Fast Methods for Simulation of Biomolecule Electrostatics*

Shihhsien S. Kuo† skuo@mit.edu
Michael D. Altman‡ maltman@mit.edu
Jaydeep P. Bardhan† jbardhan@mit.edu
Bruce Tidor§,† tidor@mit.edu
Jacob K. White† white@mit.edu
Massachusetts Institute of Technology
† Department of Electrical Engineering and Computer Science
‡ Department of Chemistry
§ Biological Engineering Division
77 Massachusetts Avenue, Cambridge, MA 02139

ABSTRACT

Computer simulation is an important tool for improving our understanding of biomolecule electrostatics, in part to aid in drug design. However, the numerical techniques used in these simulation tools do not exploit fast solver approaches widely used in analyzing integrated circuit interconnects. In this paper we describe one popular formulation used to analyze biomolecule electrostatics, present an integral formulation of the problem, and apply the precorrected-FFT method to accelerate the solution of the integral equations.

1. INTRODUCTION

Biomolecular structure and interactions in an aqueous environment are determined by a complicated interplay between physical and chemical forces including solvation, electrostatics, van der Waals forces, the hydrophobic effect, and covalent bonding. Electrostatic forces have received a great deal of study due to their long-range nature and the tradeoff between desolvation and interaction effects [1, 2, 3, 4]. In addition, electrostatic interactions play a significant role within a biomolecule as well as between biomolecules, making the balance between the two vital to the understanding of macromolecular systems. As a result, much effort has been devoted to accurate modeling and simulation of biomolecule electrostatics. One important application of this work is to compute the strength of electrostatic interactions for a biomolecule in an electrolyte solution, as well as the potential that the molecule generates in space. There are two valuable uses for these simulations. First, it provides a full picture of the electrostatic energetics of a biomolecular system, improving our understanding of how electrostatics contribute to stability, function, and molecular interactions [5]. Second, these simulations serve as a tool for molecular design, since electrostatic complementarity is an important feature of interacting molecules [6]. Through examination of the electrostatics and potential field generated by a protein molecule, for example, it may be possible to suggest improvements to other proteins or drug molecules that interact with it, or perhaps even design new interacting molecules *de novo* [7, 8, 9].

There are two approaches to simulating a protein macromolecule in an aqueous solution with nonzero ionic strength. Discrete, atomistic approaches based on Monto-Carlo or molecular dynamics simulations treat the macromolecule and solvent explicitly at the atomic level [2, 10, 11, 12, 13, 14]. An enormous number of solvent molecules are often required to provide reasonable accuracy, particularly when the electric fields of interest are far away from the macroscopic surface. In addition, free ions within the solvent are difficult to model with this approach. In this paper, we adopt instead a mixed discrete-continuum approach based on combining a continuum description of the macromolecule and solvent with a discrete description of the atomic charges [2, 15, 16, 17, 18].

Solutions to the mixed discrete-continuum model are mostly computed numerically, using schemes based on finite-difference discretizations of the model's underlying partial differential equations [1, 19, 20, 21, 22, 23]. In this paper we demonstrate that a more efficient procedure can be developed by combining a carefully chosen integral formulation of the mixed discrete-continuum model with one of the recently developed fast integral equation solvers [24, 25, 26]. In the following section we will briefly describe a widely used approximate model of biomolecule electrostatics, and then in Section Three we will derive a coupled integral formulation for the problem [27]. The numerical schemes used for computing solutions to the coupled integral formulation will be described in Section Four, and in particular we will give a brief overview of the precorrected-FFT (pFFT) accelerated method [28]. Computational results are given in Section Five, and these results will demonstrate both the strengths and the weaknesses of our current implementation. Conclusions are presented in Section Six.

2. MIXED DISCRETE-CONTINUUM FORMULATION

One commonly used simplified model for biomolecule electro-

*This work was supported by the Singapore-MIT Alliance, the National Science Foundation and the Defense Advanced Research Projects Agency.

0-7803-7607-2/02/$17.00 ©2002 IEEE

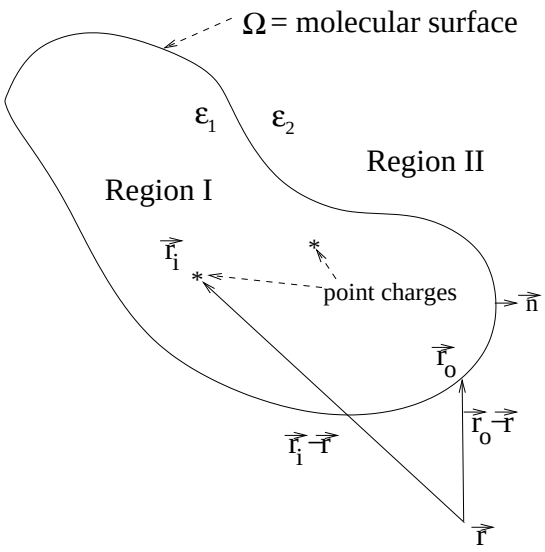

Figure 1: The continuum model of a solvated protein.

statics was introduced by Tanford and Kirkwood in 1957 [15]. In this model the interior of a protein molecule is approximated as a collection of point charges in a uniform dielectric material, where the dielectric constant is typically two to four times larger than the permittivity of free space. Any surrounding solvent is modeled as a much higher permittivity electrolyte whose behavior is described by the Debye-Hückel theory. The interface between the protein and the solvent is defined by determining how close the solvent molecules can approach the biomolecule [29, 30].

