A Fast Adaptive Multigrid Boundary Element Method for Macromolecular Electrostatic Computations in a Solvent

YURY N. VOROBJEV,* HAROLD A. SCHERAGA

Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853-1301

ABSTRACT: A fast multigrid boundary element (MBE) method for solving the Poisson equation for macromolecular electrostatic calculations in a solvent is developed. To convert the integral equation of the BE method into a numerical linear equation of low dimensions, the MBE method uses an adaptive tesselation of the molecular surface by BEs with nonregular size. The size of the BEs increases in three successive levels as the uniformity of the electrostatic field on the molecular surface increases. The MBE method provides a high degree of consistency, good accuracy, and stability when the sizes of the BEs are varied. The computational complexity of the *unrestricted* MBE method scales as $O(N_{at})$, where N_{at} is the number of atoms in the macromolecule. The MBE method is ideally suited for parallel computations and for an integrated algorithm for calculations of solvation free energy and free energy of ionization, which are *coupled* with the conformation of a solute molecule. The current version of the 3-level MBE method is used to calculate the free energy of transfer from a vacuum to an aqueous solution and the free energy of the equilibrium state of ionization of a 17-residue peptide in a given conformation at a given pH in ~ 400 s of CPU time on one node of the IBM SP2 supercomputer. © 1997 by John Wiley & Sons, Inc. *J Comput Chem* **18**: 569−583, 1997

Keywords: fast multigrid boundary element method; macromolecular electrostatic calculations; Poisson equation

Introduction

Electrostatic interactions of a macromolecule with a polar solvent and solvent-mediated screening are essential factors in protein folding.[1-3] Proteins in aqueous solution are involved in equilibrium interactions (binding–releasing) with aqueous protons, which depend on the protein conformation. These interactions provide a *coupling* between the state of ionization **x** of the ionizable groups and the conformation of the protein at a given pH in the solution.[1,2,4] A realistic treatment of the protein folding problem, namely the prediction of chain-folding initiation sites, or the conformations of surface loops of a protein in aqueous solution, can be achieved by considering the total energy $E(\mathbf{r}_p, \text{pH})$ of a protein in the conformation $\mathbf{r}_p$ in aqueous solution at a given pH as a sum,

$$E(\mathbf{r}_p, \text{pH}) = E_{\text{int}}(\mathbf{r}_p) + F_{\text{cav}}(\mathbf{r}_p) + F_{\text{solv}}(\mathbf{r}_p)$$
$$+ F_{\text{inz}}(\mathbf{r}_p, \text{pH}), \quad (1)$$

where $E_{\text{int}}(\mathbf{r}_p)$ is the internal conformational energy of the protein or the conformational energy in a vacuum; $F_{\text{cav}}(\mathbf{r}_p)$ is the free energy of creation of the molecular cavity in an aqueous solution; $F_{\text{solv}}(\mathbf{r}_p)$ is an electrostatic solvation free energy, or the free energy of transfer of the protein from a vacuum into an aqueous solution; and $F_{\text{inz}}(\mathbf{r}_p, \text{pH})$ is the free energy for changing the initial state of ionization of the ionizable groups in a vacuum when the protein is moved into an aqueous solvent at a given pH. Methods to treat multiple-site titration phenomena have recently been developed.[5,6] The essential part of these methods is an accurate calculation of the desolvation free energy Δg_{solv_i} of the ionizable groups, when they are transferred from a solvent to the protein environment, and the pairwise potential of mean force (PMF), W_{ij}, between the ionized groups in a solvent.

Research has shown that the continuum electrostatic model provides a reliable method for obtaining a physically reasonable description of the electrostatic polarization of an aqueous solution by the charges of a solute and provides a quantitative estimation of the electrostatic free energy of solvation.[7-10] The continuum electrostatic model assumes that both the solvent and solute can be treated as linear dielectric media with different uniform dielectric constants, which are separated by a smooth dielectric boundary surface. The electrostatic potential is then described by the Poisson equation.[7-18] The solution of the Poisson equation can be achieved by either of two classes of numerical methods, the finite-difference (FD) method[7,11,16] or the boundary element (BE) method.[9,12-14,19]

The FD method, first used for biomolecular systems by Warwicker and Watson,[11] restricts the 3-dimensional (3-D) space to a rectangular 3-D box, which confines the solute and a portion of the solvent around the solute. The 3-D box is represented as a 3-D lattice; and the spatial electrostatic variables, i.e., the dielectric constant and charge density, are mapped onto the 3-D lattice and then the Poisson equation can be solved by an iterative method.[16,20] Faster and stabler results are provided by a multigrid iterative method.[20-24] For large molecules, the FD method has an accuracy problem because of the limitations inherent in mapping physical quantities onto a 3-D lattice,[25-27] i.e., the dependence of the solvation free energy on the position and orientation of the molecule on the 3-D lattice. An experimental accuracy of about kT in the calculation of the electrostatic potential can be achieved for a very fine lattice spacing, with at least three grids per Angstrom. The computational complexity of the FD method greatly depends on the lattice dimension N_g, i.e., the number of grids per box length, because the number of variables is N_g^3 and the convergence rate is about $(1 - 1/N_g)$.[22-24]

The BE method has some inherent advantages[9,12-14,19] over the FD method because it works directly with a surface variable, i.e., the polarization surface charge density on the dielectric boundary, and presents a numerical solution of an integral equation over the dielectric boundary, to which the original Poisson equation is analytically converted.[12,14,27] The BE method is invariant to rotations and translations of the molecule because the molecular surface and the distribution of the surface BEs depend only on the distances between the atoms. A comprehensive comparison of the BE and FD methods was made by Bharadwaj et al.,[27] who made use of an advanced numerical implementation of the FD method; they found that the BE method exhibits a higher degree of consistency because of the absence of grid dependencies.

We developed the BE method to solve the nonlinear Poisson–Boltzmann equation[19] and applied the BE method to study electrostatic effects in a highly charged polypeptide, poly-(L-lysine).[6,19,28,29]

In this study new algorithmic advances were added to the BE method,[19] which improve its numerical accuracy and performance. A new adaptive multigrid BE (MBE) algorithm implements the multisize representation of the boundary elements[19] and provides a fast calculation of the electrostatic field in the multigrid representation of the dielectric border surface. We show that the MBE method is suitable for designing a *rational* algorithm for calculating the solvation free energy, F_{solv}, and the free energy of ionization, F_{inz}, of a peptide in a given conformation. The calculation of the energy of transfer from the vacuum to water and the free energy of equilibrium proton binding–releasing for the 17-residue peptide Ac-ETGTKAEL-LAKYEATHK-NMe with eight ionizable residues requires about 400 sec of CPU time in serial code on one node of the IBM SP2 supercomputer.

Continuum Dielectric Model

The continuum dielectric model assumes that a solvated molecule is comprised of a cavity of low dielectric constant, D_i, embedded in a continuum solvent medium of high dielectric constant, D_o. The solute is separated from the solvent by a boundary, which is defined to be the smooth surface traced by the inward-facing part of a probe (with the radius of a water molecule) as it rolls over the solute molecule. An analytical description of this surface was presented by Connolly.[30,31] This boundary confines the solute molecular charge distribution, which is typically represented as a set of fixed charges, $\{q_k\}$, located at atomic centers, $\{r_k\}$. The total electrostatic potential, Φ_{total}, in the absence of added salt, can be written as

$$\Phi_{total}(\mathbf{r}) = \Phi_{mol}(\mathbf{r}) + \Phi_r(\mathbf{r}), \tag{2}$$

where Φ_{mol} is the potential due to the solute charge distribution and Φ_r is the reaction potential due to the surface polarization charge density $\sigma(\mathbf{s})$ on the molecular surface arising from the response of the solvent dielectric medium to the solute charge distribution. The sum of the potentials, $\Phi_{mol} + \Phi_r$, satisfies the Poisson equation,

$$\nabla \cdot D(\mathbf{r}) \nabla [\Phi_{mol}(\mathbf{r}) + \Phi_r(\mathbf{r})] + 4\pi \sum_k q_k \delta(\mathbf{r} - \mathbf{r}_k) = 0, \tag{3}$$

where $\mathbf{r}_k$ is the position of a set of fixed charges within the solute cavity and $D(\mathbf{r})$ is the dielectric constant at point $\mathbf{r}$. The Poisson equation can be converted into an integral equation[12,27] for the surface charge density $\sigma(\mathbf{s})$ on the molecular surface S,

