Incorporating Solvent and Ion Screening into Molecular Dynamics Using the Finite-Difference Poisson–Boltzmann Method

Kim Sharp

Department of Biochemistry and Molecular Biophysics, Columbia University, 630 W. 168th St., New York, New York 10032

Received 1 August 1990; accepted 8 November 1990

The finite difference method for solving the Poisson–Boltzmann equation is used to calculate the reaction field acting on a macromolecular solute due to the surrounding water and ions. Comparisons with analytical test cases indicate that the solvation forces can be calculated rapidly and accurately with this method. These forces act to move charged solute atoms towards the solvent where they are better solvated, and to screen interactions between charges. A way of combining such calculations with conventional molecular dynamics force fields is proposed which requires little modification of existing molecular dynamics programs. Simulations on the alanine dipeptide show that solvent forces affect the conformational dynamics by reducing the preference for internal H-bonding forms, increasing the *R*-alpha helix preference and reducing transition barriers. These solvent effects are similar to previous explicit solvent simulations, but require little more computation than vacuum simulations. The method should scale up with little increase in computational cost to larger molecules such as proteins and nucleic acids.

INTRODUCTION

Molecular dynamics has become a powerful tool for analyzing the structure and function of proteins and nucleic acids. Early applications focussed on the method as a way of studying actual atomic motions of molecules. In more recent applications the method has also been used as a way to explore conformational space and to generate energetically reasonable structures, often in conjunction with restraints provided by structural data or as an adjunct to model building. For a review see the recent monograph by McCammon and Harvey.[1]

While considerable progress has been made in developing various potential functions for proteins and nucleic acid simulations, the nonbonded terms unfortunately remain the least well understood, yet most computationally demanding part. These nonbonding terms are central to some of the most important interactions in macromolecules including folding and stability, solvation of charged and polar groups, hydrogen bonding with solvent, interaction with mobile solvent ions, and the hydrophobic effect. The greatest difficulties arise in treating the electrostatic term because of its long range effects, the importance of the solvent, and because it is not easily decomposed into a sum series of pairwise interactions. A detailed discussion of these difficulties, and of the various electrostatic treatments is given in a recent review by Harvey.[2]

Early simulations, because of limitations in potentials functions and computer power, were often performed in vacuum. In some cases, particularly for highly charged nucleic acids, it was necessary to either omit[3] or scale[4] the charges to obtain stable simulations. Subsequently, two other ways of treating electrostatics were developed. The first is to use a dielectric function to mimic the effect of solvent screening on charge-charge interactions[5,6] There are two severe drawbacks with this approach. No single function will accurately represent the interactions between all pairs of charges in a molecule, or be transferrable to a different molecule having for example a different surface accessibility to solvent. More importantly, charge-solvent effects are not accounted for. Such solvation forces for example keep the majority of charged amino-acids on the protein surface during folding, where they are better solvated by the water, play a role in the energetics of charge burial or transfer across a membrane, contribute to ion and charged ligand binding, and indeed are important in any phenomenon where the solvation state of a charged or polar group is altered.

The second approach is to include the solvent explicitly in the simulation[7–10] Computational resources and water potential functions have advanced to the stage where many simulations can now be performed with at least one or two layers of explicit waters around the molecule, although the computational cost is still considerable. However there are

Journal of Computational Chemistry, Vol. 12, No. 4, 454–468 (1991)

CCC 0192-8651/91/040454–15$04.00

still many difficulties in developing reliable water-water and water-molecule potential functions. Electronic polarizability effects in water, while a subject of active study in pure solvent simulations,[11-14] are not yet well enough understood for inclusion in macromolecular simulations. Even with some explicit water included around a molecule, long-range electrostatic interactions are still missing, and have to be approximated, e.g., by periodic[15] or stochastic[16] boundary conditions and image charge methods.[17] The bulk dielectric properties of the water in such simulations are uncertain, depending on the water model cutoffs, and boundary condition[18,19]

Simulating ionic solutions and modeling ionic strength effects, particularly for nucleic acids raises even more difficulties. One difficulty is that ion-water potentials are in early stages of development. For example the errors in solvation energy of even a simple system such single sodium ion in water are considerable.[20,21] While studies have been made on DNA fragments that include hydrated counterions,[22] and both explicit waters and several ions,[23] the full treatment of ionic strength effects through explicit solvent representation is currently not possible. For example consider the simulation of two turns of B-DNA at ionic strengths down to 0.01 M. To meanfully simulate the ionic strength behavior the solvent layer around the molecule must be at least two Debye lengths (ca. 60 Å) thick. The simulation must include some 30,000 odd water molecules, 20–30 ions, and some 1200 DNA atoms. The simulation must be carried out for at least a thousand picoseconds for the ion-water solvent system to come to equilibrium since the characteristic relaxation time of the electrical double layer is of the order of hundreds of picoseconds.[24] Even with accurate potential functions simulation of all these ion-water, water-DNA-ion interactions, including the long-range electrostatic and dielectric interactions, with enough accuracy and for a long enough time to establish convergence will be a formidable task.

Fortunately in many cases we are not interested in the details of the solvent itself, but only in its effects on the macromolecule. Therefore it is wasteful to spend large amounts of computational time on simulating explicit water molecules and ions, particularly those some distance form the molecule, and especially if the solvent model has uncertain bulk dielectric and long range properties. In this case what is required is a way of calculating the electrostatic effects of the solvent without putting in its molecular detail.

Several strategies can be used to simulate the interaction with missing solvent. The simplest use some kind of boundary condition such as periodic and stochastic boundary conditions or Ewald summations, but these generally do not treat long range electrostatic interactions well. Warshel and co-workers have developed a model for water involving a lattice of Langevin dipoles,[25] which has been combined with molecular dynamics simulations.[26] Gilson and Honig[27] have recently proposed the addition of a $1/r^4$ term to the standard potential function in order to represent solvation energy terms. While this function has not yet been tested in molecular dynamics simulations, it could provide a very efficient way of including solvation effects in a simulation.

An alternate approach incorporates the missing solvent interactions through its reaction field, calculated using the method of image charges.[17,28] This has been applied to water simulations, using spherical simulation spaces for which the image calculation is easy. The image method has the advantage that it specifically treats the long-range electrostatic interaction. A more general reaction field formalism which can deal with arbitrary geometries and which can also incorporate ionic strength effects, is based on the Poisson–Boltzmann (PB) equation (see Sharp and Honig[29] for a review). While a continuum treatment of the solvent may appear to be a drastic simplification in that it treats the water close to the solute like bulk solvent, such models have proven to be a quick and reliable way to obtain the electrostatic component of solvation free energy. For example the Born model fits the hydration energy of a wide variety of ions[30] and in fact it can be parameterized to agree more closely with experiment than current all atom simulations.[20,21] Direct comparison of explicit atomic solvent and continuum solvent models have been made for a number of polar solutes. The calculated solvation energies show agreement to within a couple of kcal/mole.[31,32] Agreement of the PB method with experimental solvation energies for small polar molecules is also very good.[33] The PB equation has also been used for solvation energy calculations in proteins[34-36] with good results. Recent numerical advances[37,38] have rendered the method extremely rapid, so that it is more than two orders of magnitude faster than explicit solvent models. The reasons for the success of the continuum representation of the solvent have been examined in a series of molecular simulations,[21,39,40] and may be attributed largely to cancellation of effects in the behavior of water at the molecular level.

Since the PB equation can be used to calculate electrostatic solvation energies of proteins and solutes reliably, it may be expected that it can be used to calculate solvation forces for molecular dynamics as well. This article examines a way to combine the PB model of water and ion screening with the standard bonded and intramolecular nonbonded potential within the framework of a molecular dynamics simulation. This is done using the solvent reaction field. The first part of the article describes the theory and numerical procedures for calculating the reaction field. The second part describes how the forces resulting from this field are combined with the standard intramolecular electrostatic dynamics force

field components. The results section describes tests of the numerical accuracy of the reaction field calculations, followed by an application to the conformational dynamics of the alanine dipeptide, a common model for studying protein conformational energetics and dynamics.

THEORY

The Solvent Reaction Field

Consider a charged or polar solute in a particular instantaneous conformation, immersed in a solvent. The solute may be considered as forming a cavity in the solvent. This cavity contains some atomic charge distribution represented by some potential function. The electric field arising from these charges induces a response in the solvent. This response arises from three physical processes: electronic polarization, permanent dipole reorientation, and redistribution of mobile ions (if present). The first two processes determine the dielectric constant of the solvent, and in the continuum model they result in a mean induced polarization, P, or dipole moment/unit volume given by

$$P(r) = \langle \mu \rangle / V = (\varepsilon - 1)E(r)/4\pi \quad (1)$$

where $\langle \mu \rangle$ is the mean dipole moment of some small volume V of solvent at position r, with a dielectric of ε, and E is the mean Maxwell field.

