Computational Methods for Biomolecular Electrostatics

Feng Dong,
Dept. of Biochemistry and Molecular Biophysics, Center for Computational Biology, Washington University in St. Louis. 700 S. Euclid Ave., Campus Box 8036, St. Louis, MO 63110. E-mail: fdong@ccb.wustl.edu

Brett Olsen, and
Dept. of Biochemistry and Molecular Biophysics, Center for Computational Biology, Washington University in St. Louis. 700 S. Euclid Ave., Campus Box 8036, St. Louis, MO 63110. E-mail: bnolsen@artsci.wustl.edu

Nathan A. Baker*
Dept. of Biochemistry and Molecular Biophysics, Center for Computational Biology, Washington University in St. Louis. 700 S. Euclid Ave., Campus Box 8036, St. Louis, MO 63110. Tel: 314-362-2040, Fax: 314-362-0234, E-mail: baker@ccb.wustl.edu. Corresponding author.

Abstract
An understanding of intermolecular interactions is essential for insight into how cells develop, operate, communicate and control their activities. Such interactions include several components: contributions from linear, angular, and torsional forces in covalent bonds, van der Waals forces, as well as electrostatics. Among the various components of molecular interactions, electrostatics are of special importance because of their long range and their influence on polar or charged molecules, including water, aqueous ions, and amino or nucleic acids, which are some of the primary components of living systems. Electrostatics, therefore, play important roles in determining the structure, motion and function of a wide range of biological molecules. This chapter presents a brief overview of electrostatic interactions in cellular systems with a particular focus on how computational tools can be used to investigate these types of interactions.

I. Introduction
An understanding of intermolecular interactions is essential for insight into how cells develop, operate, communicate and control their activities. Such interactions include several components: contributions from linear, angular, and torsional forces in covalent bonds, van der Waals forces, as well as electrostatics. Among the various components of molecular interactions, electrostatics are of special importance because of their long range and their influence on polar or charged molecules, including water, aqueous ions, and amino or nucleic acids, which are some of the primary components of living systems. Electrostatics, therefore, play important roles in determining the structure, motion and function of a wide range of biological molecules; for example, in ion-induced RNA folding (Cole et al., 1972; Römer and Hach, 1975; Stein and Crothers, 1976), low pH protein unfolding, electrostatic steering effects on biomolecule-small molecule binding, protein-biomembrane interactions (Wang et al., 2001a; Rauch et al., 2002; Wang et al., 2002), protein-nucleic acid binding (Bajaj et al., 1990; Record et al., 1991; Senear and Batey, 1991; Overman and Lohman, 1994), etc. This chapter presents a brief overview of electrostatic interactions in cellular systems (Sec. II) with...
a particular focus on how computational tools can be used to investigate these types of interactions (Secs. III and IV).

II. Electrostatics in cellular systems

Electrostatic interactions are ubiquitous for any system of charged or polar molecules, such as biomolecules in their aqueous environment. For example, proteins are made up of 20 types of amino acids, 11 of which are charged or polar in neutral solution. Nucleic acids contain long stretches of negative charges from the phosphate groups in nucleotides. Finally, sugars and related glycosaminoglycans can possess some of the highest charge densities of any biomolecules due to the presence of numerous negative functionalities, including carboxylate and sulfate groups. We will focus on a few specific examples of electrostatic interactions in cellular systems: biomolecule-ion, biomolecule-ligand, and biomolecule-biomolecule interactions. Each of these interactions will be discussed in more detail in the following sections.

II.A. Biomolecule-ion interactions

In cellular settings, biomolecules are immersed in solution along with water, ions, and numerous other small molecules and macromolecules. Ions influence biomolecular processes and interactions in several different ways, including long-range screening, site-specific ion binding, and preferential hydration effects. Long-range screening is a phenomenon in which the strength of electrostatic interactions within and between biomolecules is reduced by the presence of aqueous ions. This is a non-specific ion effect and is described well, at low salt charge and concentration, by Debye-Hückel theory (Debye and Hückel, 1923) and the related implicit solvent models described in Sec. III.B. In site-specific ion binding, ions interact with biomolecules by binding to specific sites in a manner similar to ligand binding (see Sec. II.B) (Draper et al., 2005). Preferential hydration or Hofmeister effects are species-specific competitions between ions and water for binding to non-specific sites on biomolecules (Hofmeister, 1888; Boström et al., 2006; Collins, 2006). This competition is between weak biomolecule-solvent and biomolecule-ion interactions and therefore observed only at very high salt concentrations (Eisenberg, 1976; Anderson and Record Jr., 2004). A similar effect involves competition between ionic species around charged biomolecules (Moore and Lohman, 1994; Reuter et al., 2005). Note that, although these effects can be important, preferential hydration and ion-ion competition are not routinely considered in simulations, mainly due to limitations in current computational methodology. While there is active work in improving the theoretical and computational treatment of these effects (Boström et al., 2003; Shimizu, 2004; Shimizu and Smith, 2004; Broering and Bommarius, 2005; Zhou, 2005; Boström et al., 2006), they are currently beyond the scope of this chapter.

