JCTC Journal of Chemical Theory and Computation

An Analytical Electrostatic Model for Salt Screened Interactions between Multiple Proteins

Itay Lotan* and Teresa Head-Gordon*,†

Department of Bioengineering, University of California, Berkeley, Berkeley, California 94720

Received October 24, 2005

Abstract: We present a new general analytical solution for computing the screened electrostatic interaction between multiple macromolecules of arbitrarily complex charge distributions, assuming they are well described by spherical low dielectric cavities in a higher dielectric medium in the presence of a Debye−Hückel treatment of salt. The benefits to this approach are 3-fold. First, by exploiting multipole expansion theory for the screened Coulomb potential, we can describe direct charge−charge interactions and all significant higher-order cavity polarization effects between low dielectric spherical cavities containing their charges, while treating these higher order terms correctly at all separation distances. Second, our analytical solution is general to arbitrary numbers of macromolecules, is efficient to compute, and can therefore simultaneously provide on-the-fly updates to changes in charge distributions due to protein conformational changes. Third, we can change spatial resolutions of charge description as a function of separation distance without compromising the desired accuracy. While the current formulation describes solutions based on simple spherical geometries, it appears possible to reformulate these electrostatic expressions to smoothly increase spatial resolution back to greater molecular detail of the dielectric boundaries.

1. Introduction

Atomistic molecular dynamics simulations routinely and beneficially address materials problems in the condensed phase. However, there is another set of problems on the supramolecular scale where the limitations of size and time scales are reached, examples being the recognition events and subsequent complexation of multiple proteins in explicit solvent environments, or the study of phase behavior and interfacial properties of colloid systems. Atomistic modeling is too computationally demanding to evaluate for times long enough to measure the macromolecules' traversal over large spatial domains and to do this with enough statistical confidence to analyze complex phase behavior, association kinetics, or mechanism of assembly.

Fortunately, coarse-grained models may actually be the more sensible approach when large-scale spatial organization or dynamic events occurring over long time scales are operative. Spatial coarse-graining would involve, for example, removing explicit solvent molecules and ions as well as ignoring an individual macromolecule's internal motion for some period of time.

At large spatial separations between charged macromolecules in solution, electrostatic interactions will dominate, so that an appropriate coarse-grained model could describe them as complex charge distributions imbedded in a low dielectric medium surrounded by a high dielectric solvent continuum with salt screening defined by explicit microions or implicitly through a Debye−Hückel treatment. The electrostatic potential and forces and torques are found by solving the full Poisson−Boltzmann equation (PBE) or when expanding the exponential and keeping only linear terms, by solving the corresponding linearized PBE. Both linearized and the full PBE are typically solved numerically using either a finite-difference (FD) method, boundary-element (BE) meth-

* Corresponding authors e-mail: itayl@berkeley.edu (I.L.) and TLHead-Gordon@lbl.gov (T.H.-G.).

† Schlumberger Fellow, Chemistry Department, Cambridge University, Lensfield Road, Cambridge CB2 1EW, United Kingdom.

10.1021/ct050263p CCC: $33.50
Published on Web 03/16/2006

ods or based on more recent work with an adaptive multilevel finite element solution,[1] and the associated electrostatic fields are evaluated by a central difference approximation using the numerical potential. An additional feature to solving the full or linearized PBE numerically is that solutions are not restricted to simple geometries but can describe greater levels of molecular detail of the macromolecular cavity shape.

In this work, we have developed a new analytical solution for computing the electrostatic interaction between multiple macromolecules of arbitrarily complex charge distributions in aqueous salt solutions, assuming the macromolecules and their environment are well described by spherical low dielectric cavities in the presence of a Debye−Hückel treatment of salt.[2] By exploiting multipole expansion theory for the screened Coulomb potential, we are able to describe direct charge−charge interactions and all significant higher-order cavity polarization effects between low dielectric spherical cavities containing their charges, treating all higher order polarization terms correctly at all separation distances. Hitherto this level of completeness was only available through numerical solutions to the PBE, such as the grid-based FD solutions[3−6] and BE methods[7−9] (See ref 1 for a survey.). While the numerical methods are able to represent the shape of the molecule to high resolution, they are expensive to compute. They have an inherent tradeoff between spatial resolution (grid discretization for the FD methods and choice of boundary elements for the BE methods) and memory and time requirements, that limits them to system of few macromolecules. In contrast, our method is efficient and fast to compute even for many macromolecules, including frequent updates of changes in their charge distributions due to induced conformational changes. Unlike the effective charge approximation of Gabdoulline and Wade,[10] our formalism allows for automatically changing spatial resolutions of charge description on the fly as a function of separation distance so the desired accuracy is always achieved with no waste in computation.

Previous analytical solutions to the linearized PBE under spherical geometries have fallen short of providing complete and useable solutions for multiple macromolecules of arbitrarily complex charge distribution. When the system is comprised of only one macromolecule, a complete solution was offered by Kirkwood[11] more than 70 years ago, and a similar solution has recently been derived by Hoffman et al.[12] The interaction of two macromolecules has proven to be more difficult, and many different partial and approximate solutions have been proposed. Glendinning and Russel[13] proposed an analytical solution using multipole expansion for two equal spheres with uniform charge density. Assuming weak interaction between the two molecules, a superposition approximation was used, first by Verwey and Overbeek[14] for uniformly charged molecules and later by Sader and Lenhoff[15] for arbitrary charge distributions. Allen and Hansen[16] proposed a solution based on variational charge-density functional method for a system of two macromolecules, each with one charge on the line connecting their centers, and with no salt. Phillies[17] was the first to offer a complete analytical solution for two molecules with arbitrary charge distributions in solution with salt; however, his use of implicit re-expansion operators made the solution cumbersome and impractical. Later McClurg and Zukoski[18] improved on this solution by using, among other things, explicit re-expansion operators. While their solution is similar to ours, it is limited to the first few multipole orders and to only two molecules and thus is not as complete and general as the solution we provide here.

In devising this new formulation, much attention and care were directed at providing a computationally practical and reliable method. Our method is numerically stable even for large molecular radii and high multipole orders, which makes it usable even at short separation distances, independent of the Debye screening length. Our formulation also provides easy access to many quantities of interest, such as forces, torques, and interaction energies, and affords an intuitive physical understanding of the different factors that contribute to them. Our approach takes advantage of recent advances in the computation of Yukawa (screened Coulomb) potential[19,20] and the Helmholtz equation,[21,22] that makes the derivation and use of the re-expansion operator straightforward and computationally efficient. While the current analytical approach exhibits a reliance on simple spherical geometries, it appears possible to reformulate these electrostatic expressions to increase spatial resolution back to molecular level descriptions of cavity geometries.

2. Theory

2.1. Mathematical Preliminaries. In this section we will establish some definitions and recount some identities that will be used in developing the theory needed for computing the electrostatic interaction between charged spherical molecules. First, a definition of the *spherical harmonics*, following the work of Gumerov and Duraiswami[21]

$$Y_{n,m}(\theta,\phi) = (-1)^m \sqrt{\frac{(n-m)!}{(n+m)!}} P_{n,m}(\cos\theta) e^{im\phi} \qquad (1)$$

where $P_{n,m}(x)$ are the *Associated Legendre Polynomials*.[23] Note, this is not the standard definition of spherical harmonics, as it differs by a $(-1)^m$ factor. The resulting expressions, however, are simpler. This definition can be used for all $m \geq 0$. When m is negative we rely on the simple identity

$$Y_{n,-m}(\theta,\phi) = \overline{Y_{n,m}}(\theta,\phi) \qquad (2)$$

where $\overline{Y_{n,m}}$ is the complex conjugate of $Y_{n,m}$.

We also require the use of the *modified spherical Bessel functions* (MSBF) defined as

$$i_n(z) = \sqrt{\frac{\pi}{2z}} I_{n+1/2}(z)$$

$$k_n(z) = \sqrt{\frac{\pi}{2z}} K_{n+1/2}(z) \qquad (3)$$

where $I_n(z)$ and $K_n(z)$ are the *modified Bessel functions* of the first and second kinds, respectively.[23] We will make use of two addition theorems that use the MSBFs and the spherical harmonics.[23] For convenience we copy them below. Given two points in 3-D space specified by their spherical coordinates $\mathbf{p} = [\rho,\vartheta,\varphi]$ and $\mathbf{t} = [r,\theta,\phi]$, with $\rho < r$, and the Euclidean distance between them $R = |\mathbf{t} - \mathbf{p}|$, we have

$$\frac{1}{R} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{\rho^n}{r^{n+1}} \overline{Y_{n,m}}(\vartheta,\varphi) Y_{n,m}(\theta,\phi) \quad (4)$$

$$\frac{e^{-\kappa R}}{R} = \frac{2\kappa}{\pi} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} (2n+1) i_n(\kappa\rho) k_n(\kappa r) \overline{Y_{n,m}}(\vartheta,\varphi) Y_{n,m}(\theta,\phi) \quad (5)$$

Note that while in eq 4 the exponentiated ratio of distances ρ^n/r^{n+1} appears explicitly; in eq 5 it is implicit in the MSBFs. Namely, $k_n(r) \propto 1/r^{n+1}$ and $i_n(r) \propto \rho^n$.

Kirkwood developed a general analytic solution to the linearized Poisson−Boltzmann equation (PBE) around a spherical solvent-excluding dielectric cavity

$$\Phi(\mathbf{t}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{A_{n,m}}{r^{n+1}} e^{-\kappa r} \hat{\mathrm{k}}_n(\kappa r) Y_{n,m}(\theta,\phi) \quad (6)$$

where the coefficients $A_{n,m}$ are determined by enforcing appropriate boundary conditions. His solution uses an adapted definition of the MSBF suitable for the conditions of the problem:

$$\hat{\mathrm{k}}_n(z) = \frac{2}{\pi} \frac{e^z z^{n+1}}{(2n-1)!!} k_n(z) \quad (7)$$

Based on this definition we can now rewrite the addition theorem in eq 5

$$\frac{e^{-\kappa R}}{R} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{\rho^n}{r^{n+1}} \hat{\mathrm{i}}_n(\kappa\rho) e^{-\kappa r} \hat{\mathrm{k}}_n(\kappa r) \overline{Y_{n,m}}(\vartheta,\varphi) Y_{n,m}(\theta,\phi) \quad (8)$$

where

$$\hat{\mathrm{i}}_n(z) = \frac{(2n+1)!!}{z^n} i_n(z) \quad (9)$$

Besides their usefulness in solving the Poisson−Boltzmann equation around spherical cavities, these *adapted* MSBFs are also more numerically stable than the *standard* MSBFs since the dependence on exponents of r and ρ that may grow rapidly with the order n of the functions has been taken out. A recursive method for directly computing $\hat{\mathrm{k}}_n$ and $\hat{\mathrm{i}}_n$ without first computing k_n and i_n is given in the Appendix.

2.2. Boundary-Value Problem. The system we are solving is comprised of N spherical molecules of radii a_i, whose centers are positioned at points $\mathbf{c}^{(i)}$. The molecules are immersed in a solvent, containing charged salt ions. The dielectric constant (permittivity) in the interior of each molecule is ϵ_p, and the dielectric constant of the solvent is ϵ_s. For simplicity we will assume all molecules have the same dielectric constant, although that is not a restriction of the solution. The inverse Debye screening length of the ions in the solution is κ.