The redistribution of mobile ions in response to the solute field results in a nonuniform distribution of mobile ion, or creation of an electrical double layer. Within the continuum model, this can be represented as a mean net charge density, ρ^m, induced in the solvent, given in the Debye–Huckel/Boltzmann model (for a 1:1 electrolyte) by

$$\rho^m(r) = -(\kappa^2 kT/4\pi e)\sinh(e\phi(r)/kT) \quad (2)$$

where $\phi(r)$ is the mean potential, e is the proton charge, k is Boltzmann's constant, T is the absolute temperature and κ is the Debye–Huckel parameter given by

$$\kappa^2 = 8\pi N_a e^2 I/1000\varepsilon kT \quad (3)$$

where N_a is Avogadro's number and I is the bulk ionic strength in moles/L ($1/\kappa$ is the Debye length). The reaction potential acting back at any point s in the solute is given by

$$\phi^x(s) = \iiint \{P(r) \cdot \nabla(1/(s - r)) + \rho^m(r)/|s - r|\}\delta r^3 \quad (4)$$

Where the integration is taken over the volume of the solvent. The electrostatic contribution to the solvation free energy of the solute is then given by

$$\Delta G^0_{el} = \tfrac{1}{2} \sum q^f_k \phi^x_k \quad (5)$$

where the summation is over all atoms of the solute, and q^f_k is the charge assigned to the kth atom in the particular force field used (eq. (5) is valid for the linearized equation. The correction due to the nonlinearity of eq. (2) has been discussed elsewhere[41]).

To determine the polarization of the solvent it is necessary to determine the electrostatic potential distribution in the solvent, $\phi(r)$. The potential is related to the charge distribution by the Poisson equation

$$\nabla \cdot \varepsilon(r)\nabla\phi(r) + 4\pi(\rho^f(r) + \rho^m(r)) = 0 \quad (6)$$

where $\rho^f(r)$ is the charge distribution of the solute. Combining the linearized form of eq. (2) with (6) yields the linear Poisson–Boltzmann (PB) equation

$$\nabla \cdot \varepsilon(r)\nabla\phi(r) - \varepsilon(r)\kappa^2(r)e\phi(r)/kT + 4\pi\rho^f(r) = 0 \quad (7)$$

Equation (7) applies both to the solvent, where $\rho^f = 0$, $\varepsilon = 78.6$ (for water), and κ is given by the bulk ionic strength, and to the solute, where $\rho^m = \kappa = 0$, and $\varepsilon = 1$. A solute dielectric of 1 is used to be consistent with molecular dynamics potential functions which do not include electronic polarizability. The dipole reorientation within the solute is being modeled explicitly in the dynamics simulation, so there is no remaining dielectric response to account for. This is in contrast to some previous applications of the PB equation, where the molecule's electronic polarizability is accounted for by using a dielectric constant of 2, and the dipole reorientation with a higher dielectric constant.[29] The spatial distributions of ε, κ, and ρ^f are defined by the conformation of the solute. To obtain a general solution we apply the finite difference method[42] to the PB equation (FDPB) whereby we divide a region of space containing the solute and surrounding solvent into a cubic lattice, yielding an expression for the potential at each grid point.[43,44]

$$\left(\sum \varepsilon_j \phi_j + 4\pi q^f_o/h\right) \Big/ \left(\sum \varepsilon_j + \varepsilon_o \kappa^2_o h^2\right) \quad j = \pm x, \pm y, \pm z \quad (8)$$

where h is the lattice spacing, the potential is expressed in units of kT/e, the subscript o refers to quantities at each grid point, and the subscript j refers to quantities at the six neighboring points. For an Nth order lattice N^3 simultaneous equations like eq. (8) replace the partial differential eq. (7). These simultaneous equations are then solved by the relaxation method, as described previously[43–45] to obtain the potential at each lattice point. The finite difference form of eq. (1) for the polarization vector components is

$$P_{oj} = (\varepsilon_j - 1)(\phi_j - \phi_o)/4\pi h \quad j = x, y, z \quad (9)$$

The polarization P_o is zero inside the solute, since $\varepsilon = 1$. Using eqs. (8) and (6) an expression for the

mobile charge density induced at any grid point in the solvent, ρ_o^m, is

$$\rho_o^m = \varepsilon \left(6\phi_o - \sum_{j = \pm x, \pm y, \pm z} \phi_j\right) \Big/ 4\pi h^2 \quad (10)$$

where ε is the solvent dielectric. The volume integral for the reaction potential, eq. (4), now becomes a summation over the solvent grid points

$$\phi^x(s) = \sum_o \{P_o \cdot \nabla(1/(s - r_o)) + \rho_o^m/|s - r_o|\}h^3 \quad (11)$$

While eq. (11) can be used to calculate the solvent reaction field for any given solute conformation, an alternative way of representing the contribution of the polarization, P, to the reaction potential is more convenient. At the boundary between two regions of different dielectric, e.g., the solute surface, the following condition holds

$$\varepsilon_1 E_1 \cdot n = \varepsilon_2 E_2 \cdot n \equiv D_1 \cdot n = D_2 \cdot n \quad (12)$$

where n is the normal surface vector, and D is the displacement vector, and the subscripts 1 and 2 refer to the two regions. If the two dielectrics are replaced by a single dielectric, $\varepsilon = \varepsilon_1$, with the surface charge density, σ given by

$$E_1 \cdot n - E_2 \cdot n = 4\pi\sigma/\varepsilon \quad (13)$$

then the same potential distribution in region 1 will be obtained (for example see Jackson[46]). Alternatively if we use $\varepsilon = \varepsilon_2$, the potential in region 2 is reproduced instead. This equivalence allows us to represent the solvent reaction field either as arising from a dipole moment distribution induced in a high dielectric solvent, or as an induced surface charge distribution at the solute/solvent boundary. The finite difference expression for the induced surface charge at each grid point, σ_o, corresponding to eq. (13) is

$$\sigma_o = \varepsilon \left(6\phi_o - \sum_{j = \pm x, \pm y, \pm z} \phi_j\right) \Big/ 4\pi h \quad (14)$$

Since we require the reaction field within the solute, we chose $\varepsilon = 1$. Expressing the solvent polarization as an induced surface charge distribution replaces the first volume integral in eq. (4) by a surface integral, leading to a more succinct form. It should be noted that replacement of boundary condition (12) by (13) has a physical interpretation as well as a mathematical one. Remembering that the solvent polarization arises primarily from dipole reorientation, the induced surface charge at the solvent boundary really represents the oriented water dipoles at the solute surface.

Equations (10) and (14) have the same form, and may be combined to give

$$q_o^i = h \left(6\phi_o - \sum_{j = \pm x, \pm y, \pm z} \phi_j\right) \Big/ 4\pi - q_o^f \quad (15)$$

q_o^i being the induced charge at the oth grid point arising from either the dielectric boundary or the mobile ions. (For completeness the solute charge assigned to that grid point, q_o^f if any, must be subtracted). The electrostatic solvation free energy is now

$$\Delta G_{el}^0 = \tfrac{1}{2} \sum_k q_k^f \sum_o q_o^i / r_{ko} \quad (16)$$

A comparison of eqs. (8) and (15) shows that q_o^i will be zero except where $\kappa_o \neq 0$, or where $\varepsilon_j \neq \varepsilon_k$ for some j, $k = 1, 6$. Another way of stating the relationship between eqs. (8) and (15) is that the latter is the solution to the question: What distribution of charges within the solvent region and on the solvent/solute boundary, using Coulomb's Law and a vacuum dielectric, will give the same reaction field at the solute as would be obtained with a solvent of dielectric 78.6, and an ionic strength of I?

Determination of Forces from the Reaction Field

Expression (16) can be applied directly in Monte Carlo methods, which require only the relative energies of configurations to be evaluated. Molecular dynamics is more difficult since atomic forces are also needed. One could in principle obtain the force on the kth atom, f_k; it is only necessary to find the gradient of the energy with respect to that atom's coordinate,

$$f_k = \partial(\Delta G_{el}^0)/\partial r_k \quad (17)$$

at each time step of the simulation and add this force to the standard *in vacuo* intramolecular force.

Equation (17) is exact within the continuum model, subject only to the approximations inherent in a finite difference solution to the PB equation. However it cannot be applied directly for several reasons. (1) To update the magnitude and positions of q_o^i would require the PB equation to be resolved at every time step. This is still prohibitively expensive. (2) Exact analytical expressions for the derivatives of eq. (16) are difficult to obtain since both the induced charges and the shape of the dielectric boundary are functions of the atomic charge positions (See appendix). (3) We also have to determine how to accumulate the forces that act on each atom. Consider a surface atom with some atomic charge, i.e., one whose surface forms part of the solute/solvent boundary. Charge will be induced on its surface, producing a force on the atom. This represents the force exerted by the aligned water dipoles acting on

that atom. Since the atomic charge cannot move with respect to its surface however, this force does no work and should not be accumulated into that atom's force. To do so would produce a large fictitious force constantly pulling charged atoms toward their endlessly receding surface.