The Tanford and Kirkwood model for a single protein in a solvent is depicted in Figure 1, where Region I corresponds to the interior of the protein and Region II corresponds to the surrounding solvent. The electrostatic behavior in Region I, the protein interior, is governed by a Poisson equation

$$\nabla^2\varphi_1(\vec{r}) = -\sum_{i=1}^{n_c}\frac{q_i}{\varepsilon_1}\delta(\vec{r}-\vec{r}_i) \quad \text{(Region I)} \tag{1}$$

where φ_1 is the electrostatic potential, $\vec{r}$ is an evaluation position, $\vec{r}_i$ is the location of the i^{th} protein point charge, q_i is the point charge strength, n_c is the number of point charges, and ε_1 is the dielectric constant in the protein interior. Note also that δ is the standard Dirac-Delta function.

To determine the electrostatic potential in the solvent, Debye-Hückel theory suggests that the electrostatic potential should satisfy a nonlinear Poisson-Boltzmann equation, but the nonlinearity generates an unnecessarily complicated model. Instead, the simpler linearized Poisson-Boltzmann equation, which is also a Helmholtz equation, is more commonly used, and has been tested extensively and shown to accurately predict biomolecular properties under a variety of conditions. Therefore, the electrostatic potential in the solvent, Region II of Figure 1, is presumed to satisfy the Helmholtz equation

$$\nabla^2\varphi_2(\vec{r}) - \kappa^2\varphi_2(\vec{r}) = 0 \quad \text{(Region II)} \tag{2}$$

where κ is the inverse Debye screening length.

A wide variety of numerical techniques can be used to compute solutions to the combination of (1) and (2). For the biomolecule application, the most commonly used approach is based on the finite-difference method for discretizing partial differential equations, with researchers frequently making use of the DelPhi software package [1, 19, 20, 21, 22, 23]. Although finite-difference methods have proven to be effective, there are several characteristics of the biomolecule application which are problematic for such methods. Inaccuracies can be generated when projecting the discrete charges, which appear in (1), on to finite-difference grids. The problem is particularly troublesome when attempting to compute reaction forces at those point charge locations [31]. In addition, the large jump in dielectric constant across the irregularly-shaped protein-solvent boundary must be treated carefully. Finally, the solvent region is unbounded, at least formally, and must be somehow truncated before applying a finite-difference method. Modifications of the basic finite-difference method have been developed to resolve many of these difficulties [19, 20, 23, 32, 33], though often at considerable computational cost.

3. INTEGRAL EQUATION FORMULATION

As this section will make clear, numerical methods based on solving an integral formulation of (1) and (2) can treat point charges, irregularly shaped regions with large jumps in parameters, unbounded domains, and the reaction force computation much more naturally than finite-difference methods. For this reason, a number of researchers have developed integral formulations [27, 34, 35, 36, 37, 38], but most efforts have only addressed systems with zero ionic strength ($\kappa = 0$ in (2)). In Juffer et al. [35], an integral formulation was presented which allows for a general κ, but the formulation uses integrals with hypersingular kernels, and those integrals are challenging to evaluate accurately. In this work we followed the approach of Yoon and Lenhoff [27], as their approach allows for a general κ and avoids hypersingular kernels.

Even though integral formulations have many advantages for this application, they are not often used; the available numerical techniques for solving integral equations were too computationally expensive to use on complicated problems, but recently developed fast algorithms have changed that situation considerably. In this section we will describe an integral formulation for (1) and (2), and in the next section we will describe a fast numerical technique for computing the integral equation solutions.

To begin the formulation derivation, first consider that the well-known fundamental solutions to (1) and (2) are, respectively,

$$G_1(\vec{r};\vec{r}') = \frac{1}{4\pi|\vec{r}-\vec{r}'|} \tag{3}$$

$$G_2(\vec{r};\vec{r}') = \frac{e^{-\kappa|\vec{r}-\vec{r}'|}}{4\pi|\vec{r}-\vec{r}'|}. \tag{4}$$

