Macromolecular electrostatics: continuum models and their growing pains

Thomas Simonson

Theoretical understanding of macromolecular electrostatics has advanced substantially over the past year. Continuum models have given promising results for calculating protein–ligand binding free energy differences, as well as pK_as and redox properties, particularly with explicit treatment of multiple conformers. Generalized Born and other techniques have led to the first molecular dynamics simulations of proteins and RNA with continuum solvent. Continuum and microscopic descriptions of dielectric relaxation have been critically compared.

Addresses
Laboratory for Structural Biology and Genomics, CNRS, IGBMC, 1 rue Laurent Fries, 67404 Strasbourg-Illkirch, France; e-mail: simonson@igbmc.u-strasbg.fr

Current Opinion in Structural Biology 2001, 11:243–252

0959-440X/01/$ – see front matter

Abbreviations

BPTI	bovine pancreatic trypsin inhibitor
FDPB	finite-difference Poisson–Boltzmann
GB	generalized Born
MD	molecular dynamics
PB	Poisson–Boltzmann
rms	root mean square

Introduction

Theoretical treatments of electrostatics go back to Linderstrom–Lang, Tanford, Kirkwood, Hill and the days when proteins were spherical. Long before molecular dynamics (MD) simulations, they had expanded into ellipsoids, truncated spheres, cylinders and slabs. Today, when one speaks of continuum electrostatics models in structural biology, one usually refers to the semi-microscopic numerical approaches pioneered by Warwicker and Watson [1], and represented by household software packages such as GRASP [2]. These are atomistic models that have the ambition to provide not only qualitative insights but also, in some cases, quantitative predictions. For many applications, continuum models do surprisingly well, yielding at least semi-quantitative accuracy for aspects such as redox properties, pH-dependent protein stability and conformational thermodynamics of peptides and nucleic acids [3]. Unfortunately, their use often involves adjustment of parameters, such as the solute dielectric constant(s). Indeed, although they can accurately model long-range electrostatic interactions, continuum models tend to break down when short-range effects become important, such as individual water molecules bridging functional groups. To address this, theories have evolved from considering spheres and slabs to increasingly complex descriptions in which more and more degrees of freedom are treated in atomic detail. As a result, continuum electrostatic models now span an enormous area. At one end are the 'physicists', with semi-analytical treatments of large macromolecular assemblies in nonideal solutions [4]. At the other are the 'chemists': only the solvent is implicit — the solute is simulated by MD and/or quantum mechanics [5,6]. This review covers a broad intermediate area, which should be most relevant to structural biology. It includes the time period between early 1999 and the end of 2000. I focus almost exclusively on continuum electrostatic models, referred to below simply as 'continuum models'.

Despite the continuing flow of interesting applications, I believe that the most important recent contributions are methodological and my main emphasis is on these. Most fall into two categories: steps to improve the continuum model by making it more microscopic and critical studies of the model's limitations and its relation to the underlying microscopic structure and dynamics. I start with studies of molecular recognition, including calculations of protein–ligand binding free energies and very fast methods for docking. Next, I consider techniques to couple continuum models with MD simulations. Several of these rely on the generalized Born (GB) model. I then describe critical comparisons between continuum models and microscopic approaches, including calculations of protein dielectric constants from MD simulations. Finally, I describe several recent studies of acid–base equilibria, and discuss them in the light of the critical comparisons above.

This review is far from exhaustive, even for the time period covered, and I refer the reader to several other review articles on this and related areas. Implicit solvent models, including lattice models, integral equation models and surface area models, have been reviewed recently [5–8]. Electrostatic free energy calculations [3], protein–protein association and interactions [9], spectroscopic applications [10], binding free energies [11,12] and the GB model [13] have all been reviewed. For reasons of space, calculations with a quantum mechanical solute and continuum solvent are beyond the scope of this review: they are described in [6]. Important new theories of the hydrophobic effect could also not be covered (see [14•,15••] and references therein).

Molecular recognition studies

Continuum models are being increasingly used to study binding affinities. This is a challenging area, in which electrostatic interactions play an important role, but are not necessarily predominant. Therefore, continuum models have traditionally been used to compute relative binding free energies of series of ligands or receptors in such a way that nonelectrostatic contributions approximately cancel

out. Attempts to increase the scope of continuum models by including 'nonelectrostatic' terms have a long history (see e.g. [5,6,11]). The use of such hybrid models has increased over the past year and they are now being applied to macromolecule–ligand association. Other recent developments include free energy component analyses, which provide insights into the structural basis of ligand binding and specificity, and the emergence of very fast continuum models for docking and ligand design.

Binding free energies and free energy differences

Many competing effects can contribute to ligand–receptor binding free energies: changes in the rotational, translational, conformational and vibrational entropy of the partners; entropy changes associated with solvent ordering around hydrophobic or charged groups; solute conformational strain; changes in electrostatic and van der Waals interactions within and between the partners and the solvent; and counterion reorganization. All of these effects can be accounted for automatically in explicit solvent simulations, as long as sufficient conformational sampling is performed. They have all been included in simplified continuum or semi-microscopic models in the past [11], with different flavors and varying degrees of success.