In spite of the difficulties in applying eq. (17) rigorously, approximate expressions for the forces that still encapsulate the basic physical effects of solvation may still be of great use. First it is assumed that the PB calculation will be used to update the induced charges every n steps, where n is to be determined empirically for each molecule in a compromise between speed and accuracy. If the positions as well as the magnitudes of the induced charges were fixed between updates, then large errors would result due to small motions of surface charged atoms, as the distance between the nearest pairs of q^f and q^i charges is of the order of atomic radii. Since the atom surfaces themselves define the dielectric boundary where the induced charge is always located, it makes sense to somehow link the motion of each induced charge to the atom whose surface it is on. In this way the positions of the induced charges are effectively updated every time step. One way is to maintain lists of induced charge "belonging" to each atom, and apply the atomic displacement calculated in each molecular dynamics time step to all an atom's charges. A further simplification has been adopted here, with the aim of reducing the amount of extra computation and bookkeeping, and also a minimise the modification of existing molecular dynamics code, to produce the following algorithm:

The potentials are calculated using eq. (8), and the induced charge q_o^i is back-calculated using eq. (15). Each induced surface charge is assigned to the atom whose van der Waals surface is nearest. The total surface charge per atom, q_k^s, is added to the solute charge, q_k^f to give the 'effective' atomic charge q_k^e (Fig. 1). The forces are then calculated from the three resulting charge lists: the atom charges, q_k^f, the effective atom charges, q_k^e and the induced grid charges from the mobile ions, q_o^m, using (eq. (A7) from the appendix):

$$f_k = q_k^f E_k = q_k^f \sum_{j \neq k} q_j^e \nabla(1/r_{kj}) + \sum_o q_o^m \nabla(1/r_{io}) \quad (18)$$

where E_k is the total field at the kth atom. The first summation contains both the standard intramolecular electrostatic forces arising from $q_k^f q_j^f$ terms, plus $q_k^f q_j^s$ terms representing the mean reaction field from the solvent polarization. The second summation over all o solvent grid points represents the contribution from mobile ions. It can be seen from eqs. (12) and (13) that with a high external dielectric,

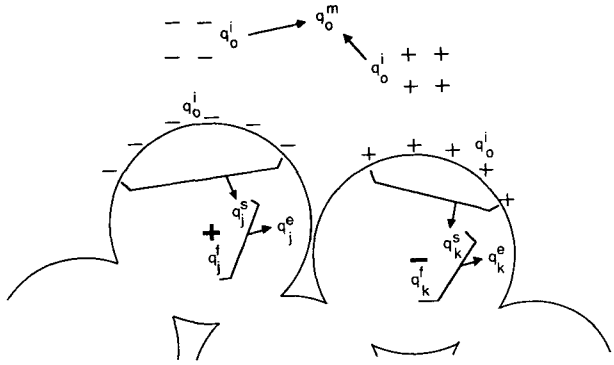

Figure 1. Schematic illustration of the assignment of induced charge, q_o^i, calculated from eq. (15). A fragment of the solute containing two atoms with partial atomic charges q_j^f and q_k^f is shown. Induced charge at the surface grid points arises from water polarization and is assigned to the nearest atom, and summed to give q_j^s and q_k^s. This is added to the atomic charges to give the effective charges q_j^e and q_k^e. Induced charge at the solvent grid points, q_o^m arises from mobile ions. This is not assigned to any atom.

a charge will induce surface of opposite sign, and so q_j^e will generally be smaller in magnitude than q_i^f, i.e., the forces will be reduced, or screened as expected.

Coalescing the induced surface charge at the atom centers and combining it with the atomic charges has several consequences: (1) The induced surface charge list is no larger than the atom charge list, adding almost no time to the force computation. (2) The updating of the surface charge positions is handled automatically with the atom positions, with no extra computation. (3) Since the molecular dynamics pairwise summation includes only $j \neq k$ terms the exclusion of the force between each atom and its surface charge, which can produce no relative motion, is automatically handled. Solvation forces between 1–2 and 1–3 bonded atoms will also be excluded in this scheme due to the way standard potential functions are evaluated.[1] However these distances will not vary much in a normal simulation since these interactions are determined by the bond stretch and angle terms, so the electrostatic component may safely be omitted. Forces between 1–4 bonded atoms pose more of a problem. Some early forms of force fields omitted 1–4 nonbonded interactions completely, while one current force field, DISCOVER[7] includes the full nonbonded interaction. Both CHARMM[47] and AMBER[48] scale the 1–4 electrostatic interaction, by 0.4 and 0.5 respectively. What this means is that part of the *in vacuo* 1–4 electrostatic interaction is effectively included in the corresponding 1–2 torsion parameter in the latter two force fields, i.e., it will not be affected by electrostatic screening. It is possible, however, with the solvent screening method described here to treat the 1–4 electrostatic interactions of the "fixed" and "induced" charge differently, maintaining separate

charge and exclusion lists and scaling only the former charges, when using CHARMM or AMBER, but this complicates the algorithm. From this point of view DISCOVER is a more convenient force field with which to implement this method. It should be pointed out however that this difficulty is not unique to the electrostatic screening treatment described here: It also applies to the use of the distant dependent dielectric, and to use of constant dielectrics greater than one. (4) Finally there is an additional approximation in the reaction field at the remaining atom pair interactions since the charge has been moved from the atom surfaces to the center, a shift of 1–2 Å.

METHODS

Calculation of the Reaction Field Charges

Finite-Difference Poisson–Boltzmann calculations were performed as described previously using the electrostatics package DelPhi.[43–45] The lattice order is 65 and the molecule is scaled to occupied 75% of the lattice. The molecular surface[49] is defined using a 1.4 Å probe radius. Grid points in the interior of the molecule are assigned a dielectric of 1, while those in the solvent are set to 80. Forces arising from ionic strength effects were only examined in the static test comparisons with analytical solutions for 0.15 M (physiological level). Implementation of the ionic force in dynamics is currently underway and will be discussed in a future paper. The positions of the fixed atomic charges for alanine dipeptide are determined from the current atomic coordinates during a molecular dynamics trajectory, and assigned to the grid using a linear interpolation scheme. Potentials on the lattice boundary are approximated with a summation of Debye–Huckel type potentials. Using eq. (8) the potentials are iterated to a maximum change of less than $10^{-4}\ kT/e$ between successive iterations, using the over-relaxation method. The optimum over-relaxation parameter is automatically set by the method of Nicholls and Honig.[38] The entire algorithm is highly vectorizable, and once the molecule is mapped onto the lattice, the method is independent of the number of atoms, taking about 15 s. of cpu time on a Convex C220. Once the potentials at the lattice have been determined, the induced charge at each lattice point is back-calculated using eq. (15) by a program called BACKQ. The expected total induced surface charge, from Gauss's law is

$$\sum_o q_o^i = (\varepsilon_m - \varepsilon_s)/\varepsilon_s \sum_k q_k^f \qquad (19)$$

Where ε_m and ε_s are the molecule and solvent dielectrics respectively. The calculated total charge in all cases agreed to within 0.5%, providing a check on the numerical accuracy of the back calculation procedure. The whole process, starting from the atomic coordinates and ending with the charges lists takes about 45 s. to 1 min. on a Convex C220.

Analytical Tests Cases

Series solutions to the Poisson-Boltzmann equation for charges in a spherical cavity are obtained using the Tanford–Kirkwood series solutions.[50,51]

Integration into Molecular Dynamics

The procedure for calculating the solvent reaction field component of the atomic forces is combined with the molecular dynamics program XPLOR.[52] This is done by adding an additional nonbonded parameter option and subroutine to XPLOR to read in the charge lists produced by BACKQ. The nonbonded energy routine is also modified to calculate the electrostatic forces using eq. (18) instead of the standard expression. The procedure for performing a simulation consists of the following cycle: (1) XPLOR writes out the atomic coordinates of the current configuration. (2) The coordinates are read by DelPhi, which outputs the lattice potentials. (3) The lattice potentials are read by BACKQ, which outputs the effective and mobile charge lists. (4) XPLOR reads in the charge lists and continues with the molecular dynamics simulations. This cycle is referred to here as dielectrically modified Newtonian dynamics (DIEMOND), and is repeated as necessary until the simulation is finished. The energy conservation properties of the DIEMOND procedure were checked using the "test first" command in XPLOR, this compares the analytical expression for the energy of the system to that obtained by numerical integration of the forces. The rms fluctuation in the total energy obtained from the logfiles was also monitored. Variations were similar to those obtained in standard dynamics, i.e., less than 0.5%. Only the zero ionic strength form of the algorithm has been examined to date.

Tests on Alanine Dipeptide

The alanine dipeptide (*N*-acetylalanyl-*N*-methylamide) was constructed in XPLOR using the CHARMM force field[47] with explicit amide hydrogens and united atom C-alpha and methyls carbon atoms (The CHARMM 19 parameter set supplied with XPLOR was used). The fixed atomic charges and radii are also required as input to DelPhi, and are: Methyls: (0, 1.9 Å), C: (0.48, 1.8 Å), O: (−0.48, 1.48 Å), Cα: (0.1, 1.8 Å), *N*-ala: (−0.36, 1.55 Å), *N*-acetamide: (−0.26, 1.55 Å), H: (0.26, 0.8 Å) respectively. Dynamics simulations were performed with a time step of 1 fs, with no cutoff on nonbonded interactions. Simulations were done with four treatments of electro-

static screening: (1) No electrostatic energy terms. (2) In vacuum, i.e., with a constant dielectric of 1. (3) With a constant dielectric of 80. (4) Including solvent screening with the DIEMOND procedure, using a molecule dielectric of 1, a solvent dielectric of 80, and zero ionic strength. Electrostatic interactions between 1–4 bonded atoms in CHARMM are scaled by 0.4 (See Theory section).