The electrostatic potential of the system is governed by the PBE. For systems under physiological conditions the PBE can be safely linearized to give[2]

$$\nabla[\epsilon(\mathbf{r})\nabla\Phi(\mathbf{r})] - \epsilon(\mathbf{r})\kappa^2\Phi(\mathbf{r}) = 4\pi\rho(\mathbf{r}) \quad (10)$$

In the general case, when the dielectric boundary has arbitrary shape, only numerical solutions exist. For the case of spherical cavities, however, one can write general parametric expressions for the potential in different parts of the system. The values of the coefficients are then determined by enforcing a set of boundary conditions. These conditions stipulate the continuity of the electrostatic potential and the electrostatic field at the surface of each molecule. These conditions take the form of the following equations

$$\Phi_{out}^{(i)} = \Phi_{in}^{(i)}|_{r=a_i}$$
$$\epsilon_s \frac{\partial\Phi_{out}^{(i)}}{\partial r} = \epsilon_p \frac{\partial\Phi_{in}^{(i)}}{\partial r}\Big|_{r=a_i} \quad (11)$$

where $\Phi_{in}^{(i)}$ is the total potential inside molecule i and $\Phi_{out}^{(i)}$ is the total potential in the solvent, both expressed in the coordinate frame of molecule i.

The system at hand consists of two types of regions: the inside of each molecule and the outside solvent. The inside region is characterized by a low dielectric constant and an arbitrary distribution of free charges. In this region eq 10 reduces to $\epsilon_p\Delta\Phi(\mathbf{r}) = 4\pi\rho(\mathbf{r})$. The outside region is characterized by the high dielectric constant of the solvent and the presence of salt. In this region eq 10 reduces to $\Delta\Phi(\mathbf{r}) - \kappa^2\Phi(\mathbf{r}) = 0$. Following Kirkwood[11] we write general parametrized expressions for both regions. The electrostatic potential inside molecule i is described by

$$\Phi_{in}^{(i)}(\mathbf{t}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{E_{n,m}^{(i)}}{\epsilon_p r^{n+1}} + B_{n,m}^{(i)} r^n \right) Y_{n,m}(\theta,\phi) \quad (12)$$

where $\mathbf{E}^{(i)}$ is the multipole expansion of the charges inside molecule i. It is defined as

$$E_{n,m}^{(i)} = \sum_{j=1}^{M_i} q_j^{(i)} (\rho_j^{(i)})^n Y_{n,m}(\vartheta_j^{(i)}, \varphi_j^{(i)}) \quad (13)$$

where M_i is the number of charges in molecule i, $q_j^{(i)}$ is the magnitude of the jth charge, and $\mathbf{p}_j^{(i)} = [\rho_j^{(i)}, \vartheta_j^{(i)}, \varphi_j^{(i)}]$ is its position in spherical coordinates. The expansion in eq 13 can be used in eq 12 only when $a_i \geq r > \max_j \rho_j^{(i)}$, a condition that is always met when computing the potential at the surface of the sphere. Its use facilitates the application of the boundary conditions and the subsequent computation of the coefficient vectors $\mathbf{B}^{(i)}$. If later on one desires to compute the potential at a point inside the molecule, the following variant of eq 12 can be used:

$$\Phi_{in}^{(i)}(\mathbf{t}) = \frac{1}{\epsilon_p} \sum_{j=1}^{M_i} \frac{q_j^{(i)}}{||\mathbf{p}_j^{(i)} - \mathbf{t}||} + \sum_{n=0}^{\infty} \sum_{m=-n}^{n} B_{n,m}^{(i)} r^n Y_{n,m}(\theta,\phi) \quad (14)$$

The general form of the potential outside all molecules (in a coordinate frame whose origin is the center of molecule i) is

$$\Phi_{out}^{(i)}(\mathbf{t}) = \frac{1}{\epsilon_s} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{A_{n,m}^{(i)}}{r^{n+1}} e^{-\kappa r} \hat{\mathrm{k}}_n(\kappa r) + L_{n,m}^{(i)} r^n \hat{\mathrm{i}}_n(\kappa r) \right) Y_{n,m}(\theta,\phi) \quad (15)$$

The vectors $\mathbf{A}^{(i)}$ and $\mathbf{B}^{(i)}$ of eqs 12 and 15 are the free parameters (coefficients) that will be determined through the application of the boundary conditions of eq 11.

The coefficients $\mathbf{L}^{(i)}$ in eq 15 are a re-expansion of the external potential coefficients $\mathbf{A}^{(j)}$, $j \neq i$ of all other molecules in the system. It is defined as

$$\mathbf{L}^{(i)} = \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathrm{T}^{(i,j)} \cdot \mathbf{A}^{(j)} \tag{16}$$

where $\mathrm{T}^{(i,j)}$ is the linear re-expansion operator that transforms a multipole expansion at $\mathbf{c}^{(j)}$ to a local (Taylor) expansion at $\mathbf{c}^{(i)}$. This operator is described in detail in the Appendix. The use of the $\mathrm{T}^{(i,j)}$ operators allows us to represent the potentials due to all molecules in the coordinate frame of a single molecule. This is crucial to our ability to analytically apply the boundary conditions and derive compact expressions for the free parameters of the formulation.

Applying the boundary conditions to eqs 12 and 15 yields the following expressions for the elements of the parameter vectors

$$A_{n,m}^{(i)} = \gamma_n^{(i)} \delta_n^{(i)} L_{n,m}^{(i)} + \gamma_n^{(i)} E_{n,m}^{(i)} \tag{17}$$

$$B_{n,m}^{(i)} = \frac{1}{\epsilon_s}\left(\frac{A_{n,m}^{(i)}}{a_i^{2n+1}} e^{-\kappa a_i} \hat{\mathrm{k}}_n(\kappa a_i) + L_{n,m}^{(i)} \hat{\mathrm{i}}_n(\kappa a_i)\right) - \frac{E_{n,m}^{(i)}}{\epsilon_p a_i^{2n+1}} \tag{18}$$

where the vectors of constants $\Gamma^{(i)}$ and $\Delta^{(i)}$ are

$$\gamma_n^{(i)} = \frac{(2n+1)e^{\kappa a_i}}{(2n+1)\hat{\mathrm{k}}_{n+1}(\kappa a_i) + n\hat{\mathrm{k}}_n(\kappa a_i)(\epsilon_p/\epsilon_s - 1)} \tag{19}$$

$$\delta_n^{(i)} = \frac{a_i^{2n+1}}{2n+1}\left[\kappa^2 a_i^2 \frac{\hat{\mathrm{i}}_{n+1}(\kappa a_i)}{2n+3} + n\hat{\mathrm{i}}_n(\kappa a_i)(1 - \epsilon_p/\epsilon_s)\right] \tag{20}$$

Given a solution to all the $\mathbf{A}^{(i)}$ vectors, the $\mathbf{B}^{(i)}$ vectors are easily calculated using eq 18. We therefore concentrate on solving for $\mathbf{A}^{(i)}$. Based on eq 17 we can write a system of linear equations that describes the electrostatics of the entire problem

$$\mathbf{A} = \Gamma \cdot (\Delta \cdot \mathrm{T} \cdot \mathbf{A} + \mathbf{E}) \tag{21}$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \\ \vdots \\ \mathbf{A}^{(N)} \end{bmatrix} \quad \Gamma = \begin{bmatrix} \Gamma^{(1)} & 0 & \cdots & 0 \\ 0 & \Gamma^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Gamma^{(N)} \end{bmatrix}$$

$$\Delta = \begin{bmatrix} \Delta^{(1)} & 0 & \cdots & 0 \\ 0 & \Delta^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Delta^{(N)} \end{bmatrix}$$

$$\mathrm{T} = \begin{bmatrix} 0 & \mathrm{T}^{(1,2)} & \cdots & \mathrm{T}^{(1,N)} \\ \mathrm{T}^{(2,1)} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{T}^{(N-1,N)} \\ \mathrm{T}^{(N,1)} & \cdots & \mathrm{T}^{(N,N-1)} & 0 \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} \mathbf{E}^{(1)} \\ \mathbf{E}^{(2)} \\ \vdots \\ \mathbf{E}^{(N)} \end{bmatrix} \tag{22}$$

and

$$\Gamma^{(i)} = \begin{bmatrix} \gamma_1^{(i)} & 0 & \cdots & 0 \\ 0 & \gamma_2^{(i)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \gamma_p^{(i)} \end{bmatrix} \quad \Delta^{(i)} = \begin{bmatrix} \delta_1^{(i)} & 0 & \cdots & 0 \\ 0 & \delta_2^{(i)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \delta_p^{(i)} \end{bmatrix} \tag{23}$$

Note that in order to write the system of equations we need to truncate all vectors $\mathbf{A}^{(i)}$ and concordantly all other coefficient vectors at a finite order (the maximal value for n), which we call p. The appropriate choice for p depends on practical considerations, which we will elaborate on in the next section.

We can give intuitive physical meaning to the matrices and vectors of eq 21. The vectors $\mathbf{A}^{(i)}$ represent the *effective* multipole expansion of the charge distributions of each molecule. The effective expansion represents an equivalent system, where the low dielectric cavities have been removed and the solvent (with the salt ions) is allowed to penetrate everywhere. Namely, one can use these effective expansions to represent the now missing dielectric boundary. This effective expansion is reminiscent of the effective charges of Gabdoulline and Wade[10] in the sense that it incorporates the effect of the dielectric boundary into the charge distribution. In contrast to Gabdoulline and Wade, our effective expansion also accounts for the polarization effects caused by the dielectric boundary. The diagonal matrix Δ can be understood as a cavity polarization operator. It transforms the coefficients of an external charge distribution (re-expanded around the center of the cavity) to yield the expansion coefficients of a polarization charge distribution that forms on the surface of the cavity (due to that external charge distribution). The diagonal matrix Γ can be taken to be a dielectric boundary crossing operator. It transforms the coefficients of a charge distribution expansion inside a dielectric cavity to yield the effective expansion as if the charges were in the solution. Both the Γ and Δ operators are functions of the parameters of the system being solved, namely the radii a_i, the dielectric constants ϵ_p and ϵ_s, and the inverse Debye screening length κ . Thus eq 21 can be understood to state the intuitive fact that the external potential field induced by a molecule is the sum of the contribution of its free charges and the contribution of polarization charges induced by other molecules, transformed by the effect of its dielectric boundary.

It is instructive to note that this formulation separates the solution into a free charge distributions (the vector $\mathbf{E}^{(i)}$) and operators that represent the configuration of the system (the matrices $\mathrm{T}^{(i,j)}$) and the geometric and physical conditions (the operators $\Gamma^{(i)}$ and $\Delta^{(i)}$). The charge distributions and the operators are independent except for the fact that they all need to be defined in terms of the same molecular centers. This observation may be exploited in devising approximate schemes that can deal with nonspherical molecules.

