Proceedings of the 2005 IEEE
Engineering in Medicine and Biology 27th Annual Conference
Shanghai, China, September 1-4, 2005

An Accurate Surface Formulation for Biomolecule Electrostatics in Non-Ionic Solutions

Michael D. Altman, Jaydeep P. Bardhan, Jacob K. White, Bruce Tidor

***Abstract*— The electrostatic interactions between biomolecules and solvent are generally difficult to model because there exist an enormous number of solvent degrees of freedom. Continuum electrostatic models provide an approximate method to analyze these interactions; these models are typically solved numerically in either differential or integral form. In this paper we demonstrate the importance of using an appropriate numerical technique, called qualocation, for a popular integral formulation of the electrostatics problem. Numerical results illustrate that qualocation exhibits superior accuracy relative to naive implementations. We also show that the integral formulation is extremely well-conditioned and converges rapidly when iterative methods are used to solve the discretized integral equation.**

I. INTRODUCTION

Electrostatic interactions within and between biomolecules are known to play important structural and functional roles [1], [2]. Analyzing these interactions computationally is challenging because solvent molecules surround the biomolecules of interest, so the physical problem has an enormous number of degrees of freedom. Monte Carlo and molecular dynamics methods [1], [3]–[7] treat all or most of the solvent molecules explicitly, but for many problems the computational expense is prohibitive.

Continuum models offer an alternative approach to studying biomolecule electrostatics [1], [8]–[11]. In these models, macroscopic laws of electrostatics are assumed to hold in the molecule interior and in the solvent, and the resulting systems of partial differential equations are solved numerically on a computer. Finite difference methods, finite element methods, and boundary element methods (BEM) have all been applied to the biomolecule electrostatics problem [12]–[17]. The boundary element method offers numerical advantages such as an improved representation of the biomolecule–solvent interface and exact treatment of discrete point charges. Here we study an integral formulation and boundary element technique for solving biomolecule electrostatics problems in which the solvent ionic strength is zero.

This work was supported by the National Institutes of Health (GM065418),by the Singapore-MIT Alliance, and by the National Science Foundation. J. P. Bardhan gratefully acknowledges support from a Department of Energy Computational Science Graduate Fellowship.

M. D. Altman is with the Department of Chemistry, Massachusetts Institute of Technology. (maltman@mit.edu)

J. P. Bardhan is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

J. K. White is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

B. Tidor is with the Biological Engineering Division and the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

The integral formulation, called the equivalent charge formulation (ECF), has been previously discussed in the literature [18], [19]. In this work we demonstrate that a numerical technique called qualocation [20] substantially improves accuracy when compared to naive implementations of the integral formulation. The qualocation method can be applied to many types of BEM problems in addition to the biomolecule problem discussed here.

The following section introduces the electrostatics model and the boundary element method used to solve the model numerically. Section 3 presents the ECF–qualocation method and Section 4 illustrates the method's performance with computational results. Section 5 summarizes the paper.

II. BACKGROUND

A. Mixed Discrete-Continuum Electrostatics Model

Figure 1 illustrates the mixed discrete–continuum electrostatics model. The boundary Ω separates the molecular interior from the solvent exterior; Ω is taken to be the Richards molecular surface [21], which is formed by rolling a probe sphere around the union of van der Waals–radius spheres located at the atom centers. We treat the molecular interior as a homogeneous medium with permittivity ε_I, in which the electrostatic potential obeys the Poisson equation

$$\nabla^2 \varphi_I(r) = -\sum_{i=1}^{n_c} \frac{q_i}{\varepsilon_I} \delta(r - r_i), \tag{1}$$

where n_c is the number of discrete point charges and r_i and q_i denote the location and value of the i^{th} charge. The solvent region is treated as a homogeneous medium with a much higher permittivity ε_{II}, and in this region the Laplace equation holds:

$$\nabla^2 \varphi_{II}(r) = 0. \tag{2}$$

At the dielectric boundary, the potential and normal component of the displacement field are continuous:

$$\varphi_I(r_\Omega) = \varphi_{II}(r_\Omega) \tag{3}$$

$$\varepsilon_I \frac{\partial \varphi_I}{\partial n}(r_\Omega) = \varepsilon_{II} \frac{\partial \varphi_{II}}{\partial n}(r_\Omega). \tag{4}$$