The adiabatic phi-psi angle energy surface was calculated with the different electrostatic treatments by setting the phi and psi angles at 10° intervals, minimizing for 100 steps by the conjugate gradient method with respect to the other internal degrees of freedom and evaluating the final energy of the system. Conformational probability maps were generated by running dynamics at 300 K. One million steps of dynamics were run, the phi and psi angles were evaluated every 50 fs, and a normalized two dimensional phi-psi probability histogram generated with 10° intervals. To ensure adequate sampling of the phi-psi space umbrella sampling was used.[53] Sixteen simulations were performed with an umbrella potential of 1/kcal/rad applied to both the phi and psi angles, centered at phi/psi = −90°, 0°, 90° or 180°. Each probability map was corrected for the umbrella potential, and the maps then were combined by renormalizing the maps with respect to each other[54] so as to minimize the mean square difference in probabilities between overlap regions of the maps. This straightforward windowing method of using umbrella potentials has the advantage of being easy to implement in XPLOR, although a more efficient adaptive umbrella sampling method[55] could be used. In absence of previous experience, the frequency of charge list updating with DelPhi/BACKQ was somewhat arbitrarily chosen as 200 fs. While the optimum frequency will have to be determined by experience, this seems a reasonable compromise between keeping the additional computational burden in of a simulation small, yet responding to changes in conformation. For small molecules such as alanine dipeptide this increases the time over the vacuum simulation by about a factor of 3–4, but since the FDPB procedure is essentially independent of the molecules size the penalty will be proportionally smaller for large proteins. To check whether a 200 fs update is adequate for this molecule, selected simulations were also run at 100 fs and 50 fs updates with essentially the same probability distributions within statistical error.

RESULTS

Calculations on Spherical Test Cases

To determine the numerical accuracy of the reaction field calculations, comparisons were made between the FDBP method and analytical solutions for charges in a sphere. The energies and forces were calculated either from eqs. (16) and (17) respectively, or from the Tanford–Kirkwood series solutions.[50,51] Two charges of +3 and −3 were placed 6 Å apart at the centre of a 5 Å diameter sphere. The potential and force at one of the charges due to the other charge and the reaction field were calculated for four cases: (1) a uniform dielectric of 1, i.e., with no reaction field. (2) an interior dielectric of 1, an exterior dielectric of 80; only surface contributes to the reaction field. (3) Dielectrics of 1 and 80, ionic strength 0.15 M; with both contributions to the reaction field. The results are given in Table I. With a couple of exceptions, the errors in the numerical calculations are within a couple of percent. The largest error is in the force component arising from the mobile ion reaction field (ca. 10%), where the term is very small, but the absolute error again is small (less than 0.02 kT/eÅ).

In any finite method it is important to assess the error arising from the lattice resolution and mapping procedure. This was done by altering the scale and position of the sphere with respect to the grid.[44] The numerical results varied by less than the errors seen in Table I.

It can be seen from Table I that the reaction field induced by a charge on itself can give rise to a considerable force on that charge. This "solvation" force acts to pull the charge into the high dielectric region.

Table I. Comparison of forces on test charges in a sphere using finite difference Poisson–Boltzman (FDPB) and Tanford–Kirkwood (TK) methods.

Dielectric		Ionic		Potential (kT/e)		Force (kT/e/Å)	
Inner	Outer	strength	Component[a]	FDPB	TK	FDPB	TK
1	1	0.	Other	280.6	280.8	−140.4	−140.4
			Self	0.	0.	0.	0.
1	80	0.	Other	35.5	35.6	−75.3	−75.6
			Self	524.0	518.0	299.7	290.1
80	80	0.15	Other	1.89	1.86	−1.65	−1.64
			Self	1.88	1.93	0.186	0.168
1	80	0.15	Other	33.8	34.1	−75.3	−75.6
			Self	526.1	520.2	300.9	290.4

[a]The potential and force were evaluated at a charge of −3 e, placed at (−3,0,0) in a 5 Å diameter sphere, arising from either a charge of 3 e at (3,0,0) (Other), or from the negative charge's solvent/ion reaction field (self)

Table II. Forces on a random set of charges in a sphere: Comparison of the finite difference Poisson–Boltzmann (FDPB) and Tanford–Kirkwood (TK) methods.

Number of charges[a]	Error in force		Error in Total Energy (%)
	Angle (°)	Magnitude (%)	
4	2	5	8
8	6	6	7
16	5	10	5
32	9	11	1
64	7	11	1

[a]Charges were distributed randomly through a 10 Å radius sphere at depths between 1 Å to 2 Å below the surface. The dielectric constants were 2 inside, 80 outside, with an ionic strength of 0.15 M.

The reaction field induced by the second charge screens the Coulombic interaction reducing the charge-charge interaction when a high dielectric or ions are outside the sphere. The mobile ion component of the reaction field in this example is smaller than the effect of the dielectric boundary.

Table II shows a comparison of the FDPB and Tanford–Kirkwood methods for 4 to 64 charges placed randomly in a 10 Å sphere between 1 Å and 2 Å below the surface. This test case was chosen to be more representative of the type of charge distribution in an actual protein. Numerical errors in the FDPB procedure are largest for charges close to the dielectric boundary.[44] The range 1 Å to 2 Å was chosen to cover the typical van der Waals atomic radii, which represent the closest a charge can get to the dielectric boundary in the FDPB model. A statistical analysis was made of the error in angle and magnitude of the field and the total electrostatic energy of the system as a function of the number of charges included in the calculation. It can be seen that the total energy of the system has little error, and in fact improves with more charges due to cancellation of errors. The force magnitude and direction have a larger error, but the figures indicate that these level off at about 10% and 7 degrees respectively.

Analysis of Forces in a Model Molecule

In order to illustrate some of the properties of eq. (18) when applied to a molecule the components of the force vector were analysed in a model "macromolecule" consisting of a cubic lattice of 125 "atoms" of radius 1.9 Å spaced 3.5 Å apart. Two atoms are given a charge of +1, and the rest are neutral. Figure 2 shows the force components acting on one of the charged atoms when they are placed either in the interior or at the surface of the molecule. The Coulombic or vacuum component due to the second charged atom is the same in either case and is directed along the vector between the two charges. Considering the self reaction field term, the negative charge induced by the buried charge is greatest at the nearest surface (as is apparent from eq. (13)). Thus it is assigned primarily to the 5 surface atoms

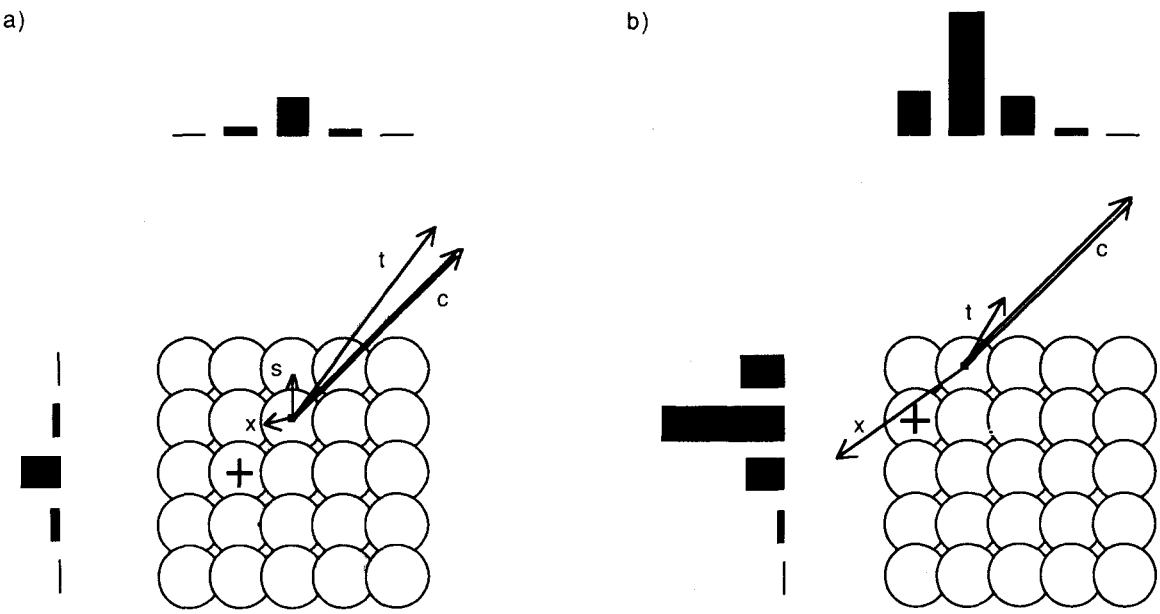

Figure 2. Analysis of force components on a charged atom in a test 'molecule' consisting of a cubic array of 125 atoms of radius 1.95 Å, spaced 3 Å apart. The middle slice of atoms containing this atom (·) and a second charged atom (+) is shown. The forces arising from the Coulombic interaction between the charges (c), the self-reaction field of the first charge (s), and the cross reaction field of the second charge (x) are shown by arrows (scale arbitrary). The relative surface charge induced by the first charge, which is mostly on the top row of surface atoms is indicated by the upper histogram, that induced by the second charge, mostly on the leftmost row of surface atoms is shown by the lower histogram. (a) Both charges are buried. (b) Both charges are at the surface.

in the top row (depicted schematically by the top histogram), and an upward force is exerted on the charged atom which acts to pull it to the surface, in order to better solvate it. When the atom is at the surface, it induces even more negative surface charge (Fig. 2(b)). However a large part of the induced charge is assigned to the atom itself, and is excluded from the force sumation in eq. (18). A physical intepretation of this is that the reaction force pulling the atom into the solvent is balanced by the dipole pressure of the oriented water pushing back on the charged atom surface. The rest of the surface charge is assigned to the neighboring atoms in the top row, which produces no upward force component. The top leftmost atom has more surface charge assigned to it, being more exposed, so a small residual lateral force on the charged atom remains (too small to show on the figure) acting to pull it into the corner position where it would be maximally solvated. However the solvating force is essentially gone once the atom is at the surface.