2.3. Interaction Energy. Plugging the expression in eq 18 into eq 12 and rearranging the terms yields

$$\Phi_{in}^{(i)}(\mathbf{t}) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left[\frac{E_{n,m}^{(i)}}{\epsilon_p}\left(\frac{1}{r^{2n+1}} - \frac{1}{a_i^{2n+1}}\right) + \frac{\gamma_n^{(i)} E_{n,m}^{(i)}}{\epsilon_s a_i^{2n+1}} e^{-\kappa a_i}\hat{\mathrm{k}}_n(\kappa a_i)\right] r^n Y_{n,m}(\theta,\phi) + \frac{1}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\gamma_n^{(i)} L_{n,m}^{(i)} r^n Y_{n,m}(\theta,\phi) \quad (24)$$

Note that the first double summation is the internal potential due to the charges inside the molecule and the effects of the external salt ions, while the second double summation is the potential due to the other charged molecules (external sources of charge), which we dub $\hat{\Phi}_{in}^{(i)}$:

$$\hat{\Phi}_{in}^{(i)}(\mathbf{t}) = \frac{1}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\gamma_n^{(i)} L_{n,m}^{(i)} r^n Y_{n,m}(\theta,\phi) \quad (25)$$

Plugging eq 17 into eq 15 yields the expression for the potential anywhere in the solvent (outside all molecules) with the coordinate frame centered at molecule i

$$\Phi_{out}^{(i)}(\mathbf{t}) = \frac{1}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\gamma_n^{(i)}(\delta_n^{(i)} L_{n,m}^{(i)} + E_{n,m}^{(i)})\frac{e^{-\kappa r}\hat{\mathrm{k}}_n(\kappa r)}{r^{n+1}} Y_{n,m}(\theta,\phi) + \frac{1}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n} L_{n,m}^{(i)} r^n \hat{\mathrm{i}}_n(\kappa r) Y_{n,m}(\theta,\phi) \quad (26)$$

Note that by analogy to eq 6, the expansion of the total charge distribution of a molecule, dubbed $\mathbf{G}^{(i)}$, taking into account both its free charges and the charges due to the polarization of its dielectric cavity by external charges, is

$$G_{n,m}^{(i)} = \delta_n^{(i)} L_{n,m}^{(i)} + E_{n,m}^{(i)} \quad (27)$$

The interaction energy $\Omega^{(i)}$ of each molecule is the product of its charge distribution (both free and polarization charges) of eq 27 with the potential due to external sources of eq 25. We can compute this as the inner product of the coefficients of the two corresponding multipole expansions

$$\Omega^{(i)} = \frac{1}{\epsilon_s}\langle \Gamma^{(i)}\cdot\mathbf{L}^{(i)}, \mathbf{G}^{(i)}\rangle = \frac{1}{\epsilon_s}\langle \mathbf{L}^{(i)}, \mathbf{A}^{(i)}\rangle \quad (28)$$

where the inner product of two coefficient vectors is defined as

$$\langle \mathbf{U},\mathbf{V}\rangle = \sum_{n=0}^{p}\sum_{m=-n}^{n} U_{n,m}\overline{V_{n,m}} \quad (29)$$

Note that using our formulation computing the interaction energy is very simple and straightforward.

2.4. Charge Distribution on the Surface of the Molecules. Because of the dielectric discontinuity, a charge distribution develops on the surface of each of the molecules in the system. These charges are a reaction to both the electrostatic field of the free charges of the molecule itself and the electrostatic field of other molecules in its vicinity. In the former case we shall call this charge distribution *self-polarization* charge, and in the latter case we shall call this charge distribution *external polarization* charge. The charged ions in the solution, whose distribution is governed by the potential field, also polarize the dielectric cavities. We will call this charge distribution the *salt polarization* charges. We can compute the surface charge density by using Gauss' Law on an infinitesimal patch on the surface of each molecule with zero thickness. Recall that Gauss' Law states that the net electrostatic field through the surface of a closed volume is proportional to the net charge inside. In the case of a small and very thin patch on the surface of a sphere, this translates to the following equality:

$$\left.\frac{\partial\Phi_{in}^{(i)}}{\partial r}\right|_{r=a_i} - \left.\frac{\partial\Phi_{out}^{(i)}}{\partial r}\right|_{r=a_i} = 4\pi\sigma^{(i)} \quad (30)$$

Plugging in the expression for $\Phi_{in}^{(i)}$ from eq 12 and for $\Phi_{out}^{(i)}$ from eq 15 we arrive at the following expression for the charge distribution on the surface:

$$\sigma^{(i)}(\theta,\phi) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{(2n+1)}{4\pi a_i^{n+2}}\left[E_{n,m}^{(i)}\left(\frac{\gamma_n^{(i)}\hat{\mathrm{k}}_{n+1}(\kappa a_i)}{\epsilon_s e^{\kappa a_i}} - \frac{1}{\epsilon_p}\right) + \frac{nL_{n,m}^{(i)}\gamma_n^{(i)} a_i^{2n+1}}{\epsilon_s(2n+1)}\left(1 - \frac{\epsilon_p}{\epsilon_s}\right)\right] Y_{n,m}(\theta,\phi) \quad (31)$$

A careful look at the terms in eq 31 reveals a clear separation between self-polarization and external polarization. The first term in the brackets describes the self-polarization charge distribution. Note, however, that it also includes the salt polarization that depends on the potential of the free charges. We can extract the *pure* self-polarization term by setting $\kappa = 0$. We then get

$$\sigma_{self}^{(i)}(\theta,\phi) = \frac{\epsilon_p - \epsilon_s}{\epsilon_p}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{(2n+1)}{4\pi a_i^{n+2}}\frac{(n+1)}{(n+1)\epsilon_s + n\epsilon_p} E_{n,m}^{(i)} Y_{n,m}(\theta,\phi) \quad (32)$$

Similarly the second term above describes the external polarization charge distribution (here too, the salt polarization governed by the external potential is included). Without the contribution of the salt-induced charges, the external polarization charge is

$$\sigma_{ext}^{(i)}(\theta,\phi) = \frac{\epsilon_s - \epsilon_p}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n} a_i^{n-1}\frac{(2n+1)}{4\pi}\frac{n}{(n+1)\epsilon_s + n\epsilon_p} L_{n,m}^{(i)} Y_{n,m}(\theta,\phi) \quad (33)$$

Note that if among the external potential sources are also other dielectric cavities that are polarized by the potential of this molecule's charges, then there would be some (usually small) portion of the external polarization charges that actually depends on the free charges. Namely, in this case, removing the free charges will have some effect on the external polarization charges.

By integrating eq 31 over the surface of the molecule we can construct a multipole expansion representation of the polarization charge:

$$S_{n,m}^{(i)} = E_{n,m}^{(i)}\left(\frac{\gamma_n(i)\hat{\mathrm{k}}_{n+1}(\kappa a_i)}{\epsilon_s e^{\kappa a_i}} - \frac{1}{\epsilon_p}\right) + \frac{nL_{n,m}^{(i)}\gamma_n^{(i)} a_i^{2n+1}}{\epsilon_s(2n+1)}\left(1 - \frac{\epsilon_p}{\epsilon_s}\right) \quad (34)$$

We can now use this expansion to describe the potential outside molecule *i* (ignoring the direct contribution of other molecules) as if there was no dielectric boundary; namely, no difference in dielectric constant between the inside and the outside:

$$\Phi_{out}^{(i)}(\mathbf{t}) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}(S_{n,m}^{(i)} + E_{n,m}^{(i)}/\epsilon_p)\frac{e^{\kappa a_i}}{\hat{\mathrm{k}}_{n+l}(\kappa a_i)}\frac{e^{-\kappa r}\hat{\mathrm{k}}_n(\kappa r)}{r^{n+1}}Y_{n,m}(\theta,\phi) \quad (35)$$

Here, the term $[e^{\kappa a_i}/\hat{\mathrm{k}}_{n+1}(\kappa a_i)]$ is a simplified version of the coefficients $\Gamma^{(i)}$ of eq 19, that result when there is no dielectric discontinuity, only a uniform dielectric, and salt cannot penetrate into the cavity.

Note that eq 35 is not equivalent to eq 15. Recall that molecule *i* not only caused a disruption in the dielectric medium of the solvent but also created a cavity devoid of salt ions. This displacement of salt affects the screening of the potential field of the external sources of charge by the salt and thus alters it. By examining the difference between the two equations we are able to extract another set of multipole expansion coefficients that accounts for this effect:

$$Z_{n,m}^{(i)} = \frac{1}{\epsilon_s}\frac{a_i^{2n+1}}{(2n+1)}\frac{\kappa^2 a_i^2 \hat{\mathrm{i}}_{n+1}(\kappa a_i)}{(2n+3)}\frac{e^{\kappa a_i}}{\hat{\mathrm{k}}_{n+1}(\kappa a_i)}L_{n,m}^{(i)} \quad (36)$$

Figure 1 shows the surface charge on a Barstar molecule (see section 3 for details) interacting with a second Barstar molecule in close proximity (1Å separation distance between bounding spheres). The charge distribution is broken into its different components using the formulas we have developed above (eqs 31−33).

2.5. Forces and Torques. The force acting on a molecule due to other molecules in its vicinity is the gradient (in Cartesian coordinates) of the interaction energy with respect to the position of the center of the molecule, namely

$$F^{(i)} = \nabla_i\Omega^{(i)} = \frac{1}{\epsilon_s}[\langle\nabla_i\mathbf{L}^{(i)},\mathbf{A}^{(i)}\rangle + \langle\mathbf{L}^{(i)},\nabla_i\mathbf{A}^{(i)}\rangle] \quad (37)$$

where ∇_i is the gradient with respect to the coordinates of the center $\mathbf{c}^{(i)}$.

We compute the gradient of the coefficient vectors $\mathbf{A}^{(i)}$ by solving the linear system

$$\nabla_i\mathbf{A} = \Gamma\cdot\Delta\cdot[\nabla_i\mathrm{T}\cdot\mathbf{A} + \mathrm{T}\cdot\nabla_i\mathbf{A}] \quad (38)$$

which we arrive at by applying the gradient operator to both sides of eq 21. The gradient of the re-expansion operator, $\nabla_i\mathrm{T}^{(i,j)}$, can be computed analytically as described in the Appendix.

The torque on a molecule is the sum of the torques due to each of the charges in the molecule. By definition, the torque on a charge is the cross product of its position relative to the center of mass of the molecule (the moment arm) with the force it experiences. It is not difficult to convince oneself that the force acting on the external polarization charges generates no net torque. This is because rotating a spherical cavity cannot change the distribution of external polarization charges on its surface. The torque on molecule *i*, due to the charge $q_j^{(i)}$ inside it, is thus

$$\tau_j^{(i)} = \mathbf{p}_j^{(i)} \times \left[\frac{1}{\epsilon_s}q_j^{(i)}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\gamma_n^{(i)}\nabla_i L_{n,m}^{(i)}(\rho_j^{(i)})^n Y_{n,m}(\vartheta_j^{(i)},\varphi_j^{(i)})\right] \quad (39)$$

where $\mathbf{p}_j^{(i)} = [x_j^{(i)}, y_j^{(i)}, z_j^{(i)}]$ in Cartesian coordinates. Due to the linearity of the cross product we can rewrite this as