B. The Boundary Element Method

Consider the problem of computing the capacitance of a conducting sphere, whose surface is S, suspended in free space. By setting the potential on the sphere to unity and solving the first kind integral equation

$$\int_S \frac{\sigma(r')dA'}{4\pi\varepsilon_0 ||r - r'||} = \Psi(r), \tag{5}$$

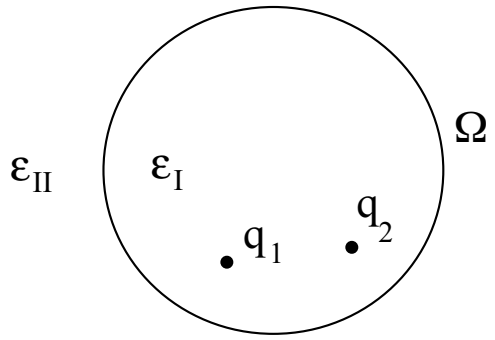

Fig. 1. Mixed discrete-continuum electrostatics model.

we can integrate $\sigma(r)$ over S to find the capacitance. To solve the problem numerically, we discretize the boundary surface into a set of n_p panels and represent the solution $\sigma(r)$ on the discretized surface as a weighted combination of compactly supported basis functions:

$$\sigma(r) = \sum_{i=1}^{n_p} y_i \chi_i(r). \tag{6}$$

Here, $\chi_i(r)$ denotes the i^{th} basis function and y_i the associated weight. In this paper, we use piecewise-constant basis functions such that each function takes value unity on a single panel and is zero everywhere else:

$$\chi_i(r) = \begin{cases} 1 & \text{if } r \text{ is on panel } i \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

In general, the span of the basis functions will not permit exact solution of the original integral equation. Instead, consider computing the basis function weights so as to reduce the residual $R(r)$, which is the difference between the known potential $\Psi(r)$ and the result of applying the integral operator to the approximate solution:

$$R(r) = \Psi(r) - \int_S G(r;r') \left(\sum_{i=1}^{n_p} y_i \chi_i(r') \right) dA'. \tag{8}$$

In the commonly used centroid collocation scheme, $R(r)$ is forced to be zero at the basis function centroids [22]. The resulting linear system is of the form $Ay = b$ with

$$A_{ij} = \int_S G(r_{c_i};r') \chi_j(r') dA' \tag{9}$$

$$b_i = \Psi(r_{c_i}), \tag{10}$$

where r_{c_i} is the centroid of panel i. Alternatively, Galerkin methods force the residual to be orthogonal to the basis functions $\{\chi_1, \chi_2, \ldots, \chi_{n_p}\}$. Galerkin methods produce linear systems of equations of the same $Ay = b$ form, though now the entries are

$$A_{ij} = \int_S \int_S \chi_i(r') G(r';r'') \chi_j(r'') dA' dA'' \tag{11}$$

$$b_i = \int_S \chi_i(r') \Psi(r') dA'. \tag{12}$$

For both the collocation and Galerkin methods, the linear systems can be solved using sparsification-accelerated iterative methods [23]–[25].

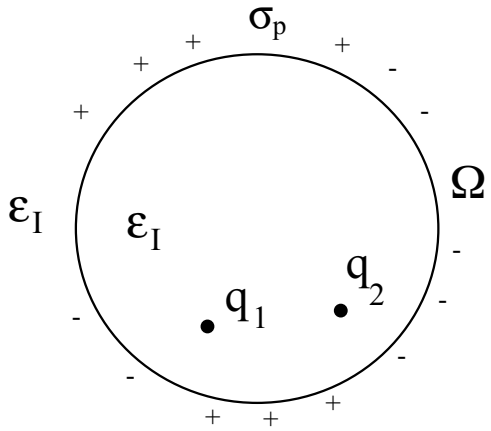

Fig. 2. Physical model of the equivalent charge formulation.

III. DERIVATION OF THE ECF–QUALOCATION METHOD

A. Integral Formulation

The essential idea of the equivalent charge formulation is to replace the original problem, which has two dielectric regions, with a simpler problem, shown in Figure 2, which is a Poisson problem with the same dielectric constant everywhere in space. In Figure 2, we have replaced the solvent dielectric ε_{II} with ε_I from the interior and introduced a fictitious layer of charge $\sigma_p(r)$ on the surface. The variables $\hat{\varphi}_I$ and $\hat{\varphi}_{II}$ denote the potential in the modified problem. Finding a surface charge layer $\sigma_p(r)$ such that the original boundary conditions (3,4) hold ensures that the solution of the homogeneous dielectric problem is equivalent to that of the original multiple dielectric region problem.

