

## The Geometry Behind Numerical Solvers of the Poisson-Boltzmann Equation

Xinwei Shi<sup>1</sup> and Patrice Koehl<sup>2,\*</sup>

<sup>1</sup> *Genome Center, University of California, Davis, CA 95616, USA.*

<sup>2</sup> *Department of Computer Science and Genome Center, University of California, Davis, CA 95616, USA.*

Received 3 October 2007; Accepted (in revised version) 9 December 2007

Available online 24 January 2008

---

**Abstract.** Electrostatics interactions play a major role in the stabilization of biomolecules: as such, they remain a major focus of theoretical and computational studies in biophysics. Electrostatics in solution is strongly dependent on the nature of the solvent and on the ions it contains. While methods that treat the solvent and ions explicitly provide an accurate estimate of these interactions, they are usually computationally too demanding to study large macromolecular systems. Implicit solvent methods provide a viable alternative, especially those based on Poisson theory. The Poisson-Boltzmann equation (PBE) treats the system in a mean field approximation, providing reasonable estimates of electrostatics interactions in a solvent treated as continuum. In the first part of this paper, we review the theory behind the PBE, including recent improvement in which ions size and dipolar features of solvent molecules are taken into account explicitly. The PBE is a non linear second order differential equation with discontinuous coefficients, for which no analytical solution is available for large molecular systems. Many numerical solvers have been developed that solve a discretized version of the PBE on a mesh, either using finite difference, finite element, or boundary element methods. The accuracy of the solutions provided by these solvers highly depend on the geometry of their underlying meshes, as well as on the method used to embed the physical system on the mesh. In the second part of the paper, we describe a new geometric approach for generating unstructured tetrahedral meshes as well as simplifications of these meshes that are well fitted for solving the PBE equation using multigrid approaches.

**AMS subject classifications:** 35J05, 35J65, 65N06, 65N30, 65N38, 65N50, 65N55

**PACS:** 02.30.Jr, 02.40.-k, 02.60.Lj, 02.70.Bf, 02.70.Dh, 31.70.Dk

**Key words:** Poisson-Boltzmann equation, biomolecular modeling, mesh generation, finite difference methods, finite element methods.

---

\*Corresponding author. *Email addresses:* xshi@ucdavis.edu (X. W. Shi), koehl@cs.ucdavis.edu (P. Koehl)

## 1 Introduction

Electrostatics interactions are central in physics, chemistry and biology: without them, molecules would not hold together. Understanding electrostatics is especially important in biology: biomolecules are usually considered as large polyelectrolytes, whose properties depend on their own charge distribution as well as on their interactions with surrounding charged neighbors. In the case of proteins for example, Kauzmann [1] foresaw the importance of electrostatics for stability, proposing that polar (charged) groups would either compensate for each other, or be solvated by water. Later, Perutz [2] confirmed these predictions based on information deduced from the first high resolution protein structures, and further emphasized the role of electrostatics in biomolecular interactions. Electrostatics is even more important for nucleic acids. RNA and DNA are polyanions that bind water and cations in order to acquire a three-dimensional structure. It is well known that in the absence of metal ions, RNA and DNA form most of their secondary structures, but only a fraction (if any) of their tertiary contacts [3]. Accurate theoretical and subsequent numerical treatments of electrostatics has therefore always been a major concern in structural biology.

Theoretical modeling of electrostatics interactions is deceptively simple and, as such, has been and remains the subject of development by both the physics and the scientific computing communities. As a first 'classical' approximation, the interaction between two charges can be described by Coulomb's law. When more than two charges interact, the total electrostatic energy of the system is derived as the sum of all pairwise Coulomb interactions (superposition principle). Applications of these simple principles imply that the positions of all charges are known. While this seems to be a simple requirement, it is unfortunately difficult to meet when modeling large molecular systems. This is mostly due to the inherent difficulties in accounting for the solvent that surrounds the molecules and the ions this solvent may contain. Explicit representation of the solvent provides an accurate treatment of electrostatics, but it increases the size of the system under study by orders of magnitude [4]. In addition, interactions involving solvent need to be averaged over relatively long time intervals before results become meaningful. As a response to these problems, there has been a continuous effort in the physics community to develop simplified models that are computationally tractable and that remain physically accurate. Most of these models include the solvent implicitly, reducing the solute-solvent interactions to their mean field characteristics, which are expressed as functions of the solute degrees of freedom alone. These approaches treat the solvent as a dielectric continuum and are therefore referred to as a continuum dielectric model. Many of these approaches use the Poisson-Boltzmann equation to describe the electrostatic potential of a system, and rely on numerical methods to solve this non linear elliptic equation. This paper is designed around two distinctive parts. Firstly, it provides an overview of the recent methodological developments in designing fast and accurate solvers of the Poisson-Boltzmann equation. We also cover recent theoretical models that modify the Poisson-Boltzmann equation to better represent the physics of the system under study.

We show how one of these modified Poisson-Boltzmann formalisms that introduce explicit solvent dipoles can be used to predict binding sites on proteins. Secondly, we focus on the geometric components of these solvers, and present our own recent developments in this field.

The review part of this paper is far from exhaustive, and the reader is referred to several recent review articles on this and related areas, including overviews of protein electrostatics [5,6], implicit [7–10], and explicit [4] solvent models, electrostatics in force fields [11], and applications of electrostatics calculations to membrane proteins [12,13].

## 2 Continuum electrostatics: Poisson equations

The Poisson-Boltzmann model is the most commonly used model to account for electrostatics interactions between charged objects. It assumes point-like charges immersed in a continuum dielectric media and treats the system in a mean-field approximation. The medium is modeled by a homogeneous dielectric constant. Here we briefly describe the theory behind the Poisson-Boltzmann model, we list some of its limitations, as well as the current solutions to circumvent these problems.

### 2.1 The Poisson-Boltzmann model

The fundamental equation of electrostatics is given by Gauss's law:

$$\nabla \cdot \vec{D}(\mathbf{r}) = 4\pi\rho(\mathbf{r}), \quad (2.1)$$

which relates spatial variation of the displacement field  $\vec{D}$  with position  $\mathbf{r}$  to the charge density distribution  $\rho$ . In general, the displacement field is defined as:

$$\vec{D} = \epsilon_0 \vec{E} + \vec{P}, \quad (2.2)$$

where  $\vec{E}$  is the electric field,  $\epsilon_0$  is the vacuum permittivity and  $\vec{P}$  is the polarization density of the material considered.

In free space,  $\vec{P} = 0$ , and Eq. (2.1) reduces to

$$\nabla \cdot \vec{E}(\mathbf{r}) = \frac{4\pi\rho(\mathbf{r})}{\epsilon_0}, \quad (2.3)$$

or, expressing  $\vec{E}$  as the gradient of the electrostatic potential  $\phi$ ,

$$\nabla^2 \phi(\mathbf{r}) = -\frac{4\pi\rho(\mathbf{r})}{\epsilon_0}, \quad (2.4)$$

which is the Poisson equation. When the charge distribution can be described in terms of  $N$  point charges, the solution to Eq. (2.4) becomes Coulomb's law:

$$\phi(\mathbf{r}) = \sum_{i=1}^N \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|}, \quad (2.5)$$

where  $r_i$  is the position, and  $q_i$  the magnitude of the  $i$ th point charge. In any particular problem if all charges are known and represented explicitly, Eq. (2.5) applies rigorously. If however it is not possible to define the charge density explicitly (such as in a complex medium), it is necessary to go back and rewrite the Gauss law in a more convenient form.