The cross reaction field term is the one induced by the second charge acting on the first charge. The negative charged induced by the second charge is assigned primarily to the surface atoms on the left column (indicated in the figure by the histogram at the left). When the atoms are buried the effect of this surface charge is small because of the distance. The Coulombic force thus remains largely unscreened, and the net force (t) is dominated by the vacuum term. When the atoms are at the surface, the relative distance between the charge and the induced charge from the other atom is smaller, and the cross reaction field cancels the vacuum term to a greater degree resulting in better screening. The net force on the charge is thus smaller.

Dynamics of the Alanine Dipeptide

Figure 3 shows the adiabatic energy surface for the peptide phi-psi space with no electrostatic component. The map is dominated by the unfavorable regions around (0, 0) and (0, −180) arising from the steric clash of the actyl oxygen with either the alanine oxygen or the methylamide proton. As expected the conformational probability map (Fig. 3(b)) shows the overall features of the well known Ramachandran plot, with large peaks in the least sterically strained regions of the map, *R*-alpha (−120/−60, −50), beta (−120/−60, 160), and smaller peaks at *L*-alpha (60, 60) and C_7^{ax} (60, −60). Results using a uniform dielectric of 80 were essentially identical to those obtained with no electrostatic interaction, and are not shown.

Figure 4 shows the energy and population plots when the electrostatic energy component is included. The vacuum electrostatic component (a) is dominated by a deep trough centered at (0, 0) due to the electrostatic attraction, or hydrogen bonding potential between the acetyl oxygen and the methylamide proton. A smaller minimum at (180, 180) arises from the competing electrostatic attraction of the alanine proton and oxygen. These two regions are separated by an unfavorable ridge running through the *L*-alpha and *R*-alpha regions which arise from hydrogen-hydrogen and oxygen-oxygen repulsions. Addition of the nonelectrostatic energy term to this map results in a large unfavorable peak in the center of the trough (b). Thus with unscreened electrostatics, the probable conformations are restricted to the two hydrogen bonded forms C_7^{eq} (−60, 60) and C_7^{ax} (Fig. 4(c)). These conformations represent a balance between the powerful electrostatic

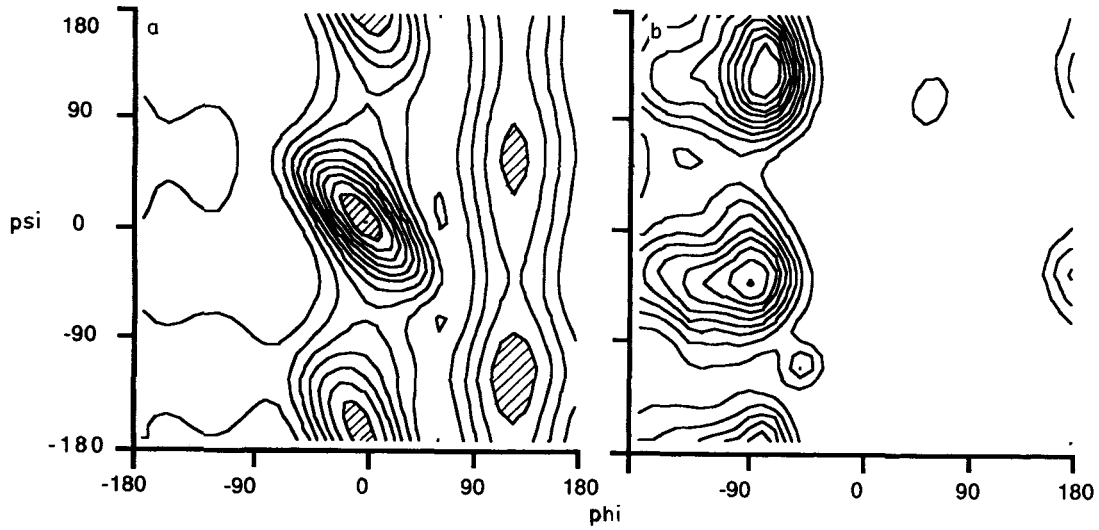

Figure 3. Alanine dipeptide conformational analysis in phi-psi space excluding electrostatic interactions. (a) Adiabatic energy map generated at 10° intervals of phi and psi by constrained minimization. Contours are at 1 kcal/mol, relative to the deepest minimum. Peaks or unfavorable energy regions are shaded. (b) Relative conformational probability map, generated at 10° intervals by sampling every 200 fs from a dynamics simulation at 300 K. Contours generated in 10 intervals from the lowest to highest probability regions.

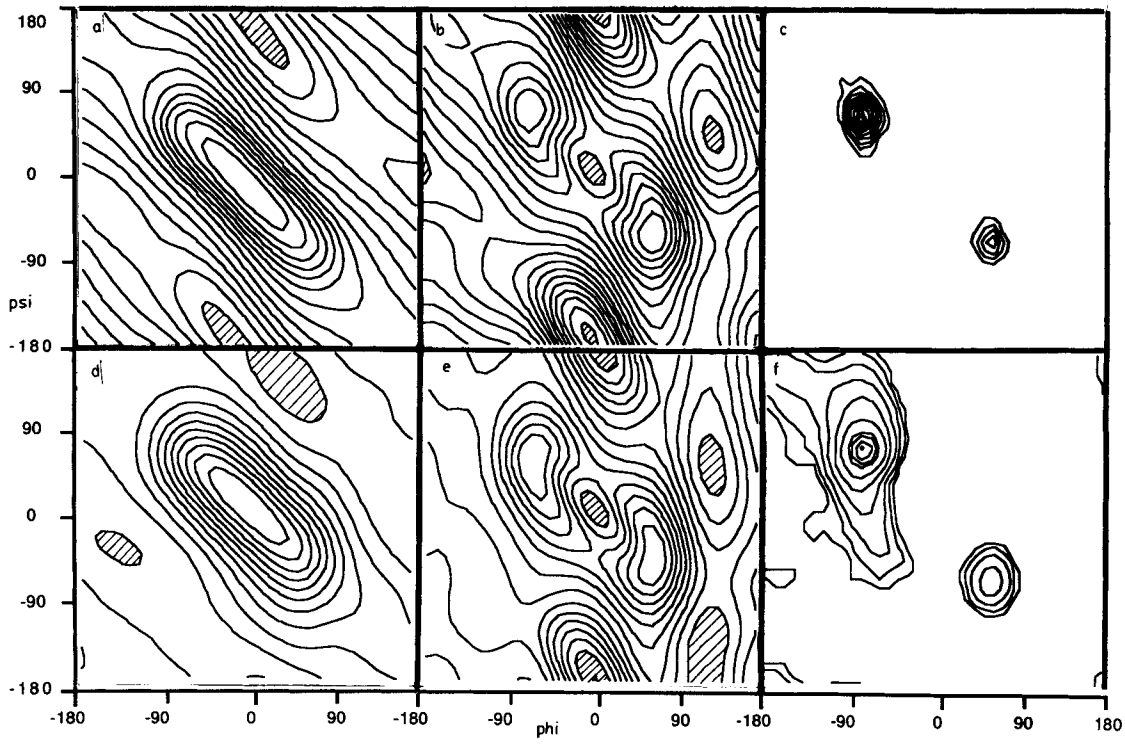

Figure 4. Alanine dipeptide conformational analysis in phi-psi space including electrostatic interactions. Adiabatic energy map for the electrostatic (a) and total (b) energies in vacuum. Relative probability map in vacuum (c). Adiabatic energy map for the electrostatic (d) and total (e) energies with DIEMOND solvent screening. Relative probability map with DIEMOND solvent screening (f). Maps were generated and diplayed as for Figure 3.

attractions, and avoidance of the severe steric clash at (0, 0). When solvent screening via the DelPhi/DIEMOND procedure is included the electrostatic component adiabatic energy map shows a number of changes (Fig. 4(d)). Overall the map is flattened, particularly the unfavorable ridge through the alpha regions, where the H-H and O-O repulsions are screened. Also the C_7^{ax} end of the trough becomes somewhat less favorable than the C_7^{eq} end. These changes are reflected in the total energy and probability plots (Figs. 4(e), 4(f)). The latter now shows a broad peak extending down from the C_7^{eq} region into the *R*-alpha region.