The two fundamental solutions can be combined with Green's second theorem to generate an integral equations for the potential and its normal derivative. In particular, the integral equation for Region I is

$$\varphi_1(\vec{r}) = \int_\Omega \left[G_1(\vec{r};\vec{r}')\frac{\partial\varphi_1}{\partial n}(\vec{r}') - \varphi_1(\vec{r}')\frac{\partial G_1}{\partial n}(\vec{r};\vec{r}') \right] d\vec{r}' + \sum_{i=1}^{n_c}\frac{q_i}{\varepsilon_1}G_1(\vec{r};\vec{r}_i), \tag{5}$$

and the equation for Region II is

$$\varphi_2(\vec{r}) = \int_\Omega \left[-G_2(\vec{r};\vec{r}')\frac{\partial\varphi_2}{\partial n}(\vec{r}') + \varphi_2(\vec{r}')\frac{\partial G_2}{\partial n}(\vec{r};\vec{r}') \right] d\vec{r}', \tag{6}$$

where $\vec{n}$ is the outward pointing normal as shown in Figure 1, and the domain of integration for the integrals, Ω, is the boundary surface separating the low permittivity protein interior from the high permittivity solvent.

The potentials φ_1 and φ_2 must satisfy a pair of matching conditions on the boundary surface Ω. In particular, the electric potential is continuous and the normal derivative of the potential jumps by an amount related to the ratio of the dielectric constants,

$$\varphi_1(\vec{r}_o) = \varphi_2(\vec{r}_o) \quad (7)$$

$$\frac{\partial\varphi_1}{\partial n}(\vec{r}_o) = \varepsilon\frac{\partial\varphi_2}{\partial n}(\vec{r}_o), \quad (8)$$

where $\vec{r}_o \in \Omega$, and $\varepsilon = \varepsilon_2/\varepsilon_1$ is the relative dielectric constant of the two regions. To enforce these matching boundary conditions, take the limit of equation (5) as $\vec{r} \to \Omega$ from the inside, and use the limit of equation (6) as $\vec{r} \to \Omega$ from the outside. In this limit, G_1, G_2, $\frac{\partial G_1}{\partial n}$, and $\frac{\partial G_2}{\partial n}$ are kernels with integrable singularities, so care must be taken in carrying out the integrations. Note that that the potential due to a monopole layer is continuous across the layer, while the potential due to a dipole layer is discontinuous across the layer [39].

The results generated by applying the limiting processes to (5) and (6) yields

$$\varphi_1(\vec{r}_o) = \lim_{\vec{r}\to\vec{r}_o}\varphi_1(\vec{r}) \quad (9)$$

$$= \int_\Omega \left[G_1(\vec{r}_o;\vec{r}')\frac{\partial\varphi_1}{\partial n}(\vec{r}') - \varphi_1(\vec{r}')\frac{\partial G_1}{\partial n}(\vec{r}_o;\vec{r}')\right] d\vec{r}' + \frac{1}{2}\varphi_1(\vec{r}_o) + \sum_{i=1}^{n_c}\frac{q_i}{\varepsilon_1}G_1(\vec{r}_o;\vec{r}_i) \quad (10)$$

and

$$\varphi_2(\vec{r}_o) = \lim_{\vec{r}\to\vec{r}_o}\varphi_2(\vec{r}) \quad (11)$$

$$= \int_\Omega \left[-G_2(\vec{r}_o;\vec{r}')\frac{\partial\varphi_2}{\partial n}(\vec{r}') + \varphi_2(\vec{r}')\frac{\partial G_2}{\partial n}(\vec{r}_o;\vec{r}')\right] d\vec{r}' + \frac{1}{2}\varphi_2(\vec{r}_o) \quad (12)$$

where $\vec{r}_o$ is the position vector of some point on the boundary Ω and the integrals are taken to be principal value integrals.

Substituting equations (7) and (8) into (10) and (12) yields a coupled pair of integral equations for φ_1 and $\frac{\partial\varphi_1}{\partial n}$ on the dielectric interface,

$$\frac{1}{2}\varphi_1(\vec{r}_o) + \int_\Omega \left[\varphi_1(\vec{r}')\frac{\partial G_1}{\partial n}(\vec{r}_o;\vec{r}') - G_1(\vec{r}_o;\vec{r}')\frac{\partial\varphi_1}{\partial n}(\vec{r}')\right] d\vec{r}' = \sum_{i=1}^{n_c}\frac{q_i}{\varepsilon_1}G_1(\vec{r}_o;\vec{r}_i), \quad (13)$$

and

$$\frac{1}{2}\varphi_1(\vec{r}_o) + \int_\Omega \left[-\varphi_1(\vec{r}')\frac{\partial G_2}{\partial n}(\vec{r}_o;\vec{r}') + G_2(\vec{r}_o;\vec{r}')\frac{1}{\varepsilon}\frac{\partial\varphi_1}{\partial n}(\vec{r}')\right] d\vec{r}' = 0. \quad (14)$$

Equations (13) and (14) can be used to compute φ_1 and $\frac{\partial\varphi_1}{\partial n}$ on Ω. Then those surface potentials and their normal derivatives can be used in (5), (6), (7), and (8) to compute the potentials anywhere. Therefore, to compute the reaction potentials at the charge locations, which are needed to determine energy changes, one need only evaluate

$$\varphi^{REAC}(\vec{r}_i) = \int_\Omega \left[G_1(\vec{r}_i;\vec{r}')\frac{\partial\varphi_1}{\partial n}(\vec{r}') - \varphi_1(\vec{r}')\frac{\partial G_1}{\partial n}(\vec{r}_i;\vec{r}')\right] d\vec{r}'. \quad (15)$$