The simplest case is the calculation of relative binding free energies in systems dominated by electrostatics. The continuum electrostatic free energies of the bound and free states are simply subtracted. The key model ingredients are the dielectric constant(s) used for the solutes and the structures used to model the various states (e.g. bound, unbound, native, mutant). Hunenberger *et al.* [16] obtained good agreement with experimental binding constants for inhibitors binding to cAMP-dependent protein kinase using a single solute dielectric of two and crystal structures of the various protein–inhibitor complexes. Archontis *et al.* [17•] found that a solute dielectric of four was appropriate to reproduce MD free energy simulation results for aspartic acid and asparagine binding to aspartyl-tRNA synthetase. Results were averaged over structures taken from MD simulations of the bound states; experimentally, the structure of the native protein hardly changes when the aspartic acid ligand is removed, so the unbound state could be modeled by simply removing the ligand from the bound-state structures. Good agreement was also obtained for a K198L mutant. But, if structures of the native protein were used to model the mutant complexes, the aspartic acid and asparagine binding free energies were in error by over 10 kcal/mol. Chong *et al.* [18•] used a solute dielectric of one and MD structures of the bound states to study hapten binding to a mature and germ-line antibody; good agreement with the experimental binding free energy difference was obtained. From these studies, it appears that the model structures should be representative of the states considered and the dielectric constant must be adjusted empirically for each system. Given the complexity of the binding reaction, this is, perhaps, less surprising than the fact that such simple models work at all.

A related 'alanine-scanning' approach was applied to a fragment of p53 binding to deletion mutants of the oncoprotein Mdm2 [19]. The key idea is to study a large number of mutations using the structure of only one variant, for example, the native protein. Although a correlation between calculated binding free energies and an experimental 'mutational tolerance' was observed, experimental binding free energies were not available. Binding of native and mutant U1A proteins to a hairpin RNA was also studied [20]. Structures of the mutants were obtained by simply replacing the mutated sidechain in the native crystal structure and energy minimizing. Although this led to reasonable binding free energy differences in most cases, in one case that disagreed with experiment, the mutation went on to be modeled more elaborately using an MD simulation, upon which the binding free energy changed by 10 kcal/mol. It appears that, although the possibility of studying multiple mutations using only one or a few structures would be very useful, the validity of the method remains to be established.

Other recent studies have attempted to calculate absolute binding free energies [21,22]. This is much more difficult, as discussed above. For example, solvent entropy contributions were modeled rather crudely using a term proportional to the surface area buried upon binding. This is appropriate for small linear alkanes [5], but is of uncertain quality for large solutes [14•,15••]; in particular, water depletion of large hydrophobic surfaces is a collective effect [15••], such that a linear model (free energy proportional to surface area) may be inappropriate. Nevertheless, one variant of the model gave good discrimination between a library of misfolded proteins and their correct folds [23] (see also Update).

Free energy component analyses

Component analyses of Poisson–Boltzmann (PB) binding free energies have been proposed by many authors, but a fully systematic treatment was developed only recently [24•] and has been applied to GCN4 leucine zipper formation and to amino acid binding by aspartyl-tRNA synthetase [17•]. It distinguishes two main effects: partial desolvation of each ligand due to its association with the other and interactions between the charges on the two ligands, screened by the surrounding dielectric media. Each of these two terms can be decomposed further into residue or group contributions. In the GCN4 dimer [24•], electrostatic effects were found to disfavor dimerization by as much as 15 kcal/mol as a result of the desolvation of charged and polar group, insufficiently compensated by direct interactions between monomers in the bound state. Desolvation always reinforces interactions within each ligand, a net favorable effect in this system, but still insufficient to compensate for the loss of interactions with the solvent. In aspartyl-tRNA synthetase [17•], the component analysis was used to identify groups that discriminate between binding of the substrate aspartic acid and the analog asparagine to the native and mutant

proteins. This was especially useful because not all of the most important interactions were evident from visual inspection of the structures. Many competing interactions exist and point mutations around the binding pocket often led to structural shifts, whereby new groups became strongly involved and old interactions were weakened. This study included the first detailed comparison with free energy components from alchemical MD free energy simulations. In another study, Roux and MacKinnon [25•] analyzed the selectivity of the KcsA potassium channel in terms of particular free energy components, including desolvation components and direct interactions of the ion with the channel helices. The partial ion desolvation that occurs in the channel favors monovalent over multivalent ions; the field of the pore helices, enhanced by the low dielectric membrane environment, favors cations over anions. Thus, the channel appears to confer selectivity for monovalent cations through simple electrostatic principles.

Fast methods for docking and ligand design

A very important application of continuum models is the rapid screening of many possible ligands and binding modes for a particular target macromolecule. The usual finite-difference Poisson–Boltzmann (FDPB) or boundary element methods are much too slow for these applications, which require additional approximations. Several current strategies rely on the GB model, reviewed recently [13], which represents an analytical approximation to the complete dielectric continuum solvent and calculates solvation energies hundreds of times faster than FDPB.