Tables III and IV show an analysis of the induced charge distribution during the simulations. Table III shows the induced charge distribution for alanine

Table III. Distribution of induced charge on Alanine Dipeptide in the C_7^{eq} conformation

		Induced charge due to field from[b]:						
Atom[a]	Charge	1C, 2C	1O, 2O	2N, 3N	1HN, 3HN	2Ca	All	Effective charge
1 CA	0	−0.14	0.049	0.047	−0.027	−0.002	−0.068	−0.068
1 C	0.48	−0.13	0.084	0.047	−0.013	−0.004	−0.019	0.46
1 O	−0.48	−0.096	0.28	0.026	−0.007	−0.004	0.20	−0.28
2 N	−0.36	−0.057	0.015	0.091	−0.089	−0.007	−0.048	−0.41
2 HN	0.26	−0.013	0.003	0.026	−0.079	−0.002	−0.065	0.20
2 CA	0.1	−0.087	0.043	0.061	−0.018	−0.031	−0.033	0.068
2 CB	0	−0.085	0.041	0.064	−0.040	−0.023	−0.043	−0.043
2 C	0.48	−0.13	0.087	0.045	−0.017	−0.011	−0.024	0.46
2 O	−0.48	−0.094	0.28	0.021	−0.007	−0.005	0.19	−0.29
3 N	−0.26	−0.046	0.014	0.081	−0.11	−0.004	−0.067	−0.33
3 HN	0.26	−0.011	0.003	0.022	−0.073	−0.002	−0.061	0.20
3 CA	0	−0.062	0.051	0.083	−0.034	−0.003	0.035	0.035
Total								
Calculated		0.951	−0.95	−0.614	0.514	0.0986	0.006	
Expected[c]		−0.948	0.948	0.612	−0.514	−0.0988	0.0	

[a]Atom to which induced surface charge is assigned. Number prefix indicates moiety: 1 = acetyl, 2 = alanyl, 3 = methylamide.

[b]Atom(s) whose charge is creating reaction field.

[c]From eq. (19).

Table IV. Effect of conformation on Alanine Dipeptide effective charges

Atom[a]	Conformation[b]		Simulation Time (fs)			
	R-alpha	beta	0	200	400	600
1 CA	−0.0767	0.0693	−0.073	−0.071	−0.077	−0.078
1 C	0.452	0.458	0.468	0.461	0.444	0.457
1 O	−0.271	−0.262	−0.273	−0.290	−0.304	−0.311
2 N	−0.409	−0.403	−0.409	−0.413	−0.443	−0.420
2 HN	0.179	0.175	0.206	0.207	0.216	0.205
2 CA	0.0818	0.0676	0.063	0.063	0.077	0.072
2 CB	−0.0593	−0.0292	−0.053	−0.063	−0.027	−0.032
2 C	0.476	0.459	0.455	0.464	0.462	0.468
2 O	−0.261	−0.272	−0.280	−0.268	−0.271	−0.276
3 N	−0.299	−0.346	−0.348	−0.325	−0.299	−0.326
3 HN	0.167	0.196	0.214	0.208	0.196	0.215
3 CA	0.021	0.0258	0.030	0.027	0.026	0.026
$\langle \lvert q \rvert \rangle$[c] (%)	na	12.5	na	4.7	12.2	8.7

[a]Atom to which induced surface charge is assigned. Number prefix indicates moiety: 1 = acetyl, 2 = alanyl, 3 = methylamide.

[b]Phi and Psi for *R*-alpha: (−60, −60) and beta: (−120, 120).

[c]Mean percentage change in induced charge. Beta conformation is relative to *R*-alpha conformation. Other figures are compared to time 0 fs of simulation. na: not applicable.

dipeptide in the C_7^{eq} conformation broken down into its contributions from individual charged atoms. In each case the net induced charge is very close to the theoretical value given by eq. (19). Atoms with a large surface accessibility, like the two carbonyl oxygens for example, which induce net charge of 0.951, are assigned a large proportion of their charge, with 40% of it being distributed over the other atoms. The total induced charge of ~0.2 assigned to each oxygen contains contributions from all the charged atoms. Less charge is assigned to atoms with lower surface accessibility, like the carbonyl carbons.

The effective charges on the atoms in the *R*-alpha and beta conformations are shown in Table IV, showing a mean difference of about 12%, indicating the type of change in effective charge expected from a large conformational difference. The degree to which the magnitudes of the effective charges change during a simulation are also shown in Table IV. The effective charge was calculated every 200 fs over several hundred femtoseconds of simulation. The mean percentage change is about 5% over the first 200 fs, increasing to 12% at 400 fs, then decreasing to 9% at 600 fs. Further changes over several picoseconds never exceed a value of 13% (data not shown).

DISCUSSION

A new method of including electrostatic effects of solvent into molecular dynamics is proposed here. The method involves treating all or part of the solvent as a continuum, and calculating the forces arising from the solvent reaction field using the PB equation. Due to numerical advances the PB equation can now be rapidly solved for arbitrary geometry, allowing for a practical combination with molecular dynamics. The contribution of ionic strength to the reaction field can also be included explicitly.

Comparison of numerical and analytical solutions to the PB equation for charges in a sphere indicates that the direction and magnitude of the forces on atomic charges due to the solvent reaction field can be calculated rapidly with reasonable accuracy, i.e., with no more than a 10% error. The procedure appears to retain its accuracy as the number of charges is increased. This is important since to be useful in a molecular dynamics simulation, one must calculate the forces on many atoms simultaneously if the procedures are to be combined efficiently. The calculations in Table II were all done on surface charges, i.e., those near the dielectric boundary, since this is where the errors in the FDPB procedure due to the finite resolution of the lattice are the largest, providing the most stringent test on accuracy. The errors on more buried charges will be smaller. With the more complex shape and charge distributions of real molecules the errors may increase, in which case the techniques of focussing and positional averaging[44] may be used to increase the accuracy. In addition, during the simulation the internal motions of the molecule will themselves alter the lattice mapping of the atoms from calculation to calculation. Thus over the course of longer simulations for which the method is designed, an effective averaging out of these errors may well occur.

The FDPB/DIEMOND method should provide a more realistic description of electrostatic solvation effects than any dielectric function model with little increase in computation since it explicitly includes the self-energy solvation term. It also includes the dependence of charge-charge interactions not only on their separation, but also on their position with respect to the molecule/solvent boundary.

Since the method of treating solvent screening described here is somewhat novel, the alanine dipeptide was initially chosen as model on which to test the integration of the force calculations and dynamics, since it has been well studied, and is simple enough that one can understand the effects of solvation in detail. The adiabatic energy surface for the dipeptide with no electrostatic interaction has minima in both the beta and *R*-alpha regions, with a small barrier in between, and these are the most favored conformations. The unscreened electrostatic energy surface makes the *R*-alpha region unfavorable. This can be thought of as the cost of aligning the two peptide dipoles in a parallel configuration, in competition with the favorable *in vacuo* intramolecular H-bond. The vacuum simulations on alanine dipeptide show the general features expected, i.e., the conformational preference is dominated by the intramolecularly H-bonded forms. On including solvation forces, there is a shift in favor of the C_7^{eq} conformer away from the C_7^{ax} conformer. This can be explained by the relatively poor solvation of the acetyl oxygen and the methyl amide hydrogen in the axial conformation arising from their proximity to the low dielectric bulk of the alanine methyl group. With solvent screening the dipole alignment is less costly, and conformations in both the beta and *R*-alpha regions are seen, with little barrier to interconversion. It might be thought that with an extensively solvated molecule such as alanine dipeptide, a uniform dielectric of 80 for the electrostatic term would be reasonable. This would scale down the electrostatic energy surface in Figure 4(a) by a factor of 80, making it negligible compared to the nonelectrostatic term. The population distribution obtained for such a case is essentially identical to that obtained with no electrostatics, Figure 3(b) (results not shown). In fact the cavity formed by the solute intensifies the interactions between charges compared to that expected in pure water, a quite general electrostatic effect.[43,45] Hence the effective screening and conformational preferences obtained are intermediate between the vacuum and pure water dielectric models. Recent simulations with explicit waters[54,56] indeed still show a decided preference for the internal H-bonded C_7^{eq} form, in spite of weakening by solvent screening/solvent H-bonding.

There have been a number of simulations of the conformational preferences of alanine dipeptide using explicit water models.[8,10,54,56] Various methods and potential functions were used, and the conformational free energy surfaces obtained in these studies show considerable variation in the relative energies, and even the existence of different minima. Detailed comparisons are thus difficult to make. However a direct comparison of the explicit TIP4P and continuum solvent models has been made for the alanine dipeptide in the *R*-alpha. C_7^{eq} and C_5 conformations[32] using the OPLS potential functions.[57] The electrostatic component of solvation energy in the explicit water model was obtained by free energy perturbation, giving 20.9, −13.1, and −12.3 kcal/mol respectively for the *R*-alpha, C_7^{eq}, and C_5 forms, while the continuum results obtained using the FDPB method were −21.1, −13.9, −12.9 kcal/mol respectively. This very good agreement between quite disparate methods suggests that differences seen in the conformational energy surface arise primarily from the different potential functions. The consensus from the explicit water studies is that solvent flattens the overall energy surface, reduces transition barriers, weakens the intramolecular H bond, especially in the C_7^{ax} form, and increases the probability of other conformations, notably the *R*-alpha helix. These general solvent effects are seen here also for the continuum solvent treatment, the population map in Figure 4(f) most closely resembling that obtained by Anderson and Hermans.[54] Taken together these results indicate that the FDPB and explicit water models provide very similar energetics of solvent screening. However the former, particularly for larger molecules, should require considerably less computation.