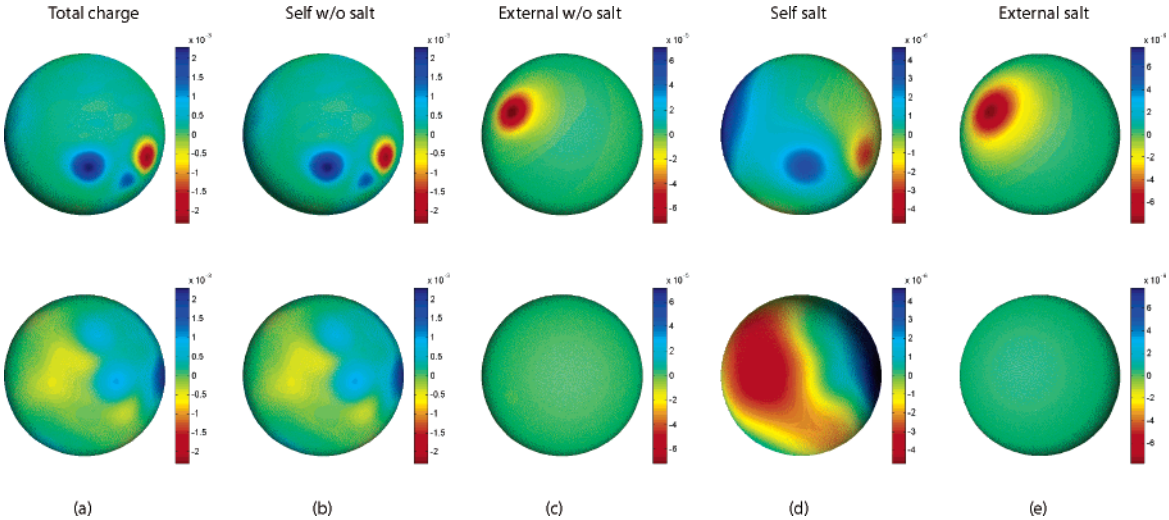

Figure 1. Surface charge distribution of a Barstar molecule broken up to its components: (a) the total charge distribution, (b) the self-polarization charge w/o the contribution of the salt ions, (c) the external-polarization charge distribution due to another molecule in close proximity w/o the contribution of the salt ions, (d) polarization charge distribution due to salt ions governed by the potential induced by the molecule's own charge, and (e) polarization charge distribution due to salt ions governed by the potential induced by external sources. Top row shows view from front and bottom row shows view from back.

$$\tau_j^{(i)} = \frac{1}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}[q_j^{(i)}\gamma_n^{(i)}(\rho_j^{(i)})^n Y_{n,m}(\vartheta_j^{(i)},\varphi_j^{(i)})\cdot\mathbf{p}_j^{(i)}] \times \nabla_i L_{n,m}^{(i)} \tag{40}$$

Summing over all charges of the molecule we get the following expression for the net torque

$$\tau^{(i)} = \frac{1}{\epsilon_s}[{}^x\mathbf{H}^{(i)},{}^y\mathbf{H}^{(i)},{}^z\mathbf{H}^{(i)}] \times [\nabla_i \mathbf{L}^{(i)}] \tag{41}$$

where

$$\begin{aligned} {}^xH_{n,m}{}^{(i)} &= \sum_{j=1}^{M_i} h_{j;n,m}^{(i)} x_j^{(i)} \\ {}^yH_{n,m}^{(i)} &= \sum_{j=1}^{M_i} h_{j;n,m}^{(i)} y_j^{(i)} \\ {}^zH_{n,m}^{(i)} &= \sum_{j=1}^{M_i} h_{j;n,m}^{(i)} z_j^{(i)} \end{aligned} \tag{42}$$

and

$$h_{j;n,m}^{(i)} = \gamma_n^{(i)} q_j^{(i)}(\rho_j^{(i)})^n Y_{n,m}(\vartheta_j^{(i)},\varphi_j^{(i)}) \tag{43}$$

The cross product for two sets of three coefficient vectors, used in eq 41, is defined in analogy to the regular cross product for vectors in 3-D, where the complex inner product of two coefficient vectors defined in eq 29 replaces the regular multiplication of two vector elements. Note that the three coefficient vectors $[{}^x\mathbf{H}^{(i)},{}^y\mathbf{H}^{(i)},{}^z\mathbf{H}^{(i)}]$ do not change as long as the charges inside the molecule do not move and do not depend on the position of the molecule or other molecules.

3. Analysis of Electrostatic Potential

Two of the benefits of our formulation are the correct inclusion of cavity polarization to all orders, whose full contribution is not always included in partial analytical solutions and in numerical approaches such as Gabdoulline and Wade's effective charges[10] and desolvation terms,[24] as well as the ability to consider multiple proteins and thereby multibody effects through the cavity polarizations. To test the consequence of neglecting cavity polarization and multibody effects, we examined model configurations of multiple copies of both Barstar and Barnase,[25] which were chosen as representing typical macromolecules in terms of size and charge distribution. Barnase is an extracellular ribonuclease, and Barstar is its intracellular inhibitor. Barstar and Barnase are modeled as spheres of radii 21.8Å and 28Å, respectively. Their charge distribution is based on assignment of OPLS[26] partial charges centered on atomic coordinates taken from their PDB file (1BRS), after inclusion of hydrogens on heavy atoms using CHARMM.[27] Barstar has a relatively large net charge of −6e, and Barnase has a net charge of +2e. These two proteins were chosen somewhat arbitrarily, and we could have chosen for the sake of the computations we present here almost any other protein. The only assumption we make is that the charge distribution inside them is not significantly different from what is normally observed in proteins. The radii were chosen so that all charges are within a sphere centered at the geometric center of each molecule. It is possible to choose a different center and as a result have somewhat different radii; however, one must remember that the correctness of the method depends on the inclusion of all charges inside the cavity, and that if the chosen center differs from the center of mass, the computed torques must be translated to the true centers of mass. In all computations that follow we use $\epsilon_p = 4$, $\epsilon_s = 80$, and the salt concentration in the solution is 50 mM, which at room temperature yields $\kappa = 0.074$.

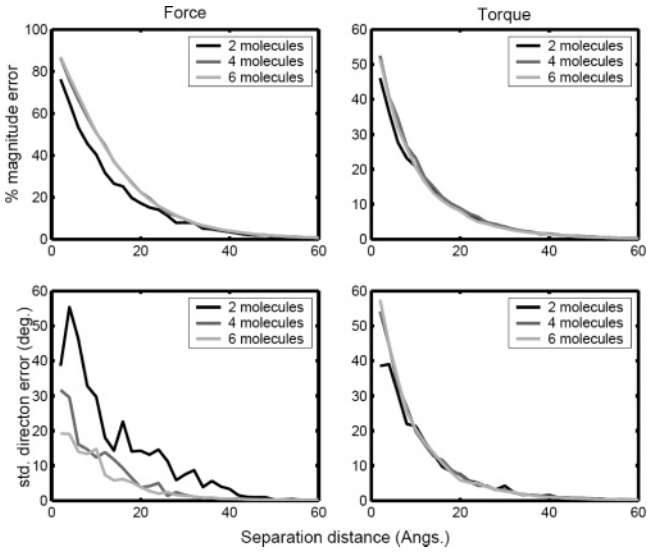

Figure 2. Higher order effect of the dielectric cavities (through polarization and salt displacement) on the forces and the torques in different configurations of Barstar molecules. The left column shows the effect on the forces acting on the different molecules, and the right column shows the effect on the torques. The top row shows the average relative error in magnitude, and the bottom row shows the standard deviation of the error in direction when ignoring the dielectric cavities. Results are shown for configurations of two, four, and six molecules, as a function of separation distance.

We considered configurations of two molecules, an equilateral tetrahedron with four molecules centered on each vertex, and six molecules configured as a 3-D cross, with the centers at the ends of the segments. In Figure 2 we show the relative error in magnitude and the standard deviation of the error in direction of both forces and torques, when cavity polarization is ignored, computed for different configurations at different separation distances for different numbers of Barstar molecules. The results for Barnase were similar and are not shown. We can see that magnitude errors as high as 85% in force and over 50% in the torque and is often as high as 60% in the direction of force or torque, at the point of closest protein contact. Although the neglect of molecular features of the interface will also be a significant error at these closest separations, at longer separations where the coarse-grained approximation is still a reasonable one, the cumulative effect of 10−30% error by neglecting cavity polarization could strongly influence, for example, a simulation of mechanism and rate of protein assembly.

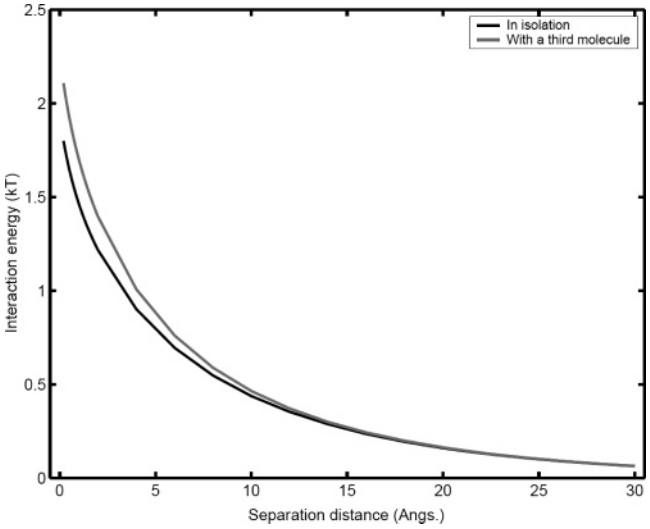

Figure 3. The effect of a third molecule on the interaction free energy of two Barstar molecules, as a function of separation distance. Results are averaged over all possible relative orientations of the molecules (by uniform sampling). In black is the interaction energy for the two molecules in isolation, and in gray the energy when a third molecule is placed an equal distance from both (forming an equilateral triangle).

In Figure 3 we compare the interaction energy of two Barstar molecules as a function of their relative separation (distance between surfaces of bounding spheres), to the effect on their potential of mean force (PMF) due to a third Barstar molecule. The three molecules are set up on the vertices of an equilateral triangle, and the PMF is calculated by uniformly sampling their relative orientations. The nonadditivity effects due to cavity polarization introduced by the third Barstar modulate the interactions of the two Barstar molecules so that there is an increased repulsion at closer distances. This effect would typically be missing in formulations of the electrostatic potential in which it is only feasible to consider pair interactions.

Figure 4 shows the power of our formulation in solving the electrostatics of very large systems. The interaction energy, force, and torque on a unit cell comprised of a single Barstar molecule in an infinite lattice is computed using neighborhoods as large as 3 surrounding layers or 7^3 unit cells (243 molecules participating in the calculation all together). As we can see the relative error going from two to three layers is smaller than 1% even at very close separations, which ensures us that a larger neighborhood is not required.