In a region of material with uniform susceptibility,  $\chi$ , the polarization density  $\vec{P}$  is given by  $\vec{P} = \chi \vec{E}$ . Defining  $\epsilon = \chi + \epsilon_0$ , the dielectric constant or permittivity of the medium, Eq. (2.1) is rewritten in the form

$$\nabla^2 \phi(\mathbf{r}) = -\frac{4\pi\rho(\mathbf{r})}{\epsilon}. \quad (2.6)$$

In general, the permittivity varies through space, and the Poisson equation becomes:

$$\nabla \cdot (\epsilon(\mathbf{r}) \vec{\nabla} \phi(\mathbf{r})) = -4\pi\rho(\mathbf{r}), \quad (2.7)$$

where  $\epsilon$  is now a function of position.

The charge density is the sum of the charges of the solute and of the ions in the solvent. The positions  $r_i$  and magnitudes  $q_i$  of the  $M$  charges of the solute are known: they represent the fixed, or "source" charges:

$$\rho_s(\mathbf{r}) = e_c \sum_{i=1}^M q_i \delta(\mathbf{r} - \mathbf{r}_i), \quad (2.8)$$

where  $e_c$  is the charge of the electron, and  $\delta(x)$  is the delta function such that  $\delta(x-y) = 1$  if  $x=y$ , and 0 otherwise.

Mobile ions in the solvent are not represented explicitly. Instead, the chemical potential of each ion is assumed to be uniform in the solvent (mean field approximation), yielding a Boltzmann distribution for the ion charge distribution:

$$\rho_{ion}(\mathbf{r}) = e_c \sum_{i=1}^m c_i z_i e^{-z_i e_c \phi(\mathbf{r}) / k_B T}, \quad (2.9)$$

where  $c_i$  and  $z_i$  are the bulk concentration and valence of ion specie  $i$  respectively,  $k_B$  is the Boltzmann constant, and  $T$  the temperature. In the case of a one-to-one electrolyte (solution with one positive and one negative ion species, both of valence 1, such as NaCl), Eq. (2.9) reduces to

$$4\pi\rho_{ion}(\mathbf{r}) = -\bar{\kappa}^2(\mathbf{r}) \sinh\left(\frac{e_c \phi(\mathbf{r})}{k_B T}\right), \quad (2.10)$$

where  $\bar{\kappa}$ , the position dependent modified Debye-Hückel coefficient is related to the ionic strength  $I$  by:

$$\begin{aligned} \bar{\kappa}^2(\mathbf{r}) &= c(\mathbf{r}) \kappa^2 \\ &= c(\mathbf{r}) \left( \frac{8\pi N_A e_c^2}{1000 k_B T} \right) I, \end{aligned} \quad (2.11)$$

where  $N_A$  is the Avogadro number and  $c$  identifies the regions in which ions intervene. Replacing Eqs. (2.8), (2.10) and (2.11) into Eq. (2.7) we obtain the Poisson-Boltzmann (PB) equation in its most standard form (for a one-to-one electrolyte)

$$\nabla \cdot \left( \epsilon(\mathbf{r}) \vec{\nabla} \phi(\mathbf{r}) \right) - c(\mathbf{r}) \kappa^2 \sinh \left( \frac{e_c \phi(\mathbf{r})}{k_B T} \right) = -4\pi e_c \sum_{i=1}^M q_i \delta(\mathbf{r} - \mathbf{r}_i). \quad (2.12)$$

A weaker form of this equation exists when the electrostatics potential is small, in which case  $\sinh(x) \approx x$ . Using this first order development of the hyperbolic sine function yields the Linearized Poisson-Boltzmann (LPB) equation:

$$\nabla \cdot \left( \epsilon(\mathbf{r}) \vec{\nabla} \phi(\mathbf{r}) \right) - c(\mathbf{r}) \kappa^2 \frac{e_c}{k_B T} \phi(\mathbf{r}) = -4\pi e_c \sum_{i=1}^M q_i \delta(\mathbf{r} - \mathbf{r}_i). \quad (2.13)$$

The linearized PB equation is simpler to solve, but is only valid under the assumption of small electrostatics potentials, which limits its field of applications.

Both the PB and LPB equations are fully defined by their coefficients  $\epsilon(\mathbf{r})$ ,  $c(\mathbf{r})$  and the charge positions  $r_i$ , which in turn are fully characterized by the biomolecular structure under study. Three different zones are usually considered:

- *Region 1: Molecular interior.* Within the interior of the molecule,  $\epsilon$  is set constant (usually  $\epsilon=1$  or 2),  $c$  is set to 0 (no ions), and the charges  $q_i$  are positioned according to the atomic coordinates.
- *Region 2: Stern region.* The Stern region is an ion-excluded region around the molecule in which  $\epsilon$  is set constant to its value in solvent (usually  $\epsilon=80$ ),  $c$  is set to 0 (no ions to account for their size) and the source charge density is 0.
- *Region 3: Solvent.* In the bulk solvent,  $\epsilon$  is set constant to usually 80, ion densities are taken into account by setting  $c=1$ , and the source charge density remains 0.

While it is easy to define the coefficients of the PB equation within each zone, special care must be taken when defining the interfaces between them, i.e. the molecular surface between Region 1 and Region 2, and the ion excluded surface between Region 2 and Region 3. These interfaces have been represented by a variety of models, described in section 3.

## 2.2 Size modified Poisson-Boltzmann equation

When dealing with ion atmospheres around molecules, the Poisson-Boltzmann theory has two limitations: it does not include ion size, nor does it account for ion-ion correlations in its treatment. Ion size plays a role in electrostatics interactions, as was recently shown for DNA in solutions containing competing cations [14]. The standard PBE usually accounts for ion size by using an exclusion layer, the so-called Stern region. Such a model however cannot be applied to a complex mixture of ions with different sizes. It

does not also account for excluded volume effects among the ions. In principle, Monte Carlo approaches can be very accurate in studying ion concentrations [15]. These methods however come at a high computational cost, making them impractical. Coalson and colleagues as well as Orland and colleagues developed an attractive solution to the problem of the influence of ion size using a lattice field theory [16–19]. This approach was recently generalized by Chu et al. [20] to deal with ions with two different sizes. In this approach, the hard core repulsion between solvated ions is approximated with an excluded term in the free energy density of a lattice gas model of the ionic solution. The domain around the charged biomolecule is treated as a 3D lattice with evenly spaced points  $a$  apart. This characteristic lattice spacing sets the volume of the larger ions at  $a^3$ ; the volume of the smaller ions is set to  $a^3/k$ , where  $k$  is a dimensionless parameter. Let us suppose that the solution contains two species with valences  $z_1$  and  $z_2$ , and bulk concentrations  $c_1$  and  $c_2$ , such that ion species one corresponds to the small ions, while ion species two corresponds to the large ions. By computing the grand partition function for each lattice site and computing the chemical potential of each species  $\mu$  as the derivative of this partition function with respect to  $\mu$ , Chu et al. derived a size modified Poisson-Boltzmann (SMPB) equation:

