COMMENTS

Analysis of Electrostatic Potential Truncation Schemes in Simulations of Polar Solvents

Johan Åqvist* and Tomas Hansson

Department of Molecular Biology, Uppsala University, Biomedical Centre, Box 590, S-751 24 Uppsala, Sweden

The treatment of long-range electrostatic interactions in computer simulations of liquid molecular systems is a central issue for accurate calculations of solvation free energies. Since truly macroscopic system sizes cannot be used in atomic level simulations, it is necessary to consider a limited number of particles and instead employ some suitable type of boundary conditions. It is then common to use periodic boundary conditions (PBC) together with either a Born-type correction, reaction field treatment, or Ewald summation to account for finite size effects or to use finite nonperiodic spherical systems with restrained boundaries also in combination with a reaction field or Born-type correction.[1]

For the important test case of charging an ion in pure water, it has recently become evident that different ways of calculating the electrostatic potential at the ion site leads to different results even for the "uncharged state", i.e., when the solute is an uncharged Lennard-Jones (LJ) sphere. In this case the solute exerts no electrostatic field on the solvent, and if the true potential $\langle\phi\rangle$ on the solute is nonzero, this must be due to a nonrandom orientation of the solvent *close to* the solute, since the dipolar field falls off rapidly and water far away must be randomly oriented with respect to the solute. However, when averaging over configurations and summing $\langle\phi\rangle$ as a function of the distance, r, from the uncharged solute, the total potential on it converges to different values depending on the summation procedure used. The differences do thus not originate from different solute−water radial distribution functions but from the boundary definition associated with r, i.e., what particles are considered to contribute to $\langle\phi\rangle$ at a given distance. Some authors have hence obtained a value of $\langle\phi\rangle \simeq -9$ kcal mol^{-1} e^{-1},[2,3] while others[4] arrived at $\langle\phi\rangle \simeq +10$ kcal mol^{-1} e^{-1} for exactly the same force field parameters, corresponding to a neutral LJ sphere with the SPC oxygen parameters[5] solvated in SPC water.

Recently, Hummer et al.[4] as well as Ashbaugh and Wood[6] have set out to resolve this fundamental problem. In both cases the authors correctly concluded that if the potential is summed on a *per molecule* basis rather than on a *per particle* basis, then its value depends on the choice of center of the water molecule-(s). That is, if we consider a cutoff distance R_c, *per molecule* (M) summation means that if a particular molecular center is inside R_c then the potential from that entire molecule is counted (otherwise not), and the total potential on the solute will then depend on the choice of center within the reference frame of each water (solvent) molecule. This is, in fact, the summation scheme used in most simulations involving cutoffs for intuitive reasons, viz. to avoid creation of net charge within the cutoff sphere by "splitting" dipolar groups (which causes oscillations
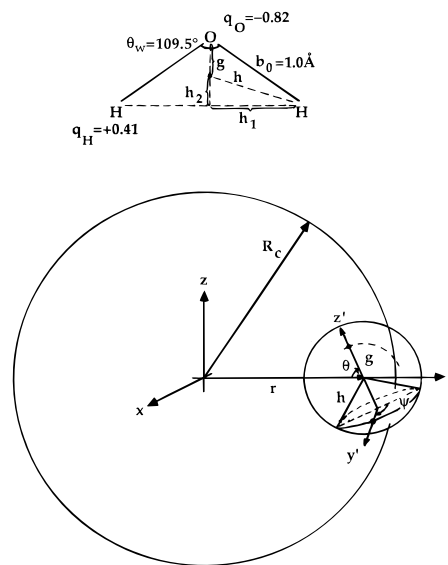
Figure 1. Geometry of the SPC water molecule[5] and of the integration problem corresponding to eq 1. The center of the water molecule is at a distance r from the origin, the oxygen is at a distance g, and the two hydrogens are at a distance h (with fixed angle to the oxygen) from the molecular center. The oxygen position is given by r and θ and one hydrogen position by r, θ, and ψ.

of $\langle\phi\rangle$ with r as shown in Figure 2 of ref 4). The other alternative, *per particle* (P) summation, corresponds to strictly applying the cutoff to particles and allowing fragments of water (solvent) molecules to contribute to ϕ at a given r.