$$\sigma(\mathbf{t}) = f\int_S \frac{\sigma(\mathbf{s})(\mathbf{t} - \mathbf{s})\mathbf{n}(\mathbf{t})\,d\mathbf{s}}{|\mathbf{t} - \mathbf{s}|^3}$$
$$+ \frac{f}{D_i} \sum_k \frac{q_k(\mathbf{t} - \mathbf{r}_k)\mathbf{n}(t)}{|\mathbf{t} - \mathbf{r}_k|^3}, \tag{4}$$

where the constant f is

$$f = \frac{1}{2\pi} \frac{D_i - D_o}{D_i + D_o}, \tag{5}$$

and $\mathbf{n}(\mathbf{t})$ is the vector normal to the surface S at the point $\mathbf{t}$. The last term in eq. (4) is the normal component of the electrostatic field at the surface point $\mathbf{t}$ due to the solute charges $\{q_k\}$. This integral equation has a unique solution if the molecular surface S is a Liapunov–Kellogg regular surface,[32] i.e., one that satisfies the equation

$$\arccos[\mathbf{n}(\mathbf{t})\mathbf{n}(\mathbf{s})] \le C|\mathbf{t} - \mathbf{s}|, \tag{6}$$

where $\mathbf{t}$ and $\mathbf{s}$ are points on the surface S and C is a constant, which is equal to the inverse local curvature radius of the surface. The smooth Connolly molecular surface[30,31] satisfies the Liapunov–Kellogg equation, eq. (6), if saddle pieces of the molecular surface do not contain cusps.[31]

The molecular potential Φ_{mol} is given straightforwardly[19] by

$$\Phi_{mol}(\mathbf{r}) = \frac{1}{D_i} \sum_k \frac{q_k \delta(\mathbf{r} - \mathbf{r}_k)}{|\mathbf{r} - \mathbf{r}_k|}, \tag{7}$$

and the reaction potential Φ_r is obtained from the surface integral[19]

$$\Phi_r(\mathbf{r}) = \int_S \frac{\sigma(\mathbf{s})\,d\mathbf{s}}{|\mathbf{r} - \mathbf{s}|}. \tag{8}$$

In the BE method, the integral appearing in the first term on the right-hand side of eq. (4) is replaced by a discrete sum over a finite number of BEs, s_α, with surface area Δs_α and constant polarization charge density σ_α. The BEs tessellate the molecular surface, so that the integral equation, eq. (4), is transformed into a set of linear numerical equations in matrix form,[33,34]

$$\sigma = \sigma \mathbf{K} + \mathbf{b}, \tag{9}$$

where these terms are identified with the terms on the left and right side, respectively, of eq. (4).

The matrix elements $K_{\alpha\beta}$ define the normal component of the electrostatic field on the BE t_α, which is produced by the uniformly charged BE s_β. The matrix elements are a function of only the geometric properties of the BEs that cover the molecular surface[12,13,33]:

$$K_{\alpha\beta} = f\frac{(\mathbf{t}_\alpha - \mathbf{s}_\beta)\mathbf{n}(\mathbf{t}_\alpha)\,\Delta s_\beta}{|\mathbf{t}_\alpha - \mathbf{s}_\beta|^3}, \qquad (10)$$

where f is given by eq. (5). For a Liapunov–Kellogg regular surface[32] S, eq. (6), the diagonal matrix elements $K_{\alpha\beta} = 0$ in the $\lim|\mathbf{t}_\alpha - \mathbf{s}_\beta| = 0$.

The last term, $\mathbf{b}$, of eq. (9) is the electrostatic field generated by the charges $\{q_k\}$ of the solute molecule and constitutes a permanent source electrostatic field. The value of the source field $\mathbf{b}$ on the BE α, i.e., b_α is given by[19,34]

$$b_\alpha = \frac{f}{D_i}\sum_k \frac{q_k(\mathbf{t}_\alpha - \mathbf{r}_k)\mathbf{n}(\mathbf{t}_\alpha)}{|\mathbf{t}_\alpha - \mathbf{r}_k|^3}. \qquad (11)$$

A solution of eq. (9), in principle, can be obtained by the variety of methods for solving a system of linear equations with a nonsymmetric matrix.[20]

Properties of BE Integral Equation

The integral equation, eq. (4), for the polarization charge density σ has several useful properties, which reflect internal self-consistency of this integral equation. By integrating both sides of eq. (4) over the molecular surface S, it is easy to show that the total polarization charge $\sigma(\mathbf{t})$ automatically satisfies Gauss' theorem,

$$\int_S \sigma(\mathbf{s})\,d\mathbf{s} = \left(\frac{1}{D_o} - \frac{1}{D_i}\right)\sum_k q_k, \qquad (12)$$

because, for any point $\mathbf{r}_k$ that is inside a closed surface S, the solid angle Ω covered by the surface S is the full solid angle 4π, i.e.,

$$\int_S \frac{(\mathbf{t} - \mathbf{r}_k)\mathbf{n}(\mathbf{t})\,d\mathbf{t}}{|\mathbf{t} - \mathbf{r}_k|^3} = 4\pi. \qquad (13)$$

If the point $\mathbf{r}_k$ moves on the surface S, and this surface is a Liapunov–Kellogg regular surface, then for any point $\mathbf{s}$ of the surface,[32] the surface S will

cover only half of a full solid angle, i.e.,

$$\int_S \frac{(\mathbf{t} - \mathbf{s})\mathbf{n}(\mathbf{t})\,d\mathbf{t}}{|\mathbf{t} - \mathbf{s}|^3} = 2\pi. \qquad (14)$$

The BE analogs of eqs. (13) and (14) can be found by using eqs. (11) and (10), respectively, and they are

$$\sum_\alpha b_\alpha = 4\pi\frac{f}{D_i}\sum_k q_k \qquad (15)$$

and

$$\sum_\beta K_{\alpha\beta}\,\Delta s_\beta = 2\pi f\Delta s_\alpha. \qquad (16)$$

The numerical equation, eq. (9), generally loses internal self-consistency because of the representation of the surface S by finite BEs. In turn, it creates an avalanche of numerical errors in the numerical solution of the linear system (9) for the charge density σ_α. The main idea of designing an accurate numerical algorithm, which is implemented in this article, is to construct a system of linear eqs. (9) that exhibit the same properties of internal self-consistency for matrix elements (16) and for the source term (15) that the initial integral equation, eq. (4), has. Rashin and Namboodiri[33] used normalization of the polarization charge σ_α by Gauss' theorem (12) to obtain the correct charge distribution. Purisima and Nilar[35] recently showed that the calculation of the diagonal matrix elements $K_{\alpha\alpha}$ on the basis of the sum rule (15) considerably improves the accuracy of the solution. Since our first implementation of the iterative BE method (IBE),[19,28,29] we have been experimenting with applying self-consistency rules (15) and (16) (in a different manner than that used by Purisima and Nilar[35]) to achieve maximum accuracy and a speedy solution of eq. (4) for a macromolecular surface.

To design an optimal numerical method, we have to reconcile three principal components of the IBE method: (i) maximum accuracy can be achieved with a small surface lattice of BEs s_α to assure a small variation of $\sigma(\mathbf{s})$ over the BE s_α; (ii) the variation of $\sigma(\mathbf{s}_\alpha)$ is proportional to the variation of the source electrostatic field over the BE s_α; (iii) to achieve maximum speed in the calculation of the polarization charge density distribution σ, the dimensions of the linear system (9) must be as small as possible. To reconcile these three components, we considered multilevel-sized BEs.