4. NUMERICAL SOLUTION

4.1 Discretization Method

A standard piecewise-constant centroid-collocation scheme is used to discretize (13) and (14) [40]. In the piecewise constant collocation method, the surface is first discretized into a set of panels, and a piecewise constant basis function, B_k, is associated with each panel. Then, the potentials are represented as a weighted combination of the panel basis functions. That is,

$$\varphi_1(\vec{r}_o) \approx \sum_k a_k B_k(\vec{r}_o) \quad (16)$$

$$\frac{\partial\varphi_1}{\partial n}(\vec{r}_o) \approx \sum_k b_k B_k(\vec{r}_o) \quad (17)$$

where k is the panel index, and a_k and b_k are weights of individual basis functions.

The basis function weights are determined by insisting that when (16) and (17) are substituted for the potential and its normal derivative in (13) and (14), the resulting equations are exactly satisfied for those values of $\vec{r}_o$ which correspond to panel centroids. The resulting system of equations can be denoted as a matrix of the form

$$\begin{bmatrix} \frac{1}{2}I + \int_{panel_k}\frac{\partial G_1}{\partial n}d\vec{r} & -\int_{panel_k}G_1 d\vec{r} \\ \frac{1}{2}I - \int_{panel_k}\frac{\partial G_2}{\partial n}d\vec{r} & 1/\varepsilon\int_{panel_k}G_2 d\vec{r} \end{bmatrix} \begin{bmatrix} a_k \\ b_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n_c}\frac{q_i}{\varepsilon_1}G_1 \\ 0 \end{bmatrix} \quad (18)$$

where n_c is the total number of charges inside the protein and $\int_{panel_k}$ corresponds to an integration over the kth panel surface. Note that the matrix is only a function of protein geometry and is independent of charge locations, making it possible to construct the matrix operator once and use it repeatedly to solve for different charge configurations.

4.2 Precorrected-FFT Method

Although the matrix equation in (18) can be readily solved with Gaussian elimination, and is used for the smaller test cases to demonstrate the validity of this formulation and to examine convergence properties, Gaussian elimination is too computationally expensive to solve practical examples of interest. An alternative approach to Gaussian elimination is to use an iterative solver such as GMRES [41], and recent advances in fast algorithms have made this approach very appealing. Most of these fast methods take advantage of the fact that an iterative solver is a matrix-implicit algorithm. No explicit matrix has to be formed or stored; only the calculation of matrix-vector products is required. An existing precorrected-FFT algorithm [28] is particularly well suited for this problem and will be described here.

As demonstrated in the above formulation, the boundary element method often involves the solution of an integral equation of the following form:

$$\varphi(\vec{r}) = \int K(\vec{r};\vec{r}')\sigma(\vec{r}')d\vec{r}', \quad (19)$$

where $K(\vec{r};\vec{r}')$ is a known kernel. Given a potential distribution $\varphi(\vec{r})$, one desires to find the corresponding charge distribution $\sigma(\vec{r})$. In the context of matrix-implicit iterative methods, what is important is the ability to efficiently compute the potential distribution

for some charge distribution $\sigma(\vec{r})$. Although charge-potential terminology has been used here, this is for illustration purposes only; they can be any general variables, such as those in the matrix equation (18), and the kernel $K(\vec{r};\vec{r}')$ does not have to be the usual $\frac{1}{|\vec{r}-\vec{r}'|}$ implied by the charge-potential relationship.

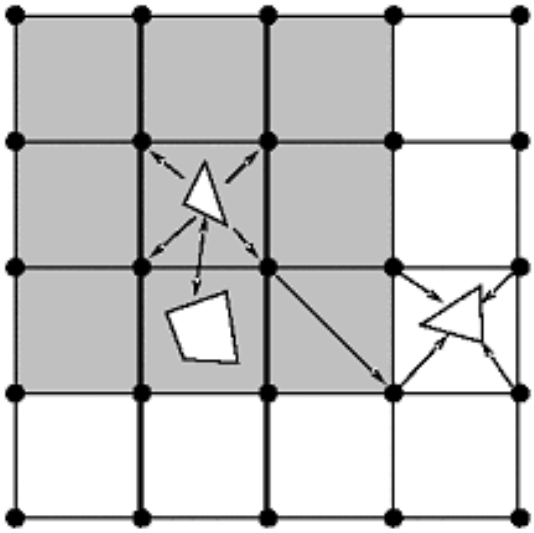

Figure 2: A pictorial representation of the precorrected FFT algorithm (image courtesy of J. Phillips)

The biomolecule electrostatic model has two integral equations with different kernels, and therefore the fast method for computing matrix-vector products must be kernel independent. Kernel independence is a key feature of the precorrected-FFT algorithm, and it is a property not shared with the more commonly used versions of the fast multipole method [25, 26, 42].