In their paper, Hummer et al.[4] show that P-summation converges to the same value $\langle\phi\rangle \simeq +10$ kcal mol^{-1} e^{-1} as that obtained by the Ewald method with conducting boundary conditions ($\epsilon = \infty$), while M-summation generally does not. The authors therefore end up recommending P-summation as a "general recipe" for electrostatic potential calculations. Ashbaugh and Wood,[6] on the other hand, attempt to calculate the error ($\Delta\phi$) induced by choosing the oxygen as center for the water molecule in M-summation, using an analytical formulation. They obtain a value of $\Delta\phi = -19$ kcal mol^{-1} e^{-1}, which is exactly the observed difference quoted above, $-9 - (+10)$ kcal mol^{-1} e^{-1}. This would then again suggest that the correct answer to the problem is provided by the P or Ewald summation result, namely $\langle\phi\rangle = +10$ kcal mol^{-1} e^{-1}.

In this Comment, we will show (i) that P-summation, in fact, induces an artificial dipole density on the cutoff sphere which yields an error in $\langle\phi\rangle$, (ii) that the equation used by Ashbaugh and Wood also contains an error with similar effect, and (iii) that there is a "best center" in M-summation which yields $\Delta\phi = 0$.

First we calculate the error ($\Delta\phi$) in the potential on an uncharged solute in SPC water caused by P-truncation at a cutoff distance R_c from the solute. It is then assumed that R_c is large enough for water molecules at that distance to be orientationally uncorrelated or random with respect to the solute. The geometry of the problem is shown in Figure 1. We start by choosing the dipolar center of the water molecule as a reference point (this

3838 *J. Phys. Chem. B, Vol. 102, No. 19, 1998*

Comments

choice will later be shown to be the appropriate one), i.e., such that $g = h_2$. Since h is the largest distance between the molecular center and an atom, water molecules with their center at $r < R_c - h$ from the origin will contribute to ϕ with all their atoms. Conversely, if $r > R_c + h$, none of their atoms will contribute. Note that any discrepancy between M- and P-summation will arise from water molecules in the $R_c - h \leq r \leq R_c + h$ spherical shell at the surface of the cutoff sphere. Water molecules closer to the center are treated equally in the two schemes, since all their three atoms then contribute to $\langle \phi \rangle$. Likewise, molecules farther away from the center are treated equally and give zero contribution.

Now, for $R_c - h \leq r \leq R_c + h$ a water molecule may be only partially inside the cutoff sphere. If the potential from all its atoms were to be summed and averaged over all orientations, its contribution to ϕ would again be zero. However, by applying the cutoff *per particle* and integrating the resulting potential over all molecules in the shell $R_c - h \leq r \leq R_c + h$ and over all their orientations we obtain

$$\Delta\phi = \frac{\rho_w}{\epsilon_0} \left\{ \frac{q_O}{2} \int_{R_c-h}^{R_c+h} \int_0^{\theta_{max}} \frac{r^2 \sin\theta \, d\theta \, dr}{\sqrt{r^2 + g^2 - 2rg\cos\theta}} + \right.$$

$$\frac{q_H}{\pi} \int_{R_c-h}^{R_c+h} \int_0^{\pi} \int_{\psi_{min}}^{\pi} \times$$

$$\left. \frac{r^2 \sin\theta \, d\psi \, d\theta \, dr}{\sqrt{h_1^2 + h_2^2 + r^2 + 2r(h_1 \sin\theta \cos\psi + h_2 \cos\theta)}} \right\} \quad (1)$$

where the first term is from the oxygen and the second from the two hydrogens. The oxygen rotation integral involves the cutoff through θ_{max} which is defined by

$$\cos\theta_{max} = \begin{cases} \pi & r < R_c - g \\ \dfrac{r^2 + g^2 - R_c^2}{2rg} & R_c - g \leq r \leq R_c + g \\ 0 & r > R_c + g \end{cases} \quad (2)$$

Likewise, the hydrogen ψ-rotation integral involves the cutoff through ψ_{min} which is defined by

$$\cos\psi_{min} = \begin{cases} 0 \\ \dfrac{R_c^2 - h_1^2 - h_2^2 - r^2 - 2rh_2 \cos\theta}{2rh_1 \sin\theta} \\ \pi \end{cases} \quad (3)$$

where $\cos\psi_{min} = 0$ or π corresponds to the "hydrogen cone" (cf. Figure 1) being entirely inside and outside the cutoff sphere, respectively, while if it intersects the sphere $\cos\psi_{min}$ is given by the geometric term.