Ion-induced RNA folding (Römer and Hach, 1975; Stein and Crothers, 1976; Cech and Bass, 1986; Dahm and Uhlenbeck, 1991; Misra and Draper, 2000, 2001; Draper et al., 2005) provides an excellent example of many of the ion-biomolecule interactions discussed above. RNA folding in the absence of salt is quite unfavorable due to a number of negative charges along its phosphodiester backbone. Bringing these negative charges together into a compact structure introduces a large energetic barrier to RNA folding. Positive ions promote folding by reducing the repulsion between these negative charges. However, some ions are more effective than others; for example, millimolar concentrations of Mg^{2+} can stabilize RNA tertiary structures that are only marginally stable in solution with a high concentration of monovalent cations, such as Na^{+} or K^{+} (Cole et al., 1972; Romer, 1975; Stein and Crothers, 1976). Accurately modeling the ion-RNA interactions is essential to explain this phenomenon. A major obstacle in modeling ion-RNA interactions is the presence of numerous different ion environments (Draper et al., 2005). Each environment is dominated by the different type of ion-biomolecule interactions described above and requires different approaches to evaluating the energies. For
example, experimental results for Mg\(^{2+}\) effects on tRNA\(^{\text{Phe}}\) folding can be modeled successfully while only considering long-range screening effects (Misra and Draper, 2000). However, the diffusive Mg\(^{2+}\) ion description provided by this model is not sufficient to describe folding of a 58-nt rRNA fragment. Instead, one Mg\(^{2+}\) ion must be explicitly included at a specific binding site (Misra and Draper, 2001). A comprehensive theoretical framework of ion-RNA interactions that accounts for the overall ion dependence of RNA folding is the aim of current RNA folding studies (Draper et al., 2005).

II.B. Biomolecule-ligand and -biomolecule interactions

Biomolecule-substrate recognition is central to nearly all biomolecular processes, including signal transduction, enzyme cooperativity, and metabolic regulation. The biomolecular binding process, from a kinetic perspective, can be reduced to two steps: diffusional association to form an initial encounter complex and non-diffusional rearrangement to form the fully bound complex. The diffusional association places an upper limit on the overall binding rate; so-called “perfect” enzymes operate at this diffusion-limited rate. Electrostatic forces have an important influence on biomolecular diffusional association: their long-range nature enables them to attract the substrate to its binding partner and orient the substrate properly for binding (Gabdoulline and Wade, 2002). It has been established that, for many biomolecular complexes, electrostatic interactions can significantly affect bimolecular association rates (Law et al., 2006). For example, by using Brownian dynamics (BD) simulations (Ermak and McCammon, 1978; Northrup et al., 1984) (Sec. IV.F) to calculate diffusional association rates, Gabdoulline and Wade demonstrated that, for fast-associating protein pairs, electrostatic interactions enhance association and are the dominant forces determining the rate of diffusional association (Radic et al., 1997; Gabdoulline and Wade, 2001). Using related methods, Sept et al. demonstrated the role of electrostatic interactions in determining the rates and polarity of actin polymerization (Sept et al., 1999; Sept and McCammon, 2001).

Electrostatic interactions also play an important role in determining thermodynamics of binding; i.e., binding affinity (Novotny and Sharp, 1992; Schreiber and Fersht, 1993, 1995; Zhu and Karlin, 1996; Chong et al., 1998; Sheinerman et al., 2000; Norel et al., 2001; Rauch et al., 2002). Substrate binding allows the formation of (potentially) favorable charge-charge interactions between the substrate and target, as well as stabilizing specific salt-bridges and hydrogen bonds (Schreiber and Fersht, 1993, 1995; Chong et al., 1998). However, at the same time, charges on the molecular binding surface must shed their bound water in order to allow close binding. This loss of water, or desolvation, is generally energetically unfavorable and offsets the favorable interactions formed upon binding. The binding affinities, from an electrostatic point of view, are determined by balance of these two energetic contributions (Xu et al., 1997; Lee and Tidor, 2001; Sheinerman and Honig, 2002; Russell et al., 2004; del Álamo and Mateu, 2005). Systematic studies of protein pairs, such as barnase and barstar (Schreiber and Fersht, 1993, 1995; Frisch et al., 1997; Dong et al., 2003), and fasciculin-2 (Radic et al., 1997), as well as protein kinase A and balanol (Wong et al., 2001), have shown that charged and polar residues at the protein-protein interfaces play important roles in binding energetics. Similarly, Sept et al. have demonstrated an important role for electrostatics in determining microtubule structure and stability (Sept et al., 2003). Finally, Wang et al. (Wang et al., 2004) have demonstrated that nonspecific electrostatic interactions can provide a driving force for recruitment of proteins to intracellular membranes, an important step in signal transduction.

However, despite the role of electrostatics in protein-protein interactions, it is important to realize that the total interaction is also strongly influenced by shape complementarity at the protein-protein interface as well as nonpolar contributions to offset the penalties of desolvation (Janin and Chothia, 1990; Lo Conte et al., 1999; Ma et al., 2003; Vasker, 2004).
III. Models for biomolecular solvation and electrostatics

As described above, computer simulations can provide atomic-scale information on energetic and dynamic contributions to biomolecular structure and interactions. However, the capabilities of computer simulations are limited by the accuracy of the underlying models describing atomic interactions and also by the computational expense of adequately exploring all the relevant conformations of the biomolecule and surrounding water and ion. Therefore, most models of biomolecular solvation and electrostatics make a trade-off between these opposing considerations of atomic accuracy and computational expense.

A variety of computational methods have been developed for studying electrostatic interactions in biomolecular systems. Popular methods for understanding electrostatic interactions in these systems can be loosely classified into two categories (see Fig. 1): explicit solvent methods (Burkert and Allinger, 1982; Jorgensen et al., 1983; Sagui and Darden, 1999; Ponder and Case, 2003; Horn et al., 2004), which treat the solvent in full atomic detail, and implicit solvent methods (Davis and McCammon, 1990b; Honig and Nicholls, 1995; Roux and Simonson, 1999; Roux, 2001; Baker, 2005b; Baker et al., 2006), which represent the solvent through its average effect on solute.

III.A. Explicit solvent methods

Explicit solvent methods offer a very detailed description of biomolecular solvation. In explicit solvent methods, interactions between mobile ions, solvent, and solute atoms are typically described by molecular mechanics force fields (Wang et al., 2001b; Ponder and Case, 2003) which use classical approximations of quantum mechanical energies to describe the Coulombic (electrostatic), van der Waals, and covalent (bond, angle) interactions. Explicit solvent methods have the obvious advantage of offering the full details of solvent-solute and solvent-solvent interactions. These details can affect some aspects of biomolecular interactions. For example, the explicit representation of solvent structure can qualitatively change the detailed features of protein side chain interactions (Masunov and Lazaridis, 2003). Similarly, Yu et al have demonstrated the importance of including first shells of solvation to correctly describe the interaction of salt bridges in solution (Yu et al., 2004).