The algorithm can be summarized in four steps, as shown in Figure 2, where a given set of panels from a discretized surface are superimposed on a uniform grid. First, panel charges are projected onto their associated grid points, in what is called the projection step. Second, given the distribution of grid charges, the grid potential can be calculated using a convolution of the Green's function (the kernel) and the grid charges; this convolution is efficiently computed using the fast Fourier transform (FFT). Third, grid potentials are interpolated back onto the panels, a step known as interpolation. In the fourth step, called precorrection, nearby interactions are computed directly, with a correction factor that removes the contributions from the grid. All four steps—projection, interpolation, FFT convolution, and precorrection—possess sparse representations, and so the algorithm is very efficient in both speed (roughly $\Theta(n \log n)$) and memory (roughly $\Theta(n)$), where n is the number of panels. This is a tremendous improvement over traditional methods for discretizing the integral equations, which generate dense matrices and therefore require n^2 memory and n^2 operations for matrix-vector multiplication.

5. COMPUTATIONAL RESULTS

Simulation results from four test cases are presented below. The first case is a hypothetical spherical molecule whose analytical results [43] for the reaction potential are known. A direct factorization of the matrix equation in (18) was implemented in Matlab [44]

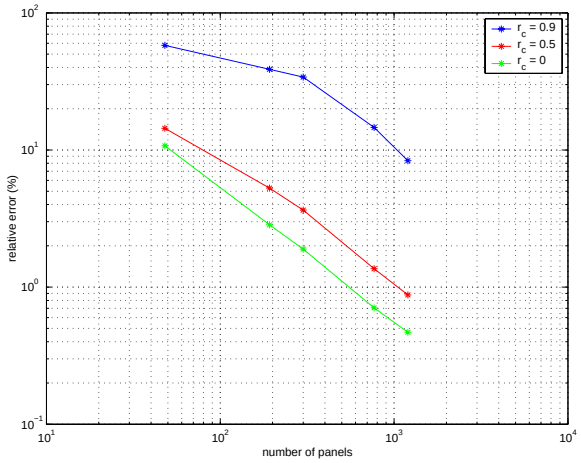

Figure 3: Convergence of the reaction potential of a spherical molecule to the analytical result as the discretization is refined.

to aid the convergence analysis and verify the validity of this formulation. However, $\Theta(n^2)$ memory requirements limited the size of this simulation problem to not much more than 1000 panels (2000 unknowns) for a computer with 1 GB of memory.

An accelerated GMRES solver with the pFFT implementation was developed to demonstrate the formulation on realistic examples and was used in each of the remaining test cases. Note that not only can the matrix equation in (18) be accelerated with pFFT, but the calculation of the reaction potential (15) can be accelerated also. The good conditioning of this integral equation formulation is illustrated by the rapid convergence of GMRES in the sphere example. The subsequent examples are simulations of a water molecule, an organic molecule in solvent, and protein macromolecules. The solvation free energy, which is simply one half of the inner product of the charge values with the vector of the potentials at the charge points, is compared with those obtained from the finite-difference solver DelPhi [1, 19, 20, 21, 22, 23].

5.1 Analytical Reaction Potential of a Spherical Molecule

A spherical molecule of radius 1 Å, in aqueous salt solution, with a single charge located at various radial distances, was simulated. A dielectric constant of 1 was used inside the molecule, and a dielectric constant of 20 was used externally; $\kappa = 3\text{Å}^{-1}$ in this example. The reaction potential calculated at the charge location was compared with the analytic result for three cases, at radial distances r_c of 0 Å, 0.5 Å, and 0.9 Å, as shown in Figure 3. As the charge moved closer to the molecular surface, the relative error also increased. All three cases exhibited reasonable convergence properties as the discretization was refined.

The number of iterations required to reach convergence with pFFT acceleration is shown in Table 1, for two charge locations, at $r_c = 0$ Å and $r_c = 0.9$ Å. Although no preconditioner was used in these test cases, GMRES converged reasonably quickly and the iteration count remained fairly constant as the number of panels increased. The conditioning of this formulation is evident and adoption of a preconditioner will further improve performance.

