JOURNAL OF MOLECULAR RECOGNITION
J. Mol. Recognit. 2002; **15**: 377–392
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI**:10.1002/jmr.577

Review

The Poisson–Boltzmann equation for biomolecular electrostatics: a tool for structural biology

F. Fogolari*, A. Brigo and H. Molinari

Dipartimento Scientifico Tecnologico, Università degli Studi di Verona, Cá Vignal 1, Strada Le Grazie 15, 37134 Verona, Italy

Electrostatics plays a fundamental role in virtually all processes involving biomolecules in solution. The Poisson–Boltzmann equation constitutes one of the most fundamental approaches to treat electrostatic effects in solution. The theoretical basis of the Poisson–Boltzmann equation is reviewed and a wide range of applications is presented, including the computation of the electrostatic potential at the solvent-accessible molecular surface, the computation of encounter rates between molecules in solution, the computation of the free energy of association and its salt dependence, the study of pKa shifts and the combination with classical molecular mechanics and dynamics. Theoretical results may be used for rationalizing or predicting experimental results, or for suggesting working hypotheses. An ever-increasing body of successful applications proves that the Poisson–Boltzmann equation is a useful tool for structural biology and complementary to other established experimental and theoretical methodologies. Copyright © 2002 John Wiley & Sons, Ltd.

Keywords: **Poisson-Boltzmann; electrostatics; free energy; Browman dynamics; molecular dynamics**

Received 22 April 2002; revised 10 June 2002; accepted 5 July 2002

INTRODUCTION

Electrostatic interactions are found to be relevant for virtually all biomolecular systems and processes. Charged and polar groups are found ubiquitously in biological macromolecules. More than 20% of all aminoacids in globular proteins are ionized under physiological conditions and polar groups are found in more than another 25% sidechains (frequencies based on McCaldon and Argos, 1988, cited by Creighton, 1993). Nucleic acids are among the strongest natural polyelectrolytes (the linear charge density of double-stranded DNA is approximately one electron charge every 1.7 Å; Manning, 1978). The vast majority of biological membranes entail small percentages of charged phospholipids which are likely to be involved in physiological functions (Langner and Kubica, 1999).

The relevance of hydrogen bonds (which are largely due to electrostatics) in determining protein and nucleic acid structures was hypothesized very early in structural biology (Watson and Crick, 1953; Pauling *et al.*, 1951) and confirmed years later, after the first protein structure and DNA molecule had been solved by X-ray crystallography (Kendrew *et al.*, 1958; Wing *et al.*, 1980).

The structures of proteins solved by X-ray crystallography during the following decades led to recognition of the richness of details of electrostatic interactions in protein behaviour, and nowadays there are many interpretations of experimental results in terms of macromolecular electrostatics.

Several reviews have been published to date (Perutz, 1978; Harvey 1989; Davis and McCammon, 1990; Honig and Nicholls 1995; Warshel and Papazyan 1998; Sheinerman *et al.*, 2000; Simonson 2001). The present review focuses on the methodology based on the Poisson–Boltzmann equation to treat electrostatics, and its aim is to present the range of possible applications.

A comprehensive account of experimental results is beyond the scope of the present review, but, before reviewing theoretical methods and applications, it is worth highlighting some of the areas where electrostatics has been most useful.

Protein structural stability

Protein structures are generally built using secondary structure elements and motifs which are held together by hydrogen bonds (Branden and Tooze, 1995). Important hydrogen bond interactions, providing so called 'helix-capping', are located at helix termini where nonbonded amide protons (at the N-terminus) and carbonyl oxygens (at the C-terminus) give rise to rather high charge densities, although in recent years the presence of flanking hydrophobic residues has also been highlighted (Aurora and Rose, 1998). The same helix termini may also interact favorably with charged groups, and this is reflected in the statistical occurrences of positively and negatively charged residues at

*Correspondence to: F. Fogolari, Dipartimento Scientifico Tecnologico, Facoltá di Scienze MM. FF. NN., Universitá di Verona, Cá Vignal 1, Strada Le Grazie 15, 37134 Verona, Italy.
E-mail: fogolari@sci.univr.it
Contract/grant sponsor: Italian MURST cofin 2000.

Abbreviations used: BD, Brownian dynamics.

C-terminal and N-terminal helix positions, respectively. Matthews has used this principle in order to design enhanced stability mutants of T4 Lysozyme (Nicholson *et al.*, 1988).

Designed salt bridges at the surface of the protein were not found to contribute to protein stability probably due to the balance of solvation and entropic contributions (Matthews, 1993). Contradicting opinions have been reported for buried electrostatic interactions, mostly due to differences in the analysis performed (e.g. Waldburger *et al.*, 1995). A large number of buried salt bridges are, however, found in proteins from thermophyles (Karshikoff and Ladenstein, 2001). A general principle is that networks of interactions, rather than single interactions, are required to balance desolvation energies. This is evident for salt bridges (Musafia *et al.*, 1995; Karshikoff and Ladenstein, 2001), but it also appears to be true for hydrogen bond interactions (e.g. Stickle *et al.*, 1992).

Indeed, in alanine scanning experiments performed on BPTI by Kim and coworkers (Yu *et al.*, 1995), mutation of Asn 43, participating in a complex hydrogen bond, was one of the few mutations preventing folding of the protein, and similarly mutation of Asn 24 gave rise to a large destabilization free energy (2.2 kcal/mol). Polar hydrogen-bonded interactions were found in this study to be as important as hydrophobic interactions and the destabilization upon residue replacement with alanine has an even higher dependence on the loss in the buried surface of polar residues rather than hydrophobic residues.

Apart from the stability of the protein, electrostatics might be important for determining the folding pathway. It has been proposed that early formation of secondary structure elements should prevent the polypeptide chain exploring the whole energy landscape (see for a review Mirny and Shaknovich, 2001). Long-range electrostatic forces are good candidates for orienting chain random motions towards correctly folded structure. Evidence has been presented by Oliveberg and Fersht (1996) that native electrostatic interactions are already formed in intermediates of barnase. Furthermore, for the kinetics of folding, the absence of electrostatic repulsion between helices in a leucine zipper has been linked to extremely fast folding (Durr *et al.*, 1999).

Molecular dynamics simulations of unfolding, which could in principle clarify the role of electrostatics in folding, are often conducted at nonphysiological temperatures, and therefore conclusions can not be mapped in a safe way to the simulated systems at much lower temperatures.

Although intramolecular hydrogen bonds or salt bridges have been mainly studied, interaction with the solvent, i.e. proper exposure towards the solvent of polar residues, might also be important in determining protein structure. Statistical preferences for solvent exposure for polar residues are not as great as the preference for the inner core of proteins for hydrophobic residues (this is seen in analyses similar to Miyazawa and Jernigan, 1996). In this respect, it is, however, interesting that in mutagenesis experiments directed towards surface residues for domain II of *Pseudomonas exotoxin*, greater sensitivity to chymotrypsin was found for a number of mutants (Kasturi *et al.*, 1992), thus witnessing the relevance to structural stability of exposed residues.

Enzyme catalysis

The origin of the lowering of reaction energy barrier in enzyme catalysis has been long debated (Bruice and Benkovic, 2000; Warshel, 1998; Cleland et al., 1998). Although phenomena that involve chemical reactions might be safely described only by a more fundamental approach (e.g. Sulpizi *et al.*, 2001; Scrutton *et al.*, 1999), it is reasonable that, among classical concepts, electrostatic effects are suited to describing several aspects of enzyme catalysis.

Electrostatics has been repeatedly proposed by Warshel and coworkers as responsible for the lowering of the energy barrier (Warshel, 1998). Although this proposal is based also on theoretical modellization and not only on experimental data, the following arguments appear rather convincing: (1) the order of magnitude sought to be explained is in the proper range for electrostatic interactions, while it is not for other (local) interactions; (2) the enzyme must provide a proper polar environment in order to compensate for large desolvation energy for charges in the transition states.

Stabilization of the transition state by the polar environment of proteins has received a powerful demonstration by catalytic antibody design by Schultz and co-workers (Schultz, 1988).

Biomolecular recognition

Principles of biomolecular recognition have been sought by analysis of protein–protein interfaces and protein–DNA interfaces found in structural databases. An early study by Janin and Chothia (1990) did not find much difference between the residue composition of protein–protein interfaces and accessible surfaces in proteins. A closer look at an enlarged contact database by Jones and Thornton (1996) reached similar conclusions, although higher hydrophobic residue propensities were found for homodimers rather than for heteromeric complexes. A more recent survey (Lo Conte *et al.*, 1999) confirmed the similarity of interfaces and accessible surfaces in proteins, although large deviations from the average were observed for different types of interfaces, such as protease inhibitor, which showed a sizeable propensity for nonpolar interface vs the more polar antigen–antibody interface.

Within interfaces an average number of 10 hydrogen bonds per interface have been found and one-third of these involve charged residues. On average one salt bridge is found per interface, although this appears to be a slight underestimate as suggested by Honig and coworkers (Sheinerman *et al.*, 2000). The large variability among complexes is to be stressed, for instance for the barnase–barstar complex four salt bridges are found. Similar conclusions have been obtained in the survey by Nussinov and coworkers, who pointed out the different role of electrostatic interactions in protein binding compared with protein folding (Xu *et al.*, 1997).

The presence of both hydrogen bonds and salt bridges at interfaces hints at electrostatic complementarity. However quantification of charge–charge complementarity for a small dataset of 12 protein complexes led to the conclusion that this is not significant. However, in the same study the

complementarity of computed surface electrostatic potential was found to be significant (McCoy *et al.*, 1997). The conclusion calls therefore for a more global approach to the study of protein–protein complexes.

For protein–DNA complexes the polyelectrolytic aspects of DNA make electrostatic interactions very strong (e.g. Manning, 1978). Backbone phosphates are contacted very frequently by several arginines and lysines, while a large number of hydrogen bonds are found between sidechains and bases, with clear biases for certain pairs (Luscombe *et al.*, 2001; Mandel-Gutfreund *et al.*, 1995; Nadassy *et al.*, 1999). The basic character of most DNA-binding protein modules supports the idea that salt bridges are used to provide nonspecific affinity of the protein for DNA.

Another role for electrostatic interactions in protein–DNA recognition is to limit the diffusion of the protein by keeping it in a small volume around DNA, and to provide proper orientation of the protein (von Hippel and Berg, 1989). Also for protein–RNA complexes many salt bridges involving arginines and lysines are found with the backbone phosphates (Draper, 1999). It is interesting that, similar to the study of McCoy *et al.* (1997), it has been noted that electrostatic potential might reveal features not recognizable from simple structural inspection (Sharp *et al.*, 1990).

Biomolecular encounter rates

Electrostatic interactions have a large force constant (332 kcal/mol) and are (in the absence of ions) long ranged. Even when reduced by the presence of ions, energies can be still comparable to thermal energy kT at large distances (say a few tens of Ångstroms) between the interacting molecules. Electrostatic fields have therefore been invoked in order to explain the experimental behavior of encounter rate constants. It must be remembered that other forces, like hydrophobic forces or van der Waals forces, may be important, for example in colloidal systems, although they act at a much shorter range than electrostatic forces.

Although simple models were worked out quite early (e.g. Davis and McCammon, 1990 and references cited therein), the kinetics of biomolecular association depends in a complex way on the electrostatic fields around the molecules and on the shape of the molecules, so that both shape and charge distribution effects should be taken into consideration. In some systems achievement of diffusion control of enzymatic reactions has been proven to be due to electrostatics (e.g. Stroppolo *et al.*, 2001).

MODELS BASED ON THE POISSON–BOLTZMANN

The Poisson–Boltzmann equation was described independently by Gouy (1910) and Chapman (1913) almost a century ago, equating the chemical potential and the force (respectively) acting on small adjacent volumes in an ionic solution between two plates at a different voltage. The approach was generalized many years later by Debye and Hückel (1923), whose work was applied to the theory of ionic solutions and led to the successful interpretation of experimental thermodynamic data. After their work, solutions to the nonlinearized equation were sought by Gronwall *et al.* (1928) in functional terms with powers of the inverse of the dielectric constant as coefficients. Quite early, after the Debye–Hückel model had been formulated, the statistical mechanics bases of the Poisson–Boltzmann framework, cast in terms of potential of mean force, were investigated and criticized (Onsager, 1933; Fowler and Guggenheim, 1939) and some inconsistencies were pointed out.

The approximations at the basis of the formalism were investigated in particular by Kirkwood (1934a), who pointed out that the Poisson–Boltzmann approach is based on the assumption that it is possible to replace the potential of mean force with the mean electrostatic potential. The latter is a somewhat less stringent approximation than any approximation on the potential derivatives. Notwithstanding criticisms, the remarkable success in explaining the behavior of ionic solutions led to the application of the theory in other fields, like colloid chemistry. In this context a dispute on the method of calculating the free energy for interacting particles started in the 1930s and saw later the emergence of a framework still popular today due to Derjaguin and Landau (1941), and Verwey and Overbeek (1948) (hence the name of DLVO theory).

Simple electrostatic models for globular proteins were put forward quite early (Kirkwood, 1934b; Linderstrom-Lang, 1924; Nozaki and Tanford, 1967), while for DNA and other linear polyelectrolytes early cylindrical symmetry models (e.g. Lifson and Katchalski, 1954; Alfrey *et al.*, 1951; Katchalski, 1971; Manning, 1978) were later specialized with proper structural parameters.

These models were based around the Poisson–Boltzmann equation or its linear approximation and led to quite accurate predictions, for instance, of protein titration behavior. Just as an example, Tanford and co-workers were able to detect an anomalously titrating acid residue ($pK_a = 7.5$) in β-lactoglobulin by comparison of model predictions and experimental behaviour (Tanford *et al.*, 1959). The same residue has been shown (40 years later) to act as a conformational switch upon titration (Qin *et al.*, 1998).

Until the early 1980s, theories were based on the conclusions derived on simple shape molecular models, namely spheres for proteins and rods for DNA. During the 1980s several methods were developed in order to solve the Poisson–Boltzmann equation (linearized or not) for any arbitrary shape and the method using the finite difference algorithm became popular through software packages like DelPhi, Grasp and UHBD (see below). This progress allowed study of electrostatics down to atomic detail.

It is worth mentioning that other (faster) ways to treat electrostatic interactions (or in general solvent effects) are available (see e.g. Lazaridis and Karplus, 1999; Simonson, 2001; Roux and Simonson, 1999) and will not be reviewed here. For many of these methods, like for example proper scaling of atomic charges, considering a distance-dependent solvent dielectric, and generalized Born solvent accessible models, the Poisson–Boltzmann equation represents the reference.

In the following section we review the theory of the Poisson–Boltzmann equation.

Theory

In the Poisson–Boltzmann approach to biomolecular electrostatics all solute atoms are often considered explicitly as particles with low dielectric constant (typical of organic molecules) with point partial charges at atomic positions. We will assume that the dielectric constant of the solute (typically a protein or a nucleic acid) is in the range 2–4. The solvent surrounding the solute is taken into account through its high dielectric constant (∼80) and sometimes through a solute–solvent surface tension coefficient (ranging typically from 5 to 70 cal Å^{-2}).

The solute dielectric value does not take into account rearrangements of polar and charged groups with external electric fields, which would lead to much larger dielectric constants (see for details on models and computation Gilson and Honig, 1986; Simonson, 1998). When group reorientations are important and are not included explicitly in the formalism, the dielectric constant of the solute should be raised (Schutz and Warshel, 2001).

We consider first a homogeneous system with dielectric constant ε and with no charges. The electrostatic potential $\psi(\vec{r})$ is described by the Laplace equation:

$$\vec{\nabla}[\vec{\nabla}\psi(\vec{r})] = 0 \tag{1}$$

where the symbolic vector $\vec{\nabla}$ has the meaning of gradient when applied to a scalar, or divergence when applied to a vector. The solution in the volume of interest depends on boundary conditions.

When a charge density $\rho(\vec{r})$ is present, a source term is added in the Laplace equation leading to Poisson equation (in the e.s.u.–c.g.s. units system):

$$\epsilon\vec{\nabla}[\vec{\nabla}\psi(\vec{r})] = -4\pi\rho(\vec{r}) \tag{2}$$

The same equation modifies to the general case of nonhomogeneous medium in order to take into account polarization charges developing at dielectric boundaries (for a detailed derivation of the following equations see Jackson, 1962).

The effect is taken into account through the derivative of the space-dependent dielectric constant:

$$\vec{\nabla}[\epsilon(\vec{r})\vec{\nabla}\psi(\vec{r})] = -4\pi\rho(\vec{r}) \tag{3}$$

While solute charges are located according to the chosen molecular model (e.g. Protein Databank (pdb) structure or theoretical model), it is difficult to hypothesize ionic charge distribution because this is due to the combined effect of solute charges, dielectric distribution and ionic distribution itself. The Poisson–Boltzmann equation accounts for this by making some reasonable assumptions.

In any complex system of interacting particles, the density of a particle at any point $[\sigma(\vec{r})]$ may be expressed relative to the density of the same particle in the absence of interactions with other particles in the system $[\sigma_0\ (\vec{r})]$.

We can therefore write:

$$\sigma(\vec{r}) = g(\vec{r})\sigma_0(\vec{r}) \tag{4}$$

The ratio between the actual density and the average density of the particle $[g(\vec{r})]$ is the distribution function of that particle (Hill, 1956). A useful concept which can be derived from the distribution function is the potential of mean force $w(\vec{r})$ (Hill, 1956) for the particle defined from the following equation:

$$g(\vec{r}) = \mathrm{e}^{[-w(\vec{r})]/kT} \tag{5}$$

In other terms we describe the observed particle distribution through a Boltzmann distribution where the potential of mean force condenses the average effect of the whole system in a single particle potential. The name 'potential of mean force' stems from the fact that the gradient of this potential, with respect to the particle coordinates, gives the mean force acting on the particle.

In an ionic system, ions will preferentially reside in regions where the average potential is high or low according to the sign of their charge. This sentence should be taken with some caution because, in a system with ions, electrostatic interactions are screened and electrostatic effects due to the solute are usually limited at a distance of 10–20 Å. Therefore a large volume (depending on the concentration of the solute) is available for ions where no relevant perturbation in distribution is present. In other words, the tendency towards regions of low potential energy is efficiently balanced by entropy.

There is a crucial but reasonable assumption which must be made in order to obtain an equation for the potential: the ionic potential of mean force is equal to the average electrostatic potential multiplied by the charge of the ion.

When this assumption is formalized in the Poisson equation for nonhomogeneous media we obtain the Poisson–Boltzmann equation (in e.s.u.–c.g.s. units system):

$$\begin{aligned}&\vec{\nabla}[\epsilon(\vec{r})\vec{\nabla}\psi(\vec{r})] = \\ &-4\pi\rho^f(\vec{r}) - 4\pi\sum_i c_i^\infty z_i q \exp\frac{-z_i q\psi(\vec{r})}{kT}\lambda(\vec{r})\end{aligned} \tag{6}$$

where $\rho^f\ (\vec{r})$ includes now only molecular charges, c_i^∞ is the concentration of ion i at an infinite distance from the molecule, z_i is its valency, q is the proton charge, k is the Boltzmann constant, T is the temperature and $\lambda(\vec{r})$ describes the accessibility to ions at point $\vec{r}$.

This equation can be linearized under the assumption that the potential is small:

$$\vec{\nabla}[\epsilon(\vec{r})\vec{\nabla}\psi(\vec{r})] = -4\pi\rho(\vec{r}) + 4\pi\frac{\sum_i c_i^\infty z_i^2 q^2\psi(\vec{r})}{kT}\lambda(\vec{r}) \tag{7}$$

An important parameter is the so called Debye screening constant (k_D), which describes the exponential decay of the potential in the solvent:

$$k_\mathrm{D}^2 = 8\pi\frac{\sum_i c_i^\infty z_i^2 q^2}{2\epsilon kT} = \frac{1}{l_\mathrm{D}^2} \tag{8}$$

l_D is called the Debye length. The linear version has the advantage that it does not lead to an inconsistency found in the nonlinear Poisson–Boltzmann equation. Consider a solution of a 2:1 salt, and consider the average electrostatic potentials $\psi_2(\vec{r})$ and $\psi_1(\vec{r})$ computed when taking the divalent or the monovalent ion as the central ion, respectively. In the nonlinear case the reciprocity condition $z_2\psi_1(\vec{r}) = z_1\psi_2(\vec{r})$ will hardly be met, while the same condition is naturally met in the linear case because the potential is proportional to the source ion valency. When the

reciprocity condition does not hold, the probability of finding a divalent ion at a distance $\vec{r}$ from a monovalent ion changes depending on which ion is used to compute the average potential, which is clearly an inconsistency of the formalism.

It is important to note that in most biologically relevant cases ψ is not very small. Notwithstanding this fact, the solution obtained from the linear Poisson–Boltzmann equation is close to the solution obtained from the nonlinear Poisson–Boltzmann equation, even when the linearization condition does not hold. A thorough comparison between the linear and nonlinear Poisson–Boltzmann equation has been performed (Fogolari *et al.*, 1999) and it has been shown that the onset of appreciable differences between the two treatments is linked to the magnitude of the electric field (therefore basically of the charge density) at the solute–solvent interface.

The solution of the Poisson–Boltzmann equation gives the electrostatic potential thoughout the space. In its most obvious meaning the electrostatic potential gives information, through the Boltzmann distribution, on local concentration of ions, which, as a result of the approximations introduced, may be by far higher than saturation concentration, and therefore should not be taken too literally.

The derivative of the potential around the solute is the electric field which may be relevant for the rates of molecular encounter and recognition.

Free energy from the Poisson–Boltzmann equation

Another important feature of the electrostatic potential is that it is sufficient to compute the free energy for the hypothetical process of charging the solute and the ions in an ionic discharged atmosphere. Although this computation might seem a rather academic exercise, by clever combination of hypothetical processes in thermodynamic cycles we can compute electrostatic free energies for real processes like solvation and association (see below).

The free energy for charging the solute in an ionic atmosphere may be computed by direct integration of the charge (Zhou, 1994; similar to the well known charging processes of Onsager), by considering a variational principle (Reiner and Radke, 1990; Sharp and Honig, 1990; but see also the discussion by Fogolari and Briggs, 1997) or by standard thermodynamics arguments (Marcus, 1955). For the sake of simplicity we will follow the latter approach.

The electrostatic energy (ΔG_{es}) may be written in different forms, obtained by integration by parts (Sharp and Honig, 1990) and entails three different terms: a classical electrostatic energy (charge times potential) term ΔG_{es}^{cp}; a term arising from the mixing of mobile species ΔG_{es}^{mob}; and a term due to the solvent ΔG_{es}^{solv}.

The classical electrostatic energy term is given by:

$$\Delta G_{es}^{cp} = \frac{1}{2}\int_V \rho(\vec{r})\psi\ (\vec{r})\ \mathrm{d}V \tag{9}$$

The mixing entropy of the ions starting from a uniform concentration (when uncharged) to the final concentration is given by the volume integral:

$$\Delta G_{es}^{mob} = kT \int_V \sum\nolimits_i c_i(\vec{r}) \ln \frac{c_i(\vec{r})}{c_i^{\infty}}\ \mathrm{d}V \tag{10}$$

Nonuniform ionic concentration leads to a so called 'osmotic' term, which is calculated as a volume integral:

$$\Delta G_{es}^{solv} = kT \int_V \sum\nolimits_i \left[c_i^{\infty} - c_i(\vec{r})\right]\ \mathrm{d}V \tag{11}$$

The Boltzmann distribution (according to the assumptions made above) is substituted in the equations above for local concentrations:

$$c_i(\vec{r}) = c_i^{\infty} \exp \frac{-z_i q \psi\ (\vec{r})}{kT} \tag{12}$$

Upon linearization several terms cancel out and the only term left is:

$$\Delta G_{es} = \frac{1}{2}\int_v \rho^{f}(\vec{r})\psi\ (\vec{r})\ \mathrm{d}V \tag{13}$$

where the superscript 'f' indicates that only fixed charges (i.e. nonionic charges) contribute the free energy.

Charging free energies are used for computing, for instance, pK_a shifts in proteins (see below). In this case the differences in ionization free energy ensuing from the protein environment are simulated (Fig. 1). All relevant free energies in Fig. 1 are obtained by subtraction of charging free energies.

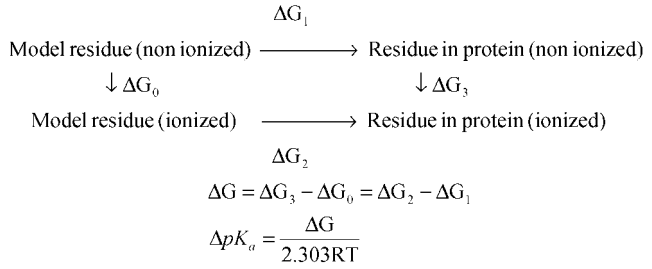


Figure 1. Scheme for calculation of pK_a shift for an isolated titratable site. The difference in the charging energy in solution between the titrated and not titrated residue is computed for the titratable group in a model compound and in a protein.

Thermodynamic cycles

As mentioned above, it is possible to calculate the free energy for the hypothetical charging process of a macromolecule in an ionic atmosphere. By combining several hypothetical processes in a thermodynamic cycle, the theoretical free energy for real processes may be computed (see e.g. Misra *et al.*, 1994a). We illustrate this principle for two processes.

The first is the computation of the electrostatic part of the solvation energy. The relevant thermodynamic cycle is reported in Fig. 2.

The sought free energy of solvation (only the electrostatic part) is then given by the difference between the free energies corresponding to two purely hypothetical charging

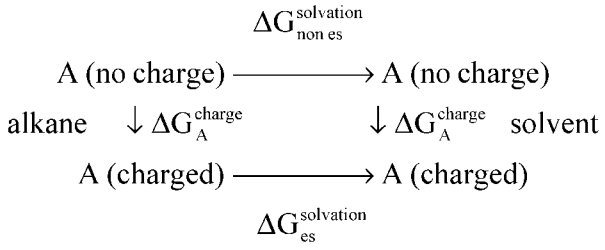


Figure 2. Thermodynamic cycle for calculation of electrostatic solvation free energy. The difference in the charging energy (i.e. ΔG_A^{charge}) in alkane phase and in solution is the electrostatic solvation free energy.

processes, in alkane phase and in ionic solvent. Much care must be taken to avoid effects due to the discretization of the charges on the grid for numerical solution of the Poisson–Boltzmann equation, which make the self-energy associated with point charges position- and grid-dependent (see Madura *et al.*, 1994 for a clear discussion of the problem).

The second example is the computation of the electrostatic free energy of association of two molecules. The relevant thermodynamic cycle is reported in Fig. 3.

As before, the free energy of association is given by the difference in the free energies for purely hypothetical charging processes.

By comparing the association free energies obtained under different environmental conditions, the dependence on the ionic strength or on medium dielectric constant can be studied.

Simplified models of proteins, DNA and membranes

Biomolecular shapes have been often assimilated (to different degrees of approximation) to sphere, cylinders and planes. For systems with simple geometry the solution of the Poisson–Boltzmann equation has been widely studied. For uniformly charged infinite planes, cylinders and spheres the Poisson–Boltzmann equation depends on a single variable and in some cases exact solutions can be obtained. For ellipsoidal shapes, analytical solutions for different boundary conditions have been also obtained (Hsu and Liu, 1996a,b; Yoon and Kim, 1989). We summarize here the solution of the linearized equation for a sphere, a cylinder and a plane.

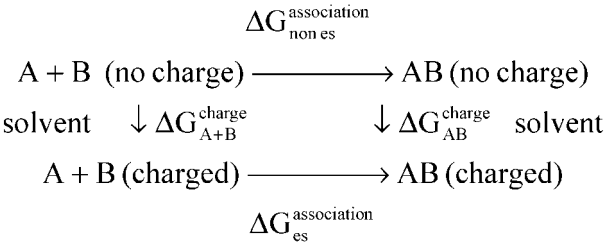


Figure 3. Thermodynamic cycle for calculation of electrostatic association free energy. The difference in the charging energy in solution between the complex and isolated molecules (i.e. $\Delta G_{AB}^{charge} - \Delta G_{A+B}^{charge}$) is the electrostatic association free energy.

For all these shapes, we assume the charge to be smeared on the surface. We define the following reduced quantities: $x = k_D\, r$ and $\phi = q\psi/kT$. With these definitions the Poisson–Boltzmann equation for these shapes reads:

$$\phi'' + \frac{m}{x}\phi' = \phi \tag{14}$$

with $m = 2, 1, 0$ for the sphere, the cylinder and the plane, respectively, and where x is a reduced distance from the center of the sphere, from the axis of the cylinder or from the plane.

The solution for these three cases reads:

$$\begin{aligned} m = 2 \quad & \phi \approx \frac{e^{-x}}{x} \\ m = 1 \quad & \phi \approx K_0(x) \rightarrow \pi\frac{e^{-x}}{\sqrt{x}} \\ m = 0 \quad & \phi \approx e^{-x} \end{aligned} \tag{15}$$

where $K_0(x)$ is the modified Bessel function.

The multiplicative coefficients depend on the boundary conditions at the solute–solvent interface, usually specified as the electric field.

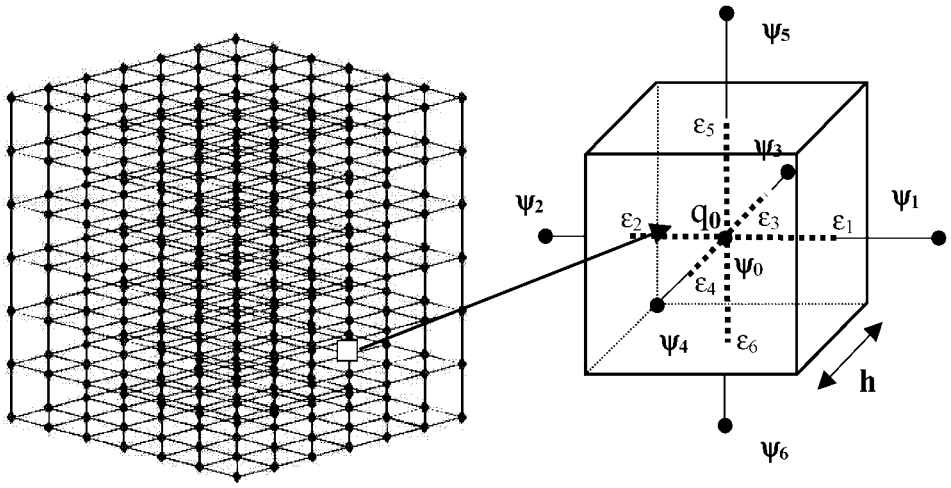
Numerical solutions of the Poisson–Boltzmann equation

Numerical solutions of the Poisson–Boltzmann equation are not easily obtained for complex shapes and charge distributions. Most available programs use different versions of the finite difference method (Warwicker and Watson, 1982), where molecular charges and dielectric are discretized on a grid, and the Poisson–Boltzmann equation is integrated and recast in a finite difference form or in a grid based extremization problem (e.g. Madura *et al.*, 1994; see Fig. 4).

The discretization procedure presents some disadvantages:

1. The free energy is largely dependent on the relative position of charges on the grid and on the dimension of the mesh. The grid-dependent self energy of charges must be treated in order to compute free energies (Madura *et al.*, 1994).
2. The mesh must be fine enough not to merge opposite charges (dipoles) at the same node, although schemes have been developed to minimize this effect (Edmonds *et al.*, 1984).
3. The mesh must be fine enough to represent properly the macromolecule–solvent interfaces, although, also here, schemes for properly smoothing the dielectric discontinuity have been developed (Davis and McCammon, 1991). When this is not done properly artifacts may result in surface potential representation.

The algorithms for solving the finite difference Poisson–Boltzmann equation are still computationally demanding, although they use state-of-art techniques for the solution of differential equations on a grid (e.g. Nicholls and Honig, 1991; Luty *et al.*, 1992; Davis and McCammon, 1989; Gilson *et al.*, 1987).



$$\psi_0 = \frac{\dfrac{4\pi q_0}{h} + \sum_{i=1}^{6} \varepsilon_i \psi_i}{\sum_{i=1}^{6} \varepsilon_i + k_D^2 \varepsilon_0 \lambda_0 h^2}$$

Figure 4. Scheme for the finite difference Poisson–Boltzmann equation. For each grid point a finite difference equation like the one reported is written.

More recently improvements to the finite difference scheme have been proposed and implemented, namely the use of multigridding techniques in order to achieve faster and more accurate solutions (Holst *et al.*, 1994a,b).

The effectiveness of multigridding techniques has been demonstrated by computation of electrostatic properties of nanosystems by McCammon and coworkers (Baker *et al.*, 2001).

Different approaches have been proposed as well, among these the boundary element method and the finite element method are appealing because they can avoid many of the problems associated with discretization.

In the boundary element method (e.g. Zauhar and Morgan, 1985; Rashin 1990; Zhou, 1993) the Poisson equation is considered only for the inner part of the macromolecule with boundary conditions in the form of polarization charges at solvent interfaces. Polarization charges are found by imposing classical electrostatics equations at solute–solvent interface. Then the electrostatic potential is computed considering all pairwise interactions among source and polarization charges. The method is not suited for considering ionic charge density and may be computationally very expensive for large systems (that can be treated efficiently in the finite difference approach using several focusing steps).

The finite element technique (Orttung, 1977; You and Harvey, 1993) seems very promising in that it allows finer meshes to be put where they are needed, such as at interfaces, and larger meshes to be put far from the molecule, where spatial changes in the electrostatic potential are small. Also, irregular shapes can be fit more easily in the finite element approach rather than in a cubic lattice. Moreover, nodes of the elements may be placed at atomic positions, which may prove advantageous when comparing different conformations of the same molecule.

Recently, impressive improvements on this scheme have been presented by Holst and coworkers (Holst *et al.*, 2000; Baker *et al.*, 2000), where a coarse space discretization is used to solve the Poisson–Boltzmann equation, then the error estimates are used to refine, using a simple scheme, the finite element space representation in an iterative way until convergence is reached.

Recently, Nielsen and Janssen (2001) proposed a quite innovative way to solve the Poisson–Boltzmann equation based on the fast Fourier transform algorithm. Applications have been reported only for small molecules with very high accuracy. It will be interesting to see whether the same approach can be applied to larger systems with the same efficiency and accuracy.

APPLICATIONS

Surface potential calculations

One of the most frequent applications of the Poisson–Boltzmann equation is the computation of the electrostatic potential at the solvent accessible surface. It is worth noting that the potential is greatly influenced by partial or net charges on the molecule, but also by molecular shape.

A demonstration of this principle is provided for instance

by the surface potential of DNA in its standard B-DNA form. In this case the most negative potential, as a consequence of overall charge and dielectric and ionic screening, is found in the minor groove of DNA.

It should be considered that partial charges (even when belonging to dipoles) may create substantial charge density (this is the case for instance for helix termini in proteins) and that accessibility to solvent and ions greatly reduces the potential, so that narrow cavities often are associated with high potential. The latter point is illustrated by the comparison between the potential computed using the Poisson–Boltzmann equation with inner dielectric constant set to 4 and in 150 mM ionic strength, and using the same methodology but with uniform dielectric constant (80) and zero ionic strength. The relative weakening of the potential at the major groove compared to the minor groove is apparent (Plate 1).

The potential at the surface should be directly related to the concentration of ions at the surfaces, while the relevance for macromolecular recognition may be questioned since the test charge approximation is not likely to apply. Few studies have surveyed the electrostatic potential at contacting surfaces in macromolecular complexes. Among these, the study of McCoy *et al.* (1997) is particularly significant because the complementarity of electrostatic potential is highlighted as a feature of protein–protein complexes, while charge complementarity is a much less conserved feature. Similar conclusions have been reached, more recently, by Eisenstein and coworkers (Heifetz *et al.*, 2002). These studies, together with the large body of evidence from specific studies, support the idea that it is possible to predict binding sites from visual inspection of electrostatic potential at a solvent-accessible surface. As a rather impressive example of these principles, the opposite faces of the Antennapedia homeodomain (four initial residues are missing in the PDB structure, pdb id. 9ANT) interacting with DNA are reported in Plate 2. A similar picture holds for other homeodomains.

One of the limitations of the approach is that electrostatic properties are computed usually on a single structure and the precise orientation of sidechains that might depend, for instance, on crystal packing, might change both the accessible surface and its potential. Such an example has been reported by us (Fogolari *et al.*, 2000a) for β-lactoglobulin (pdb id. 1BEB) where the different orientation of several charged residue sidechains has an effect on the extent and the magnitude of one of the positive potential surface patches. In this respect, averaging over an ensemble of, say, NMR structures should add reliability to the results.

Another important and emerging field of application, rather than single protein analysis, is the comparison of electrostatic properties in evolutionarily related proteins. Similar to sequence analysis, mutations in properties known to be important for a particular function hint at a different behavior of the protein (e.g. Ugolini *et al.*, 2001; Romagnoli *et al.*, 2000), whereas conserved features are likely to be important for function. In this respect, systematic investigations on protein classes, using both experimental structures and homology derived models, have appeared in recent years (Blomberg *et al.*, 1999; De Rienzo *et al.*, 2000; LiCata and Bernlohr, 1998). Results have highlighted important classification principles and conserved features which could not be assessed by sequence comparison, thus demonstrating the usefulness of this approach.

Brownian dynamics simulations

Long-range electrostatic fields, calculated using the Poisson–Boltzmann equation or straightforwardly using the Debye–Hückel potential, may be used in order to compute the diffusional trajectories of charged ligands. These are computed for a simplified model of the ligand, usually a sphere or a string of spheres, under the test charge approximation, i.e. under the assumption that the interaction of the ligand with the macromolecule is not (significantly) different from the one computed using the macromolecule electrostatic potential in the absence of the ligand. Brownian dynamics simulations are, in principle, quite demanding from a computational point of view, because the detailed hydrodynamic interactions, i.e. forces and torques arising from solute-induced movement of the solvent, between the ligand and the macromolecule implies the computation of a tensor, which depends on the radii and relative positions of all particles. Ermak and McCammon (1978) suggested the use of Oseen or Rotne–Prager (two-body) tensors that are particularly suited for a description of Brownian dynamics based on Langevin equation. When hydrodynamic interactions have been omitted, encounter rates differing by less than 20% have been computed (Antosiewicz *et al.*, 1996a; Antosiewicz and McCammon 1995).

One of the most popular programs for Brownian dynamics (BD) simulations is the University of Houston Brownian Dynamics program (Madura *et al.*, 1994, 1995), which implements the Ermak–McCammon algorithm (Ermak and McCammon, 1978) by which the Fokker–Planck equation is used, under the reasonable approximation of decoupling of diffusive motions from much shorter time scales motions, to derive an equation for the diffusing particle position in a form strictly related to Langevin equation (in the absence of hydrodynamic interactions):

$$\vec{r}(t+\Delta t) = \vec{r}(t) + \frac{D}{k_B T}\vec{F}[\vec{r}(t)]\Delta t + \vec{R} \qquad (16)$$

where $\vec{r}(t)$ is the position of the diffusing particle, D is the relative diffusion constant, $\vec{F}[\vec{r}(t)]$ is the force acting on the particle and $\vec{R}$ is a random uncorrelated displacement.

The methodology (reviewed in Davis *et al.*, 1991) then requires the definition of a 'reaction patch' on the macromolecule, logically defined by proximity to one or more atoms (see Fig. 5), the definition of a 'begin' sphere surrounding the macromolecule (with a radius typically 5–10 Debye lengths plus the molecules radii) where trajectories are started and the definition of a 'quit' sphere surrounding the macromolecule (with a radius typically 15–20 Debye lengths plus the molecules radii) that, once crossed, implies that the particle will not encounter the macromolecule. Many trajectories are started and propagated using the above equation and statistics about the number of 'reacted' trajectories (i.e. trajectories reaching the 'reaction patch') are accumulated. Then results are treated using Smoluchowski theory, using a correction for the finite probability of recrossing the 'quit' sphere and reaction rates are computed (Allison *et al.*, 1988; Zhou, 1990).

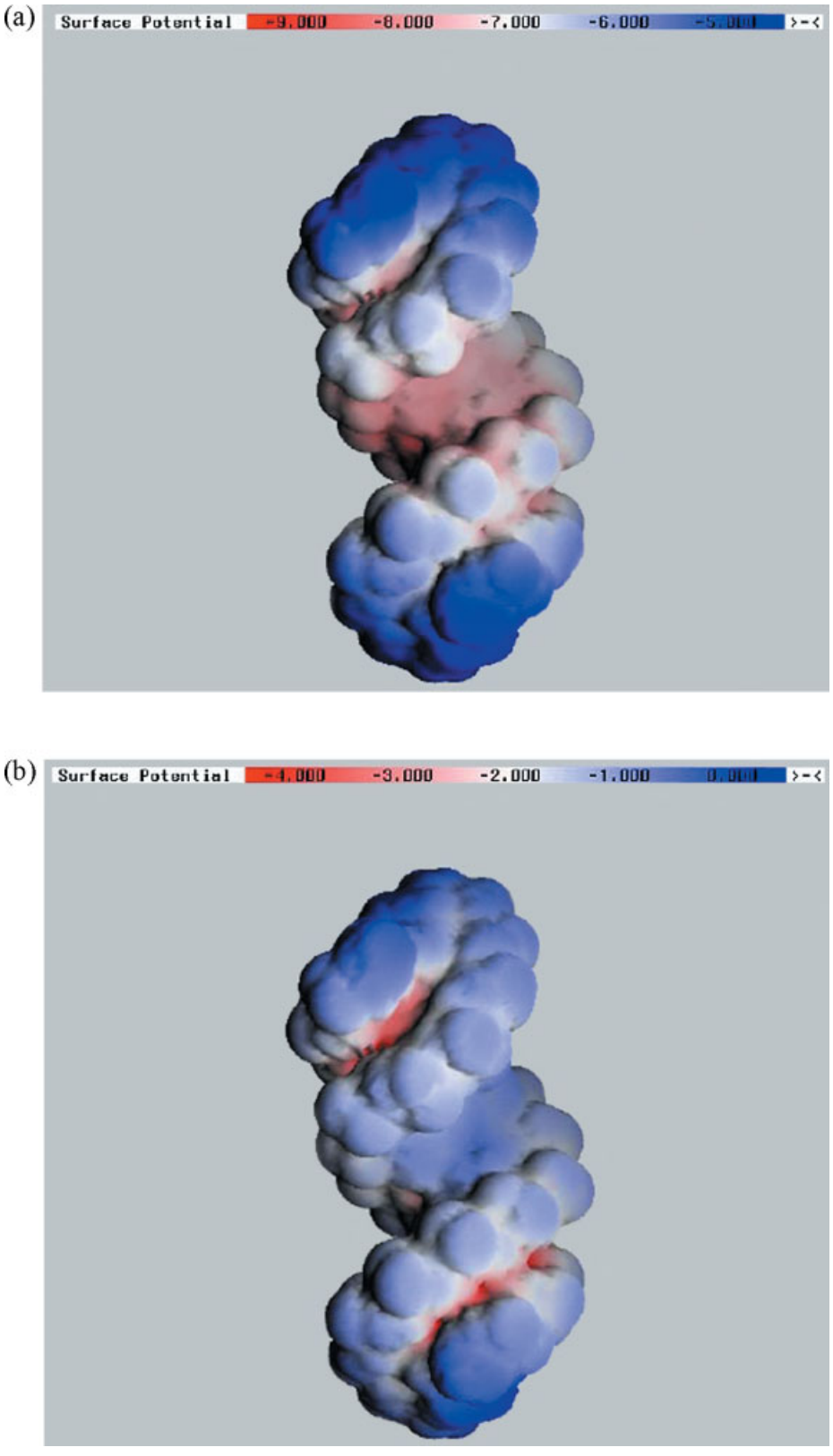


Plate 1. Electrostatic surface potential for a short B-DNA (pdb id. 9ANT, chains C and D) assuming a homogeneous ion-free medium (with dielectric constant of 80; panel a) and in 150 mM ionic strength with a solute dielectric constant of 4 (panel b). The ranges of the potential (-9.0 to -5 kcal mol^{-1} q^{-1} and -4.0 to 0.0 kcal mol^{-1} q^{-1} in the left and right panels, respectively) have been chosen for comparison.

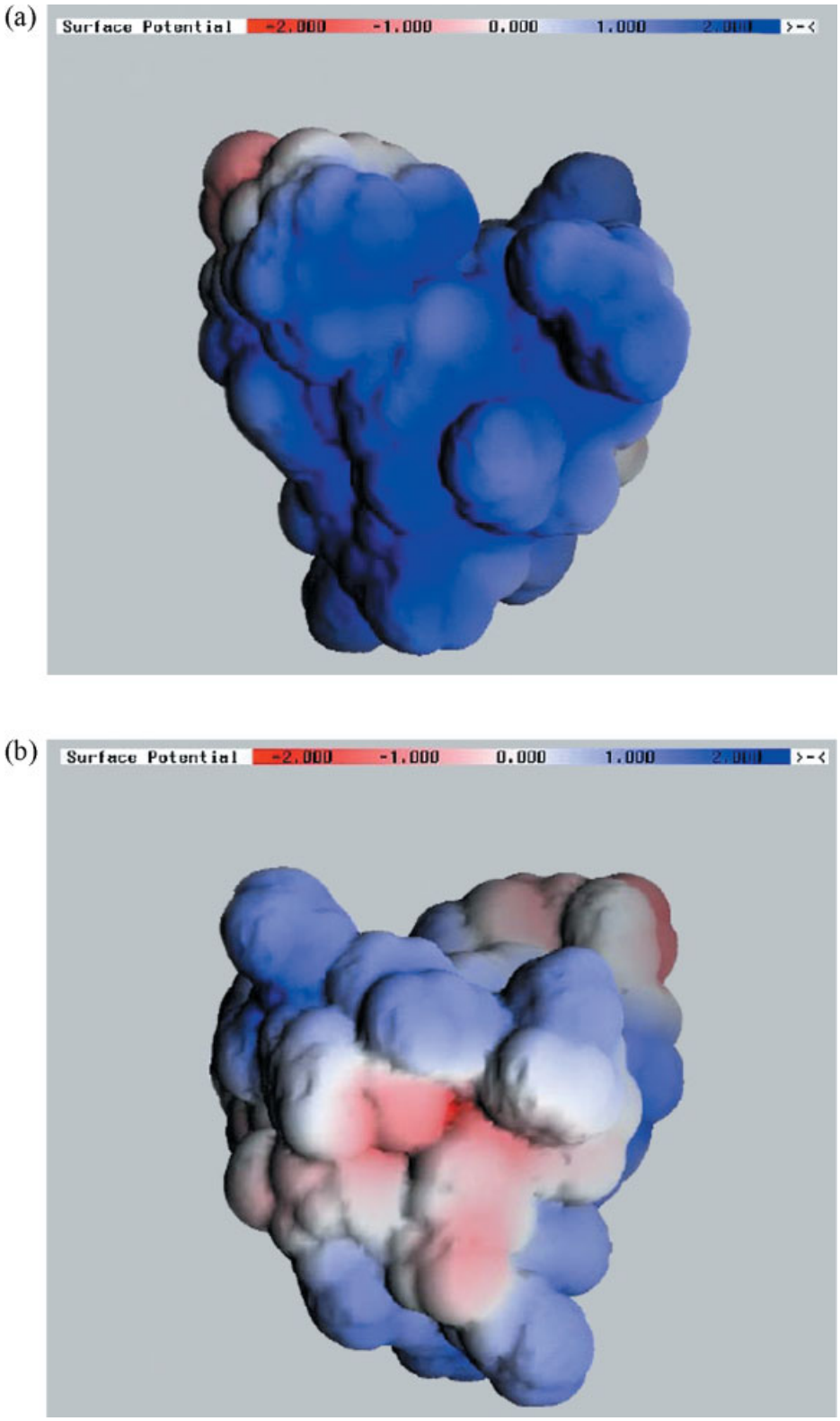


Plate 2. Electrostatic surface potential for the antennapedia homeodomain (pdb id. 9ANT, chain A). The face contacting DNA is shown in panel a and the opposite face in panel b.

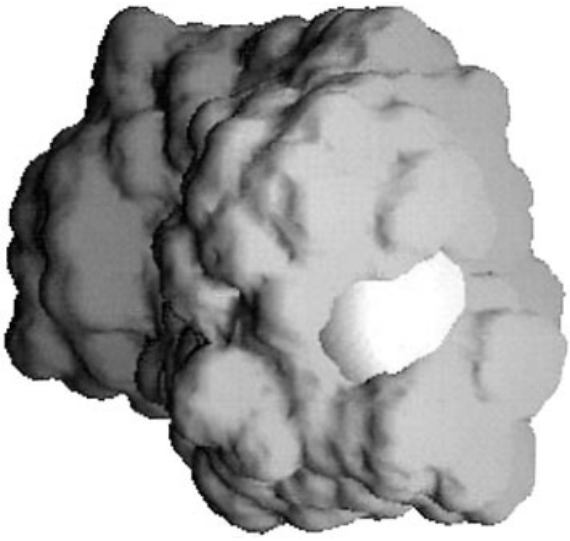


Figure 5. Definition of a reactive patch at the surface of a Fab (surface probe radius of 3 Å). The sphere has a radius of 10 Å and it is centered on an atom in the center of the binding site (reproduced from Fogolari *et al.,* 2000b).

Numerical predictions of real reaction rates are quite difficult to obtain by such a simplified model, but relative rates (for instance for mutated macromolecule models, or for different ionic strength) can be quite accurately predicted. One of the earliest success of the approach was the reproduction of the nonobvious salt dependence (Sharp *et al.*, 1987; Allison *et al.*, 1988) found in experiments for superoxide dismutase and O_2^- reaction rates (Argese *et al.*, 1987). Both substrate and enzyme carry a net negative charge, but, contrary to what could be expected, the reaction rate is enhanced by lowering the ionic strength. Another important success was the design of an enhanced superoxide dismutase enzyme by Getzoff and coworkers based on BD simulations (Getzoff *et al.*, 1992; Sines *et al.*, 1990). Some subtleties of BD simulations have been highlighted in the years (Northrup, 1994). Northrup and Erickson (1992) investigated the effect of multiple collisions on encounter rates and found that once the particles are in close contact the probability of re-encounter is high and this makes reaction rates higher than expected. Refined models for the diffusing particle have also been developed (e.g. Wade *et al.*, 1994; Kozack *et al.*, 1995). Recently McCammon and coworkers (Shen *et al.*, 2001) reported Brownian dynamics performed at atomic detail using adaptive time steps.

The reliability of the approach has been studied comparing predictions and experimental results for many molecular systems. The agreement with experimental data for several different protein–protein complexes was found to be excellent, in particular as far as comparison between data for different mutants of proteins are concerned (Gabdoulline and Wade, 2001). The absolute values of association rates appear to be sensitive to molecular flexibility.

Brownian dynamics simulations (not necessarily based on the Poisson–Boltzmann equation) are increasingly being applied also to other systems such as DNA (e.g. Klenin and Langowski, 2001) and ion-channels (e.g. Mashl *et al.*, 2001).

Free energy calculations

The computation of free energies for complex biomolecular systems from the solution of the Poisson–Boltzmann equation is a recent achievement originated from the work of Sharp and Honig (1990). The possibility of decomposing the free energy for a process in different components has been questioned by Mark and van Gunsteren (1994), although, in view of all the more drastic approximations involved, this should not be the major problem. Besides the possibility itself of disentangling the different contributions to processes like, say, bimolecular association, the estimation of contributions other than electrostatic is extremely uncertain (e.g. Baginski *et al.*, 1997 and references cited therein and for reviews Gilson *et al.*, 1997; Janin, 1997; Finkelstein and Janin, 1989). For this reason, the most successful applications have been for comparing the same system under different conditions affecting only the electrostatic part of the free energy, leading thus (possibly) to cancellation of errors. The description provided by the Poisson–Boltzmann equation framework, for instance, greatly clarified the driving forces involved in binding of charged molecules to DNA. In 1994 Sharp, Honig and coworkers (Misra *et al.*, 1994a,b) studied in great detail all salt-dependent contributions to the free energy of binding and its salt dependence for a number of ligands bound to DNA. The favorable entropic contribution to binding of multivalent charged ligands to DNA, arising from release of monovalent counterions, held as the driving force for DNA binding, was found to be more than counterbalanced by other enthalpic terms. The same authors applied the same electrostatic free energy dissection to protein–DNA complexes. Sharp (1995) also pointed out another important thermodynamic consequence of the free energy expression, i.e. that the entropy of the system does not coincide strictly with classical mixing and osmotic terms, but entails a large contribution due to the temperature dependence of the dielectric constant (Sharp, 1995). This contribution was found to be very large for protein–DNA complexes (Sharp *et al.*, 1995; Fogolari *et al.*, 1997).

Following the paper of Sharp and Honig (1990), the limits of applicability of the popular counterion condensation theory by Manning (1978, based on the Debye–Hückel theory) could also be established by comparison of the free energy values with the values afforded by the nonlinear Poisson–Boltzmann equation (Fogolari *et al.*, 1993; Sharp *et al.*, 1995; Stigter, 1995).

The very good quantitative agreement for the salt dependence of binding constants for a set of DNA ligands (Misra *et al.*, 1994a), the very accurate reproduction of solvation free energies by employing a limited set of atomic charge and radii parameters (e.g. Sitkoff *et al.*, 1994) and a better parametrization for surface tension coefficient for irregular shapes (Nicholls *et al.*, 1991) led to the expectation that the approach could provide quantitative predictions for binding constants. Although many papers have reported successful applications of the methodology, it is our own experience that predictions of absolute values of binding constants (not just dependence on environmental constants) is much more subject to errors. The problem seems more serious than just proper weighting of entropic restrictions and hydrophobic free

energy gain upon complexation, because even using weights as adjustable parameters leads to almost 100% uncertainty in the binding energy of FAB-antigen complexes (Fogolari *et al.*, 2000b).

Generally, predictions results depend on parameter choice which must be tuned to the specific type of complexes studied (see e.g. Schapira *et al.*, 1999).

Similar considerations can be derived from the work of Novotny *et al.* (1997) which, at variance with the vast majority of other studies, aims at predicting unknown binding constants, rather than reproducing them.

One problem might be related to erroneous calculation of buried inter-molecular hydrogen bonds and salt bridges which are damped, compared with standard force fields, by a factor corresponding to the relative dielectric constant. Therefore their favorable contribution to binding is overcome by unfavorable desolvation contributions, so that other model parameters must be changed. Another problem could be that entropic restrictions in backbone and sidechains rotations are not easily computed and are probably different from system to system. The best results are probably obtained for ligands which lack extensive flexibility and for which buried electrostatic interactions are not too relevant (e.g. Misra *et al.*, 1994a; Baginski *et al.*, 1997).

Notwithstanding all the above-mentioned problems Poisson–Boltzmann calculations have been used also for docking studies (e.g. Mandell *et al.*, 2001; Zacharias *et al.*, 1994). Often the largest molecule is taken as the source of electrostatic potential, while the smaller molecule is treated in the test charge approximation framework. This approximation is necessary in order to avoid as much as possible lengthy Poisson–Boltzmann calculations.

Problems associated with entropic effects and dielectric constant for macromolecules should not be as important for molecules not in direct contact. In this case we may assume that the entropy associated with degrees of freedom not explicitly considered are not much affected by the presence of the other molecule. Poisson–Boltzmann equation-derived interaction-free energies should be more accurate in this case. As an example of nonobvious results obtainable by this kind of analysis, in a study on protein–DNA interaction, very different interaction energies (by $\sim$10 kcal/mol) were obtained for different orientations of the protein at more than 10 Å distance from the DNA (Fogolari *et al.*, 1997).

In this respect, even faster and reasonably accurate procedures have been developed, in which, instead of solving the Poisson–Boltzmann equation for each configuration of the solutes, the interaction energy is computed by using a set of effective charges in the test charge approximation (Gabdoulline and Wade, 1996). Based on the required accuracy, a reduced set of effective charges may be used, thus speeding up the computation even more.

Recently, the capability of solvent continuum models to provide potentials of mean force with all degrees of freedom belonging to the solvent integrated out has been exploited by Kollman and coworkers for post-processing molecular dynamics trajectories without having to consider solvent coordinates (Kollman *et al.*, 2000; Wang *et al.*, 2001 and references cited therein) providing strong evidence for the accuracy of the methodology (Lee *et al.*, 2001)

pH dependent effects

The computation of free energies may be used for pK_a and redox potential predictions in proteins. We will discuss here the computation of pK_a shifts in proteins, although the same concepts apply to the computation of redox potentials (Ullmann and Knapp, 1999). The rationale behind this kind of predictions is that the difference in behaviour between the titrating group in model compounds and in the context of the macromolecule is due to classical electrostatics. This idea is quite old (see Davis and McCammon, 1990 and references cited therein), but a refined treatment which went beyond simple uniformly charged spheres (Linderstrom-Lang, 1924; Nitta and Sugai, 1972) was determined only later by Tanford and Kirkwood (1957), based on a previous model of proteins as low dielectric spheres with embedded charges (Kirkwood, 1934b). The Tanford–Kirkwood approach to pK_a computation was later used and corrected by Bashford and Karplus (1990, 1991) in conjuction with numerical solution of the Poisson–Boltzmann equation. Similar approaches have been used by other authors (e.g. Warshel, 1981; Beroza *et al.*, 1991; Yang *et al.*, 1993).

Antosiewicz *et al.* (1994) used the same approach of Bashford and Karplus and tested several values for the inner dielectric constant and found remarkable agreement (within $\sim$0.7 pK_a units) between theoretical predictions and experiments for a set of seven proteins using a value of 20. The use of a high solute dielectric constant should take into account molecular rearrangements following titration and increasing solvation with raising or lowering the pH. Obviously the definition of dielectric constant depends on what is or is not explictly taken into consideration (see for a recent and clear review Schutz and Warshel, 2001). A dielectric constant of *ca* 25 for cytochrome C is indeed computed, based on the Frohlich theory (Gilson and Honig, 1986), from long molecular dynamics runs (Simonson, 1998), and it appears to arise mainly from reorientation of charged side chains. Another reason for using solute dielectric constants higher than those typical of organic compounds is that crystal structures, often used to perform these calculations, are usually less solvated than ensembles of NMR structures, especially if these are obtained at acidic pH, as is often the case in order to reduce amide proton exchange. This point is confirmed by the better agreement obtained with experimental data using ensembles of NMR structures rather than crystal structures (see e.g. Antosiewicz *et al.*, 1996b; Fogolari *et al.*, 2000a). Averaging predictions over an ensemble of structures (Bashford and Gerwert, 1992; Zhou and Vijayakumar, 1997) or using an average structure (van Vlijmen *et al.*, 1998) from experimental or simulation data is seen to improve predictions, although, in the case of molecular dynamics simulations, pK_a predictions appear to depend on the starting conformation even after 500 ps molecular dyamics runs (Koumanov *et al.*, 2001).

A more fundamental approach has been proposed by Beroza and Case (1996) where the ensemble of sidechain conformation is generated driven by the computed free energy of protonation thus modelling explictly at least sidechain rearrangements following titration. These methods are, however, computationally demanding.

Other methods try to adjust solute parameters for pK_a

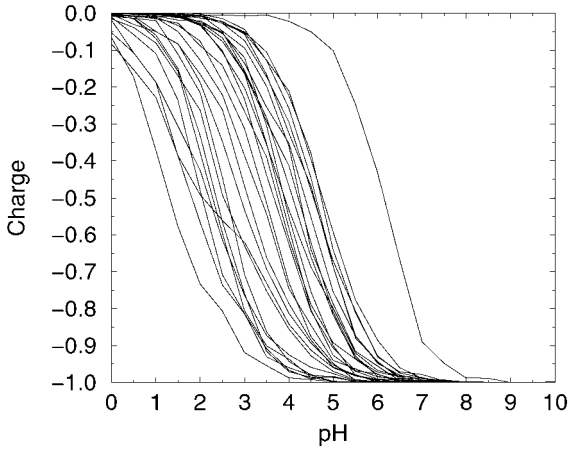


Figure 6. Titration curves obtained by Poisson–Boltzmann calculations for all the acidic residues of β-lactoglobulin (monomer B, pdb id. 1BEB). The anomalous titrating residue is GLU 89.

calculation. A simple way to perform this task has been proposed by Demchuk and Wade (1996), who partition titration sites in inner and solvent exposed (the majority of them). For the first a low (in the range of 15) dielectric constant applies, while for the latter aqueous solvent dielectric constant applies. The two types of sites can be distinguished by a simple test based on the magnitude of the reaction field. The average error in predicted pK_as is as low as 0.5 pK_a units.

It is worth to remark that predictions, even with the optimizations described, still produce a few large errors. In this respect, note that much simpler approaches have been proven to afford quantitatively comparable results (e.g. Warwicker, 1999; Sandberg and Edholm, 1999).

We wish to note here that, besides the accuracy of the predictions, the usefulness of the approach is that it can provide suggestions for the identity of anomalously titrating groups, for which the exact pK_a value is known from experiments (see Fig. 6), and it can provide a rationale in term of desolvation free energy and electrostatic interaction energies for experimental observations, often allowing discrimination between alternative conformational models by comparison with experimental data (e.g. Fogolari *et al.*, 2000a; Bashford and Gerwert, 1992; Yang *et al.*, 1993).

As discussed by Antosiewicz *et al.* (1994), titration properties are linked with pH-dependence of protein stability and titration predictions may be used to obtain, for instance, a quantitative prediction of the pH-dependent part of the free energy of unfolding. As judged by data reported by Antosiewicz *et al.* for a few proteins, effects tend to be overestimated (usually by 10–20 kcal/mol) at extreme pH conditions. This has been observed by us on β-lactoglobulin, a protein stable at low pH. Most likely this is due to the assumption of a rigid molecular model, which is not able to account for molecular relaxation following titration. In this respect, it is instructive that the pH dependence of the stability of the same protein could not be reproduced by any continuum method, thus confirming earlier suggestions about protein rearrangements (Kella and Kinsella, 1988).

Including Poisson–Boltzmann equation forces in molecular mechanics and dynamics protocols

In a recent review Kollman *et al.* (2000) gave a historical sketch of molecular dynamics pointing out the early *in vacuo* era, a second 'free energy perturbation era', a third 'explicit solvent' era and a last fourth era where advances in the last two eras are combined. In this phase, combining explicit and implicit solvent models will play an important role. One way to do this was the extensive application of the MM/PBSA (molecular mechanics/Poisson–Boltzmann solvent accessible) methodology for analyzing molecular dynamics (MD) trajectories.

In the last decade, however, several authors proposed a more radical approach, i.e. generation of MD trajectories using an implicit (most often continuum) solvent model (e.g. based on the Poisson–Boltzmann equation: Sharp, 1990; Niedermeier and Schultzen, 1992; Gilson *et al.*, 1995; Smart *et al.*, 1997; David *et al.*, 2000; Fogolari *et al.*, 2001; Huber, 1998). In this respect, early distance dependent dielectric constant usage (McCammon *et al.*, 1977) for molecular dynamics simulations was a computationally efficient way to represent implicitly the solvent.

There are several advantages in using continuum solvent models as we recently pointed out (Fogolari *et al.*, 2001), namely: (1) finely tunable solvent parameters (e.g. viscosity) without the need to resort to any molecular model; (2) no need for post-processing MD trajectories; (3) possibility to perform unphysical (but informative) simulations, such as for protein unfolding at high temperature keeping solvent properties at room temperature. The possibility of tuning solvent viscosity could also provide enhanced sampling.

Many models have been put forward (see the recent review by Simonson, 2001) and for most of them the Poisson–Boltzmann equation is the reference model. We will specifically discuss here only models that employ the Poisson–Boltzmann equation for electrostatic interactions. We will focus only on electrostatics, although in order to account for other solvent effects usually a stochastic excitation term and a damping term are added to Newtonian equations of motion and 'hydrophobic forces' are computed using the derivative of the solute–solvent accessible surface area (Sridharan *et al.*, 1995).

Gilson *et al.* (1993) derived force expressions, shown to be compatible with a Maxwell stress tensor formalism (Jackson, 1962), by taking variations of the free energy from the linear Poisson–Boltzmann equation and implemented these forces in the program UHBD. In recent years, Roux and coworkers proposed a finite difference version of the same expressions (Im *et al.*, 1998) and Friesner and coworkers (Friedrichs *et al.*, 1999) implemented the same expressions in a finite element framework.

Electrostatic forces entail three terms:

$$\vec{F}_i^{\mathrm{el}} = q_i \vec{E}_i - \int_S \frac{1}{8\pi} \left|\vec{E}\right|^2 \vec{\nabla}_i \epsilon(\vec{r}) - \int_S \frac{1}{8\pi} \epsilon(\vec{r}) \psi^2 k_{\mathrm{D}}^2 \vec{\nabla}_i \lambda(\vec{r}) \tag{17}$$

where the subscript i refers to atom i. The first term is a classical electric field term, and the last two terms describe the tendency of high dielectric and salts to displace low

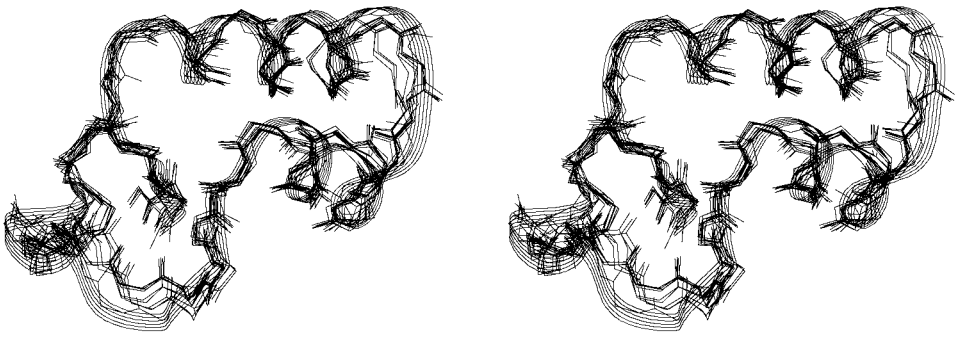


Figure 7. Ten snaphots (taken at 100 ps time intervals) of a molecular dynamics run of a small protein (viscotoxin A3) using Poisson–Boltzmann and solvent accessibility forces. The structures are superimposed on the NMR experimental structure (pdb id. ED0) shown as a ribbon.

dielectric regions inaccessible to ions, i.e. a dielectric and ionic boundary pressure term.

In the early 1990s, Sharp (1991), and later Niedermeier and Schultzen (1992), proposed how to use Poisson–Boltzmann equation-derived forces for an implicit solvent representation in standard MD protocols. Apart from different details in computation of forces, the protocol adopted by Sharp has been used also in all the following similar studies. In particular (computationally demanding), Poisson–Boltzmann electrostatic calculations are performed every 200 fs (250 fs in Niedermeier and Schultzen, 1992) and forces are kept constant in the dynamics before they are recalculated. In later studies (Gilson *et al.*, 1995; Smart *et al.*, 1997; David *et al.*, 2000; Fogolari *et al.*, 2001), this update time was substantially shorter (∼50 fs) in order to avoid changes in conformation that can lead to large changes in solvation forces. There is an important issue which would justify substantially longer intervals, i.e. the dielectric relaxation time of water which is approximately 10 ps (Harvey, 1989). A more correct way to calculate solvation forces would be then by introducing a memory function which would anyway smooth large changes in computed forces. Fogolari *et al.* (2001) used a similar approach with a very short decay time (∼50 fs).

Although promising, earlier approaches were not much explored. We suspect that the main reason is that the approach in its straightforward implementation is affected by artifacts. In particular, for consistency with current forcefields, the solute inner dielectric constant must be 1.0, thus producing very large reaction forces. On the other hand, introducing higher dielectric constants would weaken hydrogen bonds accordingly in all of those forcefields (most of the current ones), where hydrogen bonds are reproduced through electrostatic interactions. Indeed, we were able to obtain stable trajectories, comparable to typical results obtained in explicit solvent simulations, using only dielectric constants of 4–8 (Fogolari *et al.*, 2001).

Recently, we have introduced a smoothing function that switches the dielectric constant in Coulombic terms from 1.0 (for distance less than 6.0 Å) to 4.0 (for distances larger than 8.0 Å). A stable, 1 ns, unrestrained MD trajectory was obtained for a small protein, moreover, similar results are obtained by calculating Poisson–Boltzmann equation forces at 1.0 ps intervals and introducing a decay function with time constant of 10 ps (Fogolari *et al.*, in preparation, Fig. 7). Although refinement of parameters and methodology will be needed, the protocol proved to be fast and reliable, with a reduction in computation times, with respect to explicit solvent simulation, of one order of magnitude.

CONCLUSIONS

Electrostatics plays a fundamental role in virtually all processes involving biomolecules in solution. One of the most fundamental approaches to treat electrostatic effects is based on the Poisson–Boltzmann equation. There is a large body of evidence that simulating biomolecular electrostatics within the Poisson–Boltzmann equation framework is able to reproduce, explain and predict many experimental observations, and to provide working hypotheses to test. The main fields of application have been the computation of the electrostatic potential at the solvent-accessible molecular surface, the computation of reaction rates between molecules in solution, the computation of the free energy of association and its salt dependence, and the study of pK_a shifts in proteins. A relatively new field of application is the combination of classical molecular mechanics and dynamics with solvation forces computed from the Poisson–Boltzmann equation and the derivative of solvent accessible surface area. Recent results have added reliability to the approach. Further applications are expected in the field of structural genomics where electrostatic properties of homologous proteins could reveal nonobvious conserved features.

Acknowledgements

We wish to thank Professor McCammon for providing us with the UHBD software, and Professor Nicholls and Professor Honig for making GRASP available. This work has been supported by Italian MURST Cofin 2000.

REFERENCES

Alfrey T Jr, Berg PW, Morawetz H. 1951. The counterion distribution in solutions of rod-shaped polyelectrolytes. *J. Polym. Sci.* **7**: 543–547.

Allison SA, Bacquet RJ, McCammon JA. 1988. Simulation of the diffusion-controlled reaction between superoxide and superoxide dismutase. II. Detailed models. *Biopolymers* **27**: 251–269.

Antosiewicz J, McCammon JA, Gilson MK. 1994. Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* **238**: 415–436.

Antosiewicz J, McCammon JA. 1995. Electrostatic and hydrodynamic orientational steering effects in enzyme-substrate association. *Biophys. J.* **69**: 57–65.

Antosiewicz J, Briggs JM, McCammon JA. 1996a. Orientational steering in enzyme-substrate association: ionic strength dependence of hydrodynamic torque effects. *Eur. Biophys. J.* **24**: 137–141.

Antosiewicz J, McCammon JA, Gilson MK. 1996b. The determinants of pKas in proteins. *Biochemistry* **35**: 7819–7833.

Argese E, Viglino P, Rotilio G, Scarpa M, Rigo A. 1987. Electrostatic control of the rate-determining step of the copper, zinc superoxide dismutase catalytic reaction. *Biochemistry* **26**: 3224–3228.

Aurora R, Rose GD. 1998. Helix capping. *Protein Sci.* **7**: 21–38.

Baginski M, Fogolari F, Briggs JM. 1997. Electrostatic and non-electrostatic contributions to the binding free energies of anthracycline antibiotics to DNA. *J. Mol. Biol.* **274**: 253–267.

Baker NA, Holst MJ, Wang F. 2000. Adaptive multilevel finite element solution of the Poisson–Boltzmann equation II: refinement at solvent accessible surfaces in biomolecular systems. *J. Comput. Chem.* **21**: 1343–1352.

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA* **98**: 10037–10041.

Bashford D, Gerwert K. 1992. Electrostatic calculations of the pKa values of ionizable groups in bacteriorhodopsin. *J. Mol. Biol.* **224**: 473–486.

Bashford D, Karplus M. 1990. pK_a's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry* **29**: 10219–10225.

Bashford D, Karplus M. 1991. Multiple-Site Titration Curves of Proteins: An analysis of Exact and Approximate Methods for Their Calculation. *J. Phys. Chem.* **95**: 9556–9561.

Beroza P, Case DA. 1996. Including side chain flexibility in continuum calculations of protein titration. *J. Phys. Chem.* **100**: 20156–20163.

Beroza P, Fredkin DR, Okamura MY, Feher G. 1991. Protonation of interacting residues in a protein by a Monte Carlo method: application to lysozyme and the photosynthetic reaction center of Rhodobacter sphaeroides. *Proc. Natl Acad. Sci. USA* **88**: 5804–5808.

Blomberg N, Gabdoulline RR, Nilges M, Wade RC. 1999. Classification of protein sequences by homology modeling and quantitative analysis of electrostatic similarity. *Proteins Struct. Funct. Genet.* **37**: 379–387.

Branden C-I, Tooze J. 1999. *Introduction to Protein Structure*. Garland: New York.

Bruice TC, Benkovic SJ. 2000. Chemical basis for enzyme catalysis. *Biochemistry* **39**: 6267–6274.

Chapman, DL. 1913. A contribution to the theory of electrocapillarity. *Phil. Mag.* **25**: 475–481.

Cleland WW, Frey PA, Gerlt JA. 1998. The low barrier hydrogen bond in enzymatic catalysis. *J. Biol. Chem.* **273**: 25529–25532.

Creighton T. 1993. *Proteins: Structures and Molecular Properties*. WH Freeman: San Francisco, CA.

David L, Luo R, Gilson M. 2000. Comparison of generalized Born and Poisson models: energetics and dynamics of HIV protease. *J. Comp. Chem.* **21**: 295–309.

Davis ME, McCammon JA. 1989. Solving the finite difference linearized Poisson–Boltzmann equation: a comparison of relaxation and conjugate gradient methods. *J. Comput. Chem.* **10**: 386–391.

Davis ME, McCammon JA. 1990. Electrostatics in biomolecular structure and dynamics. *Chem. Rev.* **90**: 509–521.

Davis ME, McCammon JA. 1991. Dielectric boundary smoothing in finite difference solutions of the Poisson equation: an approach to improve accuracy and convergence. *J. Comput. Chem.* **12**: 909–912.

Davis ME, Madura JD, Sines J, Luty BA, Allison SA, McCammon JA. 1991. Diffusion-controlled enzymatic reactions. *Meth. Enzymol.* **202**: 473–497.

De Rienzo F, Gabdoulline RR, Menziani MC, Wade RC. 2000. Blue copper proteins: A comparative analysis of their molecular interaction properties. *Protein Sci.* **9**: 1439–1454.

Debye P, Hückel E. 1923. Zur Theorie der Elektrolyte. *Phys. Zeitschr.* **24**: 185–206.

Demchuk E, Wade RC. 1996. Improving the continuum dielectric approach to calculating pK_as of ionizable groups in proteins. *J. Phys. Chem.* **100**: 17373–17387.

Derjaguin B, Landau L. 1941. A theory of the stability of strongly charged lyophobic sols and the coalescence of strongly charged particles in electrolytic solution. *Acta Phys. Chim. USSR* **14**: 633–662.

Draper DE. 1999. Themes in RNA–protein recognition. *J. Mol. Biol.* **293**: 255–270.

Durr E, Jelesarov I, Bosshard HR. 1999. Extremely fast folding of a very stable leucine zipper with a strengthened hydrophobic core and lacking electrostatic interactions between helices. *Biochemistry* **38**: 870–880.

Edmonds DT, Rogers NK, Sternberg MJE. 1984. Regular representation of irregular charge distributions: application to the electrostatic potential of globular proteins. *Mol. Phys.* **52**: 1487–1494.

Ermak DL, McCammon JA. 1978. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* **69**: 1352–1360.

Finkelstein AV, Janin J. 1989. The price of lost freedom: entropy of bimolecular complex formation. *Protein Engng* **3**: 1–3.

Fogolari F, Briggs JM. 1997. On the variational approach to the Poisson–Boltzmann free energies. *Chem. Phys. Lett.* **281**: 135–139.

Fogolari F, Cattarinussi S, Esposito G, Viglino P. 1993. Free energies calculated according to Manning's polyelectrolyte model compared with Poisson Boltzmann predictions. *J. Biomol. Struct. Dyn.* **11**: 629–635.

Fogolari F, Elcock AH, Esposito G, Viglino P, Briggs JM, McCammon JA. 1997. Electrostatic effects in homeodomain-DNA interaction. *J. Mol. Biol.* **267**: 368–381.

Fogolari F, Zuccato P, Esposito G, Viglino P. 1999. Biomolecular electrostatics with the linearized Poisson–Boltzmann equation. *Biophys. J.* **76**: 1–16.

Fogolari F, Ragona L, Licciardi S, Romagnoli S, Michelutti R, Ugolini R, Molinari H. 2000a. Electrostatic properties of bovine beta-lactoglobulin. *Proteins Struct. Funct. Genet.* **39**: 317–330.

Fogolari F, Ugolini R, Molinari H, Viglino P, Esposito G. 2000b. Simulation of electrostatic effects in Fab-antigen complex formation. *Eur. J. Biochem.* **267**: 4861–4869.

Fogolari F, Esposito G, Viglino P, Molinari H. 2001. Molecular mechanics and dynamics of biomolecules using a solvent continuum model. *J. Comput. Chem.* **22**: 1830–1842.

Fowler RH, Guggenheim EA. 1939. *Statistical Thermodynamics*. Cambridge University Press: Cambridge.

Friedrichs M. Zhou R, Edinger SR, Friesner RA. 1999. Poisson–Boltzmann analytical gradients for molecular modeling calculations. *J. Phys. Chem. B* **103**: 3057–3061.

Gabdoulline RR, Wade RC. 1996. Effective charges for macromolecules in solvent. *J. Phys. Chem.* **100**: 3868–3878.

Gabdoulline RR, Wade RC. 2001. Protein–protein association: investigation of factors influencing association rates by brownian dynamics simulations. *J. Mol. Biol.* **306**: 1139–1155.

Getzoff ED, Cabelli DE, Fisher CL, Parge HE, Viezzoli MS, Banci L,

Hallewell RA. 1992. Faster superoxide dismutase mutants designed by enhancing electrostatic guidance. *Nature* **358**: 347–351.
Gilson MK, Honig B. 1986. The dielectric constant of a folded protein. *Biopolymers* **25**: 2097–2119.
Gilson MK, Sharp KA, Honig B. 1987. Calculating the electrostatic potential of molecules in solution: method and error assessment. *J. Comput. Chem.* **9**: 327–335.
Gilson MK, Davis ME, Luty BA, McCammon JA. 1993. Computation of electrostatic forces on solvated molecules using the Poisson–Boltzmann equation. *J. Phys. Chem.* **97**: 3591–3600.
Gilson MK, McCammon JA, Madura JD. 1995. Molecular dynamics simulation with a continuum electrostatic model of the solvent. *J. Comput. Chem.* **16**: 1081–1095.
Gilson MK, Given JA, Bush BL, McCammon JA. 1997. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* **72**: 1047–1069.
Gouy M. 1910. Sur la constitution de la charge électrique a la surface d'un électrolyte. *J. Phys.* **9**: 457–468.
Gronwall TH, La Mer VK and Sandved K. 1928. Über den Einfluß der sogenannten höheren Glieder in der Debye-Hückelschen Theorie der Lösungen starker Elektrolyte. *Phys. Zeitschr.* **29**: 358–393.
Harvey SC. 1989. Treatment of electrostatic effects in macromolecular modeling. *Proteins Struct. Funct. Genet.* **5**: 78–92.
Heifetz A, Katchalski-Katzir E, Eisenstein M. 2002. Electrostatics in protein-protein docking. *Protein Sci.* **11**: 571–587.
Hill TL. 1956. *An Introduction to Statistical Thermodynamics*. Dover: New York.
Holst M, Kozack RE, Saied F, Subramaniam S. 1994a. Protein electrostatics: rapid multigrid-based Newton algorithm for solution of the full nonlinear Poisson–Boltzmann equation. *J. Biomol. Struct. Dyn.* **11**: 1437–1445.
Holst M, Kozack RE, Saied F, Subramaniam S. 1994b. Treatment of electrostatic effects in proteins: multigrid-based-Newton iterative method for solution of the full nonlinear Poisson–Boltzmann equation. *Proteins Struct. Funct. Genet.* **18**: 231–245.
Holst MJ, Baker NA, Wang F. 2000. Adaptive multilevel finite element solution of the Poisson–Boltzmann equation I: algorithms and examples. *J. Comput. Chem.* **21**: 1319–1342.
Honig B, Nicholls A. 1995. Classical electrostatic in biology and chemistry. *Science* **268**: 1144–1149.
Hsu J-P, Liu B-T. 1996a. Exact solution to the linearized Poisson–Boltzmann equation for spheroidal surfaces. *J. Colloid Interface Sci.* **175**: 785–788.
Hsu J-P, Liu B-T. 1996b. Solution to the linearized Poisson–Boltzmann equation for a spheroidal surface under a general surface condition. *J. Colloid Interface Sci.* **183**: 214–222.
Huber GA. 1998. Future directions for combining molecular and continuum models in protein simulations. *Prog. Biophys. Mol. Biol.* **69**: 483–496.
Im W, Beglov D, Roux B. 1998. Continuum solvation model: Electrostatic forces from numerical solutions to the Poisson–Bolztmann equation. *Comput. Phys. Commun.* **111**: 59–75.
Jackson JD. 1962. *Classical Electrodynamics*. John Wiley: New York.
Janin J. 1997. Angstroms and calories. *Structure* **5**: 473–479.
Janin J, Chothia C. 1990. The structure of protein-protein recognition sites. *J. Biol. Chem.* **265**: 16027–16030.
Jones S, Thornton JM. 1996. Principles of protein–protein interactions. *Proc. Natl Acad. Sci. USA* **93**: 13–20.
Karshikoff A, Ladenstein R. 2001. Ion pairs and the thermotolerance of proteins from hyperthermophiles: a "traffic rule" for hot roads. *Trends Biochem. Sci.* **26**: 550–556.
Kasturi S, Kihara A, FitzGerald D, Pastan I. 1992. Alanine scanning mutagenesis identifies surface amino acids on domain II of *Pseudomonas* exotoxin required for cytotoxicity, proper folding, and secretion into periplasm. *J. Biol. Chem.* **267**: 23427–23433.
Katchalski A. 1971. Polyelectrolytes. *Pure Appl. Chem.* **26**: 327–371.
Kella NK, Kinsella JE. 1988. Enhanced thermodynamic stability of beta-lactoglobulin at low pH. A possible mechanism. *Biochem. J.* **255**: 113–118.
Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. 1958. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**: 662–666.
Kirkwood JG. 1934a. On the theory of strong electrolyte solutions. *J. Chem. Phys.* **2**: 767–781.
Kirkwood JG. 1934b. Theory of solutions of molecules containing widely separated charges with special applications to zwitterions. *J. Chem. Phys.* **7**: 351–361.
Klenin KV, Langowski J. 2001. Diffusion-controlled intrachain reactions of supercoiled DNA: Brownian dynamics simulations. *Biophys. J.* **80**: 69–74.
Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE III. 2000. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **33**: 889–897.
Koumanov A, Karshikoff A, Friis EP, Borchert TV. 2001. Conformational Averaging in pK Calculations: Improvement and Limitations in Prediction of Ionization Properties of Proteins. *J. Phys. Chem. B* **105**: 9339–9344.
Kozack RE, d'Mello MJ, Subramaniam S. 1995. Computer modeling of electrostatic steering and orientational effects in antibody–antigen association. *Biophys. J.* **68**: 807–814.
Langner M, Kubica K. 1999. The electrostatics of lipid surfaces. *Chem. Phys. Lipids* **101**: 3–35.
Lazaridis T, Karplus M. 1999. Effective energy function for proteins in solution. *Proteins Struct. Funct. Genet.* **35**: 133–152.
Lee MR, Baker D, Kollman PA. 2001. 2.1 and 1.8 A average C(alpha) RMSD structure predictions on two small proteins, HP-36 and s15. *J. Am. Chem. Soc.* **123**: 1040–1046.
LiCata VJ, Bernlohr DA. 1998. Surface properties of adipocyte lipid-binding protein: Response to lipid binding, and comparison with homologous proteins. *Proteins Struct. Funct. Genet.* **33**: 577–589.
Lifson S, Katchalski A. 1954. The electrostatic free energy of polyelectrolyte solutions. II. Fully stretched macromolecules. *J. Polym. Sci.* **13**: 43–55.
Linderstrom-Lang K. 1924. On the ionization of proteins. *Compt. Rend. Trav. Lab. Carlsberg* **15**: 1–29.
Lo Conte L, Chothia C, Janin J. 1999. The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**: 2177–2198.
Luscombe NM, Laskowski RA, Thornton JM. 2001. Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level. *Nucl. Acids Res.* **29**: 2860–2874.
Luty BA, Davis ME, McCammon JA. 1992. Solving the finite-difference non-linear Poisson–Boltzmann equation. *J. Comput. Chem.* **13**: 1114–1118.
Madura JD, Davis ME, Gilson MK, Wade R, Luty BA, McCammon JA. 1994. Biological applications of electrostatics calculations and Brownian dynamics simulations. *Rev. Comp. Chem.* **5**: 229–267.
Madura JD, Briggs JM, Wade R, Davis ME, Luty BA, Ilin A, Antosiewicz J, Gilson MK, Bagheri B, Ridgway Scott L, McCammon JA. 1995. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics program. *Comput. Commun. Phys.* **91**: 57–95.
Mandel-Gutfreund Y, Schueler O, Margalit H. 1995. Comprehensive analysis of hydrogen bonds in regulatory protein–DNA complexes: in search of common principles. *J. Mol. Biol.* **253**: 370–382.
Mandell JG, Roberts VA, Pique ME, Kotlovyi V, Mitchell JC, Nelson E, Tsigelny I, Ten Eyck LF. 2001. Protein docking using continuum electrostatics and geometric fit. *Protein Engng.* **14**: 105–113.
Manning GS. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **11**: 179–246.
Marcus RA. 1955. Calculation of thermodynamic properties of polyelectrolytes. *J. Chem. Phys.* **23**: 1057–1068.
Mark AE, van Gunsteren WF. 1994. Decomposition of the free

energy of a system in terms of specific interactions. Implications for theoretical and experimental studies. *J. Mol. Biol.* **240**: 167–176.
Mashl RJ, Tang Y, Schnitzer J, Jakobsson E. 2001. Hierarchical approach to predicting permeation in ion channels. *Biophys. J.* **81**: 2473–2483.
Matthews BW. 1993. Structural and genetic analysis of protein stability. *A. Rev. Biochem.* **62**: 139–160.
McCaldon P, Argos P. 1988. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins Struct. Funct. Genet.* **4**: 99–122.
McCammon JA, Gelin BR, Karplus M. 1977. Dynamics of folded proteins. *Nature* **267**: 585–590.
McCoy AJ, Chandana Epa V, Colman PM. 1997. Electrostatic complementarity at protein/protein interfaces. *J. Mol. Biol.* **268**: 570–584.
Mirny L, Shakhnovich E. 2001. Protein folding theory: from lattice to all-atom models. *A. Rev. Biophys. Biomol. Struct.* **30**: 361–396.
Misra VK, Sharp KA, Friedman RA, Honig B. 1994a. Salt effects on ligand-DNA binding. Minor groove antibiotics. *J. Mol. Biol.* **238**: 245–263.
Misra VK, Hecht JL, Sharp KA, Friedman RA, Honig B. 1994b. Salt effects on protein-DNA interactions. The lambda cl repressor and EcoRI endonuclease. *J. Mol. Biol.* **238**: 264–280.
Miyazawa S, Jernigan RL. 1996. Residue–residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J. Mol. Biol.* **256**: 623–644.
Musafia B, Buchner V, Arad D. 1995. Complex salt bridges in proteins: statistical analysis of structure and function. *J. Mol. Biol.* **254**: 761–770.
Nadassy K, Wodak SJ, Janin J. 1999. Structural features of protein–nucleic acid recognition sites. *Biochemistry* **38**: 1999–2017.
Nicholls A, Honig B. 1991. A rapid finite difference algorithm, utililizing successive over-relaxation to solve the Poisson–Boltzman equation. *J. Comput. Chem.* **12**: 435–445.
Nicholls A, Sharp KA, Honig B. 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**: 281–296.
Nicholson H, Becktel WJ, Matthews BW. 1988. Enhanced protein thermostability from designed mutations that interact with alpha-helix dipoles. *Nature* **336**: 651–656.
Niedermeier C, Schulten K. 1992. Molecular dynamics simulations in heterogenous dielectrica and Debye-Huckel media: application to the protein bovine pancreatic trypsin inhibitor. *Mol. Simul.* **8**: 361–387.
Nielsen IMB, Janssen CL. 2001. Novel pseudospectral Fourier method for solving Poisson's equation for a solute in a non-uniform dielectric. *Comput. Phys. Commun.* **136**: 29–36.
Nitta K, Sugai S. 1972. Potentiometric titration studies on globular proteins. *Biopolymers* **11**: 1893–1901.
Northrup SH. 1994. Hydrodynamic motions of large molecules. *Curr. Opin. Struct. Biol.* **4**: 269–274.
Northrup SH, Erickson HP. 1992. Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc. Natl Acad. Sci. USA* **89**: 3338–3342.
Novotny J, Bruccoleri RE, Davis M, Sharp KA. 1997. Empirical free energy calculations: a blind test and further improvements to the method. *J. Mol. Biol.* **268**: 401–411.
Nozaki Y, Tanford C. 1967. Examination of titration behavior. *Meth. Enzymol.* **11**: 715–734.
Oliveberg M, Fersht AR. 1996. New approach to the study of transient protein conformations: the formation of a semiburied salt link in the folding pathway of barnase. *Biochemistry* **35**: 6795–6805.
Onsager L. 1933. Theories of concentrated electrolytes. *Chem. Rev.* **13**: 73–89.
Orttung WH. 1977. Direct solution of the Poisson equation for biomolecules of arbitrary shape, polarizability density and charge distribution. *Ann. NY Acad. Sci.* **303**: 22–37.
Pauling L, Corey RB, Branson HR. 1951. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl Acad. Sci. USA* **37**: 205–212.
Perutz MF. 1978. Electrostatic effects in proteins. *Science* **201**: 1187–1191.
Qin BY, Bewley MC, Creamer LK, Baker HM, Baker EN, Jameson GB. 1998. Structural basis of the Tanford transition of bovine beta-lactoglobulin. *Biochemistry* **37**: 14014–14023.
Rashin AA. 1990. Hydration phenomena, classical electrostatics and the boundary element method. *J. Phys. Chem.* **94**: 1725–1733.
Reiner ES, Radke CJ. 1990. Variational approach to the electrostatic free energy in charged colloidal suspensions: general theory for open systems. *J. Chem. Soc. Faraday Trans.* **86**: 3901–3912.
Romagnoli S, Ugolini R, Fogolari F, Schaller G, Urech K, Giannattasio M, Ragona L, Molinari H. 2000. NMR structural determination of viscotoxin A3 from Viscum album L. *Biochem. J.* **350**: 569–577.
Roux B, Simonson T. 1999. Implicit solvent models. *Biophys. Chem.* **78**: 1–20.
Sandberg L, Edholm O. 1999. A fast and simple method to calculate protonation states in proteins. *Proteins Struct. Funct. Genet.* **36**: 474–483.
Schapira M, Totrov M, Abagyan R. 1999. Prediction of the binding energy for small molecules, peptides and proteins. *J. Mol. Recogn.* **12**: 177–190.
Schultz PG. 1988. The interplay between chemistry and biology in the design of enzymatic catalysts. *Science* **240**: 426–433.
Schutz CN, Warshel A. 2001. What are the dielectric "constants" of proteins and how to validate electrostatic models? *Proteins Struct. Funct. Genet.* **44**: 400–417.
Scrutton NS, Basran J, Sutcliffe MJ. 1999. New insights into enzyme catalysis. Ground state tunnelling driven by protein dynamics. *Eur. J. Biochem.* **264**: 666–671.
Sharp KA. 1991. Incorporating solvent and ion screening into molecular dynamics using the finite-difference Poisson–Boltzmann Method. *J. Comput. Chem.* **12**: 454–468.
Sharp KA. 1995. Polyelectrolyte electrostatics: salt dependence, entropic and enthalpic contribution to free energy in the nonlinear Poisson–Boltzmann model. *Biopolymers* **36**: 227–243.
Sharp KA, Honig B. 1990. Calculating total electrostatic energies with the non-linear Poisson–Boltzmann equation. *J. Phys. Chem.* **94**: 7684–7692.
Sharp K, Fine R, Honig B. 1987. Computer simulation of the diffusion of a substrate to an active site of an enzyme. *Science* **236**: 1460–1463.
Sharp KA, Honig B, Harvey SC. 1990. Electrical potential of transfer RNAs: codon–anticodon recognition. *Biochemistry* **29**: 340–346.
Sharp KA, Friedman RA, Misra V, Hecht J, Honig B. 1995. Salt effects on polyelectrolyte-ligand binding: comparison of Poisson–Boltzmann, and limiting law/counterion binding models. *Biopolymers* **36**: 245–262.
Sheinerman FB, Norel R, Honig B. 2000. Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol.* **10**: 153–159.
Shen T, Wong CF, McCammon JA. 2001. Atomistic Brownian dynamics simulation of peptide phosphorylation. *J. Am. Chem. Soc.* **123**: 9107–9111.
Simonson T. 1998. Dielectric Constant of Cytochrome c from simulations in a water droplet including all electrostatic interactions. *J. Am. Chem. Soc.* **120**: 4875–4876.
Simonson T. 2001. Macromolecular electrostatics: continuum models and their growing pains. *Curr. Opin. Struct. Biol.* **11**: 243–252.
Sines JJ, Allison SA, McCammon JA. 1990. Point charge distributions and electrostatic steering in enzyme/substrate encounter: Brownian dynamics of modified copper/zinc superoxide dismutases. *Biochemistry* **29**: 9403–9412.
Sitkoff D, Sharp KA, Honig B. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **98**: 1978–1988.
Smart JL, Marrone TJ, McCammon JA. 1997. Conformational sampling with Poisson–Boltzmann forces and a stochastic

dynamics/Monte Carlo method: application to alanine dipeptide. *J. Comput. Chem.* **18**: 1750–1759.
Sridharan S, Nicholls A, Sharp KA. 1995. A rapid method for calculating derivatives of solvent accessible surface areas of molecules. *J. Comput. Chem.* **16**: 1038–1044.
Stickle DF, Presta LG, Dill KA, Rose GD. 1992. Hydrogen bonding in globular proteins. *J. Mol. Biol.* **226**: 1143–1159.
Stigter D. 1995. Evaluation of the counterion condensation theory of polyelectrolytes. *Biophys. J.* **69**: 380–388.
Stroppolo ME, Falconi M, Caccuri AM, Desideri A. 2001. Superefficient enzymes. *Cell Mol. Life Sci.* **58**: 1451–1460.
Sulpizi M, Schelling P, Folkers G, Carloni P, Scapozza L. 2001. The rational of catalytic activity of herpes simplex virus thymidine kinase, a combined biochemical and quantum chemical study. *J. Biol. Chem.* **276**: 21692–21697.
Tanford C, Kirkwood JG. 1957. Theory of protein titration curves. I. General equations for impenetrable spheres. *J. Am. Chem. Soc.* **79**: 5333–5339.
Tanford C, Bunville LG, Nozaki Y. 1959. The reversible transformation of β-lactoglobulin at pH 7.5. *J. Am. Chem. Soc.* **81**: 4032–4036.
Ugolini R, Ragona L, Silletti E, Fogolari F, Visschers RW, Alting AC, Molinari H. 2001. Dimerization, stability and electrostatic properties of porcine beta-lactoglobulin. *Eur. J. Biochem.* **268**: 4477–4488.
Ullmann GM, Knapp E-W. 1999. Electrostatic models for computing protonation and redox equilibria in proteins. *Eur. Biophys. J.* **28**: 533–551.
van Vlijmen HW, Schaefer M, Karplus M. 1998. Improving the accuracy of protein pKa calculations: conformational averaging versus the average structure. *Proteins Struct. Funct. Genet.* **33**: 145–158.
Verwey EJW, Overbeek JTG. 1948. *Theory of the Stability of Lyophobic Colloids*. Elsevier: Amsterdam.
von Hippel PH, Berg OG. 1989. Facilitated target location in biological systems. *J. Biol. Chem.* **264**: 675–678.
Wade RC, Luty BA, Demchuk E, Madura JD, Davis ME, Briggs JM, McCammon JA. 1994. Simulation of enzyme–substrate encounter with gated active sites. *Nat. Struct. Biol.* **1**: 65–69.
Waldburger CD, Schildbach JF, Sauer RT. 1995. Are buried salt bridges important for protein stability and conformational specificity? *Nat. Struct. Biol.* **2**: 122–128.
Wang W, Donini O, Reyes CM, Kollman PA. 2001. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein–nucleic acid noncovalent interactions. *A. Rev. Biophys. Biomol. Struct.* **30**: 211–243.
Warshel A. 1981. Calculations of enzymatic reactions: calculations of pKa, proton transfer reactions, and general acid catalysis reactions in enzymes. *Biochemistry* **20**: 3167–3177.
Warshel A. 1998. Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *J. Biol. Chem.* **273**: 27035–27038.
Warshel A, Papazyan A. 1998. Electrostatic effects in macromolecules: fundamental concepts and practical modeling. *Curr. Opin. Struct. Biol.* **8**: 211–217.
Warwicker J. 1999. Simplified methods for pKa and acid pH-dependent stability estimation in proteins: removing dielectric and counterion boundaries. *Protein Sci.* **8**: 418–425.
Warwicker J, Watson HC. 1982. Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. *J. Mol. Biol.* **157**: 671–679.
Watson JD, Crick FHC. 1953. Structure for deoxyribose nucleic acid. *Nature* **171**: 737–738.
Wing R, Drew H, Takano T, Broka C, Tanaka S, Itakura K, Dickerson RE. 1980. Crystal structure analysis of a complete turn of B-DNA. *Nature* **287**: 755–758.
Xu D, Lin SL, Nussinov R. 1997. Protein binding versus protein folding: the role of hydrophilic bridges in protein associations. *J. Mol. Biol.* **265**: 68–84.
Yang AS, Gunner MR, Sampogna R, Sharp K, Honig B. 1993. On the calculation of pKas in proteins. *Proteins Struct. Funct. Genet.* **15**: 252–265.
Yoon BJ, Kim S. 1989. Electrophoresis of spheroidal particles. *J. Colloid Interface Sci.* **128**: 275–288.
You TJ, Harvey SC. 1993. Finite element approach to the electrostatics of macromolecules with arbitrary geometries. *J. Comput. Chem.* **14**: 484–501.
Yu MH, Weissman JS, Kim PS. 1995. Contribution of individual side-chains to the stability of BPTI examined by alanine-scanning mutagenesis. *J. Mol. Biol.* **249**: 388–397.
Zacharias M, Luty BA, Davis ME, McCammon JA. 1994. Combined conformational search and finite-difference Poisson–Boltzmann approach for flexible docking. Application to an operator mutation in the lambda repressor-operator complex. *J. Mol. Biol.* **238**: 455–465.
Zauhar RJ, Morgan RS. 1985. A new method for computing the macromolecular electric potential. *J. Mol. Biol.* **186**: 815–820.
Zhou HX. 1990. On the calculation of diffusive reaction rates using Brownian dynamics simulations. *J. Chem. Phys.* **92**: 3092–3095.
Zhou HX. 1993. Boundary element solution of macromolecular electrostatics: interaction energy between two proteins. *Biophys. J.* **65**: 955–963.
Zhou HX. 1994. Macromolecular electrostatic energy within the nonlinear Poisson–Boltzmann equation. *J. Chem. Phys.* **100**: 3152–3162.
Zhou HX, Vijayakumar M. 1997. Modeling of protein conformational fluctuations in pK_a predictions. *J. Mol. Biol.* **267**: 1002–1011.