MBE Algorithm

To solve the matrix equation, eq. (9), and decrease the number of variables where appropriate, we define three sets of BEs on the molecular surface: (i) a set of small BEs with coordinates $\mathbf{s}_\alpha^s$, normal vectors $\mathbf{n}_\alpha^s$, small areas Δs_α, and polarization charge densities σ_α^s; (ii) a set of large BEs with coordinates $\mathbf{s}_\beta^l$, normal vectors $\mathbf{n}_\beta^l$, large areas Δl_β, and polarization charge densities σ_β^l; (iii) a set of patch BEs with coordinates $\mathbf{s}_\gamma^p$, large areas Δp_γ, and polarization charge densities σ_γ^p (the patches are very large irregularly shaped parts of the surface). Assume that the polarization charge densities σ_α^s, σ_β^l, and σ_γ^p are *constant* over the respective BEs. Each large BE $\mathbf{s}_\beta^l$ is a collection of small BEs $(\mathbf{s}_{\beta_1}^s, \ldots, \mathbf{s}_{\beta_n}^s)$; the area of a large BE is a sum of the areas of a collection of small BEs. Likewise, each patch BE $\mathbf{s}_\gamma^p$ is a collection of large BEs $(\mathbf{s}_{\gamma_1}^l, \ldots, \mathbf{s}_{\gamma_n}^l)$. The polarization charge density of a large BE is an average polarization charge density over the collection of small BEs, i.e.,

$$\Delta l_\beta = \sum_{\beta_i} \Delta s_{\beta_i}, \tag{17}$$

$$\Delta p_\gamma = \sum_{\gamma_k} \Delta l_{\gamma_k}, \tag{18}$$

$$\sigma_\beta^l = \frac{1}{\Delta l_\beta} \sum_{\beta_i} \sigma_{\beta_i}^s \Delta s_{\beta_i}, \tag{19}$$

$$\sigma_\gamma^p = \frac{1}{\Delta p_\gamma} \sum_{\gamma_k} \sigma_{\gamma_k}^l \Delta l_{\gamma_k}. \tag{20}$$

To take advantage of the slow variation of the source term of the *linear* equation, eq. (4), on the macromolecular surface, and consequently the slow variation of the polarization charge density, consider the solution of eq. (4) for each charged atom, or group of atoms, *separately*. If $\sigma_k(\mathbf{s})$ is the polarization charge density due to charged atom k, then the total polarization charge density of the whole molecule is the sum

$$\sigma(\mathbf{s}) = \sum_k \sigma_k(\mathbf{s}), \tag{21}$$

due to the linearity of eq. (4).

Consider charged atom k with coordinate $\mathbf{r}_k$. The macromolecular surface can be divided into three regions: (i) the *Local* surface S_{Loc}, within a distance R_{Loc} from the source $\mathbf{r}_k$ of the permanent molecular electrostatic field, is represented by the small BEs; (ii) the *Intermediate* surface S_{Int}, with BEs within a distance R_{Int} from the source $\mathbf{r}_k$ of the permanent molecular electrostatic field, is represented by the large BEs; (iii) the *Distant* surface S_{Dst}, is represented by the patch BEs, which include the rest of the macromolecular surface at large distances $> R_{\mathrm{Int}}$ form the source $\mathbf{r}_k$ of the permanent molecular electrostatic field (see Fig. 1). It should be noted that, for a large macromolecule, the distant region can be subdivided into a set of distant regions with a steadily increasing size of patches. The polarization charge density distribution can be approximated by the *low-dimension vector* $(\sigma^L, \sigma^I, \sigma^D)$; the source term can be approximated by $\mathbf{b}^L, \mathbf{b}^I, \mathbf{b}^D$, and the matrix equation, eq. (9), can be represented as a linear equation of *low dimensions*,

$$\begin{pmatrix} \sigma^L \\ \sigma^I \\ \sigma^D \end{pmatrix} = \begin{pmatrix} \mathbf{K}^{LL} & \mathbf{K}^{LI} & \mathbf{K}^{LD} \\ \mathbf{K}^{IL} & \mathbf{K}^{II} & \mathbf{K}^{ID} \\ \mathbf{K}^{DL} & \mathbf{K}^{DI} & \mathbf{K}^{DD} \end{pmatrix} \begin{pmatrix} \sigma^L \\ \sigma^I \\ \sigma^D \end{pmatrix} + \begin{pmatrix} \mathbf{b}^L \\ \mathbf{b}^I \\ \mathbf{b}^D \end{pmatrix}, \tag{22}$$

where the expression for the matrix elements $K_{\alpha\beta}^{LL}$ follows from eq. (10), and is

$$K_{\alpha\beta}^{LL} = f \frac{(\mathbf{s}_\alpha^s - \mathbf{s}_\beta^s)\mathbf{n}(\mathbf{s}_\alpha^s) \Delta s_\beta}{|\mathbf{s}_\alpha^s - \mathbf{s}_\beta^s|^3}. \tag{23}$$

As mentioned above, the matrix elements $K_{\alpha\beta}$ define the average electrostatic field[19,33,34] generated by BE β on BE α. To speed up the calculation
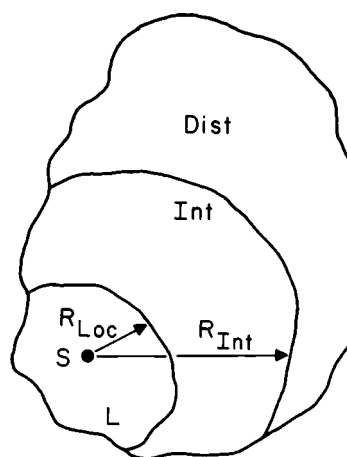

FIGURE 1. Definition of the Local, L, Intermediate, I, and Distant, D, portions of a molecular surface for the source center S (i.e., a charge or group of charges).

of the matrix elements in eq. (22), the fast multigrid approximation at the monopole level, similar to that of Greengard and Rokhlin,[36] is implemented for the calculation of the matrix elements $\mathbf{K}^{LI}$, $\mathbf{K}^{IL}$, $\mathbf{K}^{II}$, $\mathbf{K}^{LD}$, $\mathbf{K}^{DL}$, $\mathbf{K}^{ID}$, $\mathbf{K}^{DI}$, and $\mathbf{K}^{DD}$. The matrix elements between the local BEs are calculated on the basis of a fine grid, while the matrix elements for the distant BEs are calculated on the basis of a coarse grid,

$$
\begin{aligned}
K_{\alpha\beta}^{LI} &= f\sum_{\beta_k} \frac{\left(\mathbf{s}_\alpha^s - \mathbf{s}_{\beta_k}^s\right)\mathbf{n}(\mathbf{s}_\alpha^s)\,\Delta s_{\beta_k}}{|\mathbf{s}_\alpha^s - \mathbf{s}_{\beta_k}^s|^3} \\
&= \sum_{\beta_k} K_{\alpha\beta_k}^{LL} \qquad\qquad \text{if } |\mathbf{s}_\alpha^s - \mathbf{s}_\beta^l| < R_{\text{Loc}} \\
&= f\frac{\left(\mathbf{s}_\alpha^s - \mathbf{s}_\beta^l\right)\mathbf{n}(\mathbf{s}_\alpha^s)\,\Delta l_\beta}{|\mathbf{s}_\alpha^s - \mathbf{s}_\beta^l|^3} \qquad \text{if } |\mathbf{s}_\alpha^s - \mathbf{s}_\beta^l| > R_{\text{Loc}},
\end{aligned}
$$
$$(24)$$

where β_k is a collection of the small BEs that tesselate a large BE β. The matrix elements $K_{\alpha\beta}^{IL}$ are given by

$$
\begin{aligned}
K_{\alpha\beta}^{IL} &= \frac{f}{\Delta l_\alpha} \sum_{\alpha_k} \frac{\left(\mathbf{s}_{\alpha_k}^s - \mathbf{s}_\beta^s\right)\mathbf{n}(\mathbf{s}_{\alpha_k}^s)\,\Delta s_\beta}{|\mathbf{s}_{\alpha_k}^s - \mathbf{s}_\beta^s|^3} \\
&= \frac{1}{\Delta l_\alpha} \sum_{\alpha_k} K_{\alpha_k\beta}^{LL} \qquad \text{if } |\mathbf{s}_\alpha^l - \mathbf{s}_\beta^s| < R_{\text{Loc}} \\
&= f\frac{\left(\mathbf{s}_\alpha^l - \mathbf{s}_\beta^s\right)\mathbf{n}(\mathbf{s}_\alpha^l)\,\Delta s_\beta}{|\mathbf{s}_\alpha^s - \mathbf{s}_\beta^l|^3} \qquad \text{if } |\mathbf{s}_\alpha^l - \mathbf{s}_\beta^s| > R_{\text{Loc}},
\end{aligned}
$$
$$(25)$$

where Δl_α is the area of the large BE α. The matrix elements of the diagonal submatrix $\mathbf{K}^{II}$ are given by

$$
\begin{aligned}
K_{\alpha\beta}^{II} &= \frac{f}{\Delta l_\alpha} \sum_{\alpha_i} \sum_{\beta_j} \frac{\left(\mathbf{s}_{\alpha_i}^l - \mathbf{s}_{\beta_j}^s\right)\mathbf{n}(\mathbf{s}_{\alpha_i}^s)\,\Delta s_{\beta_j}\,\Delta s_{\alpha_i}}{|\mathbf{s}_{\alpha_i}^s - \mathbf{s}_{\beta_k}^s|^3} \\
&= \sum_{\alpha_i} K_{\alpha_i\beta}^{LI} \qquad\qquad \text{if } |\mathbf{s}_\alpha^s - \mathbf{s}_\beta^l| < R_{\text{Loc}} \\
&= f\frac{\left(\mathbf{s}_\alpha^l - \mathbf{s}_\beta^l\right)\mathbf{n}(\mathbf{s}_\alpha^l)\,\Delta l_\beta}{|\mathbf{s}_\alpha^l - \mathbf{s}_\beta^l|^3} \\
&\qquad\qquad\qquad \text{if } |\mathbf{s}_\alpha^l - \mathbf{s}_\beta^l| > R_{\text{Loc}}. \quad (26)
\end{aligned}
$$