5.2 Solvation Free Energy of a Water Molecule

The pFFT implementation of the linearized Poisson-Boltzmann

Number of Surface Panels	GMRES iteration count	
	$r_c = 0.0$ Å	$r_c = 0.9$ Å
192	8	25
972	19	43
4800	16	44
6912	16	46
10800	18	48

Table 1: GMRES convergence of pFFT-acccelerated implementation

equation was then applied to cases that cannot be solved analytically, such as a small polyatomic molecule like water (H_2O). The geometry of water used is that based on the TIP3P model [45]. Two hydrogen atoms of radius 1.0 Å were bonded to an oxygen atom of radius 1.4 Å with bond lengths of 0.9572 Å and a bond angle of 104.52 degrees. The oxygen atom had a charge of -0.834 and the hydrogen atoms had charges of +0.417 in units of electron charge. The molecular surface of the water molecule was triangulated with the program MSMS [46] using a probe radius of 1.4 Å. An ionic strength of 0.145 M was used, equivalent to $\kappa = 0.124$ Å^{-1} at 25° C. No Stern layer was used, allowing the ionic strength to reach the molecular surface. A dielectric constant of 4 was used for the interior of the water molecule, and a dielectric constant of 80 was used externally. Figure 4 shows the relationship between molecular surface discretization and solvation free energy.

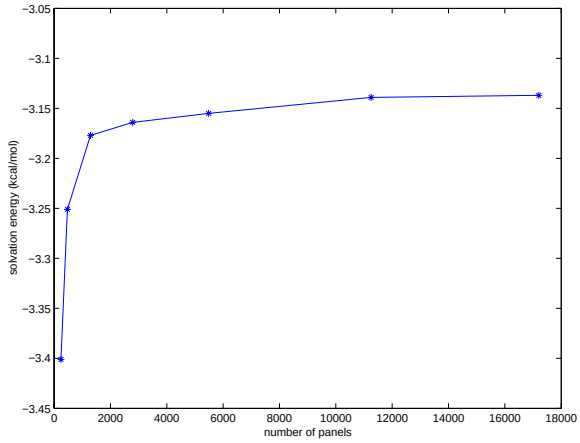

Figure 4: Simulation of a solvated water molecule.

5.3 Solvation Free Energy of a Highly Charged Small Organic Molecule

The next application of the pFFT solver was to a highly charged small organic molecule with 26 atoms, the transition state analog (TSA) of the protein enzyme chorismate mutase. The geometry of this small molecule was taken directly from an X-ray crystal structure, and can be obtained from the Protein Data Bank (PDB) as accession number 1ECM. The radii used were 1.0 Å for hydrogens, 1.4 Å for oxygens, 2.0 Å for aliphatic carbons, and 1.7 Å for carbonyl or vinyl carbons. The charges used were derived from quantum mechanical calculations. The molecular surface of the TSA molecule was triangulated with the program MSMS [46], using a probe radius of 1.4 Å for water. An ionic strength of 0.145 M was used, equivalent to $\kappa = 0.124$ Å^{-1} at 25° C. No Stern layer was used, allowing the ionic strength to reach the molecular surface. A dielectric constant of 4 was used inside the TSA molecule and a dielectric of 80 was used externally. Figure 5 shows the relationship between the surface discretization level and the solvation free energy calculated. The potential distribution on the molecular surface is shown in Figure 6, and is based on a surface mesh of 10162 panels.

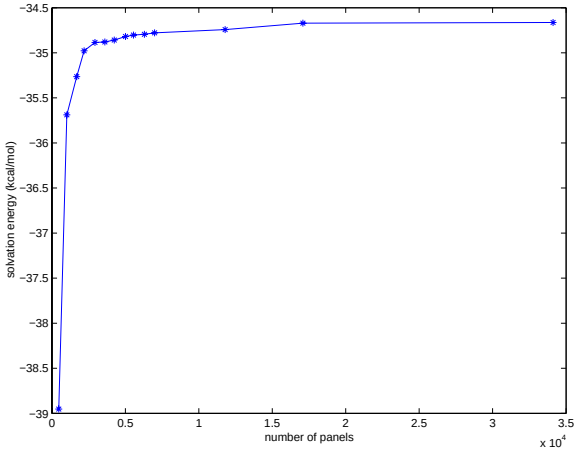

Figure 5: Simulation of a solvated TSA molecule.

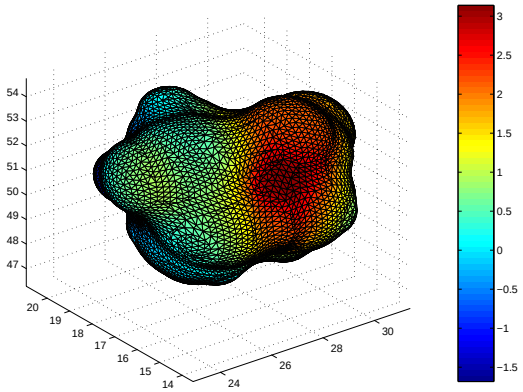

Figure 6: Potential distribution on the surface of a solvated TSA molecule.