Because the dielectric constant is homogeneous throughout space in the equivalent problem, we can write the potential as

$$\hat{\varphi}(r) = \sum_{i=1}^{n_c} \frac{q_i}{4\pi\varepsilon_I ||r - r_i||} + \int_\Omega \frac{\sigma_p(r') dA'}{4\pi\varepsilon_I ||r - r'||}. \tag{13}$$

The normal component of the electric field at a point r on the surface is therefore

$$\frac{\partial \hat{\varphi}}{\partial n}(r) = \frac{\partial}{\partial n(r)} \sum_{i=1}^{n_c} \frac{q_i}{4\pi\varepsilon_I ||r - r_i||} + \frac{\partial}{\partial n(r)} \int_\Omega \frac{\sigma_p(r') dA'}{4\pi\varepsilon_I ||r - r'||}, \tag{14}$$

and the discontinuity in the integral term implies that a side of the surface must be specified. In the homogeneous dielectric problem, the charge density determines the discontinuity of the normal component of the electric field by the relation [26]:

$$\frac{\partial \hat{\varphi}_{II}}{\partial n}(r) - \frac{\partial \hat{\varphi}_I}{\partial n}(r) = \sigma_p(r)/\varepsilon_I. \tag{15}$$

Combining (15), (4), and (14) gives

$$\frac{\varepsilon_I+\varepsilon_{II}}{2\varepsilon_I(\varepsilon_I-\varepsilon_{II})}\sigma_p(r)+\int\!\!\!\!\!\!\diagup_{\Omega}\frac{\partial}{\partial n(r)}\frac{\sigma_p(r')dA'}{4\pi\varepsilon_I||r-r'||} = -\frac{\partial}{\partial n(r)}\sum_{i=1}^{n_c}\frac{q_i}{4\pi\varepsilon_I||r-r_i||}, \quad (16)$$

which is known as the equivalent charge formulation [18], [22]; the integral over Ω is taken to be the principal value integral.

B. Qualocation Method

We now motivate the qualocation approach as it was described by Tausch *et al.* [20], and present both collocation and qualocation as simplifications of the Galerkin method. To solve (16) numerically via the Galerkin method, we discretize the molecular surface into n_p flat triangles and represent the surface charge $\sigma_p(r)$ as a weighted combination of piecewise constant basis functions. We then define a residual $R(r)$ similar to (8) and enforce $\int R(r)\chi_i(r)dA = 0$ for each basis function $\chi_i(r)$. This produces a set of equations of the form:

$$\int_{\text{panel } i}\frac{(\varepsilon_I+\varepsilon_{II})y_i dA}{2\varepsilon_I(\varepsilon_I-\varepsilon_{II})}+\int_{\text{panel } i}\int\!\!\!\!\!\!\diagup_{\text{panel } j}\frac{\partial}{\partial n(r)}\frac{y_j dA' dA}{4\pi\varepsilon_I||r-r'||} = -\int_{\text{panel } i}\frac{\partial}{\partial n(r)}\sum_k\frac{q_k dA}{4\pi\varepsilon_I||r-r_k||}, \quad (17)$$

where again y_i is the weight associated with the i^{th} basis function.

The centroid collocation method simplifies the Galerkin method by replacing each integral over panel i with a midpoint quadrature rule; the inner integral of the double integral is then evaluated analytically [27], [28]. However, the integrand of the outer integral is non-smooth for nearby panels because the normal $n(r)$ on panel i has a component in the plane of panel j. As a result, midpoint quadrature and the resulting collocation scheme are inaccurate.

In contrast, the qualocation method replaces the inner integral, which is smooth, with a midpoint quadrature rule. The resulting system has entries

$$A_{ii} = \frac{\varepsilon_I+\varepsilon_{II}}{2\varepsilon_I(\varepsilon_I-\varepsilon_{II})}\alpha_i \quad (18)$$

$$A_{ij} = \int_{\text{panel } i}\frac{\partial}{\partial n(r)}\frac{\alpha_j\, dA}{4\pi\varepsilon_I||r-r_{c_j}||} \quad (19)$$

$$b_i = -\int_{\text{panel } i}\sum_k\frac{\partial}{\partial n(r)}\frac{q_k\, dA}{4\pi\varepsilon_I||r-r_k||}, \quad (20)$$

where α_i is the area of panel i. Using qualocation, the outer, non-smooth integral can be evaluated analytically and the smooth inner integral is approximated accurately.