4. Implementation

4.1. Computing the Re-Expansion Operator and Its Derivatives. The literature describes a number of methods for computing the re-expansion operator $T^{(i,j)}$.[19−22,28,29] We based our implementation on the method of Gumerov and Duraiswami[21] for the solution of the Helmholtz equation because of its relative simplicity and efficiency. We adopt it to our formulation and extend it to the computation of partial derivatives. Note that in what follows we restrict our discussion to the practical aspects of using the re-expansion operator. The theory behind it can be found elsewhere.[21] The matrix $T^{(i,j)}$, which represents a re-expansion in an arbitrary direction, is decomposed (diagonalized) to a re-expansion along the *Z*-axis $S^{(i,j)}$. This is done using a rotation operator $R^{(i,j)}$ that orients the desired direction of re-expansion with the *Z*-axis and its inverse that rotates back to the original coordinate frame after the diagonalized re-expansion is performed. We thus have the following decomposition of $T^{(i,j)}$:

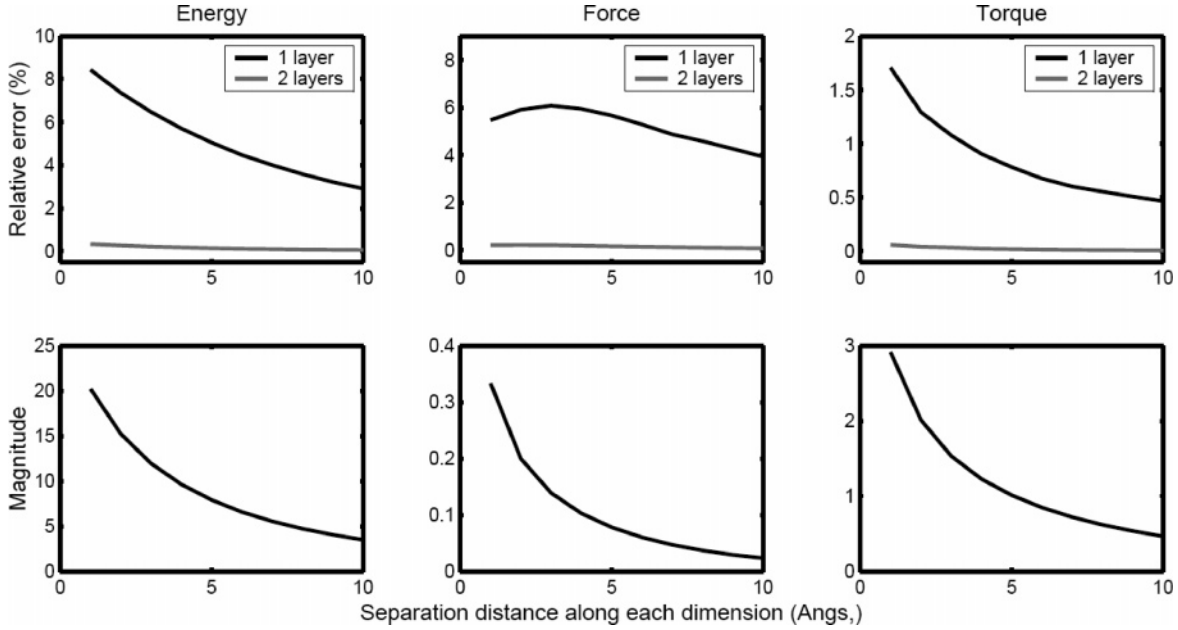

Figure 4. Computing the energy, force, and torque on a molecule in an infinite lattice. The top row shows the relative error in computing the energy, force, and torque of a unit cell comprised of a single Barstar molecule when taking into account interactions with only the first layer (a $3 \times 3 \times 3$ neighborhood) and only the second layer (a $5 \times 5 \times 5$ neighborhood). The second row shows the magnitude of the energy (kT), force (kT/Å), and torque (kT/rad) on a single Barstar molecule in the infinite lattice.

$$\mathrm{T}^{(i,j)} = (R^{(i,j)})^H \cdot S^{(i,j)} \cdot R^{(i,j)} \tag{44}$$

While $\mathrm{T}^{(i,j)}$ is a full matrix having possibly $O(p^4)$ independent elements, both $R^{(i,j)}$ and $S^{(i,j)}$ have only $O(p^3)$ nonzero elements, which can be computed recursively as described in Appendix A.1 in $O(p^3)$ time. Moreover, a simple relation exists between $\mathrm{T}^{(i,j)}$ and $\mathrm{T}^{(j,i)}$ namely, $\mathrm{T}^{(j,i)} = (\mathrm{T}^{(i,j)})^H$, which entails that

$$\mathrm{T}^{(j,i)} = (R^{(i,j)})^H \cdot (S^{(i,j)})^H \cdot R^{(i,j)} \tag{45}$$

Since both $R^{(i,j)}$ and $S^{(i,j)}$ are sparse matrices, whose nonzero elements are known in advance, we compute the re-expansion $\mathbf{Z} = \mathrm{T} \cdot \mathbf{X}$ in three steps:

$$X^1_{n,m} = \sum_{s=-n}^{n} R^{m,s}_n X_{n,s}$$

$$X^2_{n,m} = \sum_{l=|m|}^{p} S^m_{n,l} X^1_{l,m}$$

$$Z_{n,m} = \sum_{s=-n}^{n} \overline{R^{s,m}_n} X^2_{n,s} \tag{46}$$

The partial derivatives of the re-expansion matrix have a similar decomposition that can be computed by differentiating eq 44. For re-expansion along $\mathbf{v}^{(i,j)} = \mathbf{c}^{(i)} - \mathbf{c}^{(j)} = [r,\theta,\phi]$ we get

$$\frac{\partial \mathrm{T}^{(i,j)}}{\partial r} = (R^{(i,j)})^H \cdot \frac{\partial S^{(i,j)}}{\partial r} \cdot R^{(i,j)}$$

$$\frac{\partial \mathrm{T}^{(i,j)}}{\partial \theta} = \left(\frac{\partial R^{(i,j)}}{\partial \theta}\right)^H \cdot S^{(i,j)} \cdot R^{(i,j)} + (R^{(i,j)})^H \cdot S^{(i,j)} \cdot \frac{\partial R^{(i,j)}}{\partial \theta}$$

$$\frac{\partial \mathrm{T}^{(i,j)}}{\partial \phi} = \left(\frac{\partial R^{(i,j)}}{\partial \phi}\right)^H \cdot S^{(i,j)} \cdot R^{(i,j)} + (R^{(i,j)})^H \cdot S^{(i,j)} \cdot \frac{\partial R^{(i,j)}}{\partial \phi} \tag{47}$$

Note that $R^{(i,j)}$ is independent of r and $S^{(i,j)}$ is independent of θ and ϕ . To convert the derivatives with respect to the spherical coordinates to derivatives with respect to Cartesian coordinates the following 3×3 matrix is used:

$$\begin{bmatrix} \frac{\partial Z_{n,m}}{\partial x} \\ \frac{\partial Z_{n,m}}{\partial y} \\ \frac{\partial Z_{n,m}}{\partial z} \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\phi & \frac{\cos\theta\cos\phi}{r} & \frac{-\sin\phi}{r\sin\theta} \\ \sin\theta\sin\phi & \frac{\cos\theta\sin\phi}{r} & \frac{\cos\phi}{r\sin\theta} \\ \cos\theta & \frac{-\sin\theta}{r} & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial Z_{n,m}}{\partial r} \\ \frac{\partial Z_{n,m}}{\theta} \\ \frac{\partial Z_{n,m}}{\partial \phi} \end{bmatrix} \tag{48}$$

A careful examination of eqs 47 and 48 leads to the conclusion that $\nabla_j \mathrm{T}^{(i,j)} = -\nabla_i \mathrm{T}^{(i,j)}$. The derivatives with respect to θ and ϕ encounter a singularity when $\sin\theta \to 0$. In Appendix A.2 we discuss this in detail and explain how to handle the singularities.

4.2. Iterative Computation of Forces and Torques. A one-time exact solution of the forces for a given configuration of the molecular system can be computed by inverting eqs 21 and 38. We first compute the values of the $\mathbf{A}^{(i)}$ vectors by solving

$$\mathbf{A} = (\mathrm{I} - \Gamma \cdot \Delta \cdot \mathrm{T})^{-1} \cdot \Gamma \cdot \mathbf{E} \tag{49}$$

and then for every molecular center $\mathbf{c}^{(j)}$ we compute all $\nabla_j \mathbf{A}^{(i)}$ by solving

$$\nabla_i \mathbf{A} = (\mathrm{I} - \Gamma \cdot \Delta \cdot \mathrm{T})^{-1} \cdot \Gamma \cdot \Delta \cdot \nabla_i \mathrm{T} \cdot \mathbf{A} \tag{50}$$

which is an inversion of eq 38. Note that the expensive matrix inversion operation need only be performed once, since the same inverted matrix is used in both eqs 49 and 50. The values of $\mathbf{A}^{(i)}$ and $\nabla_i \mathbf{A}^{(i)}$ are then plugged into eq 37 to yield the force on molecule i and into eq 41 to yield the torque on molecule i.

As long as the distance between any pair of molecules is larger than the sum of their radii, which is a requirement of our theory, the inverted matrix of eq 49 is strictly diagonally dominant.[30] This guarantees that an iterative approach such as the Gauss-Seidel method[30] will converge quickly. Moreover, in a simulation setting, the position and orientation of each molecule changes little between time-steps, thus it makes sense to *fix* the solution computed at the previous time-step, instead of computing a new solution from scratch. Therefore instead of a costly matrix inversion we use the Gauss-Seidel method, first to solve eq 21 and then to solve eq 38 for each center $\mathbf{c}^{(j)}$. Casting eq 21 as an iteration step in a Gauss-Seidel solution of the linear system yields

$$\begin{aligned} \mathbf{A}^{(1)}_t &= \Gamma^{(1)} \cdot (\Delta^{(1)} \cdot \sum_{j=2}^{N} \mathrm{T}^{(1,j)} \cdot \mathbf{A}^{(j)}_{t-1} + \mathbf{E}^{(1)}) \\ &\vdots \\ \mathbf{A}^{(N)}_t &= \Gamma^{(N)} \cdot (\Delta^{(N)} \cdot \sum_{j=1}^{N-1} \mathrm{T}^{(N,j)} \cdot \mathbf{A}^{(j)}_t + \mathbf{E}^{(N)}) \end{aligned} \tag{51}$$

At iteration t each $\mathbf{A}^{(i)}_t$ is computed in turn, i going from 1 to N using the most recent values of all $\mathbf{A}^{(j)}$, $j \neq i$. The iterations stop when the relative change to $\mathbf{A}^{(i)}$ falls below the desired precision. The relative change is computed as

$$\text{change} = \frac{1}{2Np^2} \sum_{i=1}^{N} \sum_{k=1}^{p^2} \frac{|\mathbf{A}^{(i)}_{t,k} - \mathbf{A}^{(i)}_{t-1,k}|}{|\mathbf{A}^{(i)}_{t,k}| + |\mathbf{A}^{(i)}_{t-1,k}|} \tag{52}$$

In a simulation setting, $\mathbf{A}_0{}^{(i)} = \Gamma^{(i)} \cdot \mathbf{E}^{(i)}$ for the computation at the first step of the simulation, and for each subsequent step $\mathbf{A}_0{}^{(i)}$ is set to its final value in the previous step.

Once all $\mathbf{A}^{(i)}$ are computed, an analogous iterative form of eq 38 can be used to compute the $\nabla_j \mathbf{A}^{(i)}$:

$$\begin{aligned} \nabla_j \mathbf{A}^{(1)}_t &= \Gamma^{(1)} \cdot \Delta^{(1)} \cdot \sum_{k=2}^{N} [\nabla_j \mathrm{T}^{(1,k)} \cdot \mathbf{A}^{(k)} + \mathrm{T}^{(1,k)} \cdot \nabla_j \mathbf{A}^{(k)}_{t-1}] \\ &\vdots \\ \nabla_j \mathbf{A}^{(N)}_t &= \Gamma^{(N)} \cdot \Delta^{(N)} \cdot \sum_{k=1}^{N-1} [\nabla_j \mathrm{T}^{(N,k)} \cdot \mathbf{A}^{(k)} + \mathrm{T}^{(N,k)} \cdot \nabla_j \mathbf{A}^{(k)}_t] \end{aligned} \tag{53}$$

Here $\nabla_j \mathbf{A}_0{}^{(i)} = \mathbf{0}$ for the first step of a simulation. At each subsequent step the final value of the previous time-step is used.