The matrix elements between patches of the diagonal submatrix $\mathbf{K}^{DD}$ are given by

$$
\begin{aligned}
K_{\alpha\beta}^{DD} &= \frac{f}{\Delta p_\alpha} \sum_{\alpha_i} \sum_{\beta_j} \frac{\left(\mathbf{s}_{\alpha_i}^s - \mathbf{s}_{\beta_j}^s\right)\mathbf{n}(\mathbf{s}_{\alpha_i}^s)\,\Delta s_{\beta_j}\,\Delta s_{\alpha_i}}{|\mathbf{s}_{\alpha_i}^s - \mathbf{s}_{\beta_k}^s|^3} \\
&= \frac{1}{\Delta p_\alpha} \sum_{\alpha_i} \sum_{\beta_j} K_{\alpha_i\beta_j}^{LL}\,\Delta s_{\alpha_i} \\
&\qquad\qquad\qquad \text{if } |\mathbf{s}_\alpha^p - \mathbf{s}_\beta^p| < R_{\text{Loc}} \\
&= \frac{f}{\Delta p_\alpha} \sum_{\alpha_i} \sum_{\beta_j} \frac{\left(\mathbf{s}_{\alpha_i}^l - \mathbf{s}_{\beta_j}^l\right)\mathbf{n}(\mathbf{s}_{\alpha_i}^l)\,\Delta l_{\beta_j}\,\Delta l_{\alpha_i}}{|\mathbf{s}_{\alpha_i}^l - \mathbf{s}_{\beta_k}^l|^3} \\
&\qquad\qquad\qquad \text{if } |\mathbf{s}_\alpha^p - \mathbf{s}_\beta^p| > R_{\text{Loc}} \\
&= \frac{1}{\Delta p_\alpha} \sum_{\alpha_i} \sum_{\beta_j} K_{\alpha_i\beta_j}^{II}\,\Delta l_{\alpha_i}. \qquad (27)
\end{aligned}
$$

The matrix elements of all other submatrices $\mathbf{K}^{LD}$, $\mathbf{K}^{ID}$, etc., can be expressed by similar equations. The important properties of the expressions shown above are that the matrix element of a high level, e.g., $\mathbf{K}^{DD}$, is a sum of the subsets of low-level matrix elements, e.g., $\mathbf{K}^{II}$ or $\mathbf{K}^{LL}$.

The diagonal elements $K_{\alpha\alpha}^{LL}$ for the small BEs on the molecular surface are calculated by

$$
K_{\alpha\alpha}^{LL} = k_\alpha(\pi\,\Delta s_\alpha)^{1/2}/R_c, \qquad (28)
$$

where R_c is the local radius of curvature of the molecular surface and k_α is an empirical coefficient that is found numerically to be $+0.75$, -0.20, and 0.0 for convex, concave, and saddle faces, respectively, of the molecular surface. Equation (28) is obtained by analytical integration of eq. (10) over the small BE, which is assumed to have a circular shape, and the coefficient k_α reflects the real shape of the small BE that is generated by our program. It should be noted that the usual choice for the diagonal matrix elements in the BE method[12-14] is to equate them to zero (i.e., assuming that the small BEs are planar). Purisima and Nilar[35] obtained the diagonal matrix elements from the normalization condition of eq. (14).

The Liapunov–Kellogg condition, eq. (6), enables one to obtain the upper bound for the matrix elements $K_{\alpha\beta}$ between two BEs with centers at $\mathbf{s}_\alpha$ and $\mathbf{s}_\beta$ on the molecular surface,

$$
|K_{\alpha\beta}| \le |\mathbf{s}_\alpha - \mathbf{s}_\beta|/4R_c. \qquad (29)
$$

If condition (29) is violated for the given matrix element $K_{\alpha\beta}$, then $K_{\alpha\beta}$ is replaced by the appro-

priate upper bound value. Filtering of the matrix elements $K_{\alpha\beta}$ through this condition (29) is equivalent to smoothing singular cusps on the molecular surface. The physically real dielectric border surface does not contain singular cusps because an intramolecular dielectric constant cannot be assigned to the infinitesimally thin volume of the cusp in the vicinity of its singular point. Therefore, cusps, which appear rarely in the Connolly geometrically constructed molecular surface, have to be smoothed with a user-defined radius of curvature (about 1 Å).

As follows from the sum rule equation, eq. (16), the submatrices $\mathbf{K}^{XY}$, where $X, Y = L, I, D$, are subject to the sum rule

$$\sum_{\alpha_i^s} K_{\alpha_i^s \beta}^{LX} \Delta s_{\alpha_i^s} + \sum_{\alpha_i^l} K_{\alpha_i^l \beta}^{IX} \Delta l_{\alpha_i^l} + \sum_{\alpha_i^p} K_{\alpha_i^p \beta}^{DX} \Delta p_{\alpha_i^p}$$

$$= 2\pi \Delta s_\beta^X, \quad (30)$$

where Δs_β^X is the area of the BE X_β. The sum rule for the vector $\mathbf{b}^L, \mathbf{b}^I, \mathbf{b}^D$, follows from eq. (15),

$$\sum_{\alpha_i^s} b_{\alpha_i^s}^L \Delta s_{\alpha_i^s} + \sum_{\alpha_i^l} b_{\alpha_i^l}^I \Delta l_{\alpha_i^l} + \sum_{\alpha_i^p} b_{\alpha_i^p}^L \Delta p_{\alpha_i^p}$$

$$= 4\pi \frac{f}{D_i} \sum_k q_k. \quad (31)$$

The solution of the linear equation, eq. (22), can be found iteratively. We found that the preconditioned biconjugate gradient (PBCG) method[20] converges in all cases, whereas the Gauss–Seidel (GS) method[20] fails to converge. It should be noted that the PBCG method can be implemented only if the full matrix $\mathbf{K}$ is calculated. The iterative BE method, which uses a partially calculated matrix[19] $\mathbf{K}$ within a *local* region, or the BE method without calculation of matrix[27] $\mathbf{K}$, can use only the GS method, which often fails to attain convergence. Therefore, to obtain convergence, the iteration with underrelaxation[19,27] must be used.

Our 3-level MBE method for calculating the polarization charge density can be summarized as:

1. calculation of a set of *small* and *large* BEs on the molecular surface, using the MSEED[37] or Connolly[30] method;

2. collection of *small* BEs into the *large* BEs and collection of *large* BEs into the *patch*

BEs (it should be noted[38,†] that our method does not need the expensive surface *triangulation* used, for example, by Zauhar and Morgan[12]);

3. definition of a collection of *centers of molecular electrostatic fields* as a set of charged atoms or compact atom groups;

4. calculation of the MBE set $(\mathbf{s}^s, \mathbf{s}^l, \mathbf{s}^p)$ for the *kth center of the molecular electrostatic field*;

5. calculation of an appropriate source vector $(\mathbf{b}^s, \mathbf{b}^l, \mathbf{b}^p)$ and a uniform correction by normalization by the sum rule, eq. (31);

6. calculation of the matrix elements of eq. (22), and a uniform correction to satisfy the sum rule, eq. (30);

7. solution of the linear equation, eq. (22), by the PBCG iterative method;

8. uniform normalization of the resulting polarization charge vector $\sigma^s, \sigma^j, \sigma^p$ by Gauss' theorem, eq. (12);

9. collection of the polarization charge densities on the MBE;

10. go back to step 4 and consider the next *charged center of the molecular electrostatic field*;