$$\nabla \cdot (\epsilon(\mathbf{r}) \vec{\nabla} \phi(\mathbf{r})) = -4\pi e_c \sum_{i=1}^M q_i \delta(\mathbf{r} - \mathbf{r}_i) - c(\mathbf{r}) \frac{4\pi e_c}{D(\mathbf{r})} \times \left\{ z_1 c_1 e^{-z_1 e_c \phi(\mathbf{r})/k_B T} \left[ 1 - f_0 + (c_1 a^3/k) e^{-z_1 e_c \phi(\mathbf{r})/k_B T} \right]^{k-1} + (1 - c_2 a^3)^{k-1} z_2 c_2 e^{-z_2 e_c \phi(\mathbf{r})/k_B T} \right\}, \quad (2.14)$$

where

$$D(\mathbf{r}) = \left[ 1 - f_0 + (c_1 a^3/k) e^{-z_1 e_c \phi(\mathbf{r})/k_B T} \right]^k + (1 - c_2 a^3)^{k-1} (c_2 a^3 e^{-z_2 e_c \phi(\mathbf{r})/k_B T}) \quad (2.15)$$

and  $f_0 = (c_1 a^3/k) + c_2 a^3$  is the fractional occupancy of each lattice site.

The size-modified Poisson-Boltzmann equation was shown to be able to account for differences in monovalent ionic association on DNA for ions of different sizes and concentrations  $\leq 150$  nM [20]. Note however that this approach is currently limited to mixtures of ions with two different sizes, and that extension to include three or more ion sizes would prove cumbersome. Note also that the SMPB equation treats ion-ion interactions with a mean field approximation. As such, it still neglects ion-ion correlations. This remains an issue necessitating further advances in theory, especially for the treatment of divalent anions.

### 2.3 Dipolar Poisson-Boltzmann equation

In the PB model, the medium is modeled by a homogeneous and isotropic dielectric constant (see Eq. (2.6)). This assumption however does not take into account the strong dielectric response of water molecules around charges. The discrete dipolar moments of

water molecules orient themselves close to charges, giving rise to hydration shells and to hydrophobic interactions. These phenomena are very important in many biological processes, but are not accounted for by the PB equation. Recently, Abrashkin et al. [21] proposed a modified PB equation, the dipolar Poisson-Boltzmann (DPB) equation. In the DPB model, the solvent is represented by  $N_d$  mobile dipoles, each with a dipolar moment  $\mathbf{p}$ . The charge density created by a point dipole  $\mathbf{p}$  at point  $\mathbf{r}_0$  is given by

$$\rho_d(\mathbf{r}) = -\mathbf{p} \cdot \vec{\nabla} \delta(\mathbf{r} - \mathbf{r}_0).$$

Solving for  $\mathbf{p}(\mathbf{r})$  using a mean field approach, Abrashkin et al. derived the dipolar Poisson-Boltzmann (DPB) equation:

$$\epsilon_0 \nabla^2 \phi(\mathbf{r}) = -4\pi \rho_{ion}(\mathbf{r}) - 4\pi \rho_f(\mathbf{r}) - 4\pi \lambda_d p_0 \vec{\nabla} \cdot \left[ \frac{\vec{\nabla} \phi(\mathbf{r})}{|\vec{\nabla} \phi(\mathbf{r})|} \Gamma \left( \frac{p_0 |\vec{\nabla} \phi(\mathbf{r})|}{k_B T} \right) \right], \quad (2.16)$$

where  $\lambda_d$  is the fugacity of the dipoles,  $p_0$  is the assumed constant magnitude of the dipoles,  $\rho_f$  is the density of the fixed charges, and the function

$$\Gamma(u) = \cosh u / u - \sinh u / u^2.$$

Note that in reference to the Gauss law (see Eqs. (2.1)-(2.2)), the last term of Eq. (2.16) is the divergence of the polarization  $\vec{P}$ .

The DPBE model can be combined with the size modified PB model, to account for the finite sizes of ions and dipoles. This was recently proposed by Azuara et al. [22] in another modified PB equation, named GPBL which stands for Generalized Poisson-Boltzmann Langevin, as the water is modeled with explicit Langevin dipoles. Compared to the classical PB equation, the GPBL equation provides information on the water dipole density around a solute. Based on this information, it is possible to color the solute with a simple color code: red (solvated) for values at or above a threshold value, and blue otherwise, where the threshold is computed based on the expected bulk water density. Blue patches on the surface correspond to poorly solvated region for the protein. We have shown that such regions matches with interface regions in homodimeric proteins [22]. This is illustrated here in the case of the Trp repressor, a 25kD protein that regulates the biosynthesis of tryptophane in bacteria. Trp repressor is a fully helical protein that is present in the cell mostly as an homodimer. The structure of the dimer was solved both by X-ray crystallography and by NMR, in presence and in absence of its ligands. We chose the NMR structure, available in the PDB with the identifier 2WRP. The interface between the two monomers is large, representing approximately 29% of the total accessible surface area of a monomer. Most of this interface region is predicted to be poorly solvated if a monomer Trp repressor is put in solution (see Fig. 1.)

More generally, It will be of interest to see how the GPBL equation can help us understand of hydration forces, ionic associations and short range hydrophobic effects.

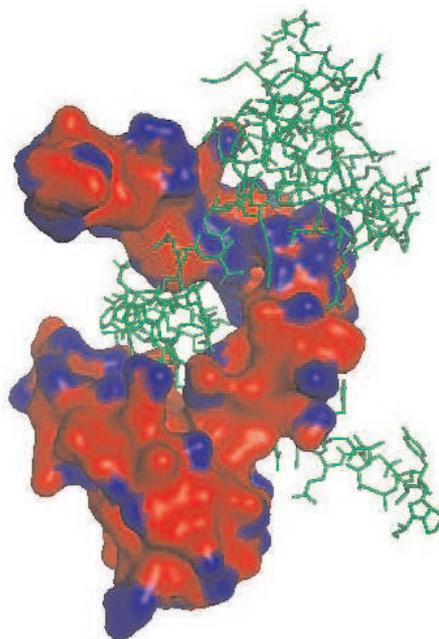


Figure 1: *Predicting protein binding sites from their computed desolvation sites.* Trp repressor is a large homodimeric protein that regulates the biosynthesis of tryptophane (W) in bacteria. We studied the electrostatics around one Trp repressor monomer, using the GPBL equation. The solution to this equation provides us with the solvation free energy as well as with the solvent density around the molecule. We color the molecular surface of the Trp repressor monomer such that red and blue regions represent solvated and poorly solvated regions, respectively. Interestingly, many of the blue patches on the surface are part of the interface between the two monomers in the Trp repressor dimer (the second monomer is shown in green, in line mode). The figure was generated using pymol (<http://www.pymol.org>).