4.3. Adaptive Control of Precision. Since we must truncate each vector of expansion coefficients at a finite order, there will always be an inherent error in the computation of the forces. However by carefully choosing the

truncation order we can guarantee the error will never exceed a preset value. In the interest of speed, we would like to use the smallest order required to give the desired precision. A theoretical upper bound on the error when computing the potential at a distance R using a truncated multipole expansion is $O(a/R)^{p+1}$, where a is the radius of the charges and $R > a$.[31] When computing the interaction between two sets of charges whose centers are a distance R apart, with radii a_i and a_j and $R > a_i + a_j$, an upper bound on the truncation error is

$$err_p = O\left(\frac{a_i}{R - a_j}\right)^{p+1} \tag{54}$$

Thus for each interaction between a pair of molecules a different truncation order should be used that depends on the respective separation and radii of the pair.

The theoretical bound of eq 54 is often not tight for proteins, because the constant hidden in the bound depends on $\sum_i |q_i|$, which for proteins is much larger than $\sum_i q_i$. Thus we use a heuristic bound that is much tighter in practice (although its correctness cannot be guaranteed). We approximate the truncation error as

$$err_p^{(i,j)} = \left| \frac{[(\mathbf{A}^{(i)})^H \cdot \mathrm{T}^{(i,j)} \cdot \mathbf{A}^{(j)}]_{p+1} - [(\mathbf{A}^{(i)})^H \cdot \mathrm{T}^{(i,j)} \cdot \mathbf{A}^{(j)}]_p}{[(\mathbf{A}^{(i)})^H \cdot \mathrm{T}^{(i,j)} \cdot \mathbf{A}^{(j)}]_p} \right| \tag{55}$$

Namely, the relative error in computing the interaction by truncating at order p is the difference between the value using maximal order $p + 1$ and the value using maximal order p, divided by the value using maximal order p. Since the values of $\mathbf{A}^{(i)}$ and $\mathbf{A}^{(j)}$ at the current time-step are not yet known, their values from the previous step are used. Assuming these values change very little between time-steps the error introduced by this approximation will be small. When $err_p^{(i,j)}$ is larger than the desired error, the truncation order is incremented and tested again. If $err_p^{(i,j)}$ falls below a second, lower threshold, the order we are using is too high and should be decremented and tested again. This heuristic bound starts to break as the ratio in eq 54 approaches 1; in this case it may be safer to also test truncation order $p + 2$ and even $p + 3$ in order to make sure the truncated terms indeed fall off and do not diverge.

4.4. Computational Complexity. The size of the computational problem at hand is governed by a number of parameters: N: the number of molecules, $M = \max_i(M_i)$: the number of partial charges in each molecule, and p: the maximal expansion order used in the calculations. To compute the forces and torques as described above, the following calculations are performed:

1. Compute the multipole expansion for each molecule. $O(Mp^2)$ per molecule, $O(NMp^2)$ in total.

2. Invert a matrix of size $Np^2 \times Np^2$ which takes $O(N^3p^6)$ time (eq 49).

3. Compute all $\mathbf{A}^{(i)}$ and all $\nabla_j\mathbf{A}^{(i)}$ (eqs 49 and 50). Involves multiplying Np^2 vectors by $Np^2 \times Np^2$ matrices which takes $O(N^2p^4)$ per molecule or $O(N^3p^4)$ in total.

4. Compute the forces and torques, which involves dot products of vectors of p^2 elements. $O(p^2)$ per molecule or $O(Np^2)$ in total.

The total computational complexity for a one time calculation is thus $O(NMp^2 + N^3p^6)$.

During a simulation, as mentioned above, the configuration of the system changes very little between time-steps, and an iterative method can be used to *fix* the forces and torques computed at the previous step. Also, as long as the charge distribution does not change, the multipole expansion of the charges of each molecule need only be rotated to represent the new orientation of the molecule, and there is no need to compute it from scratch. Thus for an update step during a simulation the calculations are as follows:

1. Rotate the multipole expansion of each molecule. $O(p^3)$ per molecule or $O(Np^3)$ in total.

2. Iteratively update all $\mathbf{A}^{(i)}$ and all $\nabla_j\mathbf{A}^{(i)}$. This requires $O(N^2p^3)$ per cycle through the lines of eq 51 as well as per cycle through the lines of eq 53. For all molecules, assuming a maximum of k cycles, the time is $O(kN^3p^3)$.

3. Compute the forces and torques. Same as above $O(Np^2)$ in total.

The total computational complexity for an update step is thus $O(kN^3p^3)$, which is independent of M, and assuming $k \leq p$ is asymptotically faster than a one time computation by at least a factor of p^2. Note that when the charges of a molecule are slightly perturbed, its multipole expansion may need to be recomputed. The update procedure, however, can still be used effectively, since the previous values of $\mathbf{A}^{(i)}$ and $\nabla_j\mathbf{A}^{(i)}$ still constitute good initial guesses.

4.5. Numerical Stability. When the use of high order poles is required to maintain the desired precision of the computations, we run the risk of exceeding the precision bounds of the machine we are using. A simple scaling scheme can be used to significantly alleviate this problem. We define a scaling factor λ to be the average molecular radius of the system we are solving. We then use scaled versions of the multipole expansions in the computation that will be numerically better behaved:

$$\check{\mathrm{E}}_{n,m}^{(i)} = \frac{1}{\lambda^n} E_{n,m}^{(i)}, \qquad \check{\mathrm{H}}_{n,m}^{(i)} = \frac{1}{\lambda^n} H_{n,m}^{(i)} \tag{56}$$

Note that the scaling should be done before the actual construction of these expansions, by multiplying the centered coordinates of each charge by λ^{-1}. Each of the parameters $\delta_n^{(i)}$, defined in eq 20, also needs to be scaled down to $\check{\delta}_n^{(i)} = \lambda^{-2n}\,\delta_n^{(i)}$. This too should be done during the construction of the $\delta_n^{(i)}$ to ensure numerical stability. Finally the translation coefficients $S^{(i,j)}$ need to be scaled up. This is explained in detail in the Appendix A.1.

4.6. Performance Evaluation. We ran a number of experiments using our implementation of the theory, to test how well it performs in realistic scenarios. First we looked at the performance of the Gauss-Seidel linear system solver. We expected the number of iterations to be low because the matrix we need to invert is diagonally dominant. We looked at the solution of configurations of 2, 4, and 6 Barstar molecules similar to those we used in section 3. The error defined in eq 52 was set to 1%. The number of iterations required for the solution of both the effective charge expansion vectors $\mathbf{A}^{(i)}$ and their gradients $\nabla_j\mathbf{A}^{(i)}$ is recorded

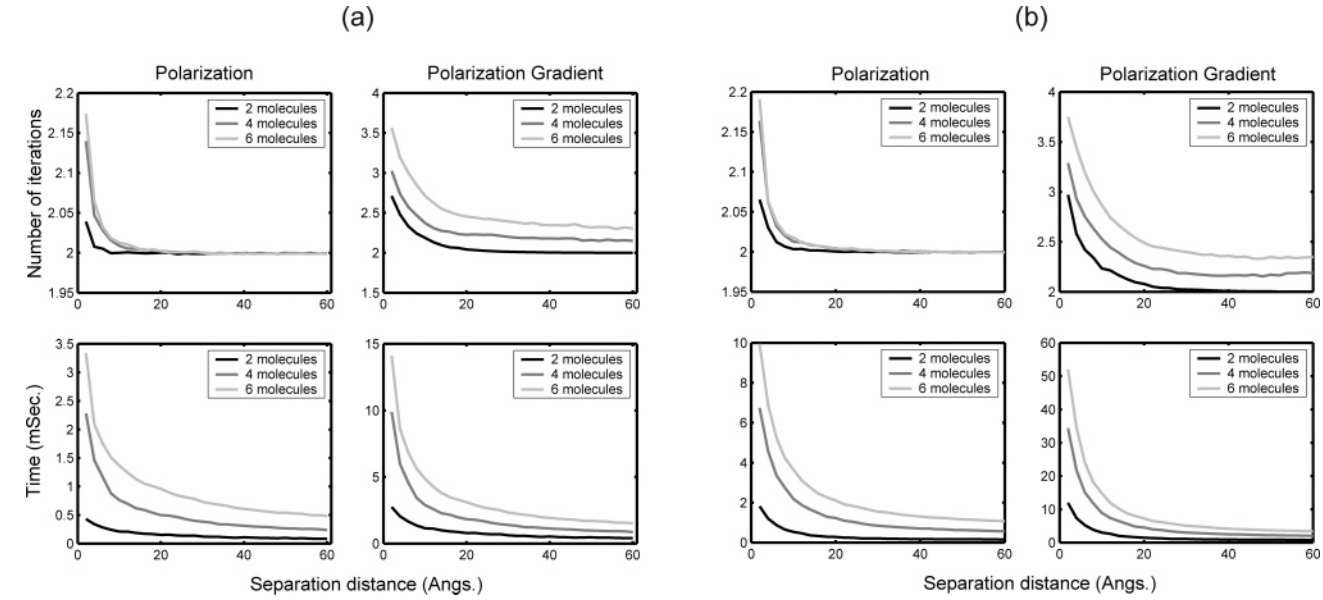

Figure 5. The computation time required to compute the energy and forces for different configurations of (a) Barstar and (b) Barnase molecules as a function of separation distance. The first column shows the number of iterations and time required to compute the polarization effects (solve the linear system of eq 51). The second column shows the number of iterations and time required to compute the polarization effects (solve the linear system of eq 51). The second column shows the number of iterations and time required to compute the partial derivatives of the polarization effects (solve the linear system of eq 53 for a single molecular center). The first row shows the number of Gauss-Seidel iterations, and the second row shows the computation time required to converge to within 1% of the correct solution.

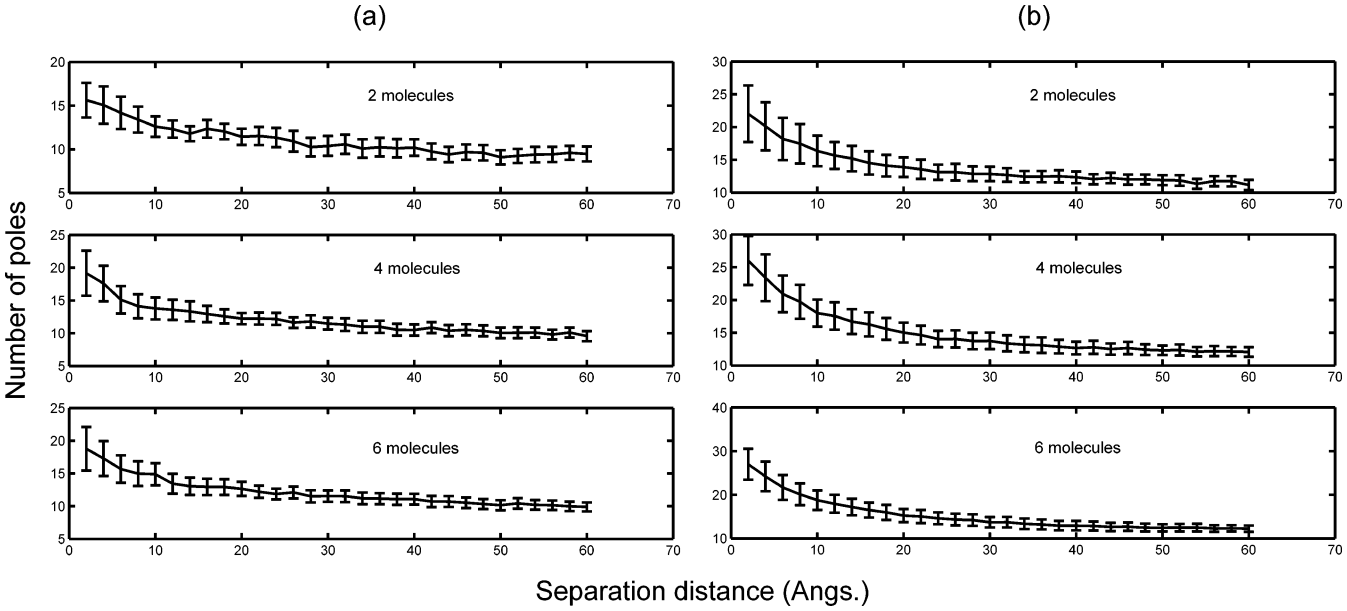

Figure 6. The average and standard deviation of the number of poles in a multipole expansion needed to compute the interaction between two molecules to within 0.1% error as a function of separation distance. Results are shown for three configurations of (a) Barstar and (b) Barnase molecules: two molecules at the top, four in the middle, and six at the bottom graph.