The present 3-level MBE implementation considers the *small grid* and *large grid* as collections of pseudoregularly distributed BEs on the molecular surface with a high density d_s, about 10 points/Å^2, and a low density d_l, about 1 point/Å^2, respectively, and each *patch* BE represents the whole surface of one surface atom. The multigrid method can be generalized for any unrestricted multilevel set of surface BEs, whose size is increased continu-

---

† Molecular surface triangulation is used to approximate the surface polarization charge density inside the triangle BE linearly on the basis of the charge densities at the vertices of the triangle, to achieve an accurate approximation of the nonsingular surface integral in eq. (4) with relatively large areas of the BEs. The same accuracy of the approximation of the surface integral can be achieved by using the average value of the polarization charge density at the center of the BEs and smaller (about 2–3 times) areas of the BEs, because the difference in the accuracy between these two approximations scales as the second power of the area of the BEs, i.e., $\sim \Delta s^2$. Yoon and Lenhoff[38] provided a numerical comparison between the linear and constant approximation and showed that comparable accuracy can be obtained in the framework of the constant approximation with twice the smaller areas of the BEs. A proper triangulation of the molecular surface is a complicated and time-consuming procedure that can produce long triangles with small areas or other unpleasant features in the molecular surface tesselation.[12]

ously, by considering a set of compact groups of surface atoms. An estimate of the number of independent surface variables n_{MBE} in the unrestricted multilevel MBE method for a large macromolecule is given by

$$n_{MBE} \sim n_L \log N_s, \quad (32)$$

where N_s is the total number of surface atoms of a macromolecule, n_L is the average number of BEs for each level of BEs, i.e., *Local*, *Intermediate*, and *Distant* regions, see Figure 1. If the size of the BEs at level i increases by any constant factor m compared to the size of the BEs at level $(i - 1)$, and the number of BEs at each size level is approximately constant, then the maximum level of the BEs is $\sim \log N_s$ and the total number of the multilevel BEs is defined by expression (32).

The computational complexity for calculating the matrix elements is on the order of $\sim n_{MBE}^2$, and the complexity for solving the linear system (22) by the PBCG method is also on the order of $\sim n_{MBE}^2$. To obtain the total polarization charge distribution, we must *repeat* the solution of the linear system (22) for each source of electrostatic field of the macromolecule (in the worst case for *each* charged atom), see eq. (21); consequently,

$$complexity \sim N_{at}(n_L \log N_s)^2, \quad (33)$$

i.e,. the *complexity* scales as N_{at}, the number of atoms in the macromolecule. The actual complexity scales as $\sim N_s$, because the surface atoms can be united into a set of compact charged groups and the atoms of the molecular core produce a more uniform polarization of a solvent on the molecular surface and, consequently, can be united in a set of large clusters. Then, each charged cluster can be considered as one effective source of an electrostatic field on the molecular surface.

MBE Algorithm Combined with Calculation of Transfer Energy and Ionization Equilibrium

The main goal of designing methods to solve the Poisson equation is to calculate the total electrostatic free energy of a macromolecule in a solution, i.e., eq. (1). The solvation (or transfer) free energy of a solute macromolecule in a solvent is a difference of polarization free energies of the macromolecule in the same conformation in the

solvent and in a vacuum environment, i.e.,

$$F_{solv}(\mathbf{r}_p) = W_{polz}(D_i, D_{sol}) - W_{polz}(D_i, D_{vac}), \quad (34)$$

where W_{polz} is the free energy of the surface polarization of the macromolecule, D_i is the dielectric constant of the macromolecule, D_{sol} is the dielectric constant of the solvent, and D_{vac} is the dielectric constant in a vacuum, which were taken as 2.0, 80.0, and 1.0, respectively, in all our calculations. The polarization free energy W_{polz} is given by

$$W_{polz} = \frac{1}{2} \sum_k \sum_\alpha \frac{q_k \sigma_\alpha}{|\mathbf{r}_k - \mathbf{s}_\alpha|}, \quad (35)$$

where σ_α is the polarization charge and s_α is the BE on the molecular surface.[12,19] For simplicity, it is reasonable to suppose that the surface, which defines the dielectric border between two dielectric media, is identical in the aqueous solvent and in a vacuum. Then, to calculate the polarization charge densities in the two environments, i.e., to obtain the solution of eq. (22), we need to calculate the matrix elements $K_{\alpha\beta}$ and the source field vector $\mathbf{b}^s, \mathbf{b}^l, \mathbf{b}^p$ only once. The solution of linear equation (22) by the PBCG method is fast and requires only about five to six iterations.

The proposed MBE method has several advantages when the calculation of the ionization free energy F_{inz} is considered. If the number N of coupled ionization sites of a macromolecule does not exceed 20, then the partition function over the ionization microstates $\mathbf{x} = (x_1, \ldots, x_N), (x_i = 0, 1)$ can be calculated directly as[5,6]

$$Z = \sum_{n=1}^{2^N} \exp\left[-\frac{\Delta G_{H_{m_n}^+}^{PS}(\mathbf{x}_n)}{kT} \right], \quad (36)$$

where $\mathbf{x_n}$ is the nth protonation microstate, the total number of protonation microstates being equal to 2^N. The free energy F_{inz} is

$$F_{inz} = -kT \ln Z. \quad (37)$$

The free energy of the ionization microstate $\mathbf{x}$ is given by

$$\Delta G(PS; \mathbf{x}) = \sum_{i=1}^{N} x_i(pH - pK_i^o - \Delta pK_i)[\ln 10]kT$$

$$+ \sum_{i, j=1; i<j}^{N} W_{ij} x_i x_j, \quad (38)$$

where K_i^o is the intrinsic dissociation constant of the ith single isolated residue in the solvent and ΔpK_i is the shift in pK of a single isolated residue

due to its transfer from the solvent to the protein environment. The last sum in eq. (38) represents the free energy of interaction between the N titratable sites in terms of the pairwise PMFs, W_{ij}. The pK shift is given by

$$\Delta pK_i[\ln 10]kT = \left[E(PS_i^+) - E(PS_i^0) \right]$$
$$- \left[E(S_i^+) - E(S_i^0) \right], \quad (39)$$

where $E(S_i^+)$ and $E(S_i^0)$ are the total electrostatic energies of the ith single isolated residue in the ionized and neutral state, respectively, in the solvent, and $E(PS_i^+)$, $E(PS_i^0)$ are the total electrostatic energies of the ith residue in the ionized and neutral state, respectively, in the protein environment in the solvent. The total electrostatic energy is given as the sum of the last two terms in eq. (1). The potential of mean force W_{ij} is given by the following equation[6]:

$$W_{ij} = \frac{1}{2}q_i \int_{S_m} \frac{\sigma_j(\mathbf{s}) \, ds}{|\mathbf{r}_{ji} + \mathbf{r}_j - \mathbf{s}|} + \frac{1}{2}q_j \int_{S_m} \frac{\sigma_i(\mathbf{s}) \, ds}{|\mathbf{r}_{ij} + \mathbf{r}_i - \mathbf{s}|}, \quad (40)$$

where $\sigma_i(\mathbf{s})$ and $\sigma_j(\mathbf{s})$ are the polarization charge densities due to charged groups i and j, respectively. It is easy to see from eq. (40) that all pK_i shifts and PMFs W_{ij} can be calculated sequentially while only *one* loop over all charged groups is carried out; therefore, these calculations are combined into the algorithm described in the previous section as step 8a, following step 8, calculation of pK_i shift and PMFs W_{ij}. These calculations of pK shifts and pairwise PMFs are carried out faster in the framework of the proposed MBE method compared to what can be achieved by a straightforward application of the FD method,[5,39] because calculations by the FD method have to be repeated for each N single ionized and $N(N-1)/2$ double ionized microstates.

Results

The MBE method has some intrinsic parameters, d_s, d_l, R_{Loc}, and R_{Int} that are defined in the previous sections and that affect the accuracy and performance of the MBE method. We compared the MBE method for a spherical cavity,[19,20,27,35] which contains a point charge and is embedded in a polar solvent, with the analytical results.[35] For real molecular systems, comparisons were made with the results of calculations by the multigrid

version of the improved DelPhi program.[40] In all calculations, we used PARSE[10] atomic charges and radii to define a molecular surface, which was calculated with the MSEED[37] or Connolly[30] methods with a probe radius of 1.4 Å.