The analysis of the induced charge arising from each of the solute charges in Table III shows that it is distributed over several nearby atoms as well as the inducing atom itself. This reflects the long range nature of the electrostatic screening. Thus forces computed during the molecular dynamics simulation from the net effective charge on each atom represent both the Coulombic intermolecular term from that atom, and that portion of the reaction field induced by all the other charges which effectively acts at that atom. Thus the effective charge on an atom should not be thought of as a property of that atom alone.

Comparison of the effective charges in two quite different conformations such as *R*-alpha and beta give an idea of just how much the solvent screening varies with conformation. In this case the difference is only about 13%, since the solvent accessibility of the atoms in a small molecule cannot change much. Similarly the change as a function of simulation time is small, being less than 5% at 200 fs, suggesting that FDPB updates need not be done more frequently than this. In general for large molecules such as proteins, changes in conformation that expose or bury atoms will be reflected in larger changes in effective charge, and a similar analysis should be done to determine the appropriate frequency for the FDPB calculations.

Clearly the recalculation of the reduced charges at every step would add a significant computational burden to simulations, but there is another reason for not updating the charges too frequently. The dielectric relaxation time of water is on the order of tens of picoseconds,[58] much longer than the simulation time step. Ion atmosphere relaxation times are even longer, on the order of tenths of nanoseconds.[24]

Recalculating the reaction field at such short time intervals assuming a fully relaxed solvent may be inappropriate. There are two other considerations. Firstly the time step of the simulation is determined by the highest frequency motions, i.e., bond stretching, whereas torsional motions that correspond to large conformational changes, and hence significant changes in solvent accessibility, charge burial etc, take place over longer time scales. It is only these latter which will greatly change the solvent reaction field, and hence the effective charges. Secondly this approach to dynamics should be conceived primarily as a way of exploring conformational space, and for generating structures with energetically reasonable electrostatic energies as part of model building. In these applications one is not trying to simulate literal motions in real time, and this issue does not arise.

For simple systems such as the alanine dipeptide, both molecular dynamics and Monte Carlo methods have been used in explicit water simulations. As mentioned earlier, the method of including solvent reaction fields described here is particularly suited to the Monte Carlo method since in its purest form it requires only energy calculations and no force calculations. With more complex systems, however, the advantage of Monte Carlo will be less: As the number of degrees of freedom increase, the problem of efficient sampling of conformational space arises. This is especially so for macromolecules, where the major problem is generating random conformations that have good sterochemistry, i.e., reasonable energies, and thus good move acceptance rates. The most efficient Monte Carlo algorithms now use some kind of force bias method anyway,[59] so that force calculations are never really eliminated. Dynamics will be more efficient because the forces bias the "sampling" toward the most populated regions of phase space. Indeed one view of dynamics is that of a multiple particle move, full force bias Monte Carlo sampling method taking to its limit.

By recasting the reaction field in terms of a list of charges to be used with Coulomb's law, through eq. (15), the results of the FDPB calculation are easily integrated with the standard molecular dynamics algorithms, which are based on atomic charge lists. This leads to a relatively minor modification of the standard form for the electrostatic forces, eq. (18), which can be calculated with minim computational overhead. The main criterion in this initial study has been to maximize the speed and simplicity of the algorithm, and to minimize the modification to the dynamics algorithm. Whether such approximations are justified will have to be determined by further tests. However the resulting force field has the correct general physical properties, as demonstrated by the analysis of the model cubical "macromolecule." This is also apparent in the similar solvation effects for alanine dipeptide obtained with continuum and explicit water solvent models. In summary: (1) A buried charge feel a solvation force pulling it to the nearest surface. This force becomes larger the nearer to the surface the charge is, but disappears once the charge is on the surface. This solvation force appears even in the absence of any other charge, and results from the solvent polarization. (2) Charge-charge interactions are screened more effectively the nearer the charges are to the surface, and the farther apart the charges are. However the solvent will still affect the interaction between two charged atoms that are completely buried.

Other more elaborate ways of combining the reaction field calculations with molecular dynamics are also possible, for example by not combining the induced surface charge with the atomic charge. Another issue is whether to replace all the water in a simulation, or to leave one or two explicit layers of water, which can for example mediate specific H bonding and van der Waals interactions with the molecule, and act to 'lubricate' motions of surface sidechains.[60] Since there is no one criterion by which to judge a molecular dynamics procedure, only application to a range of problems will demonstrate the best approach and the method's usefulness. Appropriate future applications of the method might include the analysis of conformations of charged amino acid side chains, and comparison to crystal structures, and a variety of nucleic acid simulations where charge and ionic strength effects are extremely important. Tests on the ionic strength part of the method, and applications to larger molecules are currently underway, and will be presented in a later article.

I would like to thank Dr. Axel Brunger for providing XPLOR, Drs. Steve Harvey, Barry Honig, David Norman, and Tom Simonson for many helpful discussions. This research was supported by NIH grants GM-30518 and GM-41371

APPENDIX: REACTION FIELD FORCES

Starting with the general expression for the energy of an assembly of charges in a dielectric cavity due to their reaction field (eq. (5) of the text), consider the interaction of two charges q_a^f and q_b^f

$$\Delta G_{el}^o = q_a^f \phi_a^x/2 + q_b^f \phi_b^x/2 \qquad \text{(A1)}$$

where

$$\phi_a^x = \sum_o q_o^i/r_{ao}, \qquad \phi_b^x = \sum_o q_o^i/r_{bo} \qquad \text{(A2)}$$

The summation is over all the induced charges, q_o^i, and r is the distance between solute and induced surface charges. In a linear system, the induced surface charge is the sum of that induced by each charge alone

$$q_o^i = q_o^{ia} + q_o^{ib} \qquad \text{(A3)}$$

and eq. (A1) becomes

$$\Delta G^O_{el} = \tfrac{1}{2} q^f_a \left(\sum_o q^{ia}_o / r_{ao} + \sum_o q^{ib}_o / r_{ao} \right) + \tfrac{1}{2} q^f_b \left(\sum_o q^{ia}_o / r_{bo} + \sum_o q^{ib}_o / r_{bo} \right) \quad \text{(A4)}$$

The reaction field force on the first charge is the gradient of the energy with respect to its position:

$$\begin{aligned} f_a = \nabla(\Delta G^O_{el}) = \tfrac{1}{2} q^f_a \Big\{ & \sum_o q^{ia}_o \nabla(1/r_{ao}) + \sum_o \nabla(q^{ia}_o)/r_{ao} + \sum_o q^{ib}_o \nabla(1/r_{ao}) \\ & + \sum_o \nabla(q^{ib}_o)/r_{ao} \Big\} + \tfrac{1}{2} q^f_b \Big\{ \sum_o q^{ia}_o \nabla(1/r_{bo}) \\ & + \sum_o \nabla(q^{ia}_o)/r_{bo} + \sum_o q^{ib}_o \nabla(1/r_{bo}) \\ & + \sum_o \nabla(q^{ib}_o)/r_{bo} \Big\} \end{aligned} \quad \text{(A5)}$$

where the gradient operator is taken with respect to r_a. There are two types of terms in this expression: (1) Those containing gradients of $1/r$ which arise from the gradient in potential at q_a, and/or the movement of the dielectric boundary, i.e., changes in the positions of the induced charge. (2) Terms containing gradients of q_o which arise from changes in the magnitude of the induced charge due to change in the field at the boundary. Considering each term in turn, the first summation yields the reaction field at charge **a** due to itself, plus the variation due to the boundary movement in **a**'s own reaction potential. The second gives the change in reaction potential due to the variation in charge induced by **a**. The third is the reaction field at **a** due to **b**, plus the variation due to boundary movement in **b**'s reaction potential at **a**. The fourth is the variation in reaction potential at **a** arising from the change in charge induced by **b** due to the boundary movement. The fifth is the change due to boundary movement in reaction potential at **b** arising from charge induced by **a**. The sixth is the variation in reaction potential at **b** arising from the change in q^{ia}_a with respect to the position of **a**. The seventh is the change due to boundary movement in the reaction potential at **b** arising from charge induced by **b**. The eighth is the change in reaction potential at **b** arising from variation in the charge induced by **b** due to the shift in dielectric boundary.