## 2.4 Numerical solutions to the Poisson-Boltzmann equation

The Poisson-Boltzmann equation (2.2) is a second order nonlinear elliptic partial differential equation. Analytical solution of the PBE is only available for simple geometry such as spheres and cylinders [23,24]. For the complex geometry of a biomolecule like a protein or a nucleic acid, analytical solutions are not available and the PBE must be solved using numerical methods [25]. Many such solvers have been developed [26–28]; they can be divided into two groups based on their underlying geometric model, namely, surface based methods and volume based methods.

**Integration of PBE on a surface mesh: The BEM approach.** The PB equation needs to be solved over the whole space surrounding the molecule(s) of interest. In this space however, it is easy to distinguish two regions: the “interior” of the molecule (i.e. the volume enclosed by its surface), in which electrostatics is basically described by a Poisson equation (2.4), and the “outside” of the molecule, in which electrostatics is described by the PB equation. Key to solving this problem is the requirement that the electrostatics potential and the normal component of the displacement field  $\vec{D}$  be continuous at the interface between these two regions. Boundary element methods exploit the fact that the differen-

tial equations can be rewritten into a set of boundary integral equations that need to be solved only on this interface, namely the surface of the molecule [27, 29–31]. These methods use a polygonal representation of the molecular surface, usually a triangular mesh, to express the PB (in its linearized form, Eq. (2.13)) as a collection of linear equations with potential values (or a combination of potential values with their normal derivatives) at the surface as unknowns. The electrostatics potential on the surface is subsequently derived from the solutions of this system of equations (usually by interpolation), and the volumetric potential is derived from the integral formulations of the PB equations. With the recent introduction of fast multipole techniques, BEM methods have become faster than the volumetric methods described below [29]. Their applications however remain limited to cases for which the linearized PB equation applies, though there have been attempts to combine BEM with a volumetric approach to speed up calculations while maintaining the non linear part of the PB equation [32, 33].

**Integration of the PBE over a volumetric mesh: FD and FEM approaches.** Classical methods for solving the PB equation rely on a discretization of the 3D space to project the continuous solution on a finite set of basis functions. The most popular discretization techniques use Cartesian or structured meshes to subdivide the space. Among these methods, the finite difference methods (FD) remain predominant. The most general form of FD solves the PB equation on a non-uniform Cartesian grid. A discrete version of the PBE is derived at each vertex of this mesh; the differential operators are transformed into differences involving the vertex and its neighbors on the grid, by means of a first or second order Taylor expansion [28, 34, 35]. The corresponding set of discrete equations form a large sparse matrix equation that is then solved either directly or iteratively in the case of the nonlinear equation by a variety of linear algebra techniques, among which multigrid methods prime (see the next paragraph). FD methods are usually simple to setup; in return however, they are difficult to tune to the specificity of the system under study. Atomic charges for example will generally not coincide with mesh vertices; it is also impossible to increase the accuracy of the solution locally without increasing the number of vertices across the entire grid. Adaptive finite element (FE) methods offer solutions to both problems [10, 26, 36, 37]. Finite element meshes are unstructured by nature; they are composed of vertices connected by edges, forming a series of simplices (triangles and tetrahedra) that serve as support for the basis functions that approximate the solution of the PB equation. As the set of vertices is not constrained to be a subset of a Cartesian grid, it usually contains the atomic centers. Also, the number of vertices can be arbitrarily increased in specific regions (such as in the neighborhood of the molecular surface), with no impact on other regions of the grid.

**Numerical solvers for the PB equation.** All three approaches described above, namely BE, FD and FE methods reduce the PB equation into a set of discrete algebraic equations, which needs to be solved numerically. Most PB solvers use iterative methods to solve these equations. In general, iterative methods can be grouped into two types: stationary methods such as Jacobi, Gauss-Seidel and Successive Over-Relaxation (SOR) meth-

ods, and non stationary methods such as Conjugate Gradient(CG), Generalized Minimal Residual (GMRES), and Biconjugate Gradient [38]. The choice between these two types of iterative methods depends on the properties of the coefficient matrix coupled with the linear system. As BE methods result in dense linear systems, their solvers mostly rely on non stationary methods such as GMRES [29, 31]. In opposition, as the coefficient matrix derived in FD and FE methods is usually sparse symmetric positive definite, both stationary and non stationary methods have been applied. Iterative methods start with an initial guess for the electrostatics potentials, and repeatedly modify this guess until a solution with a desired accuracy is reached. The speed at which this solution is obtained depends on the size of the problem and the nature of the equation. As BE methods usually express the PB equation using a well conditioned integral formula, iterative approaches are usually well suited as they can have a good convergence rate. In the case of FD and FE methods however, direct application of iterative methods is usually not efficient, due to their inability to reduce the effect of long range errors in the refinement of the solution [39]. This problem is overcome by multigrid methods which use a hierarchy of meshes with increasing resolution to represent the 3D space containing the system, with the finest mesh corresponding to the original mesh used in standard iterative approaches. The PB equation is first solved on the coarsest grid, and the corresponding solution serves as initial guess for the next grid level considered. This procedure is then repeated until a solution is reached on the finest grid. Multigrid methods provide significant speed up compared to standard iterative methods [26, 28, 40].

**Solving the modified Poisson-Boltzmann equation.** Despite its apparent higher complexity (see Eq. (2.14)), the size modified Poisson-Boltzmann (SMPB) equation is not more difficult to solve than the standard PB equation: the equation is still elliptical and upon discretization, the system of equations obtained remains definite positive. The only difference between the PB and SMPB equations is found in the Helmholtz term (i.e. the non-linear term that defines the influence of the ions); both the Laplacian and the source terms (i.e. fixed charges of the molecule) remain unchanged. For any grid point in the mesh used to solve the SMPB equation, the Helmholtz term remains a function of the electrostatic field at this grid point only. SMPB therefore can be solved using any of the approaches described above.

The situation is different for the DPBE, as this equation (see 2.16) includes first order terms (coming from the terms involving the electric field); there are no guarantees that its operator matrix is positive definite. The only method currently available to solve this equation applies a multigrid conjugate gradient procedure [22]. New solvers are expected soon (P. Koehl, in preparation).

### 3 Biomolecular geometry for PBE discretization

An essential component in all numerical PBE solvers is the construction of a geometric mesh that decomposes the space and defines the three regions needed to specify the

coefficients of the PBE (see Section 2.1). The geometric structure of these meshes significantly influences both the accuracy of the solution and the efficiency of the solver itself. The ideal mesh needs to be fine enough to provide an accurate representation of the interfaces between the different regions (interior of the molecule, Stern region and ion accessible region) and at the same time it should contain a small number of vertices so that the solvers that use them can be fast. In practice, meshes are usually derived as a compromise between these two conflicting constraints. In addition, multigrid techniques need a method for simplifying a fine grid into a coarse grid, and reversely a geometric interpolation technique to embed the coarse grid solution onto a finer grid.

In this section, we describe our recent work on generating meshes for biomolecular systems, with a brief overview of the related work. We focus on the algorithms used for generating high quality triangular and tetrahedral meshes for biomolecules that provide flexibility to fit complicated domains and ease of refinement. We also describe the methods for constructing a hierarchy of tetrahedral mesh for multigrid methods.