The first of these integrals (the oxygen contribution) can be solved analytically to give

$$\Delta\phi_O = -\frac{\rho_w q_H}{\epsilon_0} \left\{ 2R_c g - \frac{2}{3} g^2 + (R_c - g)^2 - (R_c - h)^2 \right\} \quad (4)$$

where we have utilized $q_O = -2q_H$. It may be noted that the dependence of $\Delta\phi_O$ on R_c here is $-2\rho_w q_H R_c/\epsilon_0$. The second (hydrogen) integral is elliptic since $1/(1 + a\cos\psi)^{1/2} \sim 1/(1 - b^2 \sin^2 \psi/2)^{1/2}$ but can be conveniently solved numerically. One then finds, as expected, that its dependence on R_c exactly cancels that of $\Delta\phi_O$ so that $\Delta\phi = \Delta\phi_O + \Delta\phi_H$ does not depend on the cutoff. Furthermore, if the cutoff is not applied, corresponding to $\theta_{max} = \pi$ and $\psi_{min} = 0$ in eq 1, $\Delta\phi$ indeed
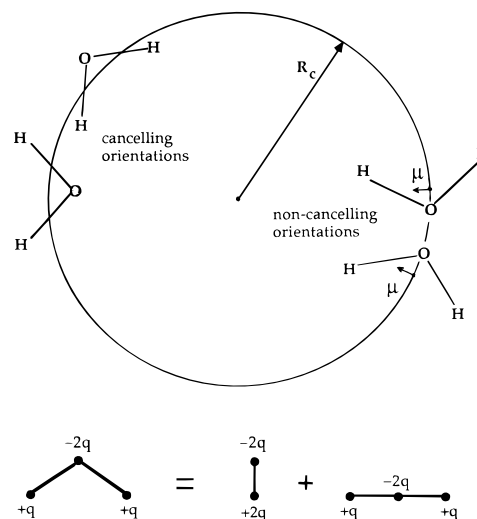
Figure 2. Schematic view of how some water orientations can give rise to a surface dipole density on the cutoff sphere due to splitting of the quadrupole when P-summation is employed.

becomes identically zero. However, application of the cutoff and insertion of the SPC water geometric parameters into eqs 1–4 yields a value of $\Delta\phi = +12.7$ kcal mol^{-1} e^{-1} independent of R_c. Since $\Delta\phi$ does not depend on the cutoff, and since the case with $\theta_{max} = \pi$ and $\psi_{min} = 0$ actually corresponds to M-summation based on the molecular rotation center (with a cutoff of $R_c + h$), what we have calculated is in fact the difference between P-summation and M-summation based on this center [i.e., the dipolar center with $g = h_2 = (b_0/2)\cos(\theta_w/2)$]. The reason for chosing this center is simple: the SPC water molecule can be represented by a superposition of a two-point dipole and a linear three-point quadrupole. The potential from the latter vanishes as $1/r^3$, i.e., more rapdily than the number of molecules in a surface shell increases, and gives no contribution to $\langle\phi\rangle$ at large distances. This can be easily verified from computer simulations where replacement of water molecules far away from the solute (after their configurations have been generated) by two-point dipoles ($\mu = 2q_H b_0 \cos\theta_w/2$) yields a quite accurate approximation of $\langle\phi\rangle$ compared to that calculated from the all-atom configurations. For example, replacing water molecules beyond 14 Å from the solute and calculating the potential up to 16 Å (100 ps PBC simulation with box length 34.15 Å) yields less than a 1% error in $\langle\phi\rangle$. Now, due to the symmetry of a simple two-point dipole there is no problem with P-summation in this case, and eq 1 will reduce to two identical and opposite terms giving $\Delta\phi = 0$. So, when the valid approximation of retaining only the dipole moment of water molecules far away is invoked, M-summation based on the dipolar center, which we will refer to as M,μ-summation, and P-summation give the same result.