For a rigid dielectric boundary, $1/r_{bo}$ and q^{ib}_o will be independent of the position of **a**, so terms four, five, seven, and eight are zero. Variations due to boundary movement in the other terms disappear, reducing the first summation to E^a_a, the reaction field at charge **a** due to itself, and reducing the third summation to E^b_a, the reaction field at **a** due to **b**. For the special cases of planar and spherical boundaries, where analytical and series solutions are available[46,50,61] it is straightforward to show that terms one and two, and terms three and six are equal. The reaction field force is then

$$f_a = q^f_a (E^a_a + E^b_a) \quad \text{(A6)}$$

and the total force on atom **a** is

$$f_a = q^f_a (E^a_a + E^b_a + q_b/r_{ab}) = q^f_a E_a \quad \text{(A7)}$$

where q^f_b/r_{ab} is the Coulombic field from **b** at **a**, and E_a is the total field at **a** arising from all charges except **a**. More generally for the case of an arbitrary shaped rigid dielectric boundary, for which analytical solutions are not available, the principle of reciprocity may be invoked to show that the third and sixth terms of (A5) are equal as follows. For a linear system, we have by reciprocity $q^f_a \phi^b_a = q^f_b \phi^a_b$, where ϕ^b_a is the total potential at **a** due to **b**, and ϕ^a_b is the total potential at **b** due to **a**. Since reciprocity holds for the Coulombic part of the potential alone, it must also hold for the reaction field part, so $q^f_a \sum_o q^{ib}_o / r_{ao} = q^f_b \sum_o q^{ia}_o / r_{bo}$. Since reciprocity holds irrespective of the positions of the charges, we may equate the gradients with respect to r_a:

$$\nabla \left(q^f_a \sum_o q^{ib}_o / r_{ao} \right) = \nabla \left(q^f_b \sum_o q^{ia}_o / r_{bo} \right) \quad \text{(A8)}$$

giving

$$q^f_a \sum_o q^{ib}_o \nabla(1/r_{ao}) = q^f_b \sum_o \nabla(q^{ia}_o)/r_{bo} \quad \text{(A9)}$$

For the interaction of **a** with its own reaction field, this may be dealt with as a special case of (A9) with $r_{ao} = r_{bo}$, and $q^f_a = q^f_b$, where **a** and **b** are considered in turn as sources of the reaction potential acting at the other charge, to yield the first and second terms of the equality respectively:

$$q^f_a \sum_o q^{ia}_o \nabla(1/r_{ao}) = q^f_a \sum_o \nabla(q^{ia}_o)/r_{bo} \quad \text{(A10)}$$

With eqs. (A9) and (A10) the total force on each atom is again the product of its charge and the total field. Note that this result is valid with or without mobile ions, so long as the system response is linear. It is also valid for any number of charges in the cavity.

In the case where the dielectric boundary is not rigid, the fourth, fifth, seventh, and eighth terms of eq. (A5) will also contribute to the force. Reciprocity may be used to show that the sum of the third and fourth terms equals the sum of the fifth and sixth, and that the first and second terms again are equal. Unfortunately there doesn't appear to be any general result analogous to eq. (A7) in the literature. Since the derivatives with respect to movements of the dielectric boundary cannot be evaluated analytically it is difficult to use eq. (A5) directly. However preliminary examinations of numerical solutions indicate that the variations due to boundary movement

are smaller than the others. Consequently, the (now approximate) eq. (A7) is used to compute the forces.

References

1. J.A. McCammon and S.C. Harvey, *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge, 1987.
2. S.C. Harvey, *Proteins*, **5,** 78 (1989).
3. M. Levitt, *Cold Spring Harbor Symp. on Quant. Biol.*, **47,** 251 (1983).
4. S.C. Harvey and J.A. McCammon, *Nature*, **294,** 286 (1981).
5. J. McCammon, P.G. Wolynes, and M. Karplus, *Biochem.*, **18,** 927 (1979).
6. B.E. Hingerty, R.H. Ritchie, T.L. Ferrell, and J.E. Turner, *Biopolymers*, **24,** 427 (1985).
7. A.T. Hagler, and J. Moult, *Nature*, **272,** 222 (1978).
8. P.J. Rossky and M. Karplus, *J. Am. Chem. Soc.*, **101,** 1913 (1978).
9. J. Hermans, H.J. Berendsen, W.F. van Gunsteren, and J.P. Postma, *Biopolymers*, **23,** 1513 (1984).
10. M. Mezei, P.K. Mehotra, and D.L. Beveridge, *J. Am. Chem. Soc.*, **107,** 2239 (1986).
11. M. Wojcik and E. Clementi, *J. Phys. Chem.*, **84,** 5970 (1986).
12. P. Barnes, J.L. Finney, J.D. Nicholas, and J.E. Quinn, *Nature*, **282,** 459 (1979).
13. T.P. Lybrand and P.A. Kollman, *J. Phys. Chem.*, **83,** 2923 (1985).
14. M. Sprik and M.L. Klein, *J. Phys. Chem.*, **89,** 7556 (1988).
15. W.F. van Gunsteren, H.J. Berendsen, J. Hermans, W.G. Hol, and J.P. Postma, *Proc. Natl. Acad. Sci.*, **80,** 4315 (1983).
16. C.L. Brooks, A. Brunger, and M. Karplus, *Biopolymers*, **24,** 843 (1985).
17. H.L. Friedman, *Mol. Physics*, **29,** 1533 (1975).
18. M. Neumann, *J Chem. Phys.*, **85,** 1567 (1986).
19. W.F. v. Gunsteren, H.J. Berendsen, and J.A. Rullmann, *Faraday Discussions*, **66,** 58 (1978).
20. J. Chandrasekhar, D.C. Spellmeyer, and W.L. Jorgensen, *J Am. Chem. Soc.*, **106,** 903 (1984).
21. B. Jayaram, R. Fine, K.A. Sharp, and B. Honig, *J Phys. Chem.*, **93,** 4320 (1989).
22. U.C. Singh, S.J. Weiner, and P.A. Kollman, *Proc. Natl. Acad. Sci.*, **82,** 755 (1985).
23. G.L. Seibel, U.C. Singh, and P.A. Kollman, *Proc. Natl. Acad. Sci.*, **82,** 6537 (1986).
24. J.O. Bockris and A.K.N. Reddy, *Modern Electrochemistry*, Plenum Press, New York, 1970.
25. A. Warshel and S. Russell, *Quart. Rev. Biophys.*, **17,** 283 (1984).
26. A. Warshel, G. Naray-Szabo, F. Sussman, and J. Hwang, *Biochemistry*, **28,** 3629 (1989).
27. M. Gilson and B. Honig, *J. Comp. Aided Molecular Design* (in press).
28. R.O. Watts, *Chem. Phys.*, **57,** 185 (1981).
29. K. Sharp and B. Honig, *Ann. Rev. Biophys. Biophys. Chem.*, **19,** 301 (1990).
30. A.A. Rashin and B. Honig, *J Phys. Chem.*, **89,** 5588 (1985).
31. A.A. Rashin, and K. Namboodiri, *J Phys. Chem.*, **91,** 6003 (1987).
32. J. Jean-Charles, A. Nicholls, K. Sharp, B. Honig, A. Tempczyk, T. Hendrickson, C. Still, *J. Am Chem. Soc.* (In Press).
33. K.A. Sharp, A. Jean-Charles, and B. Honig, *J Phys. Chem.* (In Press).
34. M.K. Gilson and B. Honig, *Proteins*, **4,** 7 (1988).
35. K. Soman, A.-S. Yang, B. Honig, and R. Fletterick, **28,** 9918 (1989).
36. D. Bashford and M. Karplus, *J Mol. Biol.* (in press).
37. M.E. Davis and J.A. McCammon, *J Comp. Chem.*, **10,** 386 (1989).
38. A. Nicholls and B. Honig, *J. Comp. Chem.*, (in press).
39. F. Hirata, P. Redfern, and R. Levy, *Int. J Quantum Chem.*, **15,** 179 (1989).
40. B. Roux, H.A. Yu, and M. Karplus, *J. Phys. Chem.*, **94,** 4683 (1990).
41. K. Sharp and B. Honig, *J. Phys. Chem.*, **94,** 7684 (1990).
42. J. Warwicker and H.C. Watson, *J. Mol. Biol.*, **157,** 671 (1982).
43. B. Jayaram, K.A. Sharp, and B. Honig, *Biopolymers*, **28,** 975 (1989).
44. M. Gilson, K.A. Sharp, and B. Honig, *J. Comp. Chem.*, **9,** 327 (1988).
45. I. Klapper, R. Hagstrom, R. Fine, K. Sharp, and B. Honig, *Proteins*, **1,** 47 (1986).
46. J.D. Jackson, *Classical Electrodynamics*, John Wiley and Sons, New York, 1962.
47. B.R. Brooks, et al., *J Comp. Chem.*, **4,** 187 (1983).
48. S.J. Weiner, P.A. Kollman, D.T. Nguyen, and D.A. Case, *J. Comp. Chem.*, **7,** 230 (1986).
49. B. Lee and F.M. Richards, *J Mol. Biol.*, **55,** 379 (1971).
50. C. Tanford and J.G. Kirkwood, *J Am. Chem. Soc.*, **79,** 5333 (1957).
51. T.L. Hill, *J. Phys. Chem.*, **60,** 253 (1956).
52. XPLOR: A. Brunger, Copyright Presidents and Fellows, Harvard University, Cambridge, Mass. 1987.
53. J.P. Valleau and G.M. Torrie, in *Modern Theoretical Chemistry*, B.J. Berne, Eds., Plenum, New York, 1977, pp. 169.
54. A.G. Anderson and J. Hermans, *Proteins*, **3,** 262 (1988).
55. S. Paine and H. Scheraga, *Biopolymers*, **68,** 1391 (1985).
56. B.M. Pettitt and M. Karplus, *Che. Phys. Lett.*, **121,** 194 (1985).
57. W. Jorgensen and J. Gao, *J Am. Chem. Soc.*, **110,** 4212 (1988).
58. C.H. Collie, J.B. Hasted, and D.M. Ritson, *Proc. Royal Soc.*, **60,** 145 (1948).
59. D. Beveridge and F.M. DiCapua, *Ann. Rev. Biophys. Biophys. Chem.*, **18,** 431 (1989).
60. F.R. Salemme, P.C. Weber, J.B. Matthew, and J.J. Wendolowski, *Abs. PAP ACS*, **196,** 276 (1988).
61. M. Gilson, A. Rashin, R. Fine, and B. Honig, *J. Mol. Biol.*, **183,** 503 (1985).