**Surface triangulation for molecules.** Surface triangulation of molecular models has been discussed in details in the literature [41–44]. It is noteworthy that there is no strong agreement on which surface definition should be used to best represent the boundary of a molecule. The three competing definitions correspond to the van der Waals surface (i.e. the boundary of the union of balls representing the molecule, centered at the atomic positions, with radii equal to the van der Waals radii of these atoms), the accessible surface (i.e. the surface generated by the center of a probe sphere rolling on the van der Waals surface), and the molecular surface (i.e. the lower envelope generated by the rolling sphere). MSMS developed by Sanner et al. [42] is probably the most widely used program to triangulate the molecular surface. It has the drawback to generate some poorly shape triangles (such as small area triangles) [29]. The problems faced by MSMS reflect the difficulties encountered when trying to generate quality triangular meshes for large complicated biomolecules. The main reason is the smoothness or rather lack of smoothness of the existing surface models. The van der Waals surface and solvent accessible surface are not smooth at the intersections of the spheres. The molecular surface is smooth in most cases but it contains self-intersections that results in cusps on the surface. To overcome these singularity problems, Edelsbrunner [45] introduced the skin surface based on the framework of the Voronoi diagram and Delaunay triangulation of a set of weighted points. The molecular skin surface is an application of this concept to biomolecules, for which the weighted points are the atoms with weights equal to their vdW radii augmented with the radius of the water probe, usually chosen to be 1.4 Å. The skin surface is smooth, free of self-intersection, and can be parameterized and triangulated with good quality and deformed freely with smooth transitions [45].

We developed efficient algorithms for triangulating the skin surface. Our approach is based on an advancing front technique combined with a Delaunay mesh generation technique. In particular, we use the restricted union of balls, which is a set of balls centered on the surface, to generate an  $\varepsilon$ -sampling of the skin surface. A subset  $P$  of points on a skin surface is an  $\varepsilon$ -sampling if for every point  $x$  on the surface there is a point  $p \in P$  such

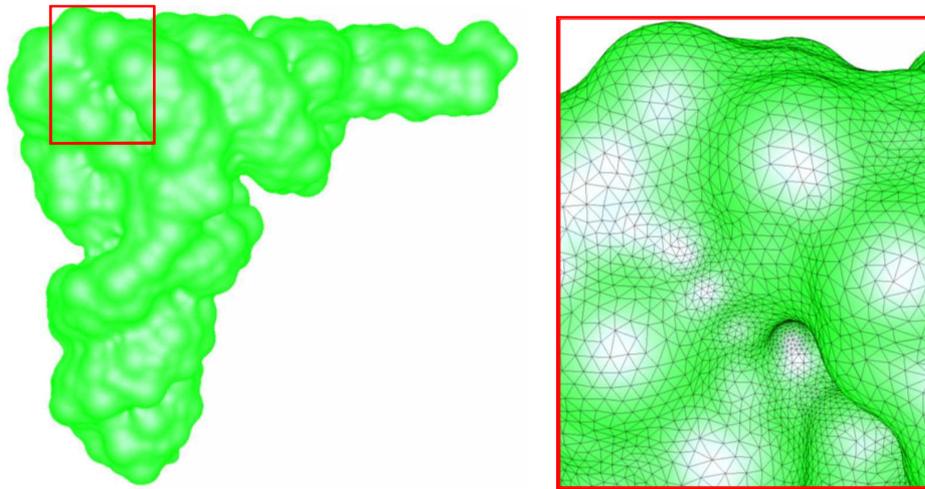


Figure 2: *Molecular skin surface mesh*. Left: The molecular skin surface of a tRNA molecule. Right: Zoomed in view of the partial mesh in the red box in the left. All the triangles in the mesh have a minimum angle of at least  $23^\circ$ .

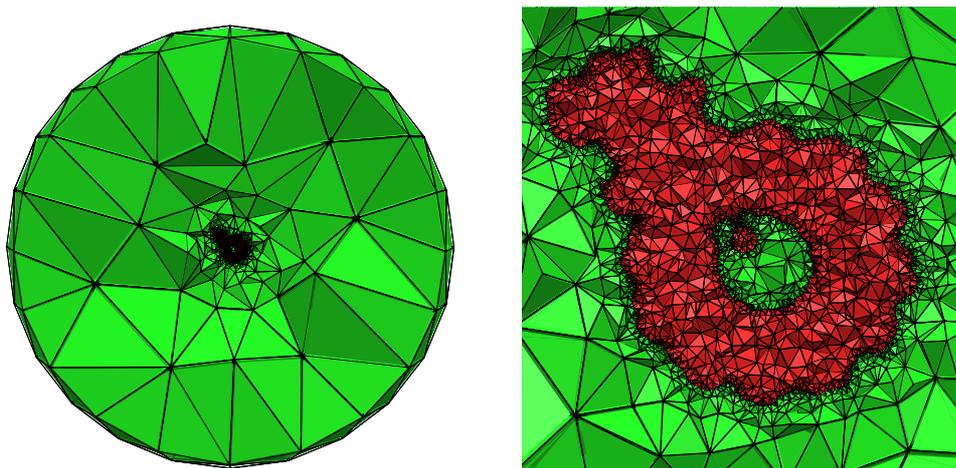


Figure 3: *Tetrahedral skin mesh of a protein*. A tetrahedral skin mesh decomposes the region bounded by a sphere into a collection of tetrahedra that conforms to the surface mesh of an ATP synthase, whose structure was solved by NMR spectroscopy and deposited in the Protein Structure DataBank (PDB, [51]) under the name 1C17. The left panel shows a partial tetrahedral mesh obtained by cutting the original mesh by a plane. Red tetrahedra are inside the molecule while green tetrahedra are outside. The right panel shows a zoomed view of the center part of the left panel. The tetrahedral mesh consists of 235,892 vertices and 1,402,859 tetrahedra. Each tetrahedron in the mesh has a guaranteed quality.

that the distance between  $x$  and  $p$  is at most  $\varepsilon q(x)$ , in which  $\varepsilon$  is a constant smaller than 1 and  $q(x)$  is the radius of the maximum curvature at the point  $x$ . In order to guarantee the topological correctness of the surface triangulation, we restrict the  $\varepsilon$  to be smaller or equal to a constant value of 0.279 [46], resulting in a dense, adaptive set of sample points on the skin surface. Our algorithm generates these sample points incrementally and guar-

antees a lower bound on the distance of each sample point to their nearest neighbors. In parallel, it builds their Delaunay triangulation. A specified subset of this Delaunay triangulation, namely the restricted Delaunay triangulation, forms a quality triangulation of the skin surface when the algorithm terminates. A Delaunay triangle is called restricted Delaunay to a skin surface if the dual Voronoi edge of the triangle intersects the skin surface [47]. A robust and efficient implementation of this algorithm was developed and will be distributed under the GNU license. Fig. 2 shows the molecular skin model of a tRNA molecule. Our implementation of these algorithms triangulates this surface within 31 seconds on a Mac PC with a 2.66GHz Dual Core Intel Xeon processor. The triangulation consists of 58,946 vertices and 117,896 triangles. Every triangle in the mesh has a minimum angle greater than  $23^\circ$ , which implies a high quality triangular mesh.