However, we find that employing P-summation strictly to the three charge centers of the water molecule, when calculating the electrostatic potential at the solute site, yields an artificial contribution of $+12.7$ kcal mol^{-1} e^{-1} from the *surface* of the cutoff sphere. This corresponds to a "P-summation-induced" surface dipole density pointing with its positive end inward, and it is basically caused by destroying the intramolecular hydrogen–hydrogen correlation at the surface as schematically in Figure 2. For example, if we consider water molecules with their C_2-axis tangential to the (cutoff) sphere surface the contribution to $\langle\phi\rangle$ from their dipole moment is zero and that from their quadrupole moment will vanish at large r, so that the overall contribution to the potential becomes zero.

Comments

J. Phys. Chem. B, Vol. 102, No. 19, 1998 **3839**

The effect of P-summation is (on average) to split the quadrupole of such molecules into two dipoles of opposite direction, where only one them is allowed to contribute to $\langle\phi\rangle$. The result is to replace the (vanishing) quadrupole potential by a (nonvanishing, since the number of such molecules is proportional to r^2) artificial dipole potential. Hence, even if the cutoff would exactly preserve charge neutrality, this would then be due to charge cancellation between atoms in different molecules rather than within the same molecules. The above discussion probably also explains why (P-based) Ewald summation yields the same error, since the induced dipole density will then persist at "infinity" even if nonconducting boundary conditions are used. In other words, the Ewald boundary at infinity is then not physical in the case of an asymmetric molecule such as water, since it consists of partial molecules with the wrong dipole moment. It would also seem rather awkward if the true value of $\langle\phi\rangle$ on an uncharged solute were determined by partial solvent molecules far away (the boundary) and not by the first solvation shells. Such a strange result is, however, the inescapable consequence of P-summation since $\langle\phi\rangle$ then is totally dominated by the boundary.

The above analysis shows that the error caused by P-summation is $\Delta\phi = +12.7$ kcal mol^{-1} e^{-1} which, when used to correct the absolute value of the potential, gives $\langle\phi\rangle = -2.7$ kcal mol^{-1} e^{-1} (using -10 for the P-summation result). This value, in turn, can be used to estimate the error caused by M-summation with the oxygen as water center: $\Delta\phi_{M,O} = -9 - (-2.7) = -6.3$ kcal mol^{-1} e^{-1}. This is also what is obtained from eq 1 if ψ_{min} is set to zero and the θ-limit set to θ_{max} in the second term. The best estimate for the correct value of $\langle\phi\rangle$ actually ranges between -3.7 and -2.7 due to differences in reported simulation results. It is clear that the above value of $\Delta\phi_{M,O}$ does not agree with that given by Ashbaugh and Wood.[6] Let us therefore examine their equation (eq 6 of ref 6) in more detail.

The central point in the Ashbaugh and Wood formulation is to use a distribution function that describes the probability that an atom at a distance δ from the M-summation center will contribute to ϕ. This distribution function is

$$p(r, \delta) = \begin{cases} 1 & r < R_c - \delta \\ \dfrac{1}{2} + \dfrac{R_c^2 - \delta^2 - r^2}{4\delta r} & R_c - \delta \le r \le R_c + \delta \\ 0 & r > R_c + \delta \end{cases} \quad (5)$$

where r is the distance from the uncharged solute to the atom. Let us now again replace the potential from a three-point SPC water molecule by that from a two-point dipole which, as demonstrated above, is a valid approximation for large r. Consider a molecular center as in Figure 1 on the water C_2-axis with distances g and h_2 to the ends of the two-point dipole. The potential from the positive and negative ends of the dipole can now be integrated independently, multiplied by the relevant probability distribution, since the orientation at large r is random. The (error) potential created by the surface layer around R_c is

$$\Delta\phi_M = \frac{2\rho_w q_H}{\epsilon_0}\Big\{-\int_{R_c-\max(g,h_2)}^{R_c+\max(g,h_2)} p(r, g)r\, dr +$$

$$\int_{R_c-\max(g,h_2)}^{R_c+\max(g,h_2)} p(r, h_2)r\, dr\Big\} \quad (6)$$

Solving these integrals gives the result

$$\Delta\phi_M = \frac{\rho_w q_H}{3\epsilon_0}(g^2 - h_2^2) \quad (7)$$

Ashbaugh and Wood considered the case with the water oxygen as molecular center, i.e., $g = 0$, $h_2 = b_0 \cos\theta_w/2$, in which case we get

$$\Delta\phi_{M,O} = -\frac{\rho_w q_H}{3\epsilon_0}b_0^2 \cos^2\frac{\theta_w}{2} \quad (8)$$