5.4 Solvation Free Energy of Protein Macromolecules

An *E. coli* chorismate mutase (ECM) protein macromolecule with 3210 atoms was also simulated with the pFFT solver. The ECM protein has two TSA molecules bound with quantum mechanical charges and two water molecules bound with TIP3P charges. Similarly to the other test cases, MSMS was used to triangulate the molecular surface of the water molecule. An ionic strength of 0.145 M was used, equivalent to $\kappa = 0.124$ Å^{-1} at 25° C. No Stern layer was used, allowing the ionic strength to reach the molecular surface. A dielectric constant of 4 was used for the interior of the TSA molecule, and a dielectric of 80 was used externally. Figure

Protein			$E_{solvation}$ (kcal/mol)		time	
	# of dielectric panels	# of salt panels	pFFT	DelPhi	pFFT	DelPhi
Water	17204	9330	-3.14	-3.17	5 minutes	3 hours
TSA	34114	5842	-34.62	-34.75	10 minutes	3 hours
ECM	82868	18596	-646.42	-653.88	2.5 hours	3.5 hours

Table 2: Solvation free energies calculated by pFFT solver and DelPhi.

Protein				Energy Calculated (kcal/mol)	
	# atoms	# of dielectric panels	# of salt panels		
ECM	3210	82868	18596	$E_{solvation}$	-646.42
Barnase	1107	43298	21284	$E_{desolvation}$	51.06
Barstar	839	35978	17434	$E_{desolvation}$	40.11
Barnase-Barstar Complex	1946	68592	31728	$E_{interaction}$	-82.65
				$E_{binding}$	8.53

Table 3: Solvation free energies calculated by pFFT solver.

7 shows the convergence of the solvation free energy with refined molecular surface discretization.

Barnase, barstar, and the barnase-barstar complex were also simulated using the pFFT solver, and the energies calculated are listed in Table 3.

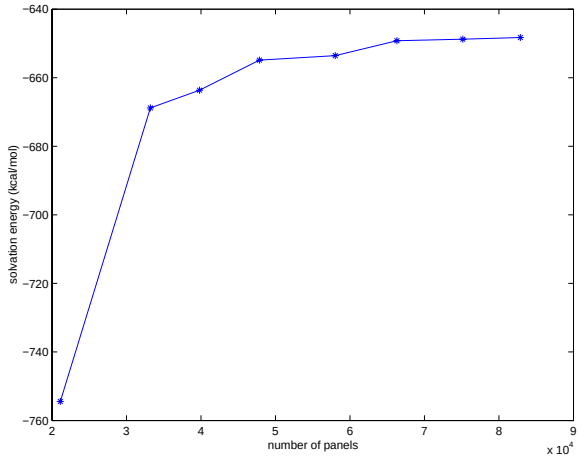

Figure 7: Simulation of a solvated ECM macromolecule.

5.5 Comparison to DelPhi

DelPhi is a popular finite-difference scheme based simulation tool for solving the linearized Poisson-Boltzmann equation, and is used both in academic and industry settings. Table 2 compares the results for the three molecules described previously and illustrates the time savings of the pFFT-accelerated integral equation formulation. An ionic strength of 0.145 M was used as before. A Stern layer of 2 Å was also used in all three cases here. The first column lists the number of discretization panels for the dielectric interface, and the second column lists the number of discretization panels for the salt interface (i.e., the Stern layer). The discretization used in DelPhi was 257 grids per Angstrom. The two solvers agree to within 1%. The pFFT solver is one order of magnitude faster than DelPhi for smaller molecules like water and TSA. The speed improvement is less significant in the ECM case for two reasons. First, no preconditioners were used in any of the examples, so convergence is slow for large macromolecules with complicated geometries (such as ECM). Work is underway to devise and implement an efficient preconditioner that can significantly improve simulation time. Second, the surface discretization software has difficulty triangulating complicated geometries such as the ECM macromolecular surface; some of the generated panels have extremely high aspect ratios, which can cause problems for the pFFT solver. This tessellation problem may also be responsible for the discrepancy in solvation free energies calculated by DelPhi and the pFFT solver. We are currently examining methods to improve the quality of the surface discretizations.

6. CONCLUSION

In this paper we presented an integral-equation based approach for computing numerical solutions to the mixed discrete-continuum model of biomolecule electrostatics. The new approach combines a carefully chosen integral formulation of the mixed discrete-continuum model with a kernel-independent precorrected-FFT accelerated integral equation solver. Computational results from our new approach, on both simple and more complicated geometries, were compared to analytic results and to the widely used finite-difference based DelPhi program. The results clearly indicate that the new simulator can be as much as thirty times faster than Delphi, though our new program generated disappointing results in some cases. These preliminary results are encouraging and indicate a potential application of this formulation. More rigorous test cases are being designed and studied, and further optimization of the pFFT implementation is under investigation.

7. REFERENCES

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

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

[3] M. E. Davis and J. A. McCammon. Electrostatics in biomolecular structure and dynamics. *Chem. Rev.*, 90:509–521, 1990.

[4] J. A. Grant, B. T. Pickup, and A. Nicholls. A smooth permittivity function for Poisson-Boltzmann solvation methods. *Journal of Computational Chemistry*, 22:608–640, 2001.

[5] L. P. Lee and B. Tidor. Barstar is electrostatically optimized for tight-binding to barnase. *Nature Structural Biology*, 8:73–76, 2001.