**Tetrahedral mesh generation.** While a surface mesh is enough for boundary element methods, a conforming tetrahedral mesh is needed to apply the finite element methods to solve the PB equation. Here conforming mesh refers to the property that the surface mesh is an exact subset of the faces in the tetrahedral mesh that divides the tetrahedra into two sets, those interior to the boundary, and those exterior. Various methods such as the advancing front method, the octree method and the Delaunay based methods have been applied to the problem of generating biomolecular tetrahedral meshes [48–50]. In particular, it is worth citing the tool chain setup by Lu et al. [33] that automates the generation of conforming tetrahedral meshes through three steps, namely generation of the surface mesh, smoothing of the latter to remove low quality triangles, and generation of a conforming mesh.

In the line of the algorithms we developed for meshing molecular skin surfaces, we developed a new algorithm to generate quality tetrahedral meshes for the volumes inside and outside the skin surface, where the outside volume is bounded by a large sphere. The algorithm builds an initial Delaunay mesh including the surface mesh and applies the Delaunay refinement to improve its quality. In particular, the algorithm inserts the circumcenters of badly shaped tetrahedra with a priority parameterized by the value of the distance function defined by the surface. The algorithm achieves an upper bound on radius-edge ratio of the tetrahedral mesh after the refinement. All the slivers are removed in a post processing procedure. The algorithm terminates with guarantees on the tetrahedral quality and an accurate approximation of the original surface boundary [52]. See Fig. 3 for an example. In Fig. 4, we plot the computing time needed for generating the quality surface and tetrahedral meshes of 20 proteins with sizes varying from 30 to 500 residues. It is noteworthy that there is no obvious relationship between the computing time and the number of the atoms. The computing time required to generate a tetrahedral mesh is mainly defined by the number of triangles in the surface mesh, which is a function of the surface area and the curvature of the surface, with no direct relationship to the number of atoms.

**Hierarchical tetrahedral meshes.** Methods that solve the PB equations using a multi-grid approach rely on a hierarchy of embedded meshes. The PBE is first solved on the

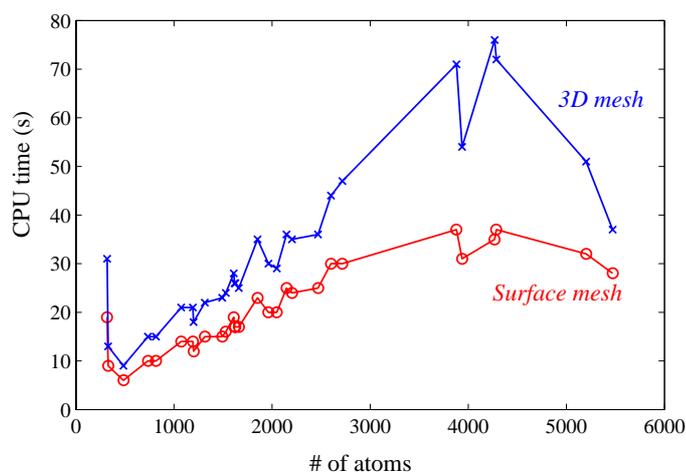


Figure 4: *Computing the molecular skin surface mesh and the corresponding conforming tetrahedral mesh.* The CPU time required to compute the skin surface mesh (lower curve) and the tetrahedral mesh (upper curve) are plotted versus the number of atoms in the molecule. Note that the CPU time for tetrahedral mesh generation is the total time, i.e. it includes the time required to generate the surface skin mesh. Computations were performed on a Mac PC with a 2.66GHz Dual Core Intel Xeon processor.

coarser grid, and the corresponding solution is used as a starting point for the next level. This procedure is propagated through the whole hierarchy such that an accurate solution is derived on the finer mesh. Key to the success of the method is the quality of the hierarchy of meshes: if a coarse grid reflects poorly the geometry of the system, the solution derived on it will be a poor approximation for solving the PBE on the following finer grid. We have developed an algorithm for generating a hierarchy of meshes starting from the tetrahedral mesh conforming with the skin surface of the molecule, in which each coarser mesh maintains the topology of the skin surface, and contains high quality tetrahedra. We aim to generate 4 levels of coarse meshes such that each level of the coarse mesh decreases the number of vertices by a factor of 2, which effectively reduces the number of the vertices from the finest mesh to the coarsest mesh by a factor of 8. Our algorithm simplifies the surface meshes using three operations, namely point deletion, edge flipping and point insertion. In the point deletion operation, we check the link conflict property in order to preserve the topology of the simplified mesh [53]. As a result, the operations performed on a fine mesh decrease its resolution and at the same time guarantee the topology and the quality of the simplified surface mesh. Once a coarse surface mesh has been derived, we apply our tetrahedral mesh generation algorithm described above to construct a quality tetrahedral mesh. See Fig. 5 for an example of the quality surface and volumetric meshes at two different levels of resolution. Our hierarchical tetrahedral meshes have a number of advantages that will facilitate fast and accurate multigrid PBE solvers. Firstly, the quality of both the surface triangulations and tetrahedral meshes is guaranteed. Secondly, the interface in the tetrahedral mesh is an accurate approximation of the molecular boundary. In particular, all the boundary points are on the surface. Thirdly, our meshes are Delaunay meshes. Finally, the meshes are