for 5000 different relative random orientations of the tested molecules. In Figure 5a we show the iterations and timing results for a range of separation distances for the molecule Barstar. The number of iterations to compute the $\mathbf{A}^{(i)}$ vectors is only slightly more than 2 and seems to vary as a function of size only for short separations. The computation of the gradient vectors $\nabla_j\mathbf{A}^{(i)}$ requires as many as 3.5 iterations for 6 molecules at short separations, and in general there is a small yet persistent difference between the different configuration sizes across separation distances, where the larger the configuration the more iterations that are required on average. Also included as a reference are the actual average running times on an apple G5 workstation. These running times have been normalized by the size of the configuration. Thus one needs to multiply the reported times for polarization by a factor of N (the size of the configuration) and the reported times for the polarization gradient by a factor of N^2 to get the total computation time for the configuration.

The iteration and timing results for Barnase are in Figure 5b. Comparing the two, one notices that the number of

iterations required for convergence is about the same for both molecules; however, the time per iteration is about 3 times larger for Barnase because higher order expansion was needed to get the same desired level of precision. The number of poles (the maximal expansion order) that are used on average to represent the $\mathbf{A}^{(i)}$ vectors determines the actual size of the computations that are performed. Recall that this number is chosen automatically by the algorithm to maintain the desired error of eq 55. In Figure 6a we show the number of poles required to reach a desired relative error of 0.001 (mean and standard deviation) for different numbers of Barstar molecules, and in Figure 6b we show the same for Barnase.

Computing these systems using numerical methods, such as DelPhi, would take many seconds, and we expect our method to be at least 2−3 orders of magnitude faster. Moreover, for the larger configurations (4 and 6 molecules) a very large grid would be needed to get high levels of accuracy. Numerical methods, in general, are designed to handle complex molecular surfaces and are unable to take advantage of simple spherical geometries, unlike our method. Thus, already from a theoretical perspective, it is always considerably better, in terms of both precision and time, to use our method when spherical geometry is assumed.

5. Conclusion

We have presented a new general analytical solution for computing the screened electrostatic interaction between multiple macromolecules of arbitrarily complex charge distributions, assuming they are well described by spherical low dielectric cavities in a higher dielectric medium in the presence of a Debye−Hückel treatment of salt. Our new formulation builds upon foundations of electrostatic theory of dielectric boundary problems laid down by Kirkwood,[11] Phillies,[17] Hansen and co-workers,[12,16] and McClurg and Zukoski[18] and unites it with the scientific computing community's recent developments in solving for the screened Coulomb (Yukawa) potential[19,20] and solution to the 3-D Helmholtz equation.[21,22] Our formulation offers the further advantage that it is general for arbitrary numbers of macromolecules with arbitrarily complex charge distributions, and the use of straightforward and computationally efficient multipole re-expansion operators to solve the screened Coulomb problem has made our formulation practical as well. In addition we present its algorithmic scaling costs and implementation (with code made available upon request to one of the corresponding authors).

Often a macromolecule in solution is represented by two concentric spheres, an inner sphere to represent the molecule and a concentric layer around it (a 3-D annulus) to represent a charge-free region due to the finite size of ions, a solvent-free region due to the finite size of solvent molecules, or an ionic Stern Layer. See, for example Kirkwood[11] and Hoffman et al.[12] When assuming that this layer has the same permittivity as the interior of the macromolecule our formulation can be used as is. The radius of the molecule is simply extended to include this outer layer as well, and any charges in it are treated as part of the molecule's charge distribution. If, however, a different permittivity is used, one would need to derive new expressions for the electrostatic potential, where now each cavity would have two boundaries, and expressions for the potential would be required for three different areas (inside, external layer, and outside). This has been done before for a single molecule.[11,12] Our approach would be applicable for this case as well; however, the expressions would be much longer and more cumbersome and thus we have chosen not to derive them. We could also model the Stern layer ions explicitly since our formulation allows for arbitrary numbers of spheres of arbitrary radii.

We plan to apply this generally formulated electrostatic model in the future to problems in protein complexation using Brownian dynamics. In typical implementations of protein−protein association using Brownian dynamics, the charge distribution of the protein in isolation is assumed to be fixed throughout the stochastic simulation of proteins in solution.[24,32,33] While this approximation is best for earlier stages of the diffusion reaction, the approximation degrades significantly as the relative separation of the two proteins become smaller, since the accumulated forces at shorter range acting on the charge sites should shift the conformational state of the two proteins, thereby altering their overall charge distribution and shape. This scenario can be further complicated by the effect of the presence of multiple proteins on the pair interactions. Our analytical solution is general to arbitrary numbers of macromolecules, is efficient to compute, and can therefore simultaneously provide on-the-fly updates to changes in charge distributions due to protein conformational changes. We can also change spatial resolutions of charge description as a function of separation distance without compromising the desired accuracy.

While for many applications the simple spherical geometry of the macromolecule is sufficient and reasonable, in some cases it falls short of capturing the desired level of atomic detail. In our future work we plan to explore a number of paths that promise to extend the usability of our formulation to cases where a more accurate description of the influence of shape on the electrostatics near the surface of proteins is required. As we discuss in section 2.2, our solution separates the description of the charge distribution from the description of the geometric and electrostatic properties (dielectric constants and salt screening length). The latter is restricted to the Γ operator that describes the behavior of the dielectric boundary between each molecule and the solution and the Δ operator that describes the polarization properties of the molecule. Thus, instead of using the analytic values for these operators computed for spherical geometry, it may be possible to compute generalized operators for more complex geometries such as unions of spheres, where more than a single expansion center is used to better approximate the electrostatic potential.

Acknowledgment. We would like to thank Rosalind Allen and Jean-Pierre Hansen for useful discussions. We would like to thank Nail Gumerov for help in the implementation of the re-expansion operator. We acknowledge financial support from DOE/LDRD and the National Institutes of Health Grant No. GM76730-01. T.H.G. gratefully acknowledges the support of a Schlumberger Fellowship while on sabbatical in the Theoretical Chemistry Sector at Cambridge University.

Appendix

A.1. Recursive Computation of Re-Expansion Coefficients. The re-expansion coefficients are computed using the decomposition of eq 34. Thus a set of rotation and translation coefficients needs to be computed. Gumerov and Duraiswami[21] propose a recursive method for the re-expansion coefficients for the Helmholtz equation, which we adopt here for the very similar Poisson−Boltzmann equation. Other formulations of the re-expansion operator also exist.[20,22,29] Below we lay out the recursive procedure for computing the rotation and translation coefficients for a re-expansion along the vector $v = [r,\theta,\phi]$.

Rotation Coefficients and Their Derivatives. The rotation coefficients $R_n^{m,s}$ $(0 \le n \le p-1,\ -n \le m,s \le n)$ are used as described in eq 46. The recursive procedure for computing $R_n^{s,m}$ is the following:

1. Set $R_n^{0,s} = Y_{n,-s}(\theta,\phi)$ for $0 \le n \le 2p-1$ and $-n \le s \le n$.

2. For each $0 \le m \le p-1$ starting at $m = 0$, fill all coefficients with $m+2 \le n \le 2p-m-1$ and $-n+1 \le s \le n-1$ using the following rule:

$$R_{n-1}^{m+1,s} = \frac{1}{b_n^m}\left[\frac{1}{2}e^{-i\phi}(1+\cos\theta)b_n^{s-1}R_n^{m,s-1} - \frac{1}{2}e^{i\phi}(1-\cos\theta)b_n^{-s-1}R_n^{m,s+1} + \sin\theta a_n^s R_n^{m,s}\right] \quad (1.1)$$

For negative values of m use $R_n^{m,s} = \overline{R_n^{-m,-s}}$.

The constants used in the recursion are

$$a_n^m = \sqrt{\frac{(n+m+1)(n-m+1)}{(2n+1)(2n+3)}}$$

$$b_n^m = \mathrm{sign}(m)\sqrt{\frac{(n-m-1)(n-m)}{(2n-1)(2n+1)}} \quad (1.2)$$

The partial derivatives of the rotation operator $\partial R_n^{m,s}/\partial\theta$ and $\partial R_n^{m,s}/\partial\phi$ can be computed in a similar fashion. For $\partial R_n^{m,s}/\partial\theta$ the following recursion should be used:

1. For $0 \le n \le 2p-1$ use the formula

$$\frac{\partial R_n^{0,-s}}{\partial\theta} = s\cot\theta Y_{n,s}(\theta,\phi) - \sqrt{(n-s)(n+s+1)}e^{-i\phi}Y_{n,s+1}(\theta,\phi) \quad (1.3)$$

when $s \ge 0$, and then use $\partial R_n^{0,-s}/\partial\theta = \overline{\partial R_n^{0,s}/\partial\theta}$ to complete the positive values of s.

2. For each value of $0 \le m \le p-1$ starting at $m = 0$ fill all coefficients with $m+2 \le n \le 2p-m-1$ and $-n+1 \le s \le n-1$ using the following rule:

$$\frac{\partial R_{n-1}^{m+1,s}}{\partial\theta} = -\frac{1}{b_n^m}\left[\frac{1}{2}e^{i\phi}b_n^{-s-1}\left(\sin\theta R_n^{m,s+1} + (1-\cos\theta)\frac{\partial R_n^{m,s+1}}{\partial\theta}\right) + \frac{1}{2}e^{-i\phi}b_n^{s-1}\left(\sin\theta R_n^{m,s-1} - (1+\cos\theta)\frac{\partial R_n^{m,s-1}}{\partial\theta}\right) - a_n^s\left(\sin\theta\frac{\partial R_n^{m,s}}{\partial\theta} + \cos\theta R_n^{m,s}\right)\right] \quad (1.4)$$

3. For negative values of m use $\partial R_n^{m,s}/\partial\theta = \overline{\partial R_n^{-m,-s}/\partial\theta}$. For $\partial R_n^{m,s}/\partial\phi$ the formula is

$$\frac{\partial R_n^{m,s}}{\partial\phi} = -isR_n^{m,s} \quad (1.5)$$

Translation Coefficients and Their Derivatives. The translation coefficients $S_{n,l}^m$ $(0 \le n,l \le p-1,\ -n \le m \le n)$ are used as in eq 46. The recursive procedure for computing $S_{n,l}^m(r)$ is the following:

1. Set $S_{0,l}^0 = (\lambda/r)^l(\hat{\mathrm{k}}_l(\kappa r)e^{-\kappa r}/r)$ for $0 \le l \le 2p-1$ and use $S_{n,0}^0 = (-1)^n S_{0,n}^0$ for $0 \le n \le 2p-1$.