[6] L. Lee and B. Tidor. Optimization of binding electrostatics: Charge complementarity in the barnase-barstar protein complex. *Protein Science*, 10:362–377, 2001.

[7] E. Kangas and B. Tidor. Electrostatic specificity in molecular ligand design. *Journal of Chemical Physics*, 112:9120–9131, 2000.

[8] E. Kangas and B. Tidor. Electrostatic complementarity at ligand binding sites: Application to chorismate mutase. *Journal of Physical Chemistry*, 105:880–888, 2001.

[9] Z. S. Hendsch, M. J. Nohaile, R. T. Sauer, and B. Tidor. Preferential heterodimer formation via undercompensated electrostatic interactions. *Journal of the American Chemical Society*, 123:1264–1265, 2001.

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

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

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

[13] 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. *J. Am. Chem. Soc.*, 113:1454–1455, 1991.

[14] S. W. Rick and B. J. Berne. The aqueous solvation of water: A comparison of continuum methods with molecular dynamics. *J. Am. Chem. Soc.*, 116:3949–3954, 1994.

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

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

[17] 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*, 157:671–679, 1982.

[18] 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*, 91:57–95, 1995.

[19] 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*, 1:47–59, 1986.

[20] M. K. Gilson, K. A. Sharp, and B. H. Honig. Calculating the electrostatic potential of molecules in solution: Method and error assessment. *Journal of Computational Chemistry*, 9:327–335, 1987.

[21] A. Nicholls and B. Honig. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. *Journal of Computational Chemistry*, 12:435–445, 1991.

[22] 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*, 105:6507–6514, 2001.

[23] W. Rocchia, S. Sridharan, A. Nicholls, E. Alexov, A. Chiabrera, and B. Honig. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. *Journal of Computational Chemistry*, 23:128–137, 2002.

[24] V. Rokhlin. Rapid solution of integral equation of classical potential theory. *Journal of Computational Physics*, 60:187–207, 1985.

[25] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *Journal of Chemical Physics*, 73:325–348, 1987.

[26] K. Nabors and J. White. FASTCAP: A multipole accelerated 3-D capacitance extraction program. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 10:1447–1459, 1991.

[27] B. J. Yoon and A. M. Lenhoff. A boundary element method for molecular electrostatics with electrolyte effects. *Journal of Computational Chemistry*, 11:1080–1086, 1990.

[28] J. R. Phillips and J. K. White. A precorrected-FFT method for electrostatic analysis of complicated 3-D structures. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 16:1059–1072, 1997.

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

[30] M. L. Connolly. Analytical molecular surface calculation. *Journal of Applied Crystallography*, 16:548–558, 1983.

[31] R. W. Hockney and J. W. Eastwood. *Computer simulation using particles*. Adam Hilger, 1988.

[32] N. K. Rogers and M. J. Sternberg. *Journal of Molecular Biology*, 174:527, 1984.

[33] Z. Zhou, P. Payne, M. Vasquez, N. Kuhn, and M. Levitt. Finite-difference solution of the Poisson–Boltzmann equation: Complete elimination of self-energy. *J. Comput. Chem.*, 11:1344–1351, 1996.

[34] R. J. Zauhar and R. S. Morgan. The rigorous computation of the molecular electric potential. *Journal of Computational Chemistry*, 9:171–187, 1988.

[35] A. H. Juffer, E. F. F. Botta, B. A. M. Van Keulen, A. V. D. Ploeg, and H. J. C. Berendsen. The electric potential of a macromolecule in a solvent: A fundamental approach. *Journal of Computational Physics*, 97:144–171, 1991.

[36] R. Bharadwaj, A. Windemuth, S. Sridharan, B. Honig, and A. Nicholls. The fast multipole boundary element method for molecular electrostatics: An optimal approach for large systems. *Journal of Computational Chemistry*, 16:898–913, 1995.

[37] R. J. Zauhar and A. Varnek. A fast and space-efficient boundary element method for computing electrostatic and hydration effects in large molecules. *Journal of Computational Chemistry*, 17:864–877, 1996.

[38] M. O. Fenley, W. K. Olson, K. Chua, and A. H. Boschitsch. Fast adaptive multipole method for computation of electrostatic energy in simulations of polyelectrolyte DNA. *Journal of Computational Chemistry*, 17:976–991, 1996.

[39] I. Stakgold. *Boundary Value Problems of Mathematical Physics*. MacMillan, 1967.

[40] R. F. Harrington. *Field Computation by Moment Methods*. MacMillan, 1968.

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

[42] L. Greengard. *The Rapid Evaluation of Potential Fields in Particle Systems*. MIT Press, 1988.

[43] J. G. Kirkwood. Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *Journal of Chemical Physics*, 2:351, 1934.

[44] *Matlab v.6*. Mathworks, Inc.

[45] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, 79:926–935, 1983.

[46] M. Sanner, A. J. Olson, and J. C. Spehner. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers*, 38:305–320, 1996.