00	0.00
propane	0.09	0.09	0.11	0.00*	0.00*
toluene	0.38	0.37	0.38	0.58	0.15*
acetamide	3.79	3.84	5.08	4.24	4.39
ethanol	1.73	1.60	2.29	2.37	1.98
methanethiol	1.55	1.54	2.27	2.33	2.05
methanol	1.69	1.61	2.17	2.09	1.97
methylethylsulfide	1.64	1.57	2.32	2.44	2.09
4-methylimidazole	3.47	3.36	4.87	3.44*	3.69
5-methylimidazole	4.05	3.92	N/A	4.18	N/A
3-methylindole	2.02	2.06	3.07	1.10*	1.81*
NMA	3.80	3.84	N/A	N/A	4.38
paracresol	1.38	1.35	1.94	2.79	1.93
propionamide	3.60	3.69	5.29	4.26	4.38

[a] Gas phase dipole moments are computed from MP2 and B3LYP electron densities. B3LYP liquid phase dipole moments were based on the B3LYP/PCM charge model[9] as discussed in the text (and are not dipole moments derived directly from B3LYP/PCM electron densities). Liquid phase dipole moments for the force fields were computed using the appropriate point charges at the gas phase (B3LYP/cc-pV(T+d)Z optimized) conformation. Asterisks indicate situations where the dipole moment determined from the charge model for the liquid phase is less than the gas phase MP2 dipole moment.

TABLE 3: Quadrupole Moment Magnitudes for Selected Amino Acid Side Chain Analogues (and *N*-Methylacetamide, NMA)[a]

molecule	MP2 $\lvert\Theta_g\rvert$	B3LYP $\lvert\Theta_g\rvert$	B3LYP/PCM $\lvert\Theta_l\rvert$	OPLS-AA $\lvert\Theta_l\rvert$	AM1-BCC $\lvert\Theta_l\rvert$
butane	0.49	0.46	0.77	0.46	0.27
isobutane	0.69	0.64	0.58	0.44	0.28
methane	0.00	0.00	0.00	0.00	0.00
propane	0.78	0.72	0.30	0.50	0.28
toluene	7.39	6.95	7.55	8.38	8.30
acetamide	6.24	6.22	7.60	6.54	6.48
ethanol	5.60	5.38	5.45	5.48	5.04
methanethiol	3.71	3.58	3.00	3.11	2.46
methanol	4.13	3.99	3.91	3.67	3.50
methylethylsulfide	4.74	4.46	5.36	6.03	4.59
4-methylimidazole	10.26	9.95	11.71	12.57	10.54
5-methylimidazole	10.28	9.86	N/A	10.26	N/A
3-methylindole	13.71	13.17	15.49	12.80	15.76
NMA	5.77	5.45	N/A	N/A	7.44
paracresol	11.55	11.23	13.47	13.28	13.34
propionamide	7.86	8.06	10.71	9.70	8.79

[a] Gas phase quadrupole moments are computed from MP2 and B3LYP electron densities. B3LYP liquid phase quadrupole moments were based on the B3LYP/PCM charge model[9] as discussed in the text (and are not quadrupole moments derived directly from B3LYP/PCM electron densities). Liquid phase quadrupole moments for the force fields were computed using the appropriate point charges at the gas phase (B3LYP/cc-pV(T+d)Z optimized) conformation.

meant to be used with the three force fields result in dipole and/or quadrupole moments that are *smaller* in magnitude than the MP2 gas phase moments. At least for the molecules in this study, this is not true for the dipole moments that are obtained for the liquid phase from B3LYP/PCM charges. Although, in principle, it is possible for the dipole moment of a particular molecular conformation to decrease in magnitude in going from the gas to the liquid phase, in general it is unphysical and probably evidence for a charge model that is too weakly polarized. In situations where a molecule has a sizable dipole moment, one expects it to polarize the surrounding medium and to produce in its vicinity a resulting (reaction) electric field that is parallel to the dipole moment. From an alternative perspective, if one considers the electric field in the region of the molecule as having been produced by some configuration of surrounding solvent molecules, one would expect the central molecule to orient so that its dipole moment is aligned parallel to this external electric field. From either perspective the external electric field will be parallel to the dipole moment of the molecule, and consequently, it will, in general, induce a dipole moment in the same direction as the permanent dipole moment. A strong example of the violation of this effect is shown for 3-methylindole, with a gas phase MP2 dipole moment of 2.02 D. The B3LYP/PCM charges give a liquid phase dipole that is over a full Debye larger. However, the OPLS-AA and AM1-BCC charge models show liquid phase dipole moments that are smaller, 1.10 and 1.81 D, respectively.

When the increase in dipole moment is large, it is not obvious what should happen to the quadrupole. It is conceivable that favorable dipole energetics could offset unfavorable, but smaller, quadrupole energetics, resulting in liquid phase quadrupole moments that could be larger, smaller, or otherwise changed without growing in magnitude. This argument changes when the dipole moment is very small, as it is with toluene. In this case, quadrupole effects may be dominant, and so a quadrupole field is induced in the surrounding medium that, in turn, produces a field that serves to polarize the molecule in a way that enhances the permanent quadrupole. This is certainly the case for molecules that have no permanent dipole moment.

The cases of methanol and methanethiol are interesting in that the dipole is increased by the polarization effect, and this is somewhat captured by the force fields. However, neither the B3LYP/PCM charges nor the force field charge models produce quadrupoles even as large as they are in the gas phase. Quadrupole moments computed directly from the B3LYP/PCM electron density for these molecules were quite different from those that were based on the B3LYP/PCM point charge model, which was derived from fitting the electrostatic potential.[9] (The fitted charges were constrained to result in a net electrically neutral molecule but were not constrained to reproduce the molecular dipole or quadrupole moments computed from the electron density.) The charge density indicates that the magnitude of the quadrupole moments should be 4.42 D-Å for methanol and 4.33 for methanethiol, compared with 3.91 and 3.00, respectively, using the fitted charges (Table 3). For these small molecules, it is probably the case that there are not enough degrees of freedom in a nuclear site-based point charge model (number of atomic sites reduced by any constraints for charge neutrality or local symmetry, such as in the hydrogen atoms of a methyl group) to reproduce both the dipole and quadrupole moments of a complex charge density with a set of reasonable point charges.

Polarization Costs for the Three Charge Models. Polarization costs computed using the gas phase polarizabilities and multipoles and the charge models described above are given in Table 4. For each of three charge models, the table lists W_{pol}^{D}, the polarization cost considering just dipole polarization computed using eq 4, as well as W_{pol}^{D+Q}, which considers both dipole and quadrupole polarization, computed using eqs 2 and 3. For the force fields and molecules considered, the polarization costs range from near zero for the alkanes to about 2 kcal/mol for amides with a maximum value of over 3 kcal/mol for propionamide using the B3LYP/PCM charge model. Note that for molecules such as acetamide and propionamide, where the B3LYP/PCM charge model has a dipole much larger than that

TABLE 4: Polarization Costs for Selected Amino Acid Side Chain Analogues (and *N*-Methylacetamide, NMA) at Their Gas Phase (B3LYP Optimized) Structures Computed Using Gas Phase Multipoles and Polarizabilities from Quantum Chemistry (MP2) and Using Liquid Phase Multipoles Computed Using the B3LYP/PCM Point Charge Model[9] as Discussed in the Text and the Force Field Point Charge Models[a]

	B3LYP/PCM		OPLS-AA		AM1-BCC	
molecule	W_{pol}^{D}	W_{pol}^{D+Q}	W_{pol}^{D}	W_{pol}^{D+Q}	W_{pol}^{D}	W_{pol}^{D+Q}
butane	0.08	0.15	0.01	0.16	0.01	0.11
isobutane	0.01	0.01	0.02	0.27	0.02	0.20
methane	0.00	0.00	0.00	0.00	0.00	0.00
propane	0.00	0.12	0.01	0.46	0.01	0.32
toluene	0.00	0.01	0.02	0.21	0.04	0.13
acetamide	1.84	2.16	0.52	0.72	0.52	0.98
ethanol	0.77	1.41	0.64	1.23	0.26	0.95
methanethiol	1.08	2.41	1.34	2.57	0.50	2.28
methanol	0.80	1.62	0.57	1.06	0.33	1.27
methylethylsulfide	0.38	0.53	0.93	1.41	0.13	0.23
4-methylimidazole	1.38	1.67	0.01	1.04	0.04	0.43
5-methylimidazole	N/A	N/A	0.10	1.12	N/A	N/A
3-methylindole	0.43	0.75	1.01	1.18	0.02	0.54
NMA	N/A	N/A	N/A	N/A	0.53	0.82
paracresol	0.24	0.51	1.08	1.50	0.15	0.39
propionamide	2.49	3.25	0.59	1.18	0.99	1.24

[a] For each of the three charge models, the first column of each pair represents the polarization cost considering only contributions from dipole polarization (eq 4) and the second column considering both dipole and quadrupole polarization (eqs 2 and 3).

of the other force fields, the polarization cost is correspondingly larger. Of the other two charge models, it is apparent in general that the dipole polarization cost computed using the OPLS-AA charge model is greater than that for AM1-BCC.

Finally, there is sometimes a large difference between polarization costs that consider only dipole polarization and those that consider also quadrupole polarization, a difference that can be as large as 1.78 kcal/mol (methanethiol with AM1-BCC). Such differences seem large in the context of a dipole correction of 0.51 kcal/mol, and examination of Table 3 shows that in this particular case the magnitude of the quadrupole moment from the AM1-BCC charge model is significantly *smaller* than that of the gas phase, suggesting, as discussed earlier, that it may be unphysical and a consequence of the fact that a nuclear site-based point charge model is not capable of representing the charge density well enough to give the quadrupole moment of methanethiol in an aqueous environment. Quadrupole contributions to polarization costs can also be quite large even if the change in the *magnitudes* of the quadrupoles are small, provided the charge model implies there is a change in the *orientation* of the quadrupolar axes. This is not conveyed in Table 3, which shows only quadrupolar magnitudes.

Hydration Free Energies Corrected for Polarization Cost. Following thermodynamic cycle 1 (Figure 1) presented earlier and subject to the approximations described in the accompanying text (see also the discussion in Appendix B), the polarization costs can be combined with computed hydration free energies or hydration enthalpies to produce quantities that can be compared more directly with experiment than those that do not account for the polarization. Using the thermodynamic cycle that results in eq 8, polarization costs from Table 4, and literature values for the hydration free energies computed using corresponding force fields and charge models, one obtains the results in Tables 5−7. Experimental values for the hydration free energy of the side chain analogues and NMA are from Wolfenden[34,35] and co-workers. More recent experimental values[36] are also

TABLE 5: Hydration Free Energies (kcal/mol) for Selected Amino Acid Side Chain Analogues from Experiment[34] and as Computed Using the B3LYP/PCM Charge Model with GAFF and a TIP3P Water Model[a]

		hydration free energy (B3LYP/PCM)		
molecule	exptl	uncorrected	dipole	quadrupole
butane	2.15	3.1	3.18	3.25
isobutane	2.28	2.5	2.51	2.51
methane	1.94	2.6	2.60	2.60
propane	1.99	2.4	2.40	2.52
toluene	−0.76	−0.8	−0.80	−0.79
MUE		0.46	0.47	0.51
acetamide	−9.68	−11.5	−9.66	−9.34
ethanol	−4.88	−4.8	−4.03	−3.39
methanethiol	−1.24	−1.7	−0.62	0.71
methanol	−5.06	−5.1	−4.30	−3.48
methylethylsulfide	−1.48	−1.2	−0.82	−0.67
4-methylimidazole	−10.27	−11.5	−10.12	−9.83
3-methylindole	−5.88	−6.6	−6.17	−5.85
paracresol	−6.11	−5.9	−5.66	−5.39
propionamide	−9.38	−11.9	−9.41	−8.65
MUE		0.82	0.43	0.90

[a] The column labeled "Uncorrected" is taken from Mobley.[9] (See Table of the Supplementary Material for this reference, in a column labeled "B3LYP/TZ SCRF".) Corrections shown include dipole polarization only, W_{pol}^{D}, or through quadrupole, W_{pol}^{D+Q}, and are from Table 4.

TABLE 6: Hydration Free Energies (kcal/mol) for Selected Amino Acid Side Chain Analogues from Experiment[34] and as Computed Using Charges from the OPLS-AA Force Field and with a TIP4P-Ew Water Model[a]

	hydration free energy (OPLS-AA)			
molecule	exptl	uncorrected	dipole	quadrupole
butane	2.15	2.89	2.90	3.05
isobutane	2.28	2.72	2.74	2.99
methane	1.94	2.27	2.27	2.27
propane	1.99	2.56	2.57	3.02
toluene	−0.76	−0.50	−0.48	−0.29
MUE		0.47	0.48	0.69
acetamide	−9.68	−8.58	−8.06	−7.86
ethanol	−4.88	−4.71	−4.07	−3.48
methanethiol	−1.24	−0.45	0.89	2.12
methanol	−5.06	−4.71	−4.14	−3.65
methylethylsulfide	−1.48	−0.05	0.88	1.36
3-methylindole	−5.88	−4.30	−3.29	−3.12
paracresol	−6.11	−5.04	−3.96	−3.54
propionamide	−9.38	−7.77	−7.18	−6.59
MUE		1.01	1.85	2.37

[a] The column labeled "Uncorrected" is taken from Hess[8] and co-workers. (See Table 3 of this reference.) Corrections shown include dipole polarization only, W_{pol}^{D}, or through quadrupole, W_{pol}^{D+Q}, and are from Table 4.

available for nine of these molecules, but these differ by 0.1 kcal/mol or less from the earlier results. Experimental uncertainties reported in these papers are generally in the range of 0.2 kcal/mol, but the agreement between the two sets of results might imply a smaller uncertainty for these molecules. Even if the uncertainty is in the range of 0.2 kcal/mol, this is still smaller than many of the polarization costs computed for these molecules.

Hydration free energies computed using the B3LYP/PCM fixed point charge model together with the Generalized AMBER Force Field (GAFF) and solvated with a TIP3P explicit water

TABLE 7: Hydration Free Energies (kcal/mol) for Selected Amino Acid Side Chain Analogues and NMA from Experiment[34,35] and as Computed Using the GAFF with Charges Derived Using the AM1-BCC Protocol and with a TIP3P Water Model[a]

	hydration free energy (AM1-BCC)			
molecule	exptl	uncorrected	dipole	quadrupole
butane	2.15	2.54	2.55	2.65
isobutane	2.28	2.74	2.76	2.94
methane	1.94	2.54	2.54	2.54
propane	1.99	2.56	2.57	2.88
toluene	−0.76	−0.71	−0.67	−0.58
MUE		0.41	0.43	0.57
acetamide	−9.68	−8.62	−8.10	−7.64
ethanol	−4.88	−3.45	−3.19	−2.50
methanethiol	−1.24	−0.26	0.24	2.02
methanol	−5.06	−3.48	−3.15	−2.21
methylethylsulfide	−1.48	0.34	0.47	0.57
4-methylimidazole	−10.27	−7.99	−7.95	−7.56
3-methylindole	−5.88	−6.55	−6.53	−6.01
NMA	−10.10	−8.39	−7.86	−7.57
paracresol	−6.11	−5.36	−5.21	−4.97
propionamide	−9.38	−9.22	−8.23	−7.98
MUE		1.24	1.59	2.05

[a] The column labeled "Uncorrected" is taken from Mobley[15] and co-workers. Corrections shown include dipole polarization only, W_{pol}^{D}, or through quadrupole, W_{pol}^{D+Q}, and are from Table 4.

model[9] are shown in Table 5; for the fixed point charge model of OPLS-AA with TIP4P-Ew[8] in Table 6; and for the AM1-BCC fixed point charge model that is recommended for use with GAFF and TIP3P[15] in Table 7. Each table shows the hydration free energies from experiment, as well as those computed using one of the charge models, and with corresponding polarization costs included from Table 4 using only the dipole term (eq 4) or including dipole−dipole, dipole−quadrupole, and quadrupole−quadrupole terms (eqs 2 and 3). The tables also include the mean unsigned error (MUE) for the computed values relative to experiment, and this is computed separately for the nonpolar and weakly polar (dipole moments less than 1 D) and for the highly polar. For the nonpolar and weakly polar molecules (butane, isobutane, methane, propane, toluene), the dipole polarization cost is small for all three charge models, as expected, and has very little effect. The situation is quite different for the more polar molecules. For the B3LYP/PCM charge model, consideration of dipole polarization *reduces* the mean unsigned error (MUE) for the hydration free energy of the polar molecules from 0.82 to 0.43 kcal/mol. In contrast, for the two other charge models, both of which are somewhat underpolarized relative to the B3LYP/PCM charges and, consequently, more insoluble, inclusion of dipole polarization makes the agreement with experiment even worse: the MUE for the OPLS-AA model goes from 1.01 to 1.85, and the MUE for the AM1-BCC model goes from 1.24 to 1.59. The degradation of results, even for the B3LYP/PCM charge model, when the quadrupole polarization is included (see the last column of Tables 5−7), is likely due to an inability of these point charge models to describe adequately the quadrupole field around the molecules.

Of particular interest are the results for acetamide, 4-methylimidazole, 3-methylindole, and propionamide, where the B3LYP/PCM charges result in much larger dipole moments than the OPLS-AA and AM1-BCC charge models and the computed hydration free energies after the polarization correction are in good agreement with experiment even though the uncorrected values are much more negative.

4. Discussion

A point that deserves discussion is whether it is meaningful to consider the quadrupole part (**A** and **C** terms) of the polarization cost in situations where the charge model is giving an unrealistic value for the magnitude of the quadrupole moment. In these cases, we recommend using eq 4, which neglects quadrupole terms in the calculation of the polarization cost for the force field. To investigate *why* point charge models might fail to produce accurate quadrupole moments, we examined point charge models obtained from least-squares fits to the electrostatic potential evaluated at points around molecules from B3LYP charge densities produced using a c-PCM implicit solvent model for water and a cc-pV(T+d)Z basis set. Using GAMESS/US, one can perform such fits to produce point charge models that are constrained to reproduce the dipole and/or quadrupole moment (as well as total charge) as obtained from the actual charge density. The fitting procedure works by optimizing a figure of merit function, which is usually sensitive to the quality of the fit of the electrostatic potential near the molecule, such as in the vicinity of the first and second solvation shells. We have seen that the dipole moments from such point charge models are generally in very close agreement to those generated from the charge density even *without* being constrained to do so. Moreover, the figure of merit used in the least-squares procedure to assess the quality of the fit does not suffer significantly when the dipole moment is constrained to agree with that from the charge density. However, in contrast to this, in many cases such point charge models do *not* produce quadrupole moments that agree well with those from the charge density, and when they are constrained to do so, there is a significant degradation in the figure of merit, implying that the fit to the electrostatic potential field in the vicinity of the molecule is being compromised to obtain the correct quadrupole moment. In some cases the difficulty in fitting the quadrupole moment is caused by having an insufficient number of degrees of freedom with a set of nuclear site-based point charges to reproduce the total charge, dipole, and quadrupole moments as well as the local field. This is the case with molecules such as methanol and methanethiol. In other cases, there may be charge density in regions "outside" of the nuclear framework of the molecule with a quadrupolar component that simply can not be described by a set of nuclear site-based point charges. For these situations, a charge model capable of a good near field fit to the potential as well as realistic dipole and quadrupole moments will probably require off-site point charges and/or some number of point dipoles or quadrupoles in addition to point charges at nuclear sites. A more thorough investigation of these effects is being performed and will be the subject of a future publication. Ideally, the charge model should be improved to give a better description of the quadrupole moment in the liquid phase, so that neglect of quadrupole polarization would not be necessary.

It has also been noted that the dipole moment magnitude for the charge model of the force field can be *less* than that determined by MP2 for the gas phase, even in cases where the gas phase dipole moment of the molecule is larger than 1 D, and for which one would actually expect a sizable *increase*. Such a choice of charge model and the resultant dipole moment may have arisen to give an uncorrected hydration free energy in close agreement with experiment. In these cases, with or without a dipole polarization correction, we would question the validity of the hydration free energy because the strength of the water−solute interaction is clearly underestimated.

5. Conclusion

We have presented a method for the computation of the effect of polarization cost on free energy inherent in the use of a force field with a fixed charge model. We have shown that when the polarization cost is fairly constant over the range of conformations that are thermally accessible at the temperatures of interest the polarization cost evaluated using the equilibrium gas phase molecular conformation can be used to make polarization corrections to *free* energies and enthalpies.

The polarization cost should be considered, not just for water (as it has been for the development of some of the water force fields) but for all molecules, before comparing computed and experimental values of certain properties, such as enthalpies of vaporization and free energies of solvation. The polarization cost is always positive. Therefore, for a force field to obtain a polarization-corrected hydration free energy that agrees with experiment, it must produce uncorrected results that appear to be too negative (i.e., too soluble in water), and it must produce enthalpies of vaporization that appear to be too positive (i.e., too strongly attractive).

Our experience suggests that the dipolar component of the polarization cost should *always* be included. However, due to the difficulties of nuclear site-based point charge models to yield accurate descriptions of the molecular quadrupole, we suggest that if there is any doubt about the accuracy of the molecular quadrupole the quadrupolar parts of the polarization cost should not be included. (Note that if quadrupolar parts are not to be included, one may use Cartesian tensor formulations of the polarization cost equations; the use of real spherical tensors is not required.)

Using our methods, polarization costs were evaluated and used to correct hydration free energies computed for small molecule amino acid side chain analogues as described by three different force fields and charge models: OPLS-AA, the Generalized AMBER Force Field (GAFF) in combination with an AM1-BCC charge model, and GAFF, but using instead a B3LYP/PCM point charge model. Comparison of the polarization-corrected results with experimental values indicates that the OPLS-AA and AM1-BCC (with GAFF) force fields are underpolarized for representing the charge density in an aqueous environment. The mean unsigned error for the *polar* side chain analogue hydration free energies computed using GAFF with the B3LYP/PCM point charges (0.43 kcal/mol) is significantly smaller than that for OPLS-AA (1.85) and GAFF with AM1-BCC charges (1.59) after the dipole polarization correction is applied. (As expected, the nonpolar and weakly polar side chain analogues do not exhibit any significant amount of polarization, so consideration of polarization energy does not substantially affect their agreement with experiment.)

These results indicate that fixed charge force fields can be improved for use in polar (e.g., aqueous) environments with a greater degree of polarization and also to the point that the molecular quadrupole moment is properly described, perhaps through the use of different charges, but more likely with the addition of extra charge sites and/or higher-order point multipoles.

Charge models based on electron densities produced using an acceptable quantum chemical method such as B3LYP or MP2, an implicit solvent model such as c-PCM, and an adequate basis set such as cc-pV(T+d)Z offer promise as a starting point for the development of improved charge models and may be better than those obtained from AM1-based approaches. This conclusion is supported by two types of observations. First, the charge model we examined that was based on B3LYP/PCM charge densities shows an increase in multipole moments, as expected, in going from the gas phase to an aqueous solvent, whereas the force field charge models sometimes show moments that are smaller than the gas phase ones, which we feel is nonphysical. (The first nonzero moment should *increase*, so molecules with small dipole moments but large quadrupole moments in the gas phase should show pronounced quadrupolar polarization upon hydration.) Second, upon examination of the hydration free energies, one can see that the charge models of both OPLS-AA and AM1-BCC result in hydration free energies that are too positive relative to experiment before applying the polarization correction, and agreement gets even worse after it is applied. In contrast, the charge model based on B3LYP/PCM, which in many cases has larger dipole moments, larger polarization corrections, and a more negative (uncorrected) hydration free energy, produces polarization-corrected hydration free energies that are in better agreement with experiment than those from either of the other two charge models.

Acknowledgment. The authors wish to thank Professor David Case for his careful reading of early drafts of these manuscripts and for many helpful suggestions. They also thank the reviewers for their comments that led to significant improvements.

Appendix A: Alternative Expressions for the Effect of Polarization Cost on Free Energy

As mentioned in Section 2, there are other thermodynamic cycles that can be used to relate experimental hydration free energies to those computed with fixed charge force fields. These differ in the choice of intermediate states used to connect the polarizable and nonpolarizable descriptions of the molecular system. This choice can affect what one computes for the effect of polarization cost on free energy. In this appendix, we will examine a second thermodynamic cycle to understand how the resulting free energy expressions differ. A key result is the development of expressions for both cycles that are computationally more tractable than those that were derived before.

A.1. Alternative Thermodynamic Cycle

Consider the thermodynamic cycle below (Figure 2).

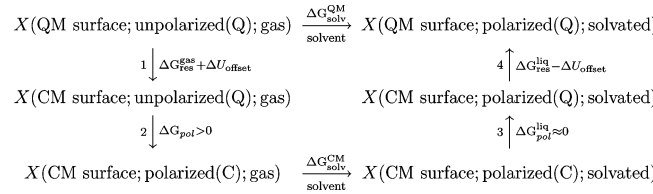

Figure 2. Thermodynamic cycle 2.

In this thermodynamic cycle (cycle 2), the process steps depicted on the top and bottom are the same as for cycle 1 from Figure 1 that was described in Section 2. However, just as for that cycle, the intermediate states in the process steps depicted on the sides are nonphysical and result in slightly different expressions and numerical values for the offset energy, for the effect of polarization cost, and for the restructuring free energies. However, the sum of the free energies involved in process steps 1, 2, 3, and 4 would agree with the corresponding sums for cycle 1 since the starting and ending states are the same. The difference is that the two cycles produce slightly different apportionments of free energy between polarization effects and restructuring effects. As before, the ultimate goal is to obtain an expression for the desired process ΔG^{QM}_{solv} in terms of ΔG^{CM}_{solv}, which is computed by classical simulations using

fixed charge force fields, with appropriate inclusion of the processes on each side of the cycle.

$$\Delta G_{\text{solv}}^{\text{QM}} = \Delta G_{\text{res}}^{\text{gas}} + \Delta G_{\text{pol}} + \Delta G_{\text{solv}}^{\text{CM}} + \Delta G_{\text{pol}}^{\text{liq}} + \Delta G_{\text{res}}^{\text{liq}}$$

A.1.1. Restructuring Free Energy, ΔG_{res}. In this thermodynamic cycle, considering steps 1 and 4, the energy offsets cancel when the two steps are considered together. As in thermodynamic cycle 1, the restructuring free energy differences capture the changes in the distribution of conformations observed using the actual (e.g., QM) rather than the fixed charge (CM) potential surface after the polarization has been taken into account. As before, for a well-designed fixed charge potential surface, the distribution of structures observed in the liquid phase would be consistent with those seen using an actual (e.g., experimental, polarizable, or QM) potential surface. Therefore, $\Delta G_{\text{res}}^{\text{liq}}$ should be close to zero. For the corresponding component of $\Delta G_{\text{res}}^{\text{gas}}$ in step 1, there may be differences in the distribution of structures observed using the fixed charge potential surface, which was designed to produce structures appropriate for a solvated molecule, and those seen in the gas phase, and so this contribution may not be zero. However, as discussed in Section 2, one expects the ratio of the gas phase partition functions to be approximately unity except in the case of large structural changes of the solute between the gas and liquid phases. In contrast to cycle 1, one should note that for this free energy difference the end state of step 1 uses the fixed charge potential surface but modified by the polarization energy to represent an *un*polarized charge model consistent with the gas phase state at the beginning.

Similar concerns, as before, are raised when there is considerable restructuring free energy in *both* the gas and liquid phases that results when the fixed charge potential surface is a poor representation of the true surfaces from a structural and/or vibrational perspective. In this situation, one hopes that the gas and liquid phase restructuring energy cancel each other and that the properties of interest are insensitive to this inaccuracy in the potential.

A.1.2. Effect of Polarization Cost on Free Energy, ΔG_{pol}. As before, for a fixed charge model that is appropriately polarized for the solvated state, $\Delta G_{\text{pol}}^{\text{liq}} \approx 0$. One must then compute ΔG_{pol}, the polarization free energy to change the charge model from an unpolarized to a polarized one in the gas phase, using the fixed charge potential surface for both the beginning and ending states. This results in the expression below

$$\begin{aligned}
\Delta G_{\text{pol}} &= G(p, T; \text{CM surface; polarized (C); gas}) \\
&\quad - G(p, T; \text{CM surface; unpolarized (Q); gas}) \\
&= -kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p)+pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p)-W_{\text{pol}}(r)+pV]}} \\
&= +kT \ln \langle e^{\beta W_{\text{pol}}(r)} \rangle_{NpT;\text{CM;polarized (C)}} \\
&= W_{\text{pol}}(r_{\text{g,min}}^{\text{CM}}) + kT \ln \langle e^{\beta[W_{\text{pol}}(r)-W_{\text{pol}}(r_{\text{g,min}}^{\text{CM}})]} \rangle_{NpT;\text{CM;polarized (C)}}
\end{aligned} \tag{9}$$

This differs from the earlier expression, eq 5, in that the thermodynamic average uses conformations generated from a simulation on the fixed charge (polarized (C)) potential surface, $U_{\text{gas}}^{\text{CM}}$, whereas eq 5 uses conformations generated on the quantum chemical (unpolarized (Q)) potential surface, $U_{\text{gas}}^{\text{QM}}$. However, the polarization cost $W_{\text{pol}}(r)$ itself is computed the same way in both cases, using polarizabilities computed from quantum chemical calculations, the difference being the choice of conformations used in the averaging process. Also, as for cycle 1, when the polarization cost is approximately constant over the set of thermally accessible conformations on the fixed charge potential surface, the second term in the expression is small, and one may use the following approximation

$$\Delta G_{\text{pol}} \approx W_{\text{pol}}(r_{\text{g,min}}^{\text{CM}}) \tag{10}$$

where $r_{\text{g,min}}^{\text{CM}}$ refers to the minimum energy structure on the polarized fixed charge potential surface. Similarly, the same expression can be used for the enthalpy of polarization.

A.2. More Easily Evaluated Free Energy Expressions

This section derives expressions for the polarization and restructuring energy for both thermodynamic cycles that are more easily evaluated because they are based on thermodynamic averages that use the classical fixed charge potential surface.

Note that the previous expression, eq 5, for the effect of polarization cost on the gas phase free energy for cycle 1 provides the free energy to change the potential surface from $U_{\text{gas}}^{\text{QM}}$ to $U_{\text{gas}}^{\text{QM}} + W_{\text{pol}}$ as a thermodynamic average over a Boltzmann distribution of conformations generated using the $U_{\text{gas}}^{\text{QM}}$ surface. In contrast, eq 9 provides the free energy for changing the potential surface from $U_{\text{gas}}^{\text{CM}} - W_{\text{pol}}$ to $U_{\text{gas}}^{\text{CM}}$, as a thermodynamic average using the $U_{\text{gas}}^{\text{CM}}$ surface. Using molecular dynamics or Monte Carlo methods, it is much easier to generate Boltzmann distributions of conformations using a fixed charge classical potential surface ($U_{\text{gas}}^{\text{CM}}$) than it is using either the quantum surface ($U_{\text{gas}}^{\text{QM}}$) or either of the two surfaces modified by the polarization cost ($U_{\text{gas}}^{\text{QM}} + W_{\text{pol}}$ or $U_{\text{gas}}^{\text{CM}} - W_{\text{pol}}$), so eq 9 is preferred from a practical perspective. This situation is reversed for the expressions for the gas phase restructuring free energy. Using cycle 1, an expression was developed, eq 7, that uses a thermodynamic average based on the fixed charge potential surface, but following the same procedure with cycle 2 produces an expression that uses a thermodynamic average based on the quantum chemical surface. However, through some extra manipulation, one may obtain an expression for the restructuring free energy for this second thermodynamic cycle that is based on the fixed charge potential surface, $U_{\text{gas}}^{\text{CM}}$

$$\begin{aligned}
\Delta G_1 &= G(p, T; \text{CM surface; unpolarized (C); gas}) \\
&\quad - G(p, T; \text{QM surface; unpolarized (Q); gas}) \\
&= -kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p)-W_{\text{pol}}(r)+pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p)+pV]}} \\
&= [U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}}) - W_{\text{pol}}(r_{\text{g,min}}^{\text{CM}}) - U_{\text{gas}}^{\text{QM}}(r_{\text{g,min}}^{\text{QM}})] \\
&\quad - kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p)-U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}})-W_{\text{pol}}(r)+W_{\text{pol}}(r_{\text{g,min}}^{\text{CM}})+pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p e^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p)-U_{\text{gas}}^{\text{QM}}(r_{\text{g,min}}^{\text{QM}})+pV]}}
\end{aligned}$$

The minimum energy conformation on the $U_{\text{gas}}^{\text{CM}} - W_{\text{pol}}$ surface is very close to that on the $U_{\text{gas}}^{\text{CM}}$ surface, represented as $r_{\text{g,min}}^{\text{CM}}$. Therefore, the first term above (in square brackets) is a good approximation to the difference in energy between the minimum on the $U_{\text{gas}}^{\text{CM}} - W_{\text{pol}}$ surface and that on the $U_{\text{gas}}^{\text{QM}}$ surface, represented as $r_{\text{g,min}}^{\text{QM}}$, and can be used as an offset

energy, ΔU_{offset}, for steps 1 and 4. This choice for the offset energy is similar to the one used in eq 7, allowing a more direct comparison of the restructuring free energies associated with the two thermodynamic cycles. This choice also results in an expression for the restructuring free energy, $\Delta G_{\text{res}}^{\text{gas}}$, that evaluates to zero if the potential surfaces $U_{\text{gas}}^{\text{CM}} - W_{\text{pol}}$ and $U_{\text{gas}}^{\text{QM}}$ are parallel since the numerator and denominator in the second term above are then equal. One may then add and subtract expressions for $G(p, T;$ CM surface; polarized (C); gas) to produce an expression for the restructuring energy in terms of thermodynamic averages that use the fixed charge potential surface

$$\begin{aligned}\Delta G_{\text{res}}^{\text{gas}} &= -kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p) - W_{\text{pol}}(r) + W_{\text{pol}}(r_{\text{g,min}}^{\text{CM}}) + pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p) + U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}}) - U_{\text{gas}}^{\text{QM}}(r_{\text{g,min}}^{\text{QM}}) + pV]}} \\ &= -kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p) - W_{\text{pol}}(r) + W_{\text{pol}}(r_{\text{g,min}}^{\text{CM}}) + pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p) + pV]}} \\ &\quad + kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p) + U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}}) - U_{\text{gas}}^{\text{QM}}(r_{\text{g,min}}^{\text{QM}}) + pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p) + pV]}}\end{aligned}$$

The first term in this expression can be related to the effect of polarization cost for the free energy associated with cycle 2. The second term can be expressed as a Boltzmann average over the fixed charge (polarized (C)) potential surface if one uses $H^{\text{QM}} = H^{\text{CM}} + (U^{\text{QM}} - U^{\text{CM}})$. The resulting expression for the restructuring free energy associated with cycle 2 is the following

$$\begin{aligned}\Delta G_{\text{res}}^{\text{gas}} &= (W_{\text{pol}}(r_{\text{g,min}}^{\text{CM}}) - \Delta G_{\text{pol}}) \\ &\quad + kT \ln \langle \mathrm{e}^{-\beta[\Delta U_{\text{gas}}^{\text{QM}}(r) - \Delta U_{\text{gas}}^{\text{CM}}(r)]} \rangle_{NpT;\text{CM;polarized (C)}}\end{aligned} \tag{11}$$

This expression is in terms of relative energies

$$\begin{aligned}\Delta U_{\text{gas}}^{\text{CM}}(r) &= U_{\text{gas}}^{\text{CM}}(r) - U_{\text{gas}}^{\text{CM}}(r_{\text{g,min}}^{\text{CM}}) \\ \Delta U_{\text{gas}}^{\text{QM}}(r) &= U_{\text{gas}}^{\text{QM}}(r) - U_{\text{gas}}^{\text{QM}}(r_{\text{g,min}}^{\text{QM}})\end{aligned}$$

Equations 9 and 11 for the effect of polarization cost on the free energy and the restructuring free energy associated with the second thermodynamic cycle are the primary results of this Appendix.

Following a similar approach leads to a more useful alternative to eq 5 for the effect of polarization cost on the free energy associated with the *first* thermodynamic cycle, cycle 1, so that it can be expressed in terms of a Boltzmann average using conformations generated using the fixed charge potential surface instead of the quantum chemical potential surface

Again, we use $H^{\text{QM}} = H^{\text{CM}} + (U^{\text{QM}} - U^{\text{CM}})$, but this time in

$$\begin{aligned}\Delta G_{\text{pol}} &= G(p, T;\text{ QM surface; polarized (C); gas}) \\ &\quad - G(p, T;\text{ QM surface; unpolarized (Q); gas}) \\ &= \{G(p, T;\text{ QM surface; polarized (C); gas}) \\ &\quad - G(p, T;\text{ CM surface; polarized (C); gas})\} \\ &\quad - \{G(p, T;\text{ QM surface; unpolarized (Q); gas}) \\ &\quad - G(p, T;\text{ CM surface; polarized (C); gas})\} \\ &= -kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p) + W_{\text{pol}}(r) + pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p) + pV]}} \\ &\quad + kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{QM}}(r,p) + pV]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p) + pV]}}\end{aligned}$$

both terms to produce two thermodynamic averages that use the fixed charge potential surface

$$\begin{aligned}\Delta G_{\text{pol}} &= -kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p) + pV]} \mathrm{e}^{-\beta[U_{\text{gas}}^{\text{QM}}(r) - U_{\text{gas}}^{\text{CM}}(r) + W_{\text{pol}}(r)]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p) + pV]}} \\ &\quad + kT \ln \frac{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p) + pV]} \mathrm{e}^{-\beta[U_{\text{gas}}^{\text{QM}}(r) - U_{\text{gas}}^{\text{CM}}(r)]}}{\int \mathrm{d}V\, \mathrm{d}r\, \mathrm{d}p \mathrm{e}^{-\beta[H_{\text{gas}}^{\text{CM}}(r,p) + pV]}} \\ &= W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}}) \\ &\quad - kT \ln \langle \mathrm{e}^{-\beta[\Delta U_{\text{gas}}^{\text{QM}}(r) - \Delta U_{\text{gas}}^{\text{CM}}(r) + W_{\text{pol}}(r) - W_{\text{pol}}(r_{\text{g,min}}^{\text{QM}})]} \rangle_{NpT;\text{CM;polarized (C)}} \\ &\quad + kT \ln \langle \mathrm{e}^{-\beta[\Delta U_{\text{gas}}^{\text{QM}}(r) - \Delta U_{\text{gas}}^{\text{CM}}(r)]} \rangle_{NpT;\text{CM;polarized (C)}}\end{aligned} \tag{12}$$

In this expression, the energies can be measured relative to any arbitrary reference on the respective potential surfaces since the effect of this choice cancels when the two terms are added. Also, the polarization cost of an arbitrary reference structure has been factored, in this case $r_{\text{g,min}}^{\text{QM}}$. This expression can be used to compute the effect of polarization cost on the free energy associated with the first thermodynamic cycle as a practical alternative to eq 5.

Appendix B: Evaluating the Effect of Polarization Cost on Free Energy With Thermodynamic Sampling

Use is made in this paper of the assumption that the polarization cost is approximately constant over the range of thermally accessible conformations on either the unpolarized quantum chemical potential surface or the polarized fixed charge classical potential surface. This assumption allows one to approximate the effect of the polarization cost in the gas phase free energy (eq 5 for thermodynamic cycle 1, shown in Figure 1, and eq 9 for cycle 2, shown in Figure 2) with the polarization cost of a single conformation, the minimum energy structure either on the unpolarized quantum chemical surface (eq 8 for cycle 1) or on the polarized classical surface (eq 10 for cycle 2). This approximation allowed the use of polarization costs (actually, *energies*) to correct hydration *free energy* results. In this Appendix we test the validity of this assumption by explicitly evaluating the dipole polarization free energy for a subset of amino acid side chain analogs studied in Section 3.

The four molecules selected for this study span a diverse subset of the side chain analogues and include an amide (acetamide), an alcohol (methanol), a sulfur-containing molecule (methylethylsulfide) and a very polar molecule, 4-methylimi-

TABLE 8: Polarization Costs and Free Energies (kcal/mol) in the Gas Phase for Four Amino Acid Side Chain Analogues[a]

molecule	acetamide		methanol		methylethylsulfide		methylimidazole	
charge model	AM1-BCC	B3LYP	AM1-BCC	B3LYP	AM1-BCC	B3LYP	AM1-BCC	B3LYP
ΔG_{pol} (cycle 1)	0.53	1.34	0.33	0.78	0.14	0.38	0.06	1.27
ΔG_{pol} (cycle 2)	0.52	1.40	0.36	0.85	0.14	0.39	0.10	1.45
W_{pol} (Q)(table)	0.52	1.84	0.33	0.80	0.13	0.38	0.04	1.38
W_{pol} (Q)	0.38	1.11	0.34	0.81	0.09	0.40	0.02	1.46
W_{pol} (C)	0.45	1.33	0.36	0.82	0.09	0.31	0.02	1.42
$\sigma(W_{pol})$	0.16	0.41	0.07	0.16	0.07	0.11	0.07	0.22

[a] In each case, the quantum chemical potential surface was computed using B3LYP with a cc-pV(T+d)Z basis set, and the fixed charge potential surface was computed using the AMBER force field but with either an AM1-BCC charge model or one derived from a B3LYP/PCM calculation.[9] ΔG_{pol} was computed using eqs 12 and 9 for cycles 1 and 2, respectively. W_{pol} are the polarization costs for specific low energy conformations.

dazole. For the force field description of these molecules we used AMBER but selected two different charge models: the AM1-BCC charge model and the model with point charges derived from B3LYP/PCM calculations.[9] These two charge models represent the most weakly and strongly polarized fixed charge representations in this study.

A set of uncorrelated and Boltzmann distributed conformations for each of the four molecules and for each of the two force fields were generated from thermally controlled molecular dynamics simulations. For the gas phase simulations, a time step size of 0.5 fs was used with a velocity−Verlet dynamical integrator,[37] and all bonds involving hydrogen were constrained using the RATTLE[38] algorithm, which enforces consistency between the constraint equations and velocities as well as the coordinates. (Constrained bond distances were converged to within 10^{-10} Å of the constraint value.) Thermal control was implemented with velocity reassignment[39] every 10 ps from a Boltzmann distribution representing a temperature of 298 K. After each velocity reassignment, the RATTLE algorithm was applied so that the constraint conditions on the velocities were satisfied. Velocity reassignment leads to faster equilibration and better decorrelation than thermal control methods based on velocity scaling because it mixes energy among vibrational modes, which is especially important for gas phase simulations. Equilibration consisted of 1 ns of simulation with this protocol, and then, with continued thermal control, molecular coordinates were saved every 10 ps during a second nanosecond of simulation to produce a set of 100 configurations. The 10 ps time between velocity reassignments was at least several periods of the slowest periodic motion of these molecules. The result was eight sets of 100 conformations (four molecules each simulated using two different force fields).

For each conformation in the set, the potential energy on the classical potential that generated it was saved, and a B3LYP energy with a cc-pV(T+d)Z basis set was computed. Also, for each conformation gas phase, multipole moments and dipole polarizabilities were computed using MP2 calculations with an aug-cc-pV(T+d)Z basis set. These were used with the dipole moments implied by the fixed charge models to compute dipole polarization costs for each conformation using eq 4.

Using this information, Boltzmann averages were performed to compute the effect of polarization cost on the free energies for cycle 1 using eq 12 and for cycle 2 using eq 9. Restructuring free energies were also computed using eqs 7 and 11 for each cycle. In these expressions, the conformations and energies with the lowest classical energies were used for $r_{g,min}^{CM}$ and $U_{gas}^{CM}(r_{g,min}^{CM})$, and the ones with the lowest B3LYP energies were used for $r_{g,min}^{QM}$ and $U_{gas}^{QM}(r_{g,min}^{QM})$. Owing to the finite sample size, these conformations do not represent true minimum energy structures on either surface, but they are close to it for the generating classical surfaces and also close for the quantum chemical surface if the classical surface is sufficiently parallel to the quantum chemical one. Equation 8 for cycle 1 and eq 10 for cycle 2 approximate the polarization free energy with the polarization cost of these low energy structures.

The results are shown in Table 8, where ΔG_{pol} was computed using eqs 12 and 9 for cycles 1 and 2, respectively. The table also includes W_{pol}, dipole polarization costs for specific low energy conformations. "W_{pol} (Q), table" represents the polarization cost for the B3LYP-optimized structure, which is also in Table 4. W_{pol} (C) and W_{pol} (Q) represent the polarization cost for the conformations with the lowest fixed charge potential energy and the lowest B3LYP energy, respectively, from the set of 100 conformations produced for the Boltzmann averaging. Also shown is the sample standard deviation, $\sigma(W_{pol})$, which conveys the degree of variability in the polarization cost observed over the set of thermally accessible conformations on the classical surface. As can be seen, there is generally good agreement in each case between the cycle 1 and cycle 2 values for the effect of polarization cost on the free energy. (As noted in Section 3, the polarization cost effect is greater with the B3LYP-based charge model than it is for the AM1-BCC charge model.)

The key observation is that for this particular set of molecules it is adequate to approximate the polarization free energy (ΔG_{pol}) with the polarization cost of a low energy structure (W_{pol}) on either the classical or quantum chemical potential surface. The largest error in doing so is for acetamide, and, in particular, for the B3LYP-based charge model. Here, the standard deviation in the polarization cost (0.41 kcal/mol) for the B3LYP-based charge model is the largest among this set of eight cases, indicating that the polarization cost is not constant over the range of thermally accessible conformations that are important for the averaging process. Furthermore, there is a large difference between the polarization cost (1.84) of the B3LYP-optimized conformation and that (1.11) of the conformation with the lowest B3LYP energy among the set of 100 conformations, indicating that the optimal B3LYP conformation is not even sampled from the simulation using the fixed charge potential surface. (The relatively large restructuring free energy, see below, is not surprising in light of this.) When the important regions of configuration space are not sufficiently sampled, one expects larger uncertainties in these measurements as well. A reasonable diagnostic for this situation would be to compute the polarization costs for both the optimal structure on the force field potential surface and that on the quantum chemical surface, and then one should perform a more careful analysis if these costs are not in good agreement.

Restructuring free energies were computed using eqs 7 and 11 for each cycle. For acetamide the restructuring energies were

−0.36 kcal/mol for cycle 1 (−0.42 for cycle 2) with the AM1-BCC charge model and 1.43 (1.59) with the B3LYP charge model. For methanol the restructuring energies were −0.55 (−0.55) with AM1-BCC and −0.29 (−0.34) with B3LYP. For methylethylsulfide the restructuring energies were 0.26 (0.26) with AM1-BCC and 0.28 (0.17) with B3LYP. For methylimidazole the restructuring energies were 1.66 (1.61) with AM1-BCC and 0.96 (0.72) with B3LYP. The agreement between the cycle 1 and cycle 2 restructuring free energies is remarkable. Positive restructuring free energies arise when use of the potential surface based on the force field results in a sampling of less configuration space than the use of one based on the quantum chemical calculations. An investigation of the origin of these effects, the degree to which they are compensated for in the condensed phase, and how this might affect computed hydration free energies is currently under study. We speculate that this is likely to be due to changes in the vibrational modes and frequencies between the two representations.

Supporting Information Available: Documents containing optimized molecular structures, multipole moments and polarizabilities for the molecules studied in this paper. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Swope, W. C.; Horn, H. W.; Rice, J. E. Accounting for Polarization Energy When Using Fixed Charge Force Fields I. Method Energy. *J. Phys. Chem. B* **2009**, DOI: 10.1021/jp911699p.

(2) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(3) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665.

(4) Abascal, J. L. F.; Vega, C. A General Purpose Model for the Condensed Phases of Water TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(5) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926.

(6) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces*; Pullmann, B., Ed.; D. Reidel Publishing: Dordrecht, 1981; p 331.

(7) Rukes, B., Ed. *Guideline on the Use of Fundamental Physical Constants and Basic Constants of Water*, 2001 ed.; The International Association for the Properties of Water and Steam, Rukes, B., president: Gaithersburg, Maryland, USA, 2001. See http://www.iapws.org (accessed December 9, 2009).

(8) Hess, B.; van der Vegt, N. F. A. Hydration Thermodynamic Properties of Amino Acid Analogues: A Systematic Comparison of Biomolecular Force Fields and Water Models. *J. Phys. Chem. B* **2006**, *110*, 17616.

(9) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. Comparison of Charge Models for Fixed-Charge Force Fields: Small-Molecule Hydration Free Energies in Explicit Solvent. *J. Phys. Chem. B* **2007**, *111*, 2242–2254, and private communications.

(10) Mobley, D. L. Erratum in reference to. Dumont, E.; Chodera, J. D.; Dill, K. A. Comparison of Charge Models for Fixed-Charge Force Fields: Small-Molecule Hydration Free Energies in Explicit Solvent. J. Phys. Chem. B 2010, submitted.

(11) We wish to emphasize that the polarization energy correction that is computed using our procedure is different from that reported by Mobley[9] and co-workers for the point charge model that they produced based on their fit to the electrostatic potentials generated by their B3LYP/cc-pVTZ/c-PCM charge densities. Our polarization energies are based on the point charges themselves or, more precisely, on the multipole moments generated by them. Mobley and co-workers determined their polarization energies from the charge density from which these charges were derived. As discussed in the companion paper,[1] these two approaches will only give the same result if the fitted charge model is capable of reproducing not only the molecule's dipole moment but the higher-order multipole moments as well. Since, in general, this is not the case, we chose to calculate the polarization energy based on the fixed charge model in a manner consistent for all the force fields. .

(12) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.

(13) One may use either classical or quantum mechanical expressions for the vibrational partition functions. Since the point of these process steps is to relate a classical molecular dynamics simulation with an experimental measurement, one should in principle include a step in the thermodynamic cycle to account for the switch from quantum dynamics to classical dynamics. These issues are outside the scope of this work, and classical dynamics and classical statistical mechanics are assumed. However, earlier work[3,14] provide examples of an approach to address this. .

(14) Horn, H. W.; Swope, W. C.; Pitera, J. W. Pitera Characterization of the TIP4P-Ew water model vapor pressure and boiling point. *J. Chem. Phys.* **2005**, *123*, 194504.

(15) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Atomistic Simulations. *J. Chem. Theory Comput.* **2009**, *5*, 350–358.

(16) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247–260.

(17) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(18) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(19) Rizzo, R. C.; Jorgensen, W. L. OPLS All-Atom Model for Amines: Resolution of the Amine Hydration Problem. *J. Am. Chem. Soc.* **1999**, *121*, 4827–4836.

(20) Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L. Gas-Phase and Liquid-State Properties of Esters, Nitriles, and Nitro Compounds with the OPLS-AA Force Field. *J. Comput. Chem.* **2001**, *22*, 1340–1352.

(21) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. Extremely Precise Free Energy Calculations of Amino Acid Side Chain Analogs: Comparison of Common Molecular Mechanics Force Fields for Proteins. *J. Chem. Phys.* **2003**, *119*, 5740.

(22) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(23) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.

(24) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the c-PCM solvation model. *J. Comput. Chem.* **2003**, *24*, 669–681.

(25) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. New Developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution. *J. Chem. Phys.* **2002**, *117*, 43.

(26) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007.

(27) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796.

(28) Woon, D. E.; Dunning, T. H., Jr. The atoms aluminum through argon. III. *J. Chem. Phys.* **1993**, *98*, 1358.

(29) Dunning, T. H., Jr.; Peterson, K. A.; Wilson, A. K. Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited. *J. Chem. Phys.* **2001**, *114*, 9244.

(30) These properties were computed relative to a reference frame coincident with the center of nuclear charge. The multipole moments were obtained from the charge density, and the polarizabilities were computed by finite difference of energies and multipole moments each with different directions of small finite perturbing electrostatic potentials (0.001 au) or potential gradients (0.001 au).

(31) Voisin, C.; Cartier, A.; Rivail, J.-L. Computation of accurate electronic molecular polarizabilities. *J. Phys. Chem.* **1992**, *96*, 7966–7971.

(32) Bak, K. L.; Gauss, J.; Helgaker, T.; Jorgensen, P.; Olsen, J. The accuracy of molecular dipole moments in standard electronic structure calculations. *Chem. Phys. Lett.* **2000**, *319*, 563–568.

(33) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M. M.; Montgomery, J. A. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347–1363. (http://www.msg.ameslab.gov/GAMESS/GAMESS.html (R5)). Modifications were made to support calculation of DFT energies and multipole moments in the presence of a field gradient.

(34) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. Affinities of Amino Acid Side Chains for Solvent Water. *Biochemistry* **1981**, *20*, 849–855.

(35) Wolfenden, R. Interaction of Peptide Bond with Solvent Water: A Vapor Phase Analysis. *Biochemistry* **1978**, *17*, 201. Hydration free energies are computed from vapor−water distribution coefficients. Coefficients of $4.1(\pm 1.3) \times 10^{-8}$ for NMA and $7.6(\pm 2.5) \times 10^{-8}$ for acetamide correspond to hydration free energies of $-10.1(\pm 0.2)$ and $-9.7(\pm 0.3)$ kcal/mol, respectively.

(36) Abraham, M. H.; Whiting, G. S.; Fuchs, R.; Chambers, E. J. Thermodynamics of Solute Transfer from Water to Hexadecane. *J. Chem. Soc., Perkin Trans.* **1990**, *2*, 291–300. Values reported are offset by 4.27 kcal/mol from ours due to use of a different reference state. This paper reports hydration free energies of 2.07 kcal/mol for *n*-butane; −5.00 for ethanol; 2.32 for isobutane; −5.10 for methanol; −1.36 for methanethiol; 2.00 for methane; −6.13 for *p*-cresol; 1.96 for propane; and −0.79 for toluene, with expected errors of approximately 0.2 kcal/mol, making them consistent with the Wolfenden data.

(37) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A Computer Simulation Method of the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters. *J. Chem. Phys.* **1982**, *76*, 637.

(38) Andersen, H. C. RATTLE: A "Velocity" Version of the SHAKE Algorithm for Molecular Dynamics Calculations. *J. Comput. Phys.* **1983**, *52*, 24–34.

(39) Andersen, H. C. Molecular Dynamics Simulations At Constant Pressure and/or Temperature. *J. Chem. Phys.* **1980**, *72*, 2384–2393.

JP911701H