2. Use the following recursion to compute $S_{n,l}^0$ for $0 \le n \le p-2$ and $n+1 \le l \le 2p-n-2$:

$$S_{n+1,l}^0 = -\frac{1}{\alpha_n^0}[\beta_{l-1}^0 S_{n,l-1}^0 + \beta_{n-1}^0 S_{n-1,l}^0 + \alpha_l^0 S_{n,l+1}^0] \quad (1.6)$$

3. For each $0 \le m \le p-2$ use the following two steps to compute all $S_{n,l}^{m+1}$:

I. For $m \le l \le 2p-2-m$ compute all $S_{m+1,l}^{m+1}$ using

$$S_{m+1,l}^{m+1} = -\frac{1}{\eta_{m+1}^{-m-1}}[\mu_l^{-m-1}S_{m,l-1}^m + \eta_{l+1}^m S_{m,l+1}^m] \quad (1.7)$$

II. For $m \le n \le p-2$ and $n+1 \le l \le 2p-n-2$ use

$$S_{n+1,l}^{m+1} = -\frac{1}{\alpha_n^{m+1}}[\beta_{l-1}^{m+1}S_{l-1}^{m+1} + \beta_{n-1}^{m+1}S_{n-1,l}^{m+1} + \alpha_l^{m+1}S_{n-1,l}^{m+1}] \quad (1.8)$$

4. To complete the set of coefficients two identities can be used:

I. For all $S_{n,l}^m$ where $n > l$ use $S_{n,l}^m = (-1)^{n+l}S_{l,n}^m$.

II. For all $S_{n,l}^m$ where $m < 0$ use $S_{n,l}^{-m} = S_{n,l}^m$.

The constants used in the recursion are the following:

$$\alpha_n^m = \sqrt{(n+m+1)(n-m+1)}$$

$$\beta_n^m = \frac{\lambda^2\kappa^2\alpha_n^m}{(2n+1)(2n+3)}$$

$$\eta_n^m = \mathrm{sign}(m)\sqrt{(n-m-1)(n-m)}$$

$$\mu_n^m = \frac{\lambda^2\kappa^2\eta_n^m}{(2n-1)(2n+1)} \quad (1.9)$$

Recall that λ is the uniform scaling factor that was defined in section 4.5.

The derivatives of the translation coefficients $\partial S_{n,l}^m/\partial r$ are computed using a similar procedure, which differs only in step #1, where the initial values are computed using

$$\frac{\partial S_{0,l}^0}{\partial r} = \left(\frac{\lambda}{r}\right)^l\frac{e^{-\kappa r}}{r^2}(l\hat{\mathrm{k}}_l(\kappa r) - (2l+1)\hat{\mathrm{k}}_{l+1}(\kappa r)) \quad (1.10)$$

A.2. Dealing with Singularities. The method we use to compute the derivatives of the re-expansion operator breaks

as the direction of the re-expansion vector $v = [r,\theta,\phi]$ approaches the *z*-axis (either the positive or negative directions). In this case the angle $\theta \rightarrow 0,\pi$ and $\sin\theta \rightarrow 0$. This is not due to a true singularity of the re-expansion operator but is rather due to the fact that the gradient of this operator is computed first in Spherical coordinates and then converted to Cartesian coordinates. This singularity affects the derivatives with respect to θ and ϕ . Note that when $\theta = 0,\pi$, the coordinate ϕ becomes meaningless and we can arbitrarily set it to 0.

For the case of $\partial R_n^{m,s}/\partial\theta$, when $\sin\theta \rightarrow 0$, eq 1.3 reduces to

$$\frac{\partial R_n^{0,-s}}{\partial\theta} = s\cot\theta Y_{n,s}(\theta,\phi) \tag{1.11}$$

which is nonzero only when $s = \pm 1$ because then the sin θ term in both $Y_{n,s}(\theta,\phi)$ and $\cot\theta$ cancels out. Examining carefully the recursion rule in eq 1.4 tells us that only the $\partial R_n^{m,s}/\partial\theta$ coefficients where $m = s \pm 1$ will be nonzero and thus the calculation and application of this operator in the singular case is significantly simplified (only $O(p^2)$ operations).

For the case of $\partial R_n^{m,s}/\partial\phi$ we also need to take into account the term $\partial\phi/\partial y = \cos\phi/(r\sin\theta)$ in the operator that converts derivatives with respect to spherical coordinates to derivatives with respect to Cartesian coordinates (see eq 48) because of its dependence on $1/\sin\theta$. Similarly to the case of the $\partial R_n^{m,s}/\partial\theta$ coefficients, here too, because we end up multiplying each $\partial R_n^{m,s}/\partial\phi$ by $\partial\phi/\partial y$ we get a cancellation of $\sin\theta$ in both. Recall that the $\partial R_n^{m,s}/\partial\phi$ coefficients have a simple dependence on the $R_n^{m,s}$ coefficients (see eq 1.5). We thus define in this case new coefficients $\partial R_n^{m,s}/\partial y = (\partial R_n^{m,s}/\partial\phi\ \partial\phi/\partial y)$, noting that when $\sin\theta = 0$, $\partial R_n^{m,s}/\partial y$ depends only on $\partial R_n^{m,s}/\partial\phi$. Here too we end up with only the $\partial R_n^{m,s}/\partial y$ coefficients, where $s = \pm 1$, being nonzero.

A.3. Computing the Adapted MSBFs. Following Kirkwood, a simple recursive formula allows for the generation of all $\hat{k}_n(z)$

$$\hat{k}_{n+1}(z) = \hat{k}_n(z) + \frac{z^2\hat{k}_{n-1}(z)}{(2n+1)(2n-1)} \tag{1.12}$$

where the starting values are $\hat{k}_0(z) = 1$ and $\hat{k}_1(z) = 1 + z$. The derivatives of this function, which are used in the derivations of section 2, can be generated using the relation

$$\frac{d\hat{k}_n(z)}{dz} = \frac{(2n+1+z)\hat{k}_n(z) - (2n+1)\hat{k}_{n+1}(z)}{z} \tag{1.13}$$

We have also derived a recursive formula for the generation of all $\hat{\imath}_n(z)$:

$$\hat{\imath}_{n+1}(z) = (2n+1)(2n+3)\frac{\hat{\imath}_{n-1}(z) - \hat{\imath}_n(z)}{z^2} \tag{1.14}$$

where the starting values are $\hat{\imath}_0(z) = \sinh(z)/z$ and $\hat{\imath}_1(z) = (3/z^2)(\cosh(z) - \hat{\imath}_0(z))$. While this formula is correct and useful in theoretical derivations, its usefulness in actual finite precision computation is limited. Instead we recommend the following formula

$$\hat{\imath}_n(z) = 1 + \sum_{j=1}^{L} t_j^n(z^2/2)$$

$$t_j^n(y) = \frac{1}{j}\, t_{j-1}^n(y)\frac{y}{2n+2j+3} \tag{1.15}$$

where the starting value is $t_0^n(y) = y/(2n+3)$. The summation in the formula should be continued until the change falls below the desired precision (we used $L = 20$). The derivatives can be generated using the relation

$$\frac{d\hat{\imath}_n(z)}{dz} = \frac{z\hat{\imath}_{n+1}(z)}{2n+3} \tag{1.16}$$

Finally a relation between the two kinds of the adapted MSBFs that is useful in simplifying complex expressions is

$$\hat{\imath}_n(z)\hat{k}_{n+1}(z) + \hat{k}_n(z)\hat{\imath}_{n+1}(z)\frac{z^2}{(2n+1)(2n+3)} = e^z \tag{1.17}$$

References

(1) Baker, N. A. In *Methods in Enzymology*; Brand, L., Johnson, M. L., Eds.; Academic Press: 2004; Vol. 383, pp 94−118.

(2) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144−1149.

(3) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. *Comput. Phys. Commun.* **1991**, *62*, 187−197.

(4) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosiewicz, J.; Gilson, M. K.; Bagheri, B.; Scott, L. R.; McCammon, J. A. *Comput. Phys. Commun.* **1995**, *91*, 57−95.

(5) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435−445.

(6) Rocchia, W.; Alexov, E.; Honig, B. *J. Phys. Chem.* **2001**, *105*, 6507−6514.

(7) Beard, D. A.; Schlick, T. *Biopolymers* **2001**, *58*, 106−115.

(8) Bordner, A. J.; Huber, G. A. *J. Comput. Chem.* **2003**, *24*, 353−367.

(9) Boschitsch, A. H.; Fenley, M. O.; Zhou, H. X. *J. Phys. Chem.* **2002**, *106*, 2741−2754.

(10) Gabdoulline, R. R.; Wade, R. C. *J. Phys. Chem.* **1996**, *100*, 3868−3878.

(11) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351−361.

(12) Hoffmann, N.; Likos, C. N.; Hansen, J. P. *Mol. Phys.* **2004**, *102*, 857−867.

(13) Glendinning, A. B.; Russel, W. B. *J. Colloid Interface Sci.* **1983**, *93*, 95−104.

(14) Verwey, E. J. W.; Overbeek, J. T. G. *Theory of the stability of lyophobic colloids*; Dover Publications: Mineola, NY, 1999.

(15) Sader, J. E.; Lenhoff, A. M. *J. Colloid Interface Sci.* **1998**, *201*, 233−243.

(16) Allen, R.; Hansen, J. P. *J. Phys.: Condens. Matter* **2002**, *14*, 11981−11997.

(17) Phillies, G. D. *J. Chem. Phys.* **1974**, *60*, 2721−2731.

(18) McClurg, R. B.; Zukoski, C. F. *J. Colloid Interface Sci.* **1998**, *208*, 529−542.

(19) Boschitsch, A. H.; Fenley, M. O.; Olson, W. K. *J. Comput. Phys.* **1999**, *151*, 212−241.

(20) Greengard, L. F.; Huang, J. F. *J. Comput. Phys.* **2002**, *180*, 642−658.

(21) Gumerov, N. A.; Duraiswami, R. *Siam J. Sci. Comput.* **2003**, *25*, 1344−1381.

(22) Rokhlin, V. *Appl. Comput. Harmonic Anal.* **1993**, *1*.

(23) Arfken, G. B.; Weber, H.-J. *Mathematical methods for physicists*, 4th ed.; Academic Press: San Diego, CA, 1995.

(24) Gabdoulline, R. R.; Wade, R. C. *J. Mol. Biol.* **2001**, *306*, 1139−1155.

(25) Buckle, A. M.; Schreiber, G.; Fersht, A. R. *Biochemistry* **1994**, *33*, 3, 8878−8889.

(26) Jorgensen, W. L.; Tiradorives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.

(27) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(28) Danos, M.; Maximon, L. C. *J. Math. Phys.* **1965**, *6*, 766−778.

(29) Epton, M. A.; Dembart, B. **1995**, *16*, 865−897.

(30) Golub, G. H.; Van Loan, C. F. *Matrix computations*, 3rd ed.; Johns Hopkins University Press: Baltimore, MD, 1996.

(31) Greengard, L.; Rokhlin, V. *J. Comput. Phys.* **1987**, *73*, 325−348.

(32) Gabdoulline, R. R.; Wade, R. C. *Methods* **1998**, *14*, 329−341.

(33) Gabdoulline, R. R.; Wade, R. C. *Curr. Opin. Struct. Biol.* **2002**, *12*, 204−213.

CT050263P