Figure 2 shows the relative errors between the MBE method and the analytical solution[35] for the polarization free energy of the off-center point charge in a spherical cavity of radius 5 Å versus the displacement D of the point charge from the center of the cavity. It can be seen that an accuracy of about 1% is achieved when the size of the small BEs is less than the distance between the point charge and the surface of the cavity. It should be noted that, in real molecular systems, the smallest distance between any point charge and the molecular surface is larger than the smallest atomic radius of ~ 1 Å, if we assume that all atomic charges are located at the centers of the atoms. Figure 3 shows the relative errors for two positions of the point charge, $D = 4.5$ and 4.75 Å, respecitvely, versus the parameter R_{Loc}. From this figure we conclude that a reasonable value of the parameter R_{Loc} is about triple the size of the large BEs. Further increase of R_{Loc} does not improve the accuracy of the free energy, because this accuracy is limited by the accuracy of the polarization charge density in the *local* region, which is defined by the size of the small BEs.

Figure 4 shows the dependence of the polarization free energy of *n*-butylamine, W_{polz}, given by eq. (35), versus $(d_s)^{-1/2}$. To define the dependence of W_{polz} on the size of the *small* BEs, the values of the parameters R_{Loc} and R_{Int} were taken larger than the maximal size of the molecule, the latter being 10 Å. It can be seen that, for values of $(d_s)^{-1/2} < 0.5$, the dependence is very flat for the MBE method; therefore, the density of the *small* BEs $d_s > 4.0$ Å^{-2} can be considered as an acceptable density, which provides a relative accuracy for the calculation of W_{solv} of better than 0.5%. As shown in our previous work,[19] a reasonable size of a *large* BE is produced with a density parameter of $d_l = 1.0$–2.0 Å^{-2}. Therefore, for further calculations, we fixed the parameters d_s and d_l in the range of 5.0–6.0 and 1.0–2.0 Å^{-2}, respectively. The set 6.0 and 1.5 Å^{-2} is our *standard set* for the parameters d_s and d_l. Figure 4 also shows that the MBE results are in excellent agreement with the DelPhi results for h (or $d_s^{-1/2}$) ≤ 0.5 Å.

Table I shows the results of the calculation of the polarization free energy W_{polz} by the MBE method for a large molecule, the 17-residue peptide Ac-ETGTKAELLAKYEATHK-NMe in the α-
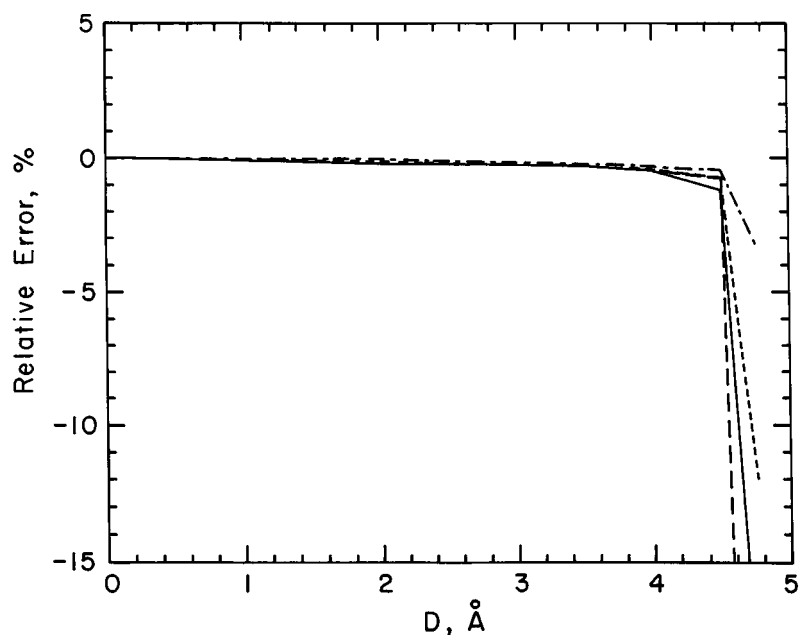
FIGURE 2. Relative error in the polarization free energy W_{polz}, eq. (35), of the point charge in the spherical cavity embedded in a polar solvent. The radius of the cavity is equal to 5 Å; D is the displacement of the point charge from the center of the cavity. The parameter $R_{Loc} = 3$ Å, and the densities d_s, d_l (in Å^{-2}) from the bottom to the top curve are (4, 1), (6, 1.5), (8, 2), (10, 2.5), respectively.
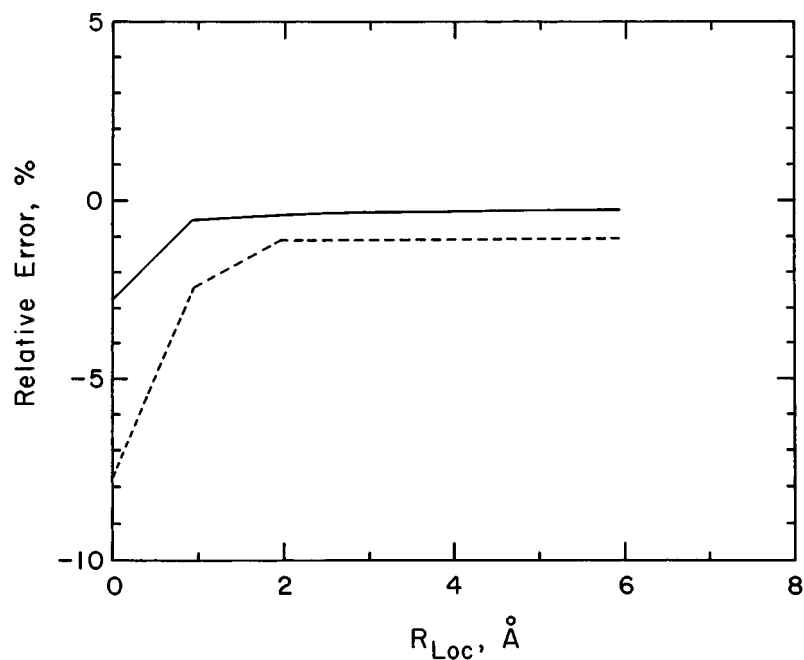

FIGURE 3. The same relative error as in Figure 2, versus the parameter R_{Loc}. The displacement D is equal to 4.5 and 4.75 Å for the top and bottom curves, respectively. The densities d_s, d_l are 6, 1.5, respectively.
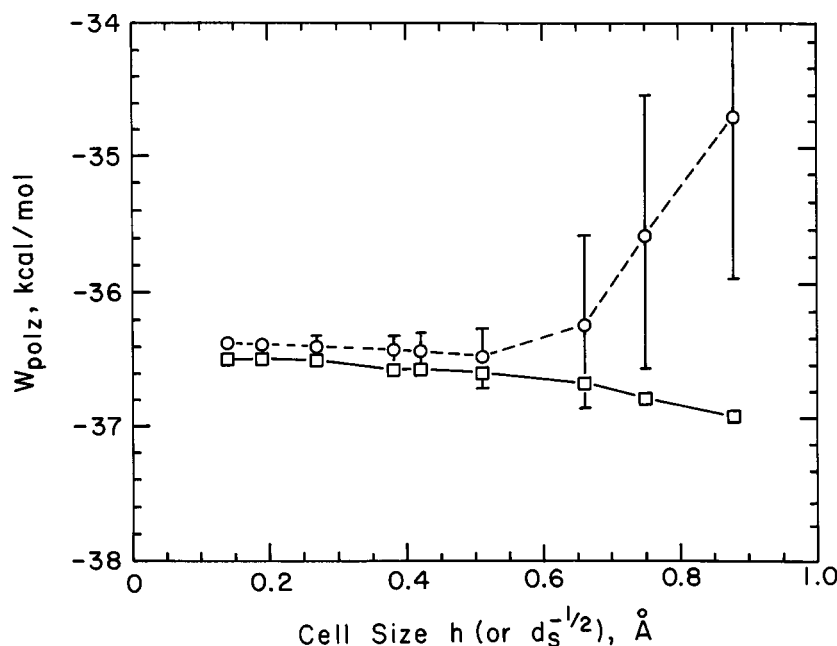
FIGURE 4. Polarization free energy W_{polz}, eq. (35), of ionized *n*-butylamine in the extended conformation in an aqueous solution, calculated by the (—) MBE method and by the (---) DelPhi method versus cell size *h* of the DelPhi method, or the size $d_s^{-1/2}$ of a *small* BE. The error symbols represent the standard deviations for the polarization free energy from the average value. The average polarization free energy is calculated over all rotations and translations of the molecule relative to the 3-D lattice in the DelPhi method.

helical conformation,[41,‡] with different sets of parameters R_{Loc} and R_{Int}. As can be seen, the MBE method demonstrates good stability as these parameters are varied. Reasonable values of the parameters R_{Loc} and R_{Int} from the point of view of accuracy and computing time are in the range of 4–5 and 8–10 Å, respectively. In these calculations, we considered the whole peptide group and each group of polar residues as the source of the molecular electrostatic field. Figure 5 shows a comparison of the polarization free energy W_{polz} of the 17-residue peptide Ac-ETGTKAELLAKYEATHK-NMe in the α-helical conformation, calculated by the MBE and DelPhi methods. As can be seen, the MBE values of W_{polz} are in good agreement (within 2.5%) with the DelPhi values of W_{polz}.