IV. RESULTS

We have implemented the ECF–qualocation formulation using the FFTSVD fast BEM algorithm [29] to rapidly apply the dense discretized integral operator. The method relies on the observation that the qualocation operator is the scaled transpose of the double layer potential operator [20]. We compare the ECF–qualocation method to ECF–collocation as well as to a more complex formulation derived from Green's theorem [14], [30]. In contrast to the ECF formulation, which has one variable per panel and one integral operator, the Green's theorem formulation has two surface variables per panel and requires two integral operators.

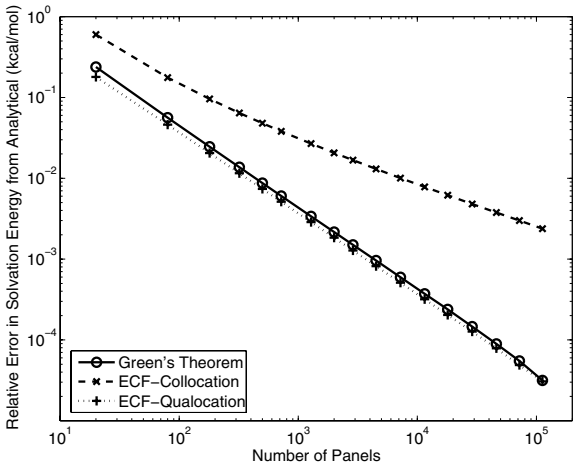

Fig. 3. Improvement in accuracy with increasing panel discretization when computing the solvation energy of a 1 Å radius sphere with a central $+1e$ charge. Results for the Green's theorem, ECF–collocation, and ECF–qualocation formulations are shown.

A. Sphere

To test the accuracy of the ECF–qualocation method, we computed the electrostatic component of the solvation free energy for a sphere with a 1 Å radius and a central $+1e$ charge. We compare the numerical results with the analytical answer as the surface discretization is refined. Figure 3 is a plot of the results computed using collocation and qualocation methods as well as those from a Green's theorem formulation [14], [30]. The qualocation method is clearly superior in accuracy to the collocation method; surprisingly, qualocation returns a slightly more accurate answer than the Green's theorem method, which has twice as many degrees of freedom.

B. Barnase–Barstar Protein Complex

We also computed the electrostatic component of the solvation free energy for the barnase–barstar protein complex (1BRS in the Protein Data Bank) [31]. Figure 4 is a convergence plot that compares the ECF–qualocation result, the ECF–collocation result, and the Green's theorem result as the surface discretization is refined.

C. Iterative Method Convergence

It is well known [20] that second-kind integral operators such as the ECF formulation (16) are well-conditioned. The discretized linear systems have tightly clustered spectra, which leads to rapid convergence when Krylov iterative methods are used instead of Gaussian elimination. The Green's theorem formulation [14], [30] is instead a mixed first-second kind equation; its poorer conditioning necessitates the development of effective preconditioners [25], [30].

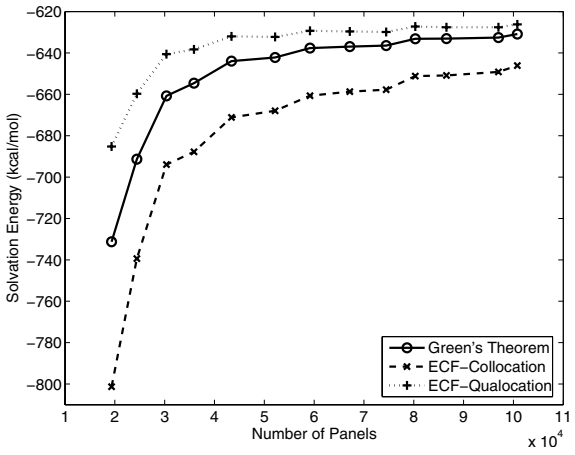

Fig. 4. Computed electrostatic components of the solvation free energy of the barnase–barstar protein complex with increasing panel discretization for the Green's theorem, ECF–collocation, and ECF–qualocation formulations.