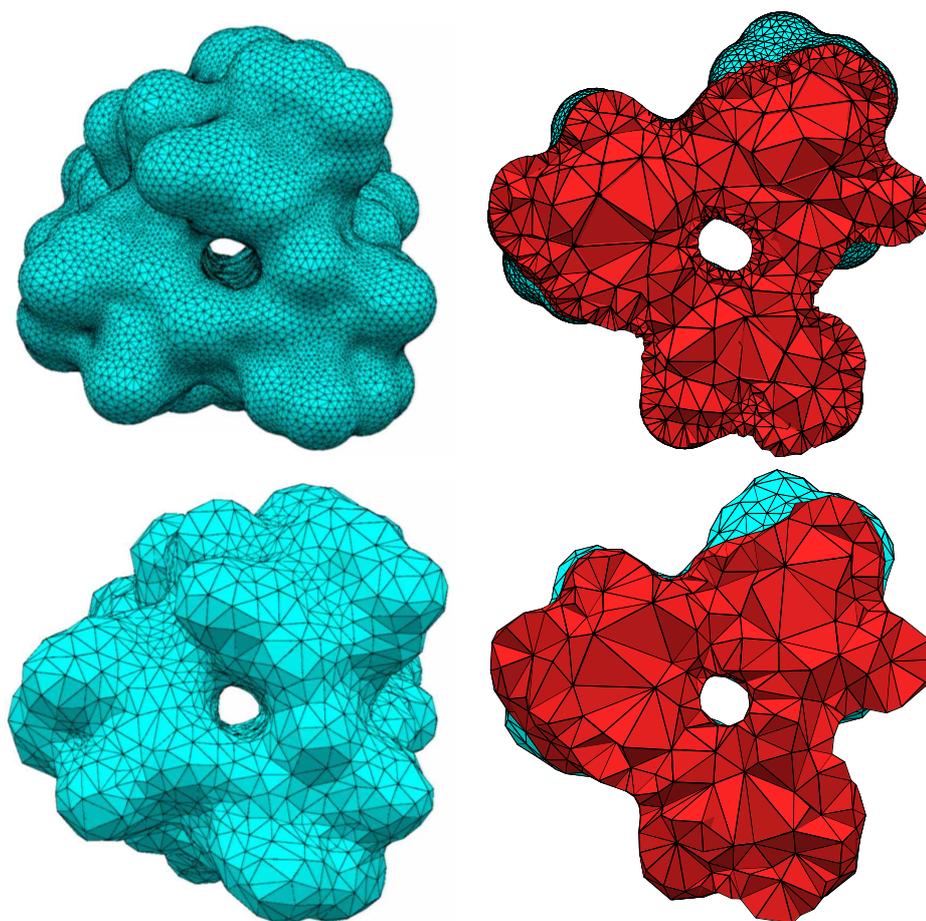


Figure 5: A two level hierarchical skin tetrahedral mesh for a protein. A fine and coarse tetrahedral meshes of the volume of a protein (PDB ID 1M1W) are showed in the top row and bottom row respectively. The left panels show the boundary of the tetrahedral meshes while the right panels show the interior tetrahedra using a cutting plane. The fine tetrahedral mesh contains 141,229 tetrahedra while the coarse mesh contains 23,161 tetrahedra.

adaptive to the geometry. Our algorithms can be easily extended to control the size of the mesh elements through a user-defined sizing field, which enables different mesh resolutions within different regions both on the surface and inside the molecular surface.

## 4 Conclusion

Computational studies of biomolecules and their interactions in solution need to take into account the effect of the environment around the molecules. Numerous approaches were developed to perform this task, and among those the ones based on the Poisson-Boltzmann equation are currently the most popular.

Despite being based on a mean field approximation with a continuum constant dielectric for the solvent and a Boltzmann distribution for the ions, the electrostatics potential obtained by solving the PB equation provides reasonable estimates of solvation energies. There are limitations however to the applications of the PB equation, as it neglects ion sizes, fluctuations in the concentration of the ions and solvent based on the presence of charges, as well as ion-ion and ion-solvent correlations. In a nutshell, PB methods work reasonably well for biomolecules with low charge density in monovalent salt solutions in low concentrations [54], and can give incorrect results for highly charged biomolecules (such as RNA or DNA), or in more concentrated solutions, or in the presence of multivalent ions. We have shown in this paper that there are many recent developments in the theory of the Poisson-Boltzmann equation that attempt to solve these problems, using lattice field theory to deal with ion size, and Langevin dipoles to account for density fluctuation of the solvent. It should be noted that the resulting modified Poisson-Boltzmann equations still ignore correlations (a drawback of using a mean field approach), and as such will still be of limited use when analyzing solutions of multivalent ions. There is therefore still room for further advances in theory.

The standard Poisson-Boltzmann equation is a non-linear elliptic equation of second order, with discontinuous coefficients, that cannot be solved analytically for any system with a shape more complicated than a sphere or a cylinder. Numerical methods for solving the PBE have been constantly refined over the past twenty years. The critical element there is probably that the field of application of PB equation is constantly evolving, with people interested in larger and larger systems, with more complicated environments. Interestingly, no consensus method has appeared and existing competing softwares still advertise their scientific computing approach as being the fastest and most robust. Solver developments focus on parallel implementation to reduce computing time and to extend to size of the problems they can handle, as well as on the iterative techniques used to solve the underlying discretized algebraic equations. While these developments are essential, we believe that efforts should also be put in developing better methods for embedding the molecular system under study onto the mesh used to discretize the solution. We have shown in this paper that geometry provides the required tools to perform such a task. We have proposed to represent the boundary of a molecule using its skin [45], i.e. a smooth surface devoid of self intersection. We have developed algorithms for generating surface triangular mesh of the skin, as well as conforming tetrahedral meshes. We anticipate that these geometric methods will prove essential for the development of more accurate PBE solvers.

## Acknowledgments

The authors acknowledge support from the National Institute of Health under contract GM080399.

## References

- [1] W. Kauzmann. Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.*, 14:1–63, 1959.
- [2] M.F. Perutz. Electrostatics effects in proteins. *Science*, 201:1187–1191, 1978.
- [3] J.R. Fresco, A. Adams, R. Ascione, D. Henley, and T. Lindahl. Tertiary structure in transfer ribonucleic acids. *Cold Spring Harbor Symp. Quant. Biol.*, 31:527–537, 1966.
- [4] C. Sagui and T.A. Darden. Molecular dynamics simulations of biomolecules: long-range electrostatics effects. *Annu. Rev. Biophys. Biomol. Struct.*, 28:155–179, 1999.
- [5] M.T. Neves-Petersen and S.B. Petersen. Protein electrostatics: a review of the equations and methods used to model electrostatics equations in biomolecules -applications in biotechnology. *Biotechnol. Annu. Rev.*, 9:315–395, 2003.
- [6] P. Koehl. Electrostatics calculations: latest methodological advances. *Curr. Opin. Struct. Biol.*, 16:142–151, 2006.
- [7] V. Tsui and D.A. Case. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers*, 56:275–291, 2000.
- [8] D. Bashford and D.A. Case. Generalized Born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.*, 51:129–152, 2000.
- [9] T. Simonson. Macromolecular electrostatics continuum models and their growing pains. *Curr. Opin. Struct. Biol.*, 11:243–252, 2001.
- [10] N.A. Baker. Poisson-Boltzmann methods for biomolecular electrostatics. *Curr. Opin. Struct. Biol.*, 383:217–224, 2004.
- [11] J.W. Ponder and D.A. Case. Force-fields for protein simulations. *Adv. Protein Chem.*, 66:27–85, 2003.
- [12] D.J. Tobias. Electrostatics calculations: recent methodological advances and applications to membranes. *Curr. Opin. Struct. Biol.*, 11:253–261, 2001.
- [13] G. Brannigan, L.C. Lin, and F.L. Brown. Implicit solvent simulation models for biomembranes. *Eur. Biophys. J.*, 35:104–124, 2005.
- [14] K. Andresen, R. Das, H.Y. Park, H. Smith, L.W. Kwok, J.S. Lamb, E.J. Kirkland, D. Herschlag, K.D. Finkelstein, and L. Pollack. Spatial distribution of competing ions around DNA in solution. *Phys. Rev. Lett.*, 93:248103, 2004.
- [15] A. Vitalis, N.A. Baker, and J.A. McCammon. ISIM: a program for grand canonical Monte Carlo simulations of the ionic environment of biomolecules. *Mol. Sim.*, 30:45–61, 2004.
- [16] R.D. Coalson and A. Duncan. Statistical mechanics of a multipolar gas: A lattice field theory approach. *J. Phys. Chem.*, 100:2612–2620, 1996.
- [17] I.D. Borukhov, D. Andelman, and H. Orland. Steric effects in electrolytes: A modified Poisson-Boltzmann equation. *Phys. Rev. Lett.*, 79:435–438, 1997.
- [18] S. Tsonchev, R.D. Coalson, and A. Duncan. Statistical mechanics of charged polymers in electrolyte solutions: a lattice field theory approach. *Phys. Rev. E*, 60:4257–4267, 1999.
- [19] R.D. Coalson, A.D. Walsh, A. Duncan, and N. Bien-Tal. Statistical mechanics of a Coulomb gas with finite size particles: a lattice field theory approach. *J. Chem. Phys.*, 102:4584–4594, 1995.
- [20] V.B. Chu, Y. Bai, J. Lipfert, D. Herschlag, and S. Doniach. Evaluation of ion binding to DNA duplexes using a size-modified Poisson-Boltzmann theory. *Biophys. J.*, 93:3202–3209, 2007.
- [21] A. Abrashkin, D. Andelman, and H. Orland. Dipolar Poisson-Boltzmann equation: ions and dipoles close to charge interfaces. *Phys. Rev. Lett.*, 99:77801, 2007.
- [22] C. Azuara, E. Lindahl, P. Koehl, H. Orland, and M. Delarue. Incorporating dipolar solvents