Table II shows free energies of desolvation Δg_i of the ionizable residues of the 17-residue peptide Ac-ETGTKAELLAKYEATHK-NMe in the α-helical conformation, when the residue is moved from

a solvent to a protein environment

$$\Delta g_i = \left[W_{polz}(PS_i^+) - W_{polz}(PS_i^0) \right]$$
$$- \left[W_{polz}(S_i^+) - W_{polz}(S_i^0) \right]. \quad (41)$$

TABLE I.
Polarization Free Energy, W_{polz}, Versus Parameters R_{Loc} and R_{Int} of MBE Method.

R_{Loc} (Å)	R_{Int} (Å)	W_{polz} (kcal/mol)	t (s)
3.0	6.0	−101.22	191
3.0	8.0	−101.58	218
5.0	8.0	−101.82	305
3.0	10.0	−101.28	280
5.0	10.0	−101.87	390
5.0	12.0	−101.89	496
DelPhi[a]		−99.51	148

The polarization free energy is calculated by eq. (35) for a 17-residue peptide Ac-ETGTKAELLAKYEATHK-NMe in the state of zero ionization for each residue in the α-helical conformation with the *standard set* of parameters d_s, d_l (see text). t is the CPU time for a calculation by serial code on one node of the IBM SP2 supercomputer.
[a] Calculation by the DelPhi method with lattice dimension $N_g = 129$, lattice size $h = 0.227$ Å.

‡ It is recognized that the α-helical form does not remain intact as the pH changes. However, this calculation was carried out only to demonstrate how to obtain ΔF_{inz} for a *given* conformation. Simulations of the *coupling* between ionization and conformational equilibria, which requires the combination of the Monte Carlo method for the total energy defined by eq. (1) with the ECEPP method for calculating a conformational energy in a vacuum, are presented elsewhere.[41]
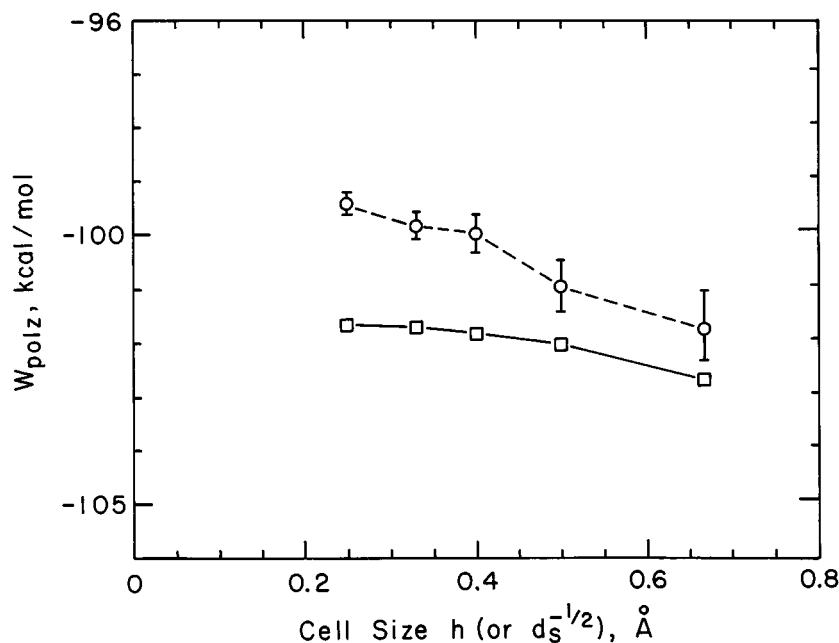
FIGURE 5. Polarization free energy W_{polz} of the 17-residue peptide Ac-ETGTKAELLAKYEATHK-NMe in the neutral form in the α-helical conformation in an aqueous solution. (—) The free energy calculated by the MBE method (see Table 1). (---) The average free energy of the DelPhi method versus the cell size. The error symbols represent the standard deviations as in Figure 4.

TABLE II.

Free Energies of Desolvation and PMFs for Ionized Groups of Peptide Ac-ETGTKAELLAKYEATHK-NMe.

	Δg_i	W_{ij}						
		K_5^+	E_7^-	K_{11}^+	Y_{12}^-	E_{13}^-	H_{16}^+	K_{17}^+
E_1^-	0.52 (0.34)	−1.07 (−1.12)	0.33	−0.23	0.25	0.27	−0.20	−0.19
K_5^+	0.13 (0.48)		−0.32 (−0.30)	0.24	−0.27	−0.29	0.22	0.20
E_7^-	0.35 (0.23)			−0.38 (−0.52)	0.31	0.35	−0.26	−0.25
K_{11}^+	−0.04 (0.13)				−0.35 (−0.27)	−0.33	0.28	0.26
Y_{12}^-	2.11 (1.86)					0.58 (0.49)	−0.93	−0.31
E_{13}^-	0.51 (0.54)						−0.40 (−0.41)	−0.40
H_{16}^+	0.96 (0.91)							0.48 (0.54)
K_{17}^+	0.00 (0.04)							

Values were calculated by the MBE and DelPhi methods and are in kcal/mol. Numbers in parentheses were calculated by the DelPhi method with lattice dimension $N_g = 129$ and lattice size $h = 0.227$ Å.
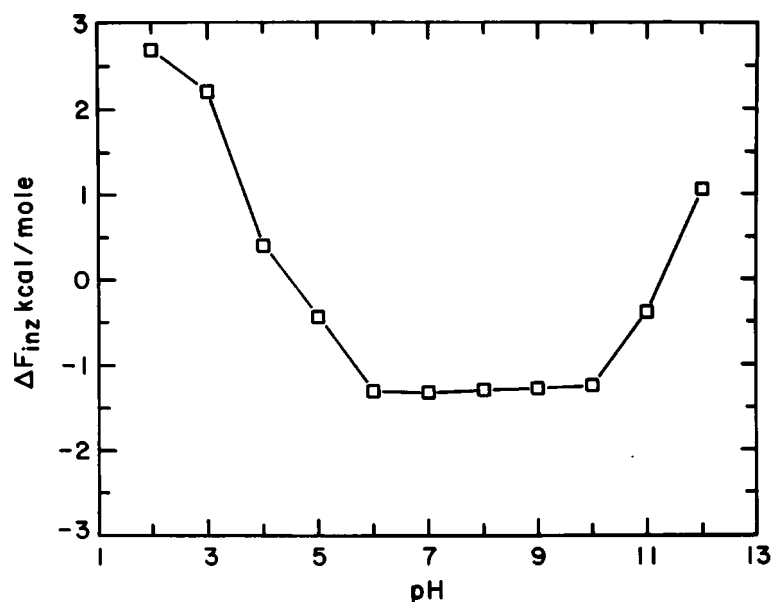
FIGURE 6. Excess free energy of ionization of the 17-residue peptide in a fixed α-helical conformation in an aqueous solution at 300 K versus pH.[‡]

As can be seen from Table II, Δg_i calculated by the MBE method is in reasonable agreement with the one computed by the DelPhi method. Calculations of the PMFs W_{ij} between charged ionizable residues of the peptide Ac-ETGTKAELLAKY-EATHK-NMe also show results in good agreement with those computed by the DelPhi method. It should be noted that these calculations of the de-solvation free energies Δg_i and the PMFs are very inexpensive with our MBE method; the results of all these calculations to obtain Δg_i and W_{ij} were obtained in one loop over charged groups for an additional CPU time of less than 120 s. Therefore, the complete calculation of the two last terms in eq. (1), i.e., the free energy F_{solv} of transfer from a vacuum to an aqueous solution and the free energy of ionization F_{inz} takes about 400 s of CPU time on one node of the IBM SP2 supercomputer.