Hence the factor $\cos^2(\theta_w/2)$ is missing in eq 6 of ref 6. It can, in fact, be trivially seen that the result of Ashbaugh and Wood corresponds to the solution for a dipole of magnitude $\mu = 2q_H b_0$ which is *not* the dipole moment of an SPC water molecule. Again, the intramolecular H–H correlation must be taken into account (in this case through $\cos^2(\theta_w/2)$), and both hydrogens cannot be treated as independent particles in addition to the oxygen.

Insertion of the SPC parameters into eq 8 gives $\Delta\phi_{M,O} = -6.3$ kcal mol^{-1} e^{-1} in perfect agreement with our result above. It is also immediately evident from eq 7 that the "best center" for which $\Delta\phi_M = 0$ is the dipolar center $[g = h_2 = (b_0/2)\cos(\theta_w/2)]$. With this choice the cutoff induces no contribution to the total potential from the surface of the R_c sphere. Again, the reason for this is that the average of the leading dipole term from the water molecules averages to zero, while the quadrupole term vanishes at large distances due to the $1/r^3$ dependence. Integration of the potential using this molecular center from a 100 ps molecular dynamics simulation with PBC and a cubic box length of 34.15 Å indeed gives a value of $\langle\phi\rangle = -3$ kcal mol^{-1} e^{-1}.

It can be noted here that by treating the three charge centers of a water molecule as independent particles the charge density inside R_c is not changed, while outside of the cutoff it becomes identically zero. It could then seem tempting to argue that since the potential from the outside is zero and the charge density inside unchanged, such a treatment must always give the correct result. This is true for linear dipoles where, as we have seen above, the result is the same as that obtained with a molecular cutoff applied to the dipolar center (M,μ-summation). In the latter case, the charge densities are reduced inside the cutoff while increased outside relative to P-summation. The positive and negative densities are, however, changed in exactly the same way wherefore no surface dipole is induced. For a three-center water molecule, on the other hand, as long as the distance from the M-center to the hydrogens is larger than to the oxygen, the positive density will be more reduced inside R_c and more increased outside of it compared to the negative density. This is, for example, true in the case of M,μ-summation where $h > g$ (cf. Figure 1). Thus, *compared to P-summation*, there will be a net dipole moment pointing with its negative end inward if M,μ-summation instead is used. However, the error associated with this dipole moment resides in the P-summation procedure and not in the M,μ-summation. It exactly compensates for the fact that splitting the linear quadrupoles (see above) on a particle basis inevitably produces a net dipole moment (positive end inward) which, by any definition, should not exist. Hence, while the dipole component of the potential can be treated by P-summation the quadrupole cannot.

The formula proposed in ref 6 (eq 6a therein) also leads to the conclusion that the "correct" center for M-summation is situated 0.866 Å away from the oxygen in the water dipole direction, i.e., "outside" the molecule. Indeed, one finds that M-summation using this center yields $\langle\phi\rangle = +9.3$ kcal mol^{-1}

e^{-1} as *P*-summation also does. Furthermore, replacing water molecules far away from the solute with two-point dipoles (after generating the configurations) yields the very same result if *M*-summation with the above center is used. This shows again that the two-point dipole approximation is valid for large distances (>14 Å was used in this case). However, we then arrive at a contradiction since for two-point dipoles even the equation in ref 6 says that the dipolar center (*M*,*μ*-summation) is the correct one. Thus, the above value of $\langle \phi \rangle$ must be wrong, and the error can again be seen to be +12.7 kcal mol^{-1} e^{-1} from eq 7.