However, the explicit solvent methods are computationally expensive. In order to extract meaningful thermodynamic and kinetic parameters, all the numerous conformations of biomolecules, as well as the solvent and ions, must be explored. The extra degrees of freedom associated with the explicit solvent and ions dramatically increase the computational cost of explicit solvent methods and limit the temporal and spatial scales of biomolecular simulations.

III.B. Implicit solvent methods

Implicit solvent methods have become popular, although lower-accuracy, alternatives to the computationally-expensive explicit solvent approaches (Davis and McCammon, 1990b; Honig and Nicholls, 1995; Gilson, 2000; Roux, 2001; Baker, 2005b; Baker et al., 2006). In implicit solvent methods, the molecules of interest are treated explicitly while the solvent is represented by its average effect on the solute (Roux and Simonson, 1999). Solute-solvent interactions are described by solvation energies; i.e., the free energy of transferring the solute from a vacuum to the solvent environment of interest (e.g., water at a certain ionic strength). This process is shown in more detail in Fig. 2, which consists of three steps: (1) solute charges are gradually reduced to zero in vacuum, (2) the uncharged solute is inserted into the solvent, and (3) solute charges are gradually increased back to their normal values in solvent. The free energy change in step (2) is called the nonpolar solvation energy. The sum of the energies associated with the step (1) and (3) is called the “charging” or polar solvation energy and represents the solvent’s effect on the solute charging process. In general, polar and nonpolar solvation terms act in
opposing directions; nonpolar solvation favors compact structures with small areas and volumes, while polar solvation favors the maximum solvent exposure for all polar groups in the solute.

III.B.1. Nonpolar solvation—One popular approximation for the nonpolar solvation free energy assumes a linear dependence between the nonpolar solvation energy $G_{\text{nonpolar}}^{\text{solv}}$ and the solvent-accessible surface area (SASA), $A$, (Chothia, 1974; Eisenberg, 1976; Spolar et al., 1989; Sharp et al., 1991; Wesson and Eisenberg, 1992; Massova and Kollman, 2000; Swanson et al., 2004):

$$G_{\text{nonpolar}}^{\text{solv}} = \gamma \cdot A$$  \hspace{1cm} (1)

where $\gamma$ is a “surface tension” which is typically chosen to reproduce the nonpolar solvation free energy of alkanes (Sharp et al., 1991; Simonson and Brunger, 1994; Sitkoff et al., 1994b) or model sidechain analogues (Eisenberg and McLachlan, 1986; Wesson and Eisenberg, 1992). The surface tension parameter may assume a single global value used for all atom types or different values may be assigned for each different type of atom. Although SASA methods have enjoyed surprising success, they are also subject to several caveats, including widely varying choices of surface tension parameter (Chothia, 1974; Eisenberg and McLachlan, 1986; Sharp et al., 1991; Sitkoff et al., 1994a; Elcock et al., 2001) as well as inaccurate descriptions of the detailed aspects of nonpolar solvation energy (Gallicchio and Levy, 2004), peptide conformations (Su and Gallicchio, 2004), and protein nonpolar solvation forces (Wagoner and Baker, 2004). Some of these problems have been fixed by new models which include the small but important attractive van der Waals interactions between solvent and solute (Gallicchio et al., 2000; Gallicchio et al., 2002; Gallicchio and Levy, 2004; Wagoner and Baker, 2006) as well as repulsive solvent-accessible volume terms (Wagoner and Baker, 2006).

III.B.2. Polar solvation—Implicit solvent methods have been used to study polar solvation and electrostatics for over 80 years, starting with work by Born on ion solvation (Born, 1920), Linderström-Lang (Linderström-Lang, 1924) and Tanford and Kirkwood (Tanford and Kirkwood, 1957) on protein titration, Manning on ion distributions surrounding nucleic acids (Manning, 1978), Flanagan et al (Flanagan et al., 1981) on the pH dependence of hemoglobin dimer assembly, and Warwicker and Watson (Warwicker and Watson, 1982) on the electrostatic potential of realistic protein geometries. Although they can be considerably different in their details and implementation, implicit solvent models generally treat the solvent as a high dielectric continuum, the aqueous ions as a diffuse cloud of charge, and the solute as a fixed array of point charges that are embedded in a lower dielectric continuum. Despite the limitations of these assumptions, implicit solvent models often give a good coarse-grained description of solvation energetics and have enjoyed widespread use over recent years.

Regardless of the particular type of implicit solvent model, the behavior of electrostatic interactions is generally determined by a few basic properties of the system, illustrated in Fig. 3: the charges, radii, and “dielectric constant” of the solute; the charges and radii of aqueous ionic species; and the radii and dielectric constant of the solvent. The relationship of these specific parameters to solvation energies and forces will be described in more detail in Secs. III.C and III.D.

III.C. Poisson-Boltzmann methods

The Poisson-Boltzmann (PB) equation is a popular continuum description of electrostatics for biomolecular system. Although there are a number of ways to derive the PB equation based
on statistical mechanics (Holm et al., 2001), the simplest derivation begins with Poisson’s equation (Jackson, 1975; Bockris and Reddy, 1998) (in SI units),

\[-\nabla \cdot \varepsilon(\vec{x}) \nabla \varphi(\vec{x}) = \rho(\vec{x})\]  

(2)

the basic equation for describing the electrostatic potential \(\varphi(\vec{x})\) at point \(\vec{x}\) generated by a charge distribution \(\rho(\vec{x})\) in an environment with a dielectric permittivity coefficient \(\varepsilon(\vec{x})\) (Jackson, 1975; Landau et al., 1982).

The coefficient \(\varepsilon(\vec{x})\) is given by \(\varepsilon(\vec{x}) = \varepsilon_0 \varepsilon_r(\vec{x})\) where \(\varepsilon_0 = 8.8542 \times 10^{-12} \text{ C}^2/(\text{N} \cdot \text{m}^2)\) is the electrostatic permittivity of a vacuum and \(\varepsilon_r(\vec{x})\) is the dielectric coefficient or the relative electrostatic permittivity. The dielectric coefficient \(\varepsilon_r(\vec{x})\) describes the local polarizability of the material; i.e., the generation of local dipole densities in response to the applied fields and changes in charge. The functional form of this coefficient depends on the shape of the biomolecule; \(\varepsilon_r(\vec{x})\) assumes lower values of 2–20 in the biomolecular interior and higher values of approximately 80, the value for water at room temperature, in solvent-accessible regions. The distinction between biomolecular “interior” and “exterior” used to assign dielectric coefficients is imprecise; as a result, a variety of different definitions for the biomolecular surface and dielectric coefficient have been developed (Lee and Richards, 1971; Warwicker and Watson, 1982; Connolly, 1985; Im et al., 1998; Grant et al., 2001).

In order to continue the derivation of the PB equation, we assume the charge distribution \(\rho(\vec{x})\) includes two contributions: the solute charges \(\rho_f(\vec{x})\) and the aqueous “mobile” ions \(\rho_m(\vec{x})\). The solute charge distribution is generally described by a collection of \(N\) point charges located at each solute atom’s position \(\vec{x}_i\) and scaled by that atom’s charge \(Q_i\); i.e., the solute charge distribution is the summation of a set of delta functions. Neglecting explicit interactions between the aqueous ions (Holm et al., 2001), the mobile charges are modeled as a continuous “charge cloud” described by a Boltzmann distribution (McQuarrie, 1976). For \(m\) ion species with charges \(q_j\), bulk concentrations \(c_j\) and steric potential \(V_j(\vec{x})\) (a potential that prevents biomolecule-ion overlap), the mobile ion charge distribution is

\[\rho_m(\vec{x}) = \sum_{j=1}^{m} c_j q_j \exp[-q_j \varphi(\vec{x})/k_B T - V_j(\vec{x})/k_T],\]

where \(k_B\) is Boltzmann’s constant and \(T\) is the absolute temperature. Combining both the solute and ion charge distributions with the Poisson equation, Eq. (2), gives the full PB equation:

\[-\nabla \cdot \varepsilon(\vec{x}) \nabla \varphi(\vec{x}) = \sum_{i=1}^{N} Q_i \delta(\vec{x} - \vec{x}_i) + \sum_{j=1}^{m} c_j q_j \exp[-q_j \varphi(\vec{x})/k_B T - V_j(\vec{x})/k_T].\]  

(3)

A common simplification is that the exponential term \(\exp[-q_j \varphi(\vec{x})/k_B T]\) can be approximated by the linear term in its Taylor series expansion \(-q_j \varphi(\vec{x})/k_B T\) for \(|q_j \varphi(\vec{x})/k_B T| \ll 1\). With this linearization and by assuming the steric occlusions are the same for all ion species \((V_j = V\) for all \(j))\), Eq. (3) reduces to the linearized PB equation:

\[-\nabla \cdot \varepsilon(\vec{x}) \nabla \varphi(\vec{x}) + \varepsilon(\vec{x}) \kappa^2(\vec{x}) \varphi(\vec{x}) = \sum_{i=1}^{N} Q_i \delta(\vec{x} - \vec{x}_i)\]  

(4)

where \(\kappa^2(\vec{x})\), related to a modified inverse Debye-Hückel screening length (Debye and Hückel, 1923), is given by

\[\kappa^2(\vec{x}) = \exp[-V(\vec{x})/k_B T] \cdot 2 l e^2/k_B T \varepsilon(\vec{x}),\]  

(5)
where $I$ is the ionic strength and $e_c$ is the unit electric charge.

Once the PB equation is solved, the electrostatic potential is known for the entire system. Given this potential, the electrostatic free energy can be evaluated by a variety of integral formulations (Sharp and Honig, 1990; Gilson, 1995; Micu et al., 1997). The simplest, for the linearized PB equation, is

$$G_{el} = \frac{1}{2} \sum_{i=1}^{N} Q_i \varphi(x_i)$$

(6)

It is also possible to differentiate integral formulations of the electrostatic energy with respect to atomic position to obtain the electrostatic or polar solvation force on each atom (Gilson et al., 1993; Im et al., 1998).

Analytical solutions of the PB equation are not available for biomolecules with realistic shapes and charge distributions. Numerical methods for solving PB equation were first introduced by Warwicker and Watson (Warwicker and Watson, 1982) to obtain the electrostatic potential at the active site of an enzyme. The most common numerical techniques for solving the PB equation are based on discretization of the domain of interest into small regions. Those methods include finite difference (Davis and McCammon, 1989; Nicholls and Honig, 1991; Holst and Saied, 1993; Holst and Saied, 1995; Baker et al., 2001), finite element (Cortis and Friesner, 1997a, 1997b; Baker et al., 2000; Baker et al., 2001; Dyshlovenko, 2002), and boundary element methods (Zauhar and Morgan, 1988; Juffer et al., 1991; Allison and Huber, 1995; Bordner and Huber, 2003; Boschitsch and Fenley, 2004), all of which continue to be developed to further improve the accuracy and efficiency of electrostatics calculations in the numerous biomolecular applications described below. The major software packages that can be used to solve the PB equation are listed in Table 1. Many of these packages are also used for visualization of the electrostatic potential around biomolecules. Such visualization can provide insight into biomolecular function and highlight regions of potential interest. Fig. 4 shows examples of the visualization of electrostatic potential calculated with APBS (Baker et al., 2001) and visualized with VMD (Humphrey et al., 1996).

III.D. Simpler models

In addition to the PB methods, simpler approximate models have also been constructed for continuum electrostatics, including distance-dependent dielectric functions (Leach, 2001; MacKerell and Nilsson, 2001), analytic continuum methods (Schaeler and Karplus, 1996), and generalized Born models (Still et al., 1990; Osapay et al., 1996; Dominy and Brooks III, 1999; Bashford and Case, 2000; Onufriev et al., 2002). Among these simpler methods, generalized Born (GB) is currently the most popular. The GB model was introduced by Still et al. in 1990 and subsequently refined by several other researchers (Osapay et al., 1996; Dominy and Brooks III, 1999; Bashford and Case, 2000; Onufriev et al., 2002). The model shares the same continuum representation of the solvent as the Poisson or PB theories. However, the GB model is based on the analytical solvation energy obtained from the solution of the Poisson equation for a simple sphere (Born, 1920). The biomolecular electrostatic solvation free energy is approximated by a modified form of the analytical solvation energy for a sphere (Still et al., 1990):

$$\Delta G_{solv}^{el} \approx -\frac{1}{2} \left( 1 - \frac{1}{\varepsilon_{sol}} \right) \sum_{i,j} Q_i Q_j f_{ij}^{GB}$$

(7)
where the self terms as $i = j, f^{GB}_{ij}$, are the “effective Born radii” and the cross terms as $i \neq j, f^{GB}_{ij}$, are the effective interaction distances. The most common form of $f^{GB}_{ij}$ (Still et al., 1990) is

$$f^{GB}_{ij} = \left[ r_{ij}^2 + R_i R_j \exp \left(-r_{ij}^2 / 4 R_i R_j \right) \right]^{\frac{1}{2}}$$

(8)

where $R_i$ are the accurate radii of the atoms and $r_{ij}$ are the distance between atom $i$ and $j$. Efficiently and accurately calculating effective radii $R_i$ is essential for GB methods. “Perfect” GB radii, which reproduce the atom $i$’s self energy obtained by solving Poisson equation for the biomolecule-solvent system with only atom $i$ charged, have demonstrated the ability to accurately follow the results of more detailed models such as PB (Onufriev et al., 2002). However using such “perfect” radii does not directly provide any computational advantage over solving the Poisson equation. In the absence of perfect radii for every biomolecular conformation, GB methods fail to capture some aspects of molecular structure included in more detailed models, such as the PB equation. Nonetheless, GB methods have become increasingly popular because of their computational efficiency.

III.E. Limitations of implicit solvent methods

Although implicit solvent methods offer simpler descriptions of the system and greater computational efficiency, it is important to recall that these reductions of complexity and effort are obtained at the cost of substantial simplification of the description of the solvent. In particular, implicit solvent methods are only capable of describing non-specific interactions between solvent and solute. In general, explicit solvent methods should be used wherever the detailed interactions between solvent and solute are important, such as solvent finite size effects in ion channels (Nonner et al., 2001), strong solvent-solute interactions (Bhattacharrya et al., 2003), strong solvent coordination of ionic species (Figueirido et al., 1994; Yu et al., 2004), and saturation of solvent polarization near a membrane (Lin J.-H. et al., 2002). Similarly, as mentioned earlier in the context of RNA-ion interactions, implicit descriptions of mobile ions can also become questionable in some cases, such as high ion valency or strong solvent coordination, specific ion-solute interactions, and high local ion densities (Holm et al., 2001), where the ions interact with each other or with the solute directly.

IV. Applications

In the previous section, we discussed the basic concepts behind the computational tools that can be used to simulate the electrostatic interactions in cellular systems. In this section, we will illustrate the use of these methods, especially PB methods, to deal with the various biomolecular problems.

IV.A. Solvation free energy

As mentioned in Sec. III.B, the solvation free energy is the free energy of transferring a solute from a uniform dielectric continuum (a constant dielectric) to an inhomogeneous medium (a low dielectric solute surrounded by a high dielectric solvent), which is often divided into two terms: a nonpolar term and polar term. The nonpolar term is usually estimated using either SASA or the improved methods discussed in Sec. III.B.1. For the polar term, as shown in Fig. 5, two PB calculations are usually performed: (1) calculating biomolecular electrostatic free energy $G_{el}^1$ in a homogeneous medium with a constant dielectric equal to the solute dielectric coefficient and (2) calculating the biomolecular electrostatic free energy $G_{el}^2$ in the inhomogeneous medium of interest; e.g., a protein in aqueous medium. The polar contribution to the solvation free energy is then given by:
Additionally, solving the PB equation twice helps to cancel the numerical artifacts which arise from the discretization used in finite difference and finite element methods, i.e., it reduces the grid size dependence. Although, in most cases, polar solvation free energy alone is not sufficient to explain the biological phenomenon, it is the foundation for the other, more complex, electrostatic calculations described below.

IV.B. Electrostatic free energy

The total electrostatic free energy can be easily obtained from the polar solvation free energy by adding the electrostatic free energy of the biomolecule in a homogeneous medium with a constant dielectric equal to the solute dielectric coefficient using Coulomb’s law:

$$G_{el} = \Delta G_{solv}^{\text{polar}} + G_{\text{coul}}$$  \hspace{1cm} (10)$$

where

$$G_{\text{coul}} = \sum_{i,j} \frac{1}{4\pi\varepsilon_0\varepsilon_p} \cdot \frac{Q_i Q_j}{r_{ij}}$$  \hspace{1cm} (11)$$

where $r_{ij}$ is the distance between charge $Q_i$ and $Q_j$ and $\varepsilon_p$ is the dielectric coefficient of the solute. The resulting electrostatic free energies are the basis for nearly all applications of continuum electrostatics methods to biomolecular systems.

As a specific example, such electrostatic free energy calculations have been used to study the electrostatic sequestration of phosphatidylinositol 4,5-bisphosphate (PIP$_2$) by membrane-adsorbed basic peptides (Wang et al., 2004). PIP$_2$ is a very important lipid in the cytoplasmic leaflet of the plasma membrane (De Camilli et al., 1996; Toker, 1998; Raucher et al., 2000; Martin, 2001; Payrastre et al., 2001; Cantley, 2002; Irvine, 2002; McLaughlin et al., 2002; Yin and Janmey, 2003) with a net charge of $-4e$ on the lipid head group. By calculating the electrostatic free energy of laterally sequestering a PIP$_2$ lipid from a region of “bulk” membrane to a region in the vicinity of a membrane-adsorbed basic peptide, Wang et al demonstrated that nonspecific electrostatic interactions provide a driving force for the lateral sequestration of PIP$_2$ by membrane-adsorbed basic peptides (Wang et al., 2001a; Rauch et al., 2002; Wang et al., 2002; Wang et al., 2004). Such lateral sequestration of PIP$_2$ is thought to contribute to the regulation of PIP$_2$ function by controlling its accessibility to other proteins (Laux et al., 2000; McLaughlin et al., 2002).

IV.C. Folding free energies

Biomolecular native (folded) structure is very important for proper performance of their biological functions. However, accurately determining the mechanism by which electrostatic interactions affect the stability of biomolecular native structure is still a challenging experimental and computational question. The electrostatic contribution to the biomolecular folding stability is usually defined as the difference in electrostatic free energy between folded ($G_{\text{el}}^{\text{folded}}$) and unfolded ($G_{\text{el}}^{\text{unfolded}}$) states:

$$\Delta G_{\text{folding}}^{\text{el}} = G_{\text{el}}^{\text{folded}} - G_{\text{el}}^{\text{unfolded}}$$  \hspace{1cm} (12)$$

If $\Delta G_{\text{folding}}^{\text{el}} < 0$, from the electrostatic point of view, the folded structure is more stable than the unfolded structure. If $\Delta G_{\text{folding}}^{\text{el}}$ reduces in response to a mutation, i.e., if

$$\Delta \Delta G_{\text{folding}}^{\text{el}} = \Delta G_{\text{folding}}^{\text{el,m}} - \Delta G_{\text{folding}}^{\text{el,w}} < 0$$

this mutation makes folded protein more stable. This
method has been widely used to study electrostatic contribution to protein folding stabilities through mutations that involve charged or polar residues. For example, Bacillus caldolyticus cold shock protein (Bc-Csp) is a thermophilic protein, which differs from Bacillus subtilis cold shock protein B (Bs-CspB), its mesophilic homolog, in 11 of its 66 residues (Mueller et al., 2000; Delbruck et al., 2001). Through mutational studies which reduced the sequence differences between these two protein molecules, both experiment (Mueller et al., 2000; Pace, 2000; Perl et al., 2000; Delbruck et al., 2001; Perl, 2001) and PB calculations (Zhou and Dong, 2003) demonstrated that the difference in stability of these two proteins arises mostly from the interactions among three residues: Arg 3, Glu 46, and Leu 66 in Bc-Csp, as compared with Glu 3, Ala 46, and Glu 66 in Bs-CspB. The removal of the repulsion between Glu 3 and Glu 66 and creation of a favorable salt bridge between Arg 3 and Glu 46 are the main reasons that Bc-Csp is more stable than Bs-CspB at higher temperatures. Moreover, the excellent agreement between PB calculations and experimental data (the correlation coefficient is 0.98) implies that electrostatic interactions dominate the thermostability of thermophilic proteins (Zhou and Dong, 2003).

IV.D. Binding free energies

The binding of biomolecules is fundamental to cellular activity. The simplest type of binding energy calculations are performed on the biomolecular complex assuming a rigid conformation; i.e., without any conformational changes upon binding, which is clearly not realistic, but often provides useful initial estimates for relative biomolecular binding affinities. Fig. 6 illustrates the procedure to calculate the polar contribution to the binding free energy \( \Delta G_{\text{binding}}^{\text{el}} \) which is given by

\[
\Delta G_{\text{binding}}^{\text{el}} = G_{\text{complex}}^{\text{el}} - (G_{\text{mol1}}^{\text{el}} + G_{\text{mol2}}^{\text{el}}),
\]

\[
= \left( \Delta G_{\text{complex}}^{\text{solv}} + \Delta G_{\text{complex}}^{\text{coul}} \right) + \left( G_{\text{mol1}}^{\text{solv}} + G_{\text{mol2}}^{\text{solv}} \right) - \left( G_{\text{mol1}}^{\text{coul}} + G_{\text{mol2}}^{\text{coul}} \right),
\]

where \( \Delta G_{\text{binding}}^{\text{solv}} = \Delta G_{\text{complex}}^{\text{solv}} - (\Delta G_{\text{mol1}}^{\text{solv}} + \Delta G_{\text{mol2}}^{\text{solv}}) \) is the polar solvation free energy change upon binding with the \( \Delta G_{\text{mol}}^{\text{solv}} \) values calculated according to Eq. (9) above. The quantity \( \Delta G_{\text{binding}}^{\text{coul}} = G_{\text{complex}}^{\text{coul}} - (G_{\text{mol1}}^{\text{coul}} + G_{\text{mol2}}^{\text{coul}}) \) is the Coulombic free energy change upon binding with components \( \Delta G_{\text{mol}}^{\text{coul}} \) calculated according to Eq. (10) above.

For the more general situation in which biomolecules experience conformational changes during the binding process, MM/PBSA and MM/GBSA methods (Kollman et al., 2000; Swanson et al., 2004) are commonly used to calculate the binding free energy. The nature of this method can be best understood through its acronym: MM stands for the molecular mechanics force fields used to calculate the intramolecular and direct intermolecular contributions to binding free energies; PB and GB refer to the implicit solvent methods used to calculate the electrostatic contributions, and SA stands for solvent accessible surface area (SASA) methods used to calculate the nonpolar contributions to binding free energies.

Binding free energy calculations using continuum solvation models have been successfully performed on many different biomolecular complexes (Eisenberg and McLachlan, 1986; Murray et al., 1997; Misra et al., 1998; Sept et al., 1999; Massova and Kollman, 2000; Wong et al., 2001; Dong et al., 2003; Sept et al., 2003; Wang et al., 2004; Green and Tidor, 2005). As specific examples, binding free energy calculations have been performed to investigate the roles of charged residues at the interface of the barnase (and extracellular ribonuclease) and barstar (a protein inhibitor), which have been a popular test case for both computational (Gabdoulline and Wade, 1997, 1998, 2001; Lee and Tidor, 2001; Steinerman and Honig, 2003).
2002; Dong et al., 2003; Wang and Wade, 2003; Spaar and Helms, 2005; Spaar et al., 2006) and experimental studies of protein-protein interactions (Schreiber and Fersht, 1993, 1995; Frisch et al., 1997). In particular, PB calculations (Dong et al., 2003) successfully reproduced the experimental result (Schreiber and Fersht, 1993, 1995; Frisch et al., 1997) that cross-interface salt-bridges and hydrogen bonds dominate the binding affinities of barnase and barstar (Dong et al., 2003).

IV.E. pKₐ calculations

The presence of ionizable sites, which can exchange protons with their environment, produces pH-dependent phenomena in proteins and has a significant influence on the protein’s function. The correct prediction of protein titration states is important for the analysis of enzyme mechanisms, protein stability, and molecular recognition. As mentioned earlier, efforts have been underway for more than 80 years (Linderström-Lang, 1924; Antosiewicz et al., 1996b; Bastyns et al., 1996; Luo et al., 1998; Nielsen and Vriend, 2001; Fitch et al., 2002; Georgescu et al., 2002; Li et al., 2002; Alexov, 2003; Nielsen and McCammon, 2003; Li et al., 2004; Jensen et al., 2005; Krieger et al., in press) to correctly predict protein titration states and understand the determinants of pKₐ s for amino acids in protein environments (see chapter by Whitten, et al. in this volume).

The free energy change, ∆G, for protonation of a single ionizable site at a given pH may be written as (Linderström-Lang, 1924; Tanford and Kirkwood, 1957)

\[ \Delta G = G_{\text{protonated}}^{\text{el}} - G_{\text{de-protonated}}^{\text{el}} = (k_b T \ln 10) \cdot (pH - pK_a) \]  

where \( pK_a = -\log_{10} K_a \) and \( K_a = \frac{[H^+][A^-]}{[HA]} \) is the equilibrium constant for dissociation of proton \( H^+ \) and it’s conjugate site \( A^- \), \( k_B \) is Boltzmann’s constant; and \( T \) is the absolute temperature. A widely-used assumption in pKₐ predictions is that any pKₐ differences of an ionizable site when located in a protein versus in a model compound are solely determined by the difference in the electrostatic free energy required to protonate that site in the protein versus the model compound. Thus, the pKₐ of the single ionizable site in protein is given by

\[ pK_a = pK_{a0} - \Delta \Delta \ln \frac{\Delta G_{\text{protein}}}{\Delta G_{\text{mod, et}}} \]  

where \( \Delta \Delta G = \Delta G_{\text{protein}} - \Delta G_{\text{mod, et}} \) and \( pK_{a0} \) is the pKₐ of the isolated ionizable site in the model compound. In general, proteins have multiple ionizable sites and the protonation energetics of these different sites are coupled, as discussed below. Single site pKₐ predictions have successfully reproduced measured pKₐ s for different residues in several different proteins (Dong and Zhou, 2002; Dong et al., 2003) and therefore have some predictive power. However, a more complete treatment of ionizable residues in proteins considers the coupling between all the ionizable sites. There are a number of techniques for treating such coupling (Tanford and Roxby, 1972; Beroza et al., 1991; Antosiewicz et al., 1996a; Antosiewicz et al., 1996b; Bashford, 2004) and thereby determining the complete titration state of the protein. Unfortunately, such methods are complex and are beyond the scope of the current discussion.

IV.F. Biomolecular association rates

Brownian dynamics (BD) calculations are popular methods to simulate the diffusional motion between two solute particles and thereby estimate the rate of diffusion-controlled rate of binding between two molecules (Ermak and McCammon, 1978; Northrup et al., 1984). Given the importance of electrostatic interactions in biomolecular association, BD simulations are usually combined with continuum electrostatic calculations to provide the most accurate estimates of diffusion-limited encounter rates (Allison and McCammon, 1985; Davis and
McCammon, 1990a; Ilin et al., 1995; Madura et al., 1995; Sept et al., 1999; Gabdoulline and Wade, 2001, 2002). Such calculations have been used in numerous diffusional encounter rate calculations, including simulations of small molecule interactions with enzymes (Allison and McCammon, 1985; Madura and McCammon, 1989; Davis et al., 1991; Sines et al., 1992; Luty et al., 1993; Tan et al., 1993; Elcock et al., 1995; Radic et al., 1997; Tara et al., 1998), simulations of protein-protein encounter (Gabdoulline and Wade, 1997; Elcock et al., 1999; Sept et al., 1999; Elcock et al., 2001; Gabdoulline and Wade, 2001; Spaar et al., 2006), as well as functional assessment of differences in protein electrostatics (Livesay et al., 2003).

V. Conclusion and future directions

Computer simulation is becoming an increasingly routine way to help with drug discovery or other applications requiring a detailed understanding of molecular interactions. A correct understanding of the energetic interactions within and between biomolecules is essential for such simulations. Among the various contributions to these energies, electrostatic interactions are of special importance because of their long range and strength. In this chapter, we have covered some of the computational methods that are currently available to model the electrostatic interactions biomolecular systems, ranging from highly detailed explicit solvent methods to simpler PB and GB methods. There are several reviews available on all of these methods which provide a more in-depth discussion of the different solvation approaches. The reviews of Ponder and Case (Ponder and Case, 2003) as well as the texts of Becker et al (Becker et al., 2001), Leach (Leach, 2001), and Schlick (Schlick, 2002) provide excellent background on explicit solvent methods. There also are several reviews available for implicit solvent methods see (Honig and Nicholls, 1995; Roux and Simonson, 1999; Bashford and Case, 2000; Simonson, 2003; Baker, 2005a), including a particularly thorough treatment by Lamm (Lamm, 2003), a discussion of current PB limitations by Baker (Baker, 2005b) and an up-to-date discussion by Feig and Brooks (Feig and Brooks III, 2004) of current challenges for GB methods. For additional background and more in-depth discussion of the principles and limitations of continuum electrostatics, interested readers should see the general volume by Jackson (Jackson, 1975) and Landau et al. (Landau et al., 1982), the electrochemistry text of Bockris et al. (Bockris and Reddy, 1998), the colloid theory treatise by Verwey and Overbeek (Verwey and Overbeek, 1999), or the excellent collection of condensed matter electrostatics articles assembled by Holm et al. (Holm et al., 2001).

Acknowledgements

The authors would like to thank Baker group members for their reading of this manuscript. This work was supported by National Institutes of Health grant R01 GM069702.

VI. Reference


Irvine RF. Nuclear lipid signaling. SciSTKE 2002;150:1–12.


Krieger E, Nielsen JE, Spronk CAEM, Vriend G. Fast empirical pKa prediction by Ewald summation. J Mol Graph Model. in press


*Methods Cell Biol. Author manuscript; available in PMC 2008 June 10.*


Methods Cell Biol. Author manuscript; available in PMC 2008 June 10.
Zhou HX. Interactions of macromolecules with salt ions: an electrostatic theory for the Hofmeister effect.
Zhou HX, Dong F. Electrostatic contributions to the stability of a thermophilic cold shock protein.
Zhu ZY, Karlin S. Clusters of charged residues in protein three-dimensional structures. Proc Natl Acad
Figure 1.
A schematic comparison of implicit and explicit solvent models: (a) in the implicit solvent model, a low dielectric solute is surrounded by a continuum of high dielectric solvent; (b) in the explicit solvent model, solvent is represented by discrete water molecules.
Figure 2.
A thermodynamic cycle illustrating the biomolecular solvation process. The steps are (1) uncharging the biomolecule in vacuum; (2) transferring the uncharged biomolecule from vacuum to solvent; (3) charging the biomolecule back to its normal value in solvent. The nonpolar solvation free energy is the free energy change in step (2). The polar solvation free energy is the sum of the free energy changes in steps (1) and (3).
Figure 3.
Description of the terms in the Poisson-Boltzmann equation: (a) the dielectric permittivity coefficient $\varepsilon(\mathbf{x})$ is much smaller inside the biomolecule than outside the biomolecule with a rapid change in value across the solvent-accessible biomolecular surface; (b) the ion-accessibility parameter $\kappa^2(\mathbf{x})$ is proportional to the bulk ionic strength outside the ion-accessible biomolecular surface; (c) the biomolecular charge distribution is defined as the collection of point charges located at the center of each atom.
Figure 4.
Examples of the visualization of the balanol electrostatic potential in the binding site of protein Kinase A as calculated by APBS (Baker et al., 2001) and visualized with VMD (Humphrey et al., 1996).
Figure 5.
Schematic of a polar solvation free energy calculation; in the initial state, the dielectric coefficient is a constant throughout the entire system and equal to the solute dielectric coefficient; in the final state, the dielectric coefficient is inhomogeneous and smaller in the solute than in the bulk solvent.
Figure 6.
Thermodynamic cycle illustrating the standard procedure for calculating the electrostatic contribution to the binding free energy of a complex with rigid-body. The steps are (1) transfer the isolated molecule from a inhomogeneous dielectric into a homogeneous dielectric, the free energy change is $-\Delta G_{\text{solv, mol. 1}}$; (2) form the complex from isolated molecules in a homogeneous dielectric, the free energy change is $\Delta G_{\text{coul, binding}}$; (3) transfer the complex from the homogeneous dielectric into the inhomogeneous dielectric, the free energy change is $\Delta G_{\text{el, complex}}$. $$\Delta G_{\text{binding}} = -\Delta G_{\text{solv, mol. 1}} - \Delta G_{\text{solv, mol. 2}} + \Delta G_{\text{coul, binding}} + \Delta G_{\text{el, complex}}$$
Table 1

<table>
<thead>
<tr>
<th>Software package</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>APBS (Baker et al., 2001)</td>
<td><a href="http://apbs.sf.net/">http://apbs.sf.net/</a></td>
</tr>
<tr>
<td>Delphi (Rocchia et al., 2002)</td>
<td><a href="http://trantor.bsc.columbia.edu/delphi/">http://trantor.bsc.columbia.edu/delphi/</a></td>
</tr>
<tr>
<td>MEAD (Bashford, 1997)</td>
<td><a href="http://www.scripps.edu/mh/bashford/">http://www.scripps.edu/mh/bashford/</a></td>
</tr>
<tr>
<td>ZAP (Grant et al., 2001)</td>
<td><a href="http://www.eyesopen.com/products/toolkits/zap.html">http://www.eyesopen.com/products/toolkits/zap.html</a></td>
</tr>
<tr>
<td>UHBD (Madura et al., 1995)</td>
<td><a href="http://mccammon.ucsd.edu/uhbd.html">http://mccammon.ucsd.edu/uhbd.html</a></td>
</tr>
<tr>
<td>Jaguar (Cortis and Friesner, 1997a, 1997b)</td>
<td><a href="http://www.schrodinger.com/">http://www.schrodinger.com/</a></td>
</tr>
<tr>
<td>CHARMM (MacKerell et al., 1998)</td>
<td><a href="http://yuri.harvard.edu">http://yuri.harvard.edu</a></td>
</tr>
<tr>
<td>Amber (Luo et al., 2002)</td>
<td><a href="http://amber.scripps.edu">http://amber.scripps.edu</a></td>
</tr>
</tbody>
</table>