- with variable density in the Poisson-Boltzmann treatment of macromolecule electrostatics. *Nucl. Acids. Res.*, 34:W34–W42, 2006.
- [23] C. Tanford and J. G. Kirkwood. Theory of protein titration curves. I. general equations for impenetrable spheres. *J. Am. Chem. Soc.*, 79:5333–5339, 1957.
- [24] C.J. Benham. The cylindrical Poisson-Boltzmann equation. I. transformations and general solutions. *J. Chem. Phys.*, 79:1969–1973, 1983.
- [25] P. Knabner and L. Angermann. *Numerical Methods for Elliptic and Parabolic Partial Differential Equations*. Springer, 2003.
- [26] M. Holst, N. A. Baker, and F. Wang. Adaptive multilevel finite element solution of the Poisson-Boltzmann equation i. algorithms and examples. *J. Comp. Chem.*, 21:1319–1342, 2000.
- [27] J. Liang and S. Subramaniam. Computation of molecular electrostatics with boundary element methods. *Biophys. J.*, 73:1830–1841, 1997.
- [28] N. A. Baker, D. Sept, J. Simpson, M. J. Holst, and J. A. McCammon. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. (USA)*, 98:10037–10041, 2001.
- [29] B. Lu, X. Cheng, J. Huang, and J. A. McCammon. Order N algorithm for computation of electrostatic interactions in biomolecular systems. *Proc. Natl. Acad. Sci. (USA)*, 103:19314–19319, 2006.
- [30] B. Lu, D. Zhang, and J. A. McCammon. Computation of electrostatic forces between solvated molecules determined by the Poisson–Boltzmann equation using a boundary element method. *J. Chem. Phys.*, 122:214102, 2005.
- [31] A.H. Boschitsch, M.O. Fenley, and H.-X. Zhou. Fast boundary element method for the linear Poisson-Boltzmann equation. *J. Phys. Chem. B.*, 106:2741–2754, 2002.
- [32] A.H. Boschitsch and M.O. Fenley. Hybrid boundary element and finite difference method for solving the nonlinear poisson-boltzmann equation. *J. Comp. Chem.*, 25:935–955, 2004.
- [33] B. Lu, Y. C. Zhou, G. A. Huber, S. D. Bond, M. J. Holst, and J. A. McCammon. Electrodifffusion: A continuum modeling framework for biomolecular systems with realistic spatiotemporal resolution. *J. Chem. Phys.*, 127:135102, 2007.
- [34] A. Nicholls and B. Honig. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. *J. Comp. Chem.*, 12:435–445, 1991.
- [35] B. A. Luty, M. E. Davis, and J. A. McCammon. Solving the finite-difference non-linear Poisson-Boltzmann equation. *J. Comp. Chem.*, 13:1114–1118, 2004.
- [36] O. Axelsson and V.A. Barker. *Finite Element solution of boundary value problems- Theory and computations*. Academic Press, London, 1984.
- [37] C. M. Cortis and R. A. Friesner. Numerical solution of the Poisson-Boltzmann equation using tetrahedral finite-element meshes. *J. Comp. Chem.*, 18:1591–1608, 1998.
- [38] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition*. SIAM, Philadelphia, PA, 1994.
- [39] M. J. Holst and F. Saied. Numerical solution of the nonlinear Poisson-Boltzmann equation: Developing more robust and efficient methods. *J. Comp. Chem.*, 16:337–364, 1995.
- [40] H. Oberoi and N. M. Allewell. Multigrid solution of the nonlinear Poisson-Boltzmann equation and calculation of titration curves. *biophys. Biophys. J.*, 65:48–55, 1993.
- [41] N. Akkiraju and H. Edelsbrunner. Triangulating the surface of a molecule 71 (1996), 5-22. *Discrete Appl. Math.*, 71:5–22, 1996.
- [42] M. F. Sanner, A. J. Olson, and J.-C. Spehner. Fast and robust computation of molecular surfaces. In *SCG '95: Proceedings of the eleventh annual symposium on Computational geometry*,

- pages 406–407, New York, NY, USA, 1995. ACM Press.
- [43] A. Varshney, F. P. Brooks, Jr. William, and V. Wright. Linearly scalable computation of smooth molecular surfaces. *IEEE Computer Graphics and Applications*, 14:19–25, 1994.
  - [44] P. Laug and H. Borouchaki. Molecular surface modeling and meshing. *Proceedings 10th International Meshing Roundtable*, pages 31–41, 2001.
  - [45] H. Edelsbrunner. Deformable smooth surface design. *Discrete Comput. Geom.*, 21:87–115, 1999.
  - [46] H. Cheng, T. K. Dey, H. Edelsbrunner, and J. Sullivan. Dynamic Skin Triangulation. *Discrete Comput. Geom.*, 25:525–568, 2001.
  - [47] H. Cheng and X. Shi. Quality mesh generation for molecular skin surfaces using restricted union of balls. In *Proceedings of IEEE Visualization*, pages 399–405, 2005.
  - [48] E. Seveno. Towards an adaptive advancing front method. *Proceedings 6th International Meshing Roundtable*, pages 349–360, 1997.
  - [49] R. Schneiders. An algorithm for the generation of hexahedral element meshes based on an octree technique. *Proceedings 6th International Meshing Roundtable*, pages 183–194, 1997.
  - [50] H. Edelsbrunner. Triangulations and meshes in computational geometry. *Acta Numerica*, pages 133–213, 2000.
  - [51] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, J.N. Shindyalov, and P.E. Bourne. The Protein Data Bank. *Nucl. Acids. Res.*, 28:235–242, 2000.
  - [52] H.-L. Cheng and X. Shi. Quality tetrahedral mesh generation for macromolecules. In *ISAAC*, pages 203–212, 2006.
  - [53] T. Dey, H. Edelsbrunner, S. Guha, and D. Nekhayev. Topology preserving edge contraction. *Publ. Inst. Math. (Beograd) (N.S.)*, 66:23–45, 1999.
  - [54] C. Holm, P. Kékicheff, and R. Podgornik, editors. *Electrostatic Effects in Soft Matter and Biophysics*, volume 46 of *NATO Science Series II - Mathematics, Physics and Chemistry*. Kluwer Academic Publishers, 2001.