So, for the problem of calculating $\langle \phi \rangle$ from (SPC) water on an uncharged solute, we can conclude that the best current value for that potential is *negative*. In particular, for a solute with LJ parameters equal to those of the SPC oxygen, we obtain a value of ~ -3 kcal mol^{-1} e^{-1}. It should be noted that this is also in agreement with the finding that the dipole vector of first shell waters is on average pointing outward from the solute,[2,3b,7,8] for which we have given an explanation earlier.[2] We have not discussed any finite-sphere simulations here, but it is worth pointing out that our obtained value of $\langle \phi \rangle$ on an uncharged solute may be useful for calibrating the surface restraints on such models. It appears, for example, that the SCAAS model[1e] presently has somewhat too weak surface polarization restraints since it gives a slightly too negative value of $\langle \phi \rangle$.[2] We have recently found that a revised spherical model (Marelius and Åqvist, in preparation) that more closely reproduces the randomly distributed surface water orientation (expected for a sphere that is part of larger system) gives convergence to $\langle \phi \rangle \simeq -3.5$ kcal mol^{-1} e^{-1} for several different sphere sizes. This therefore indicates that finite systems (clusters) with random surface polarization also converge to the correct value of $\langle \phi \rangle$. It should be noted here that the cluster simulations of ref 4 do not seem have involved any surface polarization restraints. In such a case hydrogens at the surface (and the dipole vector) tend to point outward, due to H bonding and the curvature of the hydrophobic boundary,[2] resulting in a more negative value of $\langle \phi \rangle$.

Finally, it is interesting to note that Figure 3 of ref 4 clearly indicates that the $\langle \phi \rangle$-dependence on summation procedure for the Ewald potential with the r^2 term (insulating boundary) is the same as that resulting from simulations with cutoff. This thus suggests that also Ewald calculations should employ *M*,*μ*-summation together with the insulating boundary. With a conducting Ewald boundary the system becomes truly periodic and infinite, but then the molecules at the "infinity boundary" will still be fragmented (as they are in the central box near the walls) and $\langle \phi \rangle$ becomes insensitive to summation procedure and always contains the +12.7 kcal mol^{-1} e^{-1} error. By using the insulating boundary and *M*,*μ*-summation, one can assure that the system ends with "whole molecules" without inducing any surface dipole density.

In summary, utilizing *P*-summation cannot be recommended in molecular simulations of asymmetric polar solvents since it has been shown to yield an artificial surface dipole density on the cutoff sphere. Instead, it appears that the method of choice should be *M*-summation together with the appropriate molecular center.

References and Notes

(1) (a) Straatsma, T. P.; Berendsen, H. J. C. *J. Chem. Phys.* **1988**, *89*, 5876. (b) Marrone, T. J.; Merz, K. M., Jr. *J. Phys. Chem.* **1993**, *97*, 6524. (c) Alper, H.; Levy, R. M. *J. Chem. Phys.* **1993**, *99*, 9847. (d) Hummer, G.; Pratt, L. R.; Garcia, A. E. *J. Phys. Chem.* **1996**, *100*, 1206. (e) King, G.; Warshel, A. *J. Chem. Phys.* **1989**, *91*, 3647. (f) Åqvist, J. *J. Phys. Chem.* **1990**, *94*, 8021. (g) Åqvist, J. *J. Phys. Chem.* **1994**, *98*, 8253. (h) King, G.; Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1991**, *95*, 4366. (i) Essex, W. E.; Jorgensen, W. L. *J. Comput. Chem.* **1995**, *16*, 951.

(2) Åqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512.

(3) (a) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1994**, *116*, 3949. (b) Levy, R. M.; Belhadj, M.; Kitchen, D. B. *J. Chem. Phys.* **1991**, *95*, 3627.

(4) Hummer, G.; Pratt, L. R.; Garcia, A. E.; Berne, B. J.; Rick, S. W. *J. Phys. Chem. B* **1997**, *101*, 3017.

(5) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; pp 331−342.

(6) Ashbaugh, H. S.; Wood, R. H. *J. Chem. Phys.* **1997**, *106*, 8135.

(7) Jayaram, B.; Fine, R.; Sharp, K.; Honig, B. *J. Phys. Chem.* **1989**, *93*, 4320.

(8) Roux, B.; Yu, H.-A.; Karplus, M. *J. Phys. Chem.* **1990**, *94*, 4683.