In the GB model [13,26], the interaction between two charges q_i and q_j takes the form:

$$u(r_{ij}) = \frac{q_i q_j}{r_{ij}} + \left(\frac{1}{e_w} - 1\right) \frac{q_i q_j}{\left[r_{ij}^2 + b_i b_j \exp\left(-\frac{r_{ij}^2}{4 b_i b_j}\right) \right]^{1/2}} \qquad (1)$$

where r_{ij} is the distance between the charges, e_w is the solvent dielectric and b_i and b_j are effective Born solvation radii, which are related to the Born self energy $E_s(i)$ of each charge through:

$$E_s(i) = \frac{q_i^2}{2 b_i} \left(\frac{1}{e_w} - 1\right). \qquad (2)$$

The b_i are functions of the environment of each charge and can be thought of as the distance between the charge and the solvent. The energy in Equation 1 reproduces exactly the full dielectric continuum energy in the limiting cases of a pair of very distant or very close charges. Several approximations have been proposed for the solvation radii (b_i) [13,26–28]. With these approximations, b_i and $u(r_{ij})$ become complicated but (usually) differentiable functions of the positions of all of the solute atoms, which can be used in energy minimization and MD simulations, and which are now available in many modeling programs, such as MacroModel, AMBER, CHARMM, XPLOR and TINKER.

The model agrees well with FDPB for determining solvation and conformational free energies of small molecules and peptides [13]. For large solutes such as proteins, differences between the physics of the GB model and the 'exact' dielectric continuum approach become significant, and extensive parameterization and testing of the GB model are needed to recover and verify dielectric-continuum-like behavior. Caflisch and co-workers [29•] performed extensive comparisons with the full dielectric continuum approach for small-molecule–protein binding. Other recent studies tested GB combined with a Gaussian hydrophobic energy term [30] and with nonzero ionic strength [31], and culminated in several MD simulations of proteins and nucleic acids [32•–35•,36]. Overall, the model performs surprisingly well even for large solutes. Kuntz and co-workers [37] found that it led to improved ligand scoring compared with the distant-dependent dielectric used in earlier versions of the DOCK program. Dominy and Brooks [38] also proposed a docking method using GB solvation. Another, still faster, approach simplifies the GB model further, assuming that the main contribution to electrostatic desolvation arises from the displacement of waters in the first shell around the receptor when the ligand binds [39]. The main energy components can then be precomputed on a grid surrounding the binding site, allowing the estimation of about 300 protein–fragment binding modes per second on a personal computer. A related approximation was proposed by Arora and Bashford [40].

In general, the problem of ligand design is complicated by the need for specificity, that is, the ligand should have a high affinity for the target, but should not bind to competing receptors or decoys; similarly, the receptor should not bind competing ligands. Kangas and Tidor [41] proposed an elegant analysis of specificity in a simplified mathematical framework. The ligand charges are optimized against a cost function within a space of finite dimensionality, corresponding to a finite set of predefined possible ligands and ligand positions. The cost function involves the binding free energies to the target(s) and decoy(s). Studies of model systems suggest that there is usually a trade-off between affinity and specificity, but that electrostatics can often enhance specificity with only minimal loss of affinity for the target (see also Update).

Continuum models go microscopic

'Continuum' models are becoming increasingly atomistic. Although atomic charge distributions have been used since the late 1980s, the description of dielectric relaxation was normally a 'macroscopic' one — that is, in response to a new charge or field, the system could only respond by a redistribution of induced polarization, determined by the solute and solvent dielectric constants. Today, it is the microscopic relaxation mechanisms that are increasingly included at an atomic level.

Introducing multiple conformers

A first step in this direction was a novel multiple-conformer method [42]. Selected residues or groups in the

macromolecule were each assigned a few, predetermined possible conformers. For example, charged sidechains in a protein could be assigned their most common rotamers. The different conformers play a role in the model that is analogous to the different possible protonation states of the titratable residues. Depending on interactions with their environments, each group occupies each of its available conformers with a certain Boltzmann probability. When a nearby group binds a proton or a redox electron, for example, the occupancies adjust automatically, so that a residue may shift (wholly or fractionally) from one conformer to another. Thus, a mechanism for explicit conformational relaxation is included in the model. This approach was used recently to study coupled proton and electron binding in a photosynthetic reaction center [43•,44]. Difficulties include the need to define in advance which conformations are expected to be important, the exponential growth of possible states with the number of conformers and the need for a balanced combination of explicit and implicit relaxation channels (the appropriate solute dielectric constant is likely to depend on the extent of 'explicit' relaxation allowed, that is, on the number and type of conformers included). With increasing computer power, however, large numbers of conformers can be used. Mean field and Monte Carlo methods also continue to be developed for the efficient sampling of states [45]. The main advantage of the method is the ability to treat local, anisotropic structural rearrangements, which might not be accurately described by an average, isotropic dielectric constant. By predicting which rearrangements occur as a function of pH or redox potential, the model can provide structural insights. For example, concerted rearrangements of protein sidechains and ordered waters were observed upon reducing a quinone in the reaction center [43•].

A novel method by Havranek and Harbury [46••] uses the multiple-conformer treatment with a fast approximation to the continuum model, which combines elements of the GB model and an image charge method. Following the studies of Tanford and Kirkwood [47], and more recently Abagyan and co-workers [11], the local region around each protein charge is mapped onto an equivalent low dielectric sphere and the solvent polarization induced by the protein charge is represented by an image charge. The method was applied to BPTI acid–base titration, accurately reproducing the experimental pK_as.

Molecular dynamics with continuum solvent

A more radical step towards microscopic realism is to treat the macromolecule by MD or Monte Carlo simulation, with only the solvent treated implicitly. With this approach, the macromolecule will usually be assigned a dielectric constant of one, as dielectric relaxation in response to perturbing charges or fields is modeled explicitly by structural rearrangements occurring during the simulation. Several implicit solvent models have been used in MD simulations [5–8] and models based on a dielectric continuum solvent are among the most attractive. Until very recently, they have been limited by cost and difficulties in computing accurate forces, but in the past year, simulations of both proteins and nucleic acids have been performed using the GB model, and the first protein simulation with a full dielectric continuum solvent was published [32•–35•].

Molecular dynamics with the generalized Born model

The past year has seen the first MD studies of proteins and nucleic acids using the GB model [32•–35•]. Direct comparison between GB and the full dielectric continuum solvent had already shown that conformational free energies [30,32•] and charge self energies [31,32•,34•] were in reasonable agreement, possibly within the error of the full dielectric continuum model. In MD simulations of nucleic acids, structures in good agreement with experiment and explicit solvent simulations were sampled [33•,34•]. These last two studies used two different functional forms for the analytical solvation radii b_i (Equation 1) [26,27]. Conformational sampling was found to be faster than in explicit solvent simulations, probably because of the reduced mechanical friction and the instantaneous dielectric response of the continuum solvent. In MD simulations of a small protein, a slight reparameterization of the solvation radii given in [26] led to stable structures with a mean deviation from the crystal structure of only 1.1–1.3 Å, even smaller than with explicit solvent (1.5 Å) [32•]. The computer cost was only five times greater than a vacuum simulation and over ten times less than with explicit solvent. The use of the analytical solvation radii of Schaefer *et al.* (parameterized for small peptides) [28,48] led to stable MD trajectories for two proteins [36]; the trajectories deviated by about 2 Å from the crystal structure. Several simpler solvent models gave poorer results.

These results show that the GB model, suitably parameterized, can be used for MD not only of small peptides [48], but also of proteins and nucleic acids with at least semi-quantitative accuracy. Potential applications include conformational searching, NMR structure refinement and calculation of pK_as with MD free energy simulation methods. The first such application was a study of the stability of several mesophilic and thermophilic proteins by Dominy and Brooks (B Dominy, C Brooks, personal communication). The thermophiles were found to have enhanced polarizabilities, which were proposed to play a role in their enhanced stability by reducing the penalty to partially desolvate charged groups upon folding of the protein. The corresponding reduction in favorable charge–charge interactions within the protein was estimated to make a smaller contribution, so that the overall effect of the enhanced polarizability was stabilizing.

Monte Carlo simulations with GB solvent have been extensively used by Gilson and co-workers [49] to study

molecular recognition in small host–guest systems, including helical peptides. This application illustrated the very weak contribution (<1 kcal/mol) of solvent-exposed salt bridges to helix stability.

Molecular dynamics with a full dielectric continuum solvent

This area has been very active over the past year, with several new formulations of the method, significant technical advances and the first simulation of a protein with a full dielectric continuum solvent.

In textbooks, the forces arising from a continuum solvent are derived from the Maxwell stress tensor (e.g. Equations 15.12 and 15.13 of [50]). Only recently was this formalism turned into a useful algorithm for biomolecules [35•,51]. Two new formulations have been based on quite different approaches, inspired by techniques from *ab initio* molecular orbital theory. The first involves minimization of a free energy functional, which can be done 'on the fly' during an MD simulation [52]; see also [53]. The second involves expanding either the electrostatic potential [54,55,56••] or the solute charge density [57••] in terms of a set of basis functions, leading to tractable analytical expressions for the solvation forces. These new techniques are still at an experimental stage, but they open significantly different routes that are likely to be very useful. Thus, the first MD free energy simulation of a protein in which part of the system was treated as a dielectric continuum was reported recently [58].

Most FDPB applications in the past used simple periodic grids to discretize the PB equation. Several groups are now using adaptive approaches to construct 'smarter', nonperiodic grids, which are dense close to the solute boundary and to source charges, and sparse elsewhere [55,59]. This should increase efficiency and eventually allow faster MD simulations. In one implementation on a massively parallel computer, a model of a microtubule with 600,000 particles was studied [56••].

The first MD simulation of a protein in a dielectric continuum solvent was a nanosecond simulation of HIV protease [35•]. The 'traditional' methodology [50] was improved by a careful treatment of both short-range forces and forces near the dielectric boundary, leading to a very stable trajectory in good agreement with the experimental structure. The simulation required 10 times more computer time than a vacuum simulation, still significantly faster than an explicit solvent simulation. Moreover, efficiency will presumably improve (see above). Conformational sampling also appears to be faster with a continuum solvent than with explicit solvent because of reduced friction [33•,34•]. Finally, the continuum solvent simulation directly yields the potential of mean force (the 'conformational free energy') for the solute [5,23], from which the probabilities of different conformations can be extracted much more readily than with explicit solvent.

Dielectric relaxation in macromolecules

In the applications of the 'Molecular recognition studies' section, the continuum model was mainly used to predict free energy changes. Taking a different perspective, the continuum model can provide a phenomenological interpretation of dielectric relaxation or polarization observed experimentally or in MD simulations. For example, the continuum model that 'best fits' the MD of a protein can be determined through the so-called Frohlich–Kirkwood theory, providing a 'nonempirical' estimate of the protein dielectric constant that measures the average protein polarizability (see [60,61] for reviews). Similarly, continuum models that best fit reorganization free energies in cytochrome *c* have been determined [62,63], providing a measure of the local polarizability around specific groups, such as the heme. Continuum models parameterized in this way yield useful physical insights. They also help to determine the limitations of the dielectric continuum approach and ways to improve it.

A recent advance was the demonstration that, in one protein [64••], substantially different continuum models are needed to fit different processes, depending on whether they involve only an equilibrium charge distribution ('static' processes) or relaxation in response to a perturbing charge ('relaxation' processes). This was demonstrated by studying the insertion of a charge onto protein atoms. The insertion can be performed in two steps [65]. In the first, 'static' step, the new charge is inserted, but the environment (protein and solvent) is not allowed to adjust. Specifically, polarization charge is not allowed to rearrange. In the second, 'relaxation' step, the environment adjusts to the product state charge distribution through a redistribution of polarization charge. The free energy of the second step can be shown to be independent of the charge set used for the reactant state and to depend only on the new charge and the dielectric constants. It is closely related to the reorganization energy of Marcus electron transfer theory [65,66]. The free energy of the first step depends on the equilibrium electrostatic potential at the charge insertion sites in the reactant state, which, in turn, depends strongly on the charge set used for the protein, as well as on the dielectric constants. If a charge set borrowed from a molecular mechanics force field is used, it is likely that a low protein dielectric constant will be appropriate for the first step (the charge set is probably optimized for use in simulations with a dielectric constant of one). The protein dielectric constant for the second step should accurately measure the polarizability of the medium. For charge insertion in the active site of the studied enzyme [64••], protein dielectrics of 1 and of 5–6 were optimal for the static and relaxation steps, respectively (80 was used for solvent). These values show that, in this system, the molecular mechanics charge set cannot provide a consistent description of both equilibrium potentials and dielectric relaxation using the same protein dielectric constant. The calculations also provide a phenomenological characterization of the polarizability of the active site considered:

though very polar, it is only moderately polarizable, with a local protein dielectric (for relaxation) of 5–6.

A similar analysis has been applied to other types of charge rearrangement in model systems, including photoexcitation [67,68]. Another study compared the equilibrium solvent polarization around a protein from an explicit solvent MD simulation with that from a PB continuum model [69]. No relaxation in response to perturbing charges is involved (the structure is at equilibrium with the charge distribution) and, indeed, a low protein dielectric constant of two gave qualitative agreement with the MD polarization (though the latter was probably affected by the use of an electrostatic cutoff). A continuum model was also used to probe the electrostatic potential produced in 305 proteins by the peptide backbone [70]. The potential was found to be predominantly positive because, in allowed regions of the Ramachandran diagram, the sidechain tends to come off the backbone near the amide hydrogen and because the carbonyl oxygens tend to be solvent-exposed.

Another contribution was the calculation of the Frohlich–Kirkwood dielectric constants of four proteins using multinanosecond MD simulations including all electrostatic interactions [71•]. The results are consistent with earlier studies that used somewhat shorter simulations and/or less accurate long-range electrostatics. All of the proteins have a fairly large (15–40) overall dielectric constant, which arises mainly from the motion of charged sidechains. This confirms that the protein Frohlich–Kirkwood dielectric 'constant' is spatially heterogeneous, with low values (2–4) in the interior and large values close to the surface. How do these values relate to applications of continuum models for, say, calculations of spectroscopic properties or pK_as? By construction, the Frohlich–Kirkwood dielectric constant measures the polarizability of the protein and should be appropriate to describe relaxation in response to a perturbation, such as the binding of a proton or a redox electron. In principle, the spatial heterogeneity of the system should be taken into account here (though, in practice, the low internal dielectric constant gave reasonable results for charge binding throughout cytochrome *c* [63,72]). In contrast, the Frohlich–Kirkwood dielectric constant may be completely inappropriate for the 'static' free energy component of the same charge-binding process (and for free energies of more complicated processes, such as ligand binding) [60].

Optical spectroscopy is a powerful experimental tool to study both equilibrium fields and dielectric relaxation in proteins [10,68,73,74]. Hole burning experiments on cytochrome *c* were recently compared with continuum models [73,74]. A quantum mechanical treatment of the heme is needed for accurate results [74]. A simple classical model was, however, sufficient to qualitatively describe the field heterogeneity at the heme [73]; snapshots from an MD trajectory served as input for FDPB calculations. Cytochromes from yeast and horse heart differed by the number of 'electrostatically different conformations' they sampled, accounting for the difference in observed spectral populations.

Finally, studies of time-dependent dielectric relaxation with continuum models are becoming more common, with the development of a powerful, general formalism [75] and recent applications to small chromophores [72,75] and one protein [76]. This will be useful for understanding how enzymes and photoactive proteins control the dynamics of charge binding and charge transfer.

Acid–base equilibria

The calculation of proton-binding constants continues to be one of the main applications of continuum electrostatics. Recent work has addressed the coupling of proton binding and electron transfer in redox and photoactive proteins [43•,44,77–79,80•], proton transfer steps in the reaction mechanism of beta-lactamases [81,82] and pH-dependent protein stability [83]. This area presents several difficulties. For most titratable groups in proteins, the pKa shifts relative to model compounds in solution are very small. They correspond to free energies of less than 1 kcal/mol, a formidable target for a theoretical model. If model quality is measured using absolute pK_as, it will be very hard to outperform a null model, in which all pK_a shifts are taken to be zero [84]. Models with a large protein dielectric constant have been found to perform well on average [81,82,84,85•]. It has sometimes been argued that the large dielectric constant may be needed to account for protein flexibility (but see below) or counterion binding. In fact, as the protein dielectric becomes larger — approaching that of water in some applications [81,82,85•] — the model becomes more and more like the null model. Differences arise from Coulomb interactions among protein residues, but these are reduced by a factor approaching 80, so that a 3 Å salt-bridge contact yields a pK_a shift of approximately 1 instead of about 20 in a medium with dielectric 4.

The conformational flexibility of proteins does indeed lead to large average dielectric constants, as calculated from MD simulations and Frohlich–Kirkwood theory (see section 'Dielectric relaxation in molecules'). However, the calculated dielectric constant is spatially heterogeneous, with much lower values throughout the protein bulk [60,71•,72]. In addition, although this dielectric constant should be at least roughly appropriate to describe relaxation in response to a perturbing charge, there is no direct evidence that it is appropriate for equilibrium potentials and, hence, for some of the free energy components associated with proton binding.

Instead of asking why a large protein dielectric 'works', one might ask why, in some cases, seemingly more reasonable low values fail. Spatial heterogeneity of the protein dielectric properties, which is usually neglected, has been invoked by several authors and must certainly play a role

in some cases. However, the charge-binding study discussed previously [64••] suggests an additional, more subtle, possibility. Although a protein dielectric of around four [78,79,80•] should give a reasonable description of the conformational relaxation, combining it with a molecular mechanics charge set parameterized with a dielectric of one may lead to internal inconsistency of the model, with equilibrium potentials being incorrectly scaled down. Related arguments were discussed in [66]. This inconsistency should be signaled by large differences between free energies calculated in the reactant → product and product → reactant directions [64••].

Conclusions

The low computational cost and wide availability of continuum models, especially coupled to computer graphics, have made them household items in structural biology groups. Qualitative analysis of electrostatic potentials is a standard step when a new structure is determined and structure classification based on electrostatic potentials has become a tool in bioinformatics [86•]. Although this has certainly raised 'electrostatic awareness', one should keep in mind the often simplistic assumptions of continuum models and the potential dangers of 'fixing' them by adjusting parameters.

Many applications to molecular recognition were reported in the past year. Calculations of binding free energy differences show that continuum models can be a valuable tool for ligand and protein design, as long as structural models are generated carefully and consistently, and the solute dielectric constant is calibrated and validated for a given system using experimental data and/or MD free energy simulations. Component analyses of PB free energies can then provide valuable insights into the sources of affinity and specificity. Significant progress has also been made in developing efficient treatments of electrostatics and solvation for docking.

Comparisons between continuum and microscopic descriptions of dielectric relaxation are a fruitful route for understanding the limitations of continuum models and finding ways to improve them. In one case, it was shown that different protein dielectric constants were required to describe static fields and field shifts induced by a perturbing charge. This indicates that, in some current applications, the charge set and dielectric constants may not be mutually consistent. Spatial heterogeneity of the solute dielectric properties is another, distinct effect that undoubtedly plays a role in many systems.

The most robust recipe for a fully consistent continuum model is to make the solute 'fully microscopic' by performing MD, with the continuum treatment limited to the solvent. Multiple-conformer methods for pK_a or redox calculations are another step in the same direction. MD with continuum solvent should be especially useful for conformational searching and when coupled to free energy simulation techniques. This approach will be too expensive or complex for some current applications. For example, the treatment of many simultaneously titrating sites is beyond the scope of current MD simulations and will require special techniques. Meanwhile, simpler continuum models will continue to provide essential qualitative insights into biomolecular structure and function.

Update

Lee and Tidor [87] applied their charge optimization method to the barnase–barstar complex, systematically exploring alternate charge distributions for each of the interfacial amino acid sidechains (with fixed backbone conformation). The optimal net charge for each sidechain agreed with its 'natural' charge, suggesting that the complex is electrostatically optimized for tight binding.

Lee *et al.* [88•] tested the ability of a physically based energy function to discriminate between native-like folds of two proteins. The folds were taken from a large set of *ab initio* structure predictions, generated in the CASP III structure prediction forum, whose rms deviations from the experimental structure ranged from about 1.5 to 8 Å. The energy function included a molecular mechanics term, a continuum dielectric solvation term and an accessible surface term. After a structure refinement step using MD in explicit solvent, this energy function displayed a good correlation with the rms deviation between the model structure and the experimental structure, allowing a fairly accurate ranking of the the *ab initio* predictions.

Acknowledgements

I thank the Centre National de la Recherche Scientifique for support, and B Dominy and C Brooks for communicating unpublished data.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Warwicker J, Watson H: Calculation of the electrostatic potential in the active site cleft due to alpha-helix dipoles. *J Mol Biol* 1982, 157:671-679.

2. Honig B, Nicholls A: Classical electrostatics in biology and chemistry. *Science* 1995, 268:1144-1149.

3. Schaefer M, van Vlijmen H, Karplus M: Electrostatic contributions to molecular free energies in solution. *Adv Protein Chem* 1998, 51:1-57.

4. Gelbart W, Bruinsma R, Pincus P, Parsegian V: DNA-inspired electrostatics. *Physics Today* 2000, 53:38-44.

5. Roux B, Simonson T: Implicit solvent models. *Biophys Chem* 1999, 78:1-20.

6. Orozco M, Luque F: Theoretical methods for the description of the solvent effect in biomolecular systems. *Chem Rev* 2000, 100:4187-4226.

7. Warshel A, Papazyan A: Electrostatic effects in macromolecules: fundamental concepts and practical modelling. *Curr Opin Struct Biol* 1998, 8:211-217.

8. Lazaridis T, Karplus M: Effective energy functions for protein structure prediction. *Curr Opin Struct Biol* 2000, 10:139-145.

9. Sheinerman F, Norel R, Honig B: Electrostatics aspects of protein-protein interactions. *Curr Opin Struct Biol* 2000, 10:153-159.

10. Laberge M: Intrinsic protein electric fields: basic non-covalent interactions and relationship to protein-induced Stark effects. *Biochim Biophys Acta* 1998, 1386:305-330.

11. Schapira M, Totrov M, Abagyan R: Prediction of the binding energy for small molecules, peptides and proteins. *J Mol Recognit* 1999, 12:177-190.

12. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W *et al.*: Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res* 2000, 33:889-897.

13. Bashford D, Case D: Generalized Born models of macromolecular solvation effects. *Annu Rev Phys Chem* 2000, 51:129-152.

14. • Hummer G: Hydrophobic force field as a molecular alternative to surface area models. *J Am Chem Soc* 1999, 121:6299-6305.
An effective pairwise force field for hydrophobic interactions was developed from an information theory of hydrophobic hydration. Results for a 14-methane cluster agree with MD results and disagree with a surface area model.

15. •• Huang D, Chandler D: Temperature and length scale dependence of hydrophobic effects and their possible implications for protein folding. *Proc Natl Acad Sci USA* 2000, 97:8324-8327.
A new theory of hydrophobic hydration reproduces known results at both short and mesoscopic length scales. The large difference between microscopic and macroscopic surface tensions is explained by water depletion, a collective effect at sufficiently large (nanometer) hydrophobic surfaces. This effect is missing from simple surface area models.

16. Hunenberger P, Helms V, Narayana N, Taylor S, McCammon J: Determinants of ligand binding to cAMP-dependent protein kinase. *Biochemistry* 1999, 38:2358-2366.

17. • Archontis G, Simonson T, Karplus M: Binding free energies and free energy components from molecular dynamics and Poisson-Boltzmann calculations. Application to amino acid recognition by aspartyl-tRNA synthetase. *J Mol Biol* 2001, 306:307-327.
MD with explicit solvent combined with FDPB is proposed as a tool for protein and ligand design. A component analysis of the FDPB binding free energy differences is compared with the free energy components from an MD alchemical free energy simulation for the first time. The MD simulations predict structural shifts caused by point mutations, such that new groups become involved in ligand recognition that could not be anticipated from inspecting the native structure alone.

18. • Chong L, Duan Y, Wang L, Massova I, Kollman P: Molecular dynamics and free energy calculations applied to affinity maturation in antibody 48G7. *Proc Natl Acad Sci USA* 1999, 96:14330-14335.
Antibody–hapten binding free energy differences were calculated by applying FDPB to structures sampled by MD in explicit solvent. Absolute binding free energies are also estimated qualitatively using a simple surface area model for hydrophobic effects and normal mode analysis for vibrational entropy changes.

19. Massova I, Kollman P: Computational alanine scanning to probe protein–protein interactions: a novel approach to evaluate binding free energies. *J Am Chem Soc* 1999, 121:8133-8143.

20. Reyes C, Kollman P: Investigating the binding specificity of U1A-RNA by computational mutagenesis. *J Mol Biol* 2000, 295:1-6.

21. Jayaram B, McConnell K, Dixit S, Beveridge D: Free energy analysis of protein-DNA binding: the EcoRI endonuclease-DNA complex. *J Comput Phys* 1999, 151:333-357.

22. Reyes C, Kollman P: Structure and thermodynamics of RNA–protein binding: using molecular dynamics and free energy analyses to calculate the free energies of binding and conformational change. *J Mol Biol* 2000, 297:1145-1158.

23. Vorobjev Y, Hermans J: ES/IS: estimation of conformational free energy by combining dynamics simulations with explicit solvent with an implicit solvent continuum model. *Biophys Chem* 1999, 78:195-205.

24. • Hendsch Z, Tidor B: Electrostatic interactions in the GCN4 leucine zipper: substantial contributions arise from intramolecular interactions enhanced on binding. *Protein Sci* 1999, 8:1381-1392.
The first systematic component analysis of FDPB protein–ligand binding free energies, applied to GCN4 dimerization. Electrostatic interactions are found to disfavor strong binding, because of the desolvation of charged groups upon binding, which is incompletely compensated by protein–protein interactions in the bound state.

25. • Roux B, MacKinnon R: The cavity and pore helices in the KcsA K^+ channel: electrostatic stabilization of monovalent cations. *Science* 1999, 285:100-102.
The activity and selectivity of a potassium ion channel are largely accounted for by a simple FDPB continuum model. Water in the central cavity preferentially stabilizes a mono-ion and the field of the pore helices, enhanced by the low dielectric membrane environment, selects for a cation.

26. Qiu D, Shenkin P, Hollinger F, Still W: The GBSA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. *J Phys Chem A* 1997, 101:3005-3014.

27. Hawkins G, Cramer C, Truhlar D: Pairwise descreening of solute charges from a dielectric medium. *Chem Phys Lett* 1995, 246:122-129.

28. Schaefer M, Karplus M: A comprehensive analytical treatment of continuum electrostatics. *J Phys Chem* 1996, 100:1578-1599.

29. • Majeux N, Scarsi M, Apostolakis J, Ehrhardt C, Caflisch A: Exhaustive docking of molecular fragments with electrostatic solvation. *Proteins* 1999, 37:88-105.
Methods and programs were developed for fast, large-scale protein–fragment docking. An electrostatic solvation model related to GB is tested for 2500 protein–ligand complexes.

30. Wagner F, Simonson T: Implicit solvent models: combining an analytical formulation of continuum electrostatics with simple models of the hydrophobic effect. *J Comp Chem* 1999, 20:322-335.

31. Onufriev A, Case D: Modification of the generalized Born model suitable for macromolecules. *J Phys Chem B* 1999, 104:3712-3720.

32. • Dominy B, Brooks C III: Development of a generalized Born model parameterization for proteins and nucleic acids. *J Phys Chem B* 1999, 103:3765-3773.
The authors describe the first MD simulation of a protein using GB to model the effect of solvent. Extensive testing and a slight reparameterization of the analytical solvation radii given by Qiu *et al.* [26] were done using a large database of FDPB self energies and conformational energies of small molecules, peptides, proteins and DNA. The parameters lead to an MD simulation of the protein GB1 (*sic*) that samples structures close to the crystal structure and in good agreement with an explicit solvent simulation.

33. • Williams D, Hall K: Unrestrained simulations of the UUCG tetraloop using an implicit solvation model. *Biophys J* 1999, 76:3192-3205.
The first MD simulations of an RNA using GB to model the effect of solvent. A series of long simulations with GB and several simpler models show that, for a small RNA hairpin, the structure with GB is close to the NMR structures and structures from an explicit solvent simulation, whereas a 1/r dielectric model gives poor results. Torsion angle fluctuations with GB are much larger than the spread in the NMR ensemble of structures.

34. • Tsui V, Case D: Molecular dynamics simulations of nucleic acids with a generalized Born model. *J Am Chem Soc* 2000, 122:2489-2498.
Extensive GB simulations of a DNA and an RNA duplex are described, as is an extensