Electrostatic Potentials and Free Energies of Solvation of Polar and Charged Molecules

Gerhard Hummer,* Lawrence R. Pratt, and Angel E. García

Theoretical Division, MS K710, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Bruce J. Berne

Department of Chemistry and Center for Biomolecular Simulation, Columbia University, New York, New York 10027

Steven W. Rick

Frederick Cancer Research and Development Center, National Cancer Institute, Frederick, Maryland 21702

Received: December 10, 1996; In Final Form: February 11, 1997[⊗]

Theories of solvation free energies often involve electrostatic potentials at the position of a solute charge. Simulation calculations that apply cutoffs and periodic boundary conditions based on molecular centers result in center-dependent contributions to electrostatic energies due to a systematic sorting of charges in radial shells. This sorting of charges induces a surface-charge density at the cutoff sphere or simulation-box boundary that depends on the choice of molecular centers. We identify a simple solution that gives correct, center-independent results, namely the radial integration of charge densities. Our conclusions are illustrated for a Lennard-Jones solute in water. The present results can affect the parametrization of force fields.

Accurate simulation calculations of free energies of solvation require a careful treatment of long-range electrostatic interactions. Recent computational and theoretical work on single-ion free energies[1] has converged upon a common set of ideas that are, however, discussed in slightly different ways, *i.e.*, Gaussian fluctuations of electrostatic potentials,[2] second-order perturbation theory,[3] or linear-response theory.[4,5] These approaches require the calculation of electrostatic potentials at atom positions on a solute molecule at fractional charge states (*e.g.*, uncharged or fully charged). However, a lack of consensus on how electrostatic potentials should be evaluated means that calculated partial contributions to single-ion free energies are often not fully comparable. Differences arise because of a common practice of evaluating electrostatic interactions considering whole molecules. This can lead to spurious dependences on the choice of the center of a molecule. Similar issues arose in calculations of the electrostatic potential difference of the water−vapor interface: seemingly identical calculations of electrostatic potentials can produce different final results.[6]

Discrepancies in calculated electrostatic potentials were noted recently by Åqvist and Hansson.[5] The present letter resolves the difficulties noted there. We will focus on the calculation of electrostatic potentials at the position of a solute molecule in a polar fluid, discussing the effects of different methods of summing charge interactions. This leads us to a simple, center-independent, and feasible recipe used to analyze electrostatic potentials, both in finite and infinite systems, namely spherical integration of charge densities. To illustrate our general results, we will show data for Lennard-Jones (LJ) solutes in water.

Two different center dependences will be considered (see Figure 1). The first is associated with the center of the solvent molecule denoted by M used to bin electrostatic interactions between solvent molecules and the solute molecule. The second center dependence to consider is the dependence on the solvent center P that might be used in implementing minimum-image periodic boundary conditions (PBCs) by translating a whole solvent molecule into the primary simulation box. These two centers M and P would often coincide, but they need not. The effects considered are distinct.

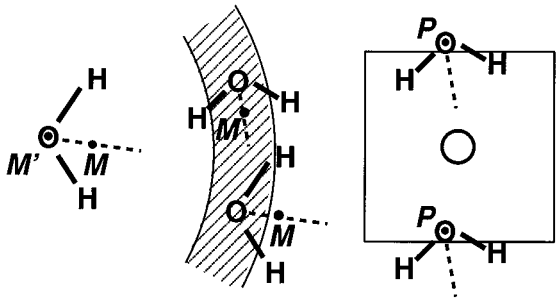

Figure 1. M- and P-center sorting of molecular partial charges. Left: Different M-centers considered for the water molecule. M and M′ coincide with the hydrogen bisector and the oxygen position. Middle: Charges of the upper molecule are counted in the shaded spherical shell (bin) but not the charges of the lower molecule. The lower molecule with an outward-pointing dipole moment is placed in a more distant bin. Right: P-center sorting, where P coincides with the oxygen position. The bottom image of the molecule is considered in the electrostatic potential calculations. For the particular choice of P = O and isotropic molecular orientations, the charge density is depleted just inside the simulation cell around the solute (outlined as square and circle, respectively) and enriched just outside.

For molecule-based summation, the electrostatic potential at the center of a spherical LJ solute molecule depends strongly on the choice of the center M of a water molecule that defines into which shell it belongs. Shown in Figure 2 are curves $\phi_M(r)$ of potential contributions of water molecules with their center M within a radius r of the solute molecule

$$\phi_M(r) = \int_0^r \mathrm{d}r \left\langle \sum_{i=1}^{N} \delta(r - r_{i,M}) \sum_{\alpha=1}^{3} \frac{q_\alpha}{r_{i,\alpha}} \right\rangle \qquad (1)$$

* Corresponding author. Phone: (505) 665-1923. Fax: (505) 665-3493. E-mail: hummer@lanl.gov.

[⊗] Abstract published in *Advance ACS Abstracts,* March 15, 1997.

S1089-5647(96)04037-0 CCC: $14.00

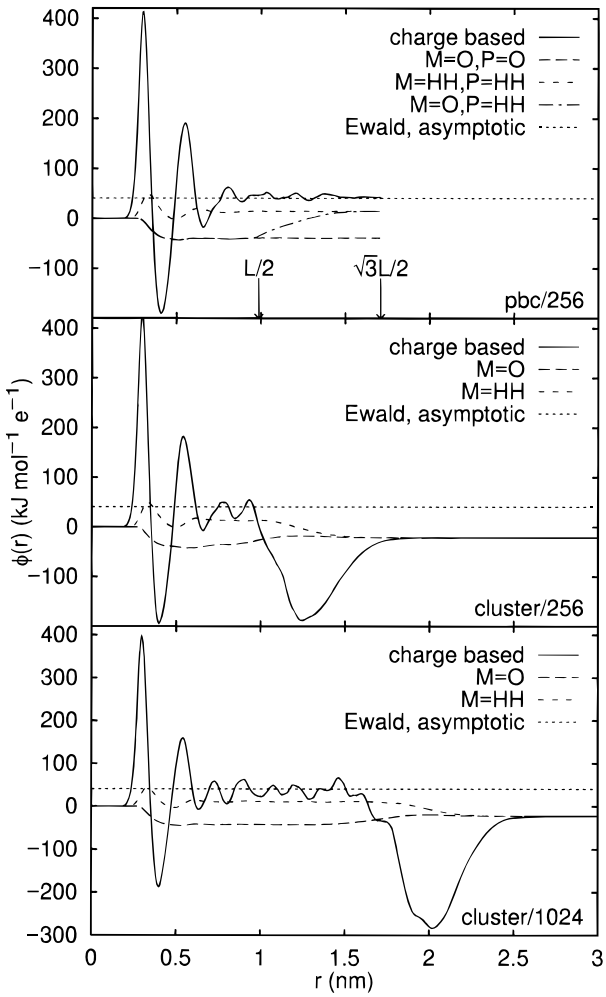

Figure 2. Integrated electrostatic potentials at the position of an uncharged LJ solute in SPC water using $1/r$ interactions. Results are shown for different ways of sorting the charges and applying PBCs (atom or molecule based). The top panel shows the results of averaging over 140 000 Monte-Carlo passes of a system with 255 SPC water molecules and one LJ solute with SPC-water LJ parameters (using Ewald summation; see ref 8 for simulation details). M and P denote the centers of sorting and applying PBCs, respectively, where O is the oxygen and HH the hydrogen-bisector position. The middle and bottom panel show the results of averaging over 100 000 and 300 000 Monte-Carlo passes of clusters of 256 and 1024 SPC water molecules, respectively, and one LJ particle at the center, again with SPC-water LJ parameters. In the cluster simulations, electrostatic interactions were calculated using $1/r$ Coulomb interactions without cutoff. The asymptotic value of charge-based integration using the Ewald potential is shown for reference.

where the α sum extends over the water oxygen atom O and hydrogen atoms H1 and H2. $\langle\cdots\rangle$ denotes a canonical ensemble average over a system of N SPC water molecules[7] with oxygen and hydrogen positions $\mathbf{r}_{i,\mathrm{O}}$, $\mathbf{r}_{i,\mathrm{H1}}$, and $\mathbf{r}_{i,\mathrm{H2}}$, respectively ($r_{i,0} = |\mathbf{r}_{i,\mathrm{O}}|$, etc.), and one uncharged LJ solute atom at position $\mathbf{r}_\mathrm{S} = 0$ with SPC-water LJ parameters. $\delta(r)$ is the Dirac delta function. q_O and q_H are the charges on the oxygen and hydrogen sites of SPC water (-0.82e and 0.41e, respectively). $\mathbf{r}_{i,\mathrm{M}}$ is the center of water molecule i, defined as $\mathbf{r}_{i,\mathrm{M}} = w\mathbf{r}_{i,\mathrm{O}} + (1-w)(\mathbf{r}_{i,\mathrm{H1}} + \mathbf{r}_{i,\mathrm{H2}})/2$. The atom positions $\mathbf{r}_{i,\mathrm{O}}$, $\mathbf{r}_{i,\mathrm{H1}}$, and $\mathbf{r}_{i,\mathrm{H2}}$ are shifted molecule-based under PBCs. (That is, the center P = M is mapped into the simulation box, leaving the molecule intact so that individual atoms can actually be outside the simulation box.) For weights $w = 1$ and 0, the center position $\mathbf{r}_{i,\mathrm{M}}$ coincides with the oxygen position and the hydrogen bisector, respectively.

The molecule-based potential defined in eq 1 contrasts with the charge-based potential $\phi_q(r)$:

$$\phi_q(r) = 4\pi \int_0^r r^2\, \mathrm{d}r\, \rho_q(r)/r \tag{2a}$$

$$\rho_q(r) = \left\langle \sum_{i=1}^{N}\sum_{\alpha=1}^{3} q_\alpha (4\pi r^2)^{-1} \delta(r - r_{i,\alpha}^{\mathrm{s}}) \right\rangle \tag{2b}$$

$\rho_q(\mathbf{r})$ is the radially averaged charge density. In eq 2b, PBCs for the positions of charges $\mathbf{r}_{i,\mathrm{O}}{}^{\mathrm{s}}$, $\mathbf{r}_{i,\mathrm{H1}}{}^{\mathrm{s}}$, and $\mathbf{r}_{i,\mathrm{H2}}{}^{\mathrm{s}}$ are applied on the basis of atoms rather than molecules.

Each of the $\phi_\mathrm{M}(r)$ curves in Figure 2 for different centers M reaches a plateau value after a 0.6−0.8 nm distance from the solute. However, the plateau values differ not only in magnitude but also in sign for different choices of M, whereas identical choices of M give agreement between simulations under PBCs and using clusters with 256 and 1024 water molecules. The differences are caused by the M-dependent sorting of molecules, even for identical configurations (positions and orientations) of the solvent molecules. If the center M is close to the oxygen atom, the first layer of molecules considered in the integration in eq 1 predominantly includes water molecules with the oxygen atoms facing the solute. Correspondingly, $\phi_\mathrm{M}(r)$ starts out negative as negative contributions of the oxygen atoms dominate. On the other hand, if the center M is close to the hydrogen atoms, the first layer of molecules considered in the integration will predominantly have the hydrogen atoms facing the solute (see Figure 1, middle). As a consequence, $\phi_\mathrm{M}(r)$ starts out positive and also reaches a positive plateau value. Results for centers M between the oxygen and the hydrogen bisector fall between the two curves.

For a finite sample, the different curves all converge to the same value when all contributions have been summed up (Figure 2, middle and bottom). Convergence is therefore reached only after crossing the interface to the exterior, so that surface-potential contributions are included. For the cluster simulations of Figure 2 (middle and bottom), the potential crosses a liquid−vacuum interface.

Similar problems arise with molecule-based cutoffs (or residue-based cutoffs for macromolecules). For instance, if the distance to the oxygen atom of a water molecule is used to determine whether a particle interacts with that water molecule, a characteristic surface-charge density is induced at the cutoff sphere. The oxygen density seen by the solute is essentially a step function. The hydrogen density is reduced just inside the cutoff and nonzero just beyond the cutoff, resulting in a net negative charge density just inside the cutoff sphere and a net positive charge density just outside. This effective surface-dipole density strongly affects the potential at the site of the particle. That effect is *independent* of the cutoff length, as the surface area and charge-dipole interaction vary with the square and the inverse square of the cutoff length, respectively.

When whole molecules are shifted under PBCs this leads to another level of ill-definition of electrostatic potentials. Shifting molecules as a whole means that PBCs are applied based on a center P with coordinates $\mathbf{r}_{i,\mathrm{P}} = u\mathbf{r}_{i,\mathrm{O}} + (1-u)(\mathbf{r}_{i,\mathrm{H1}} + \mathbf{r}_{i,\mathrm{H2}})/2$ with weight u. If that center P coincides with the center M of the ϕ_M integration, then the plateau in $\phi_\mathrm{M}(r)$ reached after the first layer of water molecules remains essentially unchanged when reaching the box boundary. However, if P and M do not coincide, $\phi_\mathrm{M}(r)$ crosses over from the M curve to the P curve when the box boundary is reached. This can be seen in Figure 2 (top) where the $\phi_\mathrm{M}(r)$ curve for M equal to O ($w = 1$, $u = 0$) crosses over to the hydrogen-bisector curve ($w = 0$, $u = 0$) when the hydrogen bisector is used for PBCs (P = HH).

Clearly, this is an unphysical behavior associated with summing electrostatic interactions and applying PBCs on the basis of molecules.

How can we eliminate these difficulties of calculating electrostatic potentials in computer simulations? The unphysical false plateaus observed for $\phi_M(r)$ in Figure 2 stem from associating partial charges with molecular centers. By choosing a center M, the water molecules were systematically sorted for analysis. For a *finite* system, integration to infinity is required to get the correct result, and that result will then contain troublesome and undesired surface-potential contributions. Under PBCs, that integration cannot be performed easily, as is manifest from the dependence of the limiting value of the potential on the choice of the molecular center P upon which PBCs are applied.

However, if we alternatively integrate over *charge densities* $\rho_q(\mathbf{r})$ rather than sum over *molecules*, we will obtain a well-defined result for the potential that coincides with taking the limit of an infinite system before extending the integral to infinity. The charge-based potential $\phi_q(r)$ is defined in eq 2a. For a finite system, eqs 1 and 2 will give identical results if the integration volume covers the whole system (extending beyond the interface to the container, vacuum *etc.*). However, unlike in eq 1 the potential $\phi_q(r)$ defined in eq 2a will reach a plateau beyond the correlation length of the charge correlation $\rho_q(r)$ independent of an arbitrary choice of the center M of a molecule. (As shown in Figure 2, that plateau is reached within about 1 nm from the neutral LJ solute. Larger correlation lengths were observed for a charged solute.[9])

These issues would be largely irrelevant with conventional Ewald treatment of electrostatic potentials, where the simulation box is replicated periodically in space. However, center dependences can arise with modifications of the standard Ewald approach. The electrostatic potentials of periodic images can be summed up using the Ewald potential $\varphi_E(\mathbf{r})$.[9,10] $\varphi_E(\mathbf{r})$ is the periodic solution of Poisson's equation $\Delta\varphi_E(\mathbf{r}) = -4\pi[\,\delta(\mathbf{r}) - 1/V\,]$ for a unit point charge and a homogeneous background in the unit cell V. The equivalents of the electrostatic potentials $\phi_M(r)$ and $\phi_q(r)$ defined in eqs 1 and 2a for periodic systems are then

$$\phi_M^E(r) = \int_0^r \mathrm{d}r \left\langle \sum_{i=1}^{N} \delta(r - r_{i,M}) \sum_{\alpha=1}^{3} q_\alpha \varphi_E(\mathbf{r}_{i,\alpha}) \right\rangle \qquad (3a)$$

$$\phi_q^E(r) = \int_0^r \mathrm{d}r \left\langle \sum_{i=1}^{N} \sum_{\alpha=1}^{3} \delta(r - r_{i,\alpha}^s) q_\alpha \varphi_E(\mathbf{r}_{i,\alpha}^s) \right\rangle \qquad (3b)$$

Again, minimum-image PBCs for $\mathbf{r}_{i,\alpha}$ and $\mathbf{r}_{i,\alpha}^s$ are applied on the basis of molecular centers P and individual atoms, respectively. Figure 3 shows that the charge-based Ewald potential and $1/r$ curves $\phi_q^E(r)$ and $\phi_q(r)$ converge but that the molecule-based curve $\phi_M^E(r)$ for periodic systems also converges to $\phi_q^E(r)$ rather than $\phi_M(r)$. This is expected because the Ewald potential is fully periodic.

Physical modifications of the Ewald potential sacrifice this periodicity. The Ewald potential is the limit of performing the lattice sum with the growing lattice embedded in a sphere cut out of a medium with infinite dielectric constant $\epsilon' = \infty$ (tin-foil boundary conditions). Total potential energies without the effect of that dielectric background $\epsilon' = \infty$ require subtraction of a term proportional to the square of the net dipole moment $\mathbf{M}$ of the simulation box.[11] Expressed as an effective potential, we can subtract a term $2\pi r^2/3V$ from $\varphi_E(\mathbf{r})$: $\varphi_{E,\epsilon'=1}(\mathbf{r}) = \varphi_E(\mathbf{r}) - 2\pi r^2/3V$. This destroys the periodicity. Use of the modified potential $\varphi_{E,\epsilon'=1}(\mathbf{r})$ in eq 3a forces $\phi_M^E(r)$ to converge to $\phi_M(r)$, as shown in Figure 3. However, the result for the potential $\phi_M^E(r)$ at the solute site then again depends on the particular choice of the molecular center P upon which PBCs are applied. Clearly, to reproduce the nonphysical effects of integrating the potential using $1/r$ with molecule-based sorting requires subtraction of a nonperiodic term from the Ewald potential and application of the potential outside the "universe," *i.e.*, the simulation box.

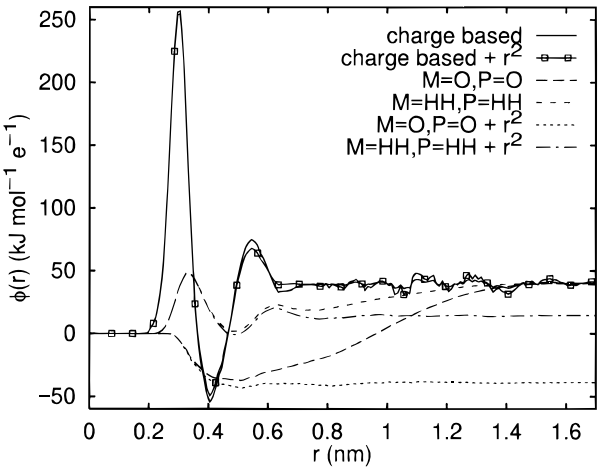

Figure 3. Integrated electrostatic potential at the position of an uncharged LJ solute in SPC water using the Ewald potential $\varphi_E(\mathbf{r})$ instead of $1/r$. Results are shown for charge and molecule-based integration with and without the r^2 modification added to $\varphi_E(\mathbf{r})$. See Figure 2 for further details.

It must be noted that subtracting the r^2 term from the Ewald potential has little effect if the integration is based on charges rather than molecules (Figure 3). We also emphasize that changing the dielectric background to a finite value $\epsilon' < \infty$ in the Ewald sum should not affect the charging of an ion at the center of the box. The dipolar field induced by a background ϵ' beyond a spherical cavity around $\mathbf{r} = 0$ is proportional to $\mathbf{r}\cdot\mathbf{M}$, which is zero at the position $\mathbf{r} = 0$ of the uncharged particle. When a point multipole is charged from zero, that contribution is also zero because of averaging over all orientations.

The results of this paper explain the differences in the sign of the electrostatic potential at the position of an uncharged LJ particle in water between Åqvist and Hansson[5] (oxygen center, $1/r$: negative potential), Rick and Berne[12] (oxygen center, Ewald with r^2 modification: negative potential), and Pratt *et al.*[3] as well as Hummer *et al.*[9] (charge based; Ewald, $1/r$ and a generalized reaction-field interaction: positive potential). The best current value for that potential is *positive*. In that context a re-examination of several results regarding free energies of charged species might be worthwhile. For instance, free energies of anions were found to be less negative in ref 13 than in ref 9 but more negative for cations. That can be explained if molecule-based summation has been used in ref 13 using a center M at or close to the oxygen atom of water. The present results also affect the parametrization of force fields involving charged species. Finally, we emphasize that the errors induced by molecule-based summation are *independent* of the cutoff length for sufficiently large cutoffs. If the induced surface-charge distribution were symmetrically distributed on a spherical shell, then it follows from Gauss's law that the correction to the induced electrostatic potential inside the spherical shell would be a constant. In that case, the contributions of M-dependent sorting would cancel each other for an overall neutral, polar solute but not for a solute with a net charge.

Our results suggest that these issues are primarily matters of analysis of configurational simulation data. A variety of methods may be used to obtain the configurational data. The center dependences considered here are introduced by the analysis of electrostatic potentials and are often larger than the secondary differences in the configurational data due to variations in their production.

The following general recipe for electrostatic-potential calculations emerges: (1) Electrostatic interactions should be integrated on the basis of *charge densities* rather than individual molecules to give correct results for atoms and molecules carrying point charges or spatially extended charge distributions. For molecule-based summation, the calculated potentials $\phi(r)$ level out nicely but the plateau values depend on the arbitrary choice of molecular centers. (2) In simulations using PBCs, all charges should be mapped into the simulation box. Molecule-based PBCs result in center-dependent surface-charge densities. (3) Under PBCs, Ewald summation provides an accurate way of summing up all interactions, minimizing finite-size effects.

Note Added in Proof. Ashbaugh and Wood[14] come to similar conclusions regarding molecule-center dependences of electrostatic potentials in their comparison of Ewald summation[9] and cutoff calculations.[15] In particular, these authors also find the potential to be positive for a neutral LJ solute in water.

References and Notes

(1) Guggenheim, E. A. *Thermodynamics. An Advanced Treatment for Chemists and Physicists*; John Wiley & Sons: New York, 1967; Chapters 8.02 and 9.03.
(2) Levy, R. M.; Belhadj, M.; Kitchen, D. B. *J. Chem. Phys.* **1991**, *95*, 3627.
(3) Pratt, L. R.; Hummer, G.; García, A. E. *Biophys. Chem.* **1994**, *51*, 147.
(4) Jayaram, B.; Fine, R.; Sharp, K.; Honig, B. *J. Phys. Chem.* **1989**, *93*, 4320.
(5) Åqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512.
(6) Wilson, M. A.; Pohorille, A.; Pratt, L. R. *J. Chem. Phys.* **1989**, *90*, 5211.
(7) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces: Proceedings of the 14th Jerusalem Symposium on Quantum Chemistry and Biochemistry*; Pullman, B., Ed.; Reidel: Dordrecht, Holland, 1981; pp 331−342.
(8) Hummer, G.; Pratt, L. R.; García, A. E. *J. Phys. Chem.* **1995**, *99*, 14188.
(9) Hummer, G.; Pratt, L. R.; García, A. E. *J. Phys. Chem.* **1996**, *100*, 1206.
(10) Figueirido, F.; Del Buono, G. S.; Levy, R. M. *J. Chem. Phys.* **1995**, *103*, 6133.
(11) de Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. London A* **1980**, *373*, 27.
(12) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1994**, *116*, 3949.
(13) Straatsma, T. P.; Berendsen, H. J. C. *J. Chem. Phys.* **1988**, *89*, 5876.
(14) Ashbaugh, H. S.; Wood, R. H. *J. Chem. Phys.*, in press.
(15) Wood, R. H. *J. Chem. Phys.* **1995**, *103*, 6177.