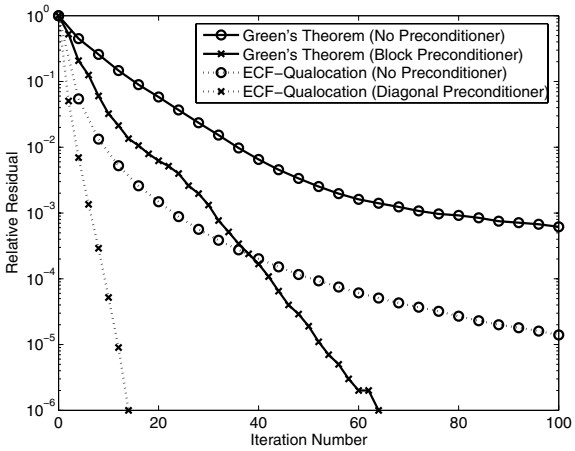

Fig. 5. Reduction in relative residual with iteration count for a 74,466 panel discretization of the barnase–barstar complex. Results are shown for the Green's theorem formulation, with and without block diagonal preconditioning, as well as for the ECF–qualocation formulation, with and without diagonal preconditioning.

To illustrate the advantageous conditioning, we have solved the barnase–barstar problem using the ECF–qualocation method using both no preconditioner and a diagonal preconditioner, and the Green's theorem formulation with no preconditioner as well as with the block diagonal preconditioner presented by Kuo *et al.* [30]. In Figure 5 we plot the relative GMRES residuals as a function of iteration count.

V. SUMMARY

We have presented a numerical technique for calculating the electrostatic component of the solvation free energy of biomolecules for solutions with zero ionic strength. The technique is based on the equivalent charge formulation [19], [22] of the electrostatics problem. Our technique differs from earlier presentations because we form a linear system of equations using qualocation [20] rather than centroid collocation or Galerkin methods. We have demonstrated that the qualocation approach exhibits superior accuracy, and that Krylov iterative methods converge rapidly for ECF–qualocation problems because the second-kind integral formulation is extremely well-conditioned. It is non-trivial to extend the ECF formulation to treat problems in which the solvent ionic strength is non-zero.

REFERENCES

[1] K. A. Sharp and B. Honig, "Electrostatic interactions in macromolecules: Theory and applications," *Annual Review of Biophysics and Biophysical Chemistry*, vol. 19, pp. 301–332, 1990.

[2] F. B. Sheinerman, R. Norel, and B. Honig, "Electrostatic aspects of protein-protein interactions," *Current Opinions in Structural Biology*, vol. 11, pp. 153–159, 2000.

[3] V. Lounnas, B. M. Pettitt, L. Findsen, and S. Subramaniam, "A microscopic view of protein solvation," *Journal of Physical Chemistry*, vol. 18, pp. 7157–7159, 1992.

[4] J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids*. Cambridge: Cambridge University Press, 1987.

[5] C. L. Brooks, III, M. Karplus, and B. M. Pettitt, "Proteins: A theoretical perspective of dynamics, structure and thermodynamics," *Adv. Chem. Phys.*, vol. 71, pp. 1–249, 1988.

[6] A. Jean-Charles, A. Nicholls, K. Sharp, B. Honig, A. Tempczyk, T. F. Hendrickson, and W. C. Still, "Electrostatic contributions to solvation energies: Comparison of free energy perturbation and continuum calculations," *Journal of the American Chemical Society*, vol. 113, pp. 1454–1455, 1991.

[7] S. W. Rick and B. J. Berne, "The aqueous solvation of water: A comparison of continuum methods with molecular dynamics," *Journal of the American Chemical Society*, vol. 116, pp. 3949–3954, 1994.

[8] C. Tanford and J. G. Kirkwood, "Theory of protein titration curves I. general equations for impenetrable spheres," *Journal of the American Chemical Society*, vol. 59, pp. 5333–5339, 1957.

[9] B. Honig and A. Nicholls, "Classical electrostatics in biology and chemistry," *Science (Washington, D.C.)*, vol. 268, pp. 1144–1149, 1995.

[10] J. Warwicker and H. C. Watson, "Calculation of the electric potential in the active site cleft due to alpha-helix dipoles," *Journal of Molecular Biology*, vol. 157, pp. 671–679, 1982.

[11] J. D. Madura, J. M. Briggs, R. C. Wade, M. E. Davis, B. A. Luty, A. Ilin, J. Antosiewicz, M. K. Gilson, B. Bagheri, L. Ridgway-Scott, and J. A. McCammon, "Electrostatics and diffusion of molecules in solution: Simulations with the University of Houston Brownian Dynamics program," *Computer Physics Communications*, vol. 91, pp. 57–95, 1995.