Figure 6 shows the dependence of the excess free energy of ionization ΔF_{inz} of the peptide Ac-ETGTKAELLAKYEATHK-NMe in the *fixed* α-helical form[41, ‡] on the pH of the aqueous solution, where the free energy ΔF_{inz} is defined as

$$\Delta F_{inz} = F_{inz} - F_{inz}^0, \qquad (42)$$

where F_{inz} is the free energy of ionization of the peptide in a given conformation and F_{inz}^0 is the free energy of ionization of the peptide when all ionized groups are considered as *isolated* and *non-interacting*, i.e., all desolvation energies $\Delta g_i = 0$ and the PMFs $W_{ij} = 0$.

Table III shows the dependence of the perfor-mance CPU time for the MBE method on molecu-lar size. The ratio t/S_m decreases as the molecular surface S_m area, i.e., the number of BEs, increases; therefore, the dependence of the CPU time t on S_m is less than a linear regression.

Discussion

We designed a 3-level MBE method that ex-hibits good accuracy and performance for short peptide molecules. As shown, the MBE method calculates the free energy of transfer and the ion-ization free energy for a 17-residue peptide Ac-ETGTKAELLAKYEATHK-NMe with eight ioniz-able residues in 400 s of CPU time on the IBM SP2 supercomputer. The size of the matrix in eq. (22) does not exceed 2000 for this 17-residue peptide; therefore, the RAM demands do not exceed 120 MB and are about the same RAM size needed by DelPhi for $N_g = 129$. We also showed that the MBE method is stable to a variation in the size and numbers of MBEs, which are tesselated on the molecular surface. This accuracy and stability are provided by the intensive use of the different kinds of sum rules for matrix elements and the source term vector. The use of the sum rules of the MBE method is important for a calculation of abso-lute solvation free energies to treat the conforma-

TABLE III.
CPU Time for MBE Method.

Molecule	N_{at}	S_m (Å^2)	t	t/S_m
Lysozyme	1309	7399	51	0.7
17-res[a]	279	1420	13	1.1
Ala14[b]	152	759	10	1.5
Ala10[b]	112	561	9	1.6
Ala6[b]	72	391	7	1.8
Ala4[b]	52	299	6	2.0
Ala1[b]	22	210	4	3.9

CPU time in seconds to calculate the surface polarization charge for one center of the molecular electrostatic field, i.e., to solve eq. (9) for one particular term **b**.

[a] Seventeen-residue peptide Ac-ETGTKAELLAKY-EATHK-NMe in the neutral form in the α-helical conformation.

[b] α-helical conformation.

tional stability of a molecule in a solvent. Using the precalculated full matrix in eq. (22) and the PBCG method, we overcame the problem of bad convergence (or nonconvergence in some cases) of the iterative solution of the BE integral equation,[19,42] which in the fast multipole approximation BE method[27] is partially solved by iterations with underrelaxation at the price of a slow convergence rate.

Our proposed MBE method can be generalized, in an obvious way, beyond the current 3-level size of the BEs, to include adaptively increasing sets of BEs tesselated on the molecular surface. As estimated, such a generalized unrestricted MBE method has a linear, computational *complexity* with respect to molecular size and can be used for large macromolecules.

Acknowledgments

References

1. W. Kauzmann, *Adv. Protein Chem.*, **14**, 1 (1959).
2. C. Tanford, *Adv. Protein Chem.*, **24**, 1 (1970).
3. K. A. Dill, *Biochemistry*, **29**, 7133 (1990).
4. J. Applequist and P. Doty, In *Polyamino Acids, Polypeptides and Proteins*, M. A. Stahmann, Ed., Univ. of Wisconsin Press, Madison, WI, 1962, p. 161.
5. A.-S. Yang and B. Honig, *J. Mol. Biol.*, **231**, 459 (1993).
6. Y. N. Vorobjev, H. A. Scheraga, B. Hitz, and B. Honig, *J. Phys. Chem.*, **98**, 10940 (1994).
7. K. A. Sharp and B. Honig, *Annu. Rev. Biophys. Biophys. Chem.*, **19**, 301 (1990).
8. M. E. Davis and J. A. McCammon, *J. Comput. Chem.*, **11**, 401 (1990).
9. A. A. Rashin, *Prog. Biophys. Mol. Biol.*, **60**, 73 (1993).
10. D. Sitkoff, K. A. Sharp, and B. Honig, *J. Phys. Chem.*, **98**, 1978 (1994).
11. J. Warwicker, and H. C. Watson, *J. Mol. Biol.*, **157**, 671 (1982).
12. R. J. Zauhar and R. S. Morgan, *J. Comput. Chem.*, **9**, 171 (1988).
13. A. A. Rashin, *J. Phys. Chem.*, **94**, 1725 (1990).
14. A. H. Juffer, E. F. F. Botta, A. M. Bert, B. A. M. van Keulen, A. Van der Ploeg, and H. J. C. Berendsen, *J. Comp. Phys.*, **97**, 144 (1991).
15. K. Sharp, *J. Comput. Chem.*, **12**, 454 (1991).
16. A. Nicholls and B. Honig, *J. Comput. Chem.*, **12**, 435 (1991).
17. C. Niedermeier and K. Schulten, *Mol. Simul.*, **8**, 361 (1992).
18. M. K. Gilson, J. A. McCammon, and J. D. Madura, *J. Comput. Chem.*, **16**, 1081 (1995).
19. Y. N. Vorobjev, J. A. Grant, and H. A. Scheraga, *J. Am. Chem. Soc.*, **114**, 3189 (1992).
20. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, Cambridge, U.K., 1988.
21. A. Brandt, In *Multi-Grid Methods*, W. Hackbusch and U. Trottenberg, Eds., Springer–Verlag, New York, 1982, p. 220.
22. M. J. Holst and F. Saied, *J. Comput. Chem.*, **14**, 105 (1993).
23. H. Oberoi and N. M. Allewell, *Biophys. J.*, **65**, 48 (1993).
24. M. J. Holst and F. Saied, *J. Comput. Chem.*, **16**, 337 (1995).
25. R. E. Bruccoleri, *J. Comput. Chem.*, **14**, 1417 (1993).
26. K. C. Smith and B. Honig, *Proteins*, **18**, 119 (1994).
27. R. Bharadwaj, A. Windemuth, S. Sridharan, B. Honig, and A. Nicholls, *J. Comput. Chem.*, **16**, 898 (1995).
28. Y. N. Vorobjev and H. A. Scheraga, *J. Phys. Chem.*, **97**, 4855 (1993).
29. Y. N. Vorobjev, H. A. Scheraga, and B. Honig, *J. Phys. Chem.*, **99**, 7180 (1995).
30. M. L. Connolly, *J. Appl. Crystallogr.*, **16**, 548 (1983).
31. M. L. Connolly, *J. Am. Chem. Soc.*, **107**, 1118 (1985).
32. M. A. Jawson and G. T. Symm, *Integral Equation Methods in Potential Theory and Elastostatics*, Academic Press, New York, 1977.

33. A. A. Rashin and K. Namboodiri, *J. Phys. Chem.*, **91**, 6003 (1987).

34. J. A. Grant, R. L. Williams, and H. A. Scheraga, *Biopolymers*, **30**, 929 (1990).

35. E. O. Purisima and S. H. Nilar, *J. Comput. Chem.*, **16**, 681 (1995).

36. L. Greengard and V. Rokhlin, *J. Comp. Phys.*, **73**, 325 (1987).

37. G. Perrot, B. Cheng, K. D. Gibson, J. Vila, K. A. Palmer, A. Nayeem, B. Maigret, and H. A. Scheraga, *J. Comput. Chem.*, **13**, 1 (1992).

38. B. J. Yoon and A. M. Lenhoff, *J. Comput. Chem.*, **11**, 1080 (1990).

39. J. Antosiewicz, J. A. McCammon, and M. K. Gilson, *J. Mol. Biol.*, **238**, 415 (1994).

40. B. Honig, Department of Biochemistry of Columbia University, personal communication (1996).

41. D. R. Ripoll, Y. N. Vorobjev, A. Liwo, J. A. Vila, and H. A. Scheraga, *J. Mol. Biol.*, **264**, 770 (1996).

42. Y. N. Vorobjev, unpublished results.