[12] M. K. Gilson, A. Rashin, R. Fine, and B. Honig, "On the calculation of electrostatic interactions in proteins," *Journal of Molecular Biology*, vol. 183, pp. 503–516, 1985.

[13] I. Klapper, R. Hagstrom, R. Fine, K. Sharp, and B. Honig, "Focusing of electric fields in the active site of Cu-Zn superoxide dismutase: Effects of ionic strength and amino-acid modification," *Proteins: Structure, Function, Genetics*, vol. 1, pp. 47–59, 1986.

[14] B. J. Yoon and A. M. Lenhoff, "A boundary element method for molecular electrostatics with electrolyte effects," *Journal of Computational Chemistry*, vol. 11, no. 9, pp. 1080–1086, 1990.

[15] A. H. Juffer, E. F. F. Botta, B. A. M. van Keulen, A. van der Ploeg, and H. J. C. Berendsen, "The electric potential of a macromolecule in a solvent: A fundamental approach," *Journal of Computational Physics*, vol. 97, pp. 144–171, 1991.

[16] W. Rocchia, E. Alexov, and B. Honig, "Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions," *Journal of Physical Chemistry B*, vol. 105, pp. 6507–6514, 2001.

[17] N. A. Baker, D. Sept, M. J. Holst, and J. A. McCammon, "Electrostatics of nanosystems: Application to microtubules and the ribosome," *Proceedings of the National Academy of Sciences, USA*, vol. 98, pp. 10 037–10 041, 2001.

[18] S. M. Rao, T. K. Sarkar, and R. F. Harrington, "The electrostatic field of conducting bodies in multiple dielectric media," *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, no. 11, pp. 1441–1448, 1984.

[19] Y. N. Vorobjev and H. A. Scheraga, "Interaction of a biomolecule with ions in aqueous solution: Comparison of three fast approximate methods with the direct solution of the nonlinear Poisson–Boltzmann equation," *Journal of Physical Chemistry*, vol. 97, pp. 4855–4864, 1993.

[20] J. Tausch, J. Wang, and J. White, "Improved integral formulations for fast 3-D method-of-moment solvers," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 20, no. 12, pp. 1398–1405, 2001.

[21] F. M. Richards, "Areas, volumes, packing, and protein structure," *Annual Review of Biophysics and Bioengineering*, vol. 6, pp. 151–176, 1977.

[22] K. Nabors, "Efficient three-dimensional capacitance extraction," Ph.D. dissertation, Massachusetts Institute of Technology, 1993.

[23] V. Rokhlin, "Rapid solution of integral equation of classical potential theory," *Journal of Computational Physics*, vol. 60, pp. 187–207, 1985.

[24] Y. Saad and M. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM Journal of Scientific and Statistical Computing*, vol. 7, pp. 856–869, 1986.

[25] K. Nabors, F. T. Korsmeyer, F. T. Leighton, and J. White, "Preconditioned, adaptive, multipole-accelerated iterative methods for three-dimensional first-kind integral equations of potential theory," *SIAM Journal on Scientific Computing*, vol. 15, no. 3, pp. 713–735, 1994.

[26] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. Wiley, 1998.

[27] J. L. Hess and A. M. O. Smith, "Calculation of non-lifting potential flow about arbitrary three-dimensional bodies," *Journal of Ship Research*, vol. 8, no. 2, pp. 22–44, 1962.

[28] J. N. Newman, "Distribution of sources and normal dipoles over a quadrilateral panel," *Journal of Engineering Mathematics*, vol. 20, pp. 113–126, 1986.

[29] M. D. Altman, J. P. Bardhan, B. Tidor, and J. K. White, "FFTSVD: A fast multiscale boundary element method solver suitable for bioMEMS and biomolecule simulation," *(to appear in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems)*, 2006.

[30] S. S. Kuo, M. D. Altman, J. P. Bardhan, B. Tidor, and J. K. White, "Fast methods for simulation of biomolecule electrostatics," *International Conference on Computer Aided Design (ICCAD)*, 2002.

[31] A. M. Buckle, G. Schreiber, and A. R. Fersht, "Protein–protein recognition: crystal structural analysis of a barnase–barstar complex at 2.0-å resolution," *Biochemistry*, vol. 33, pp. 8878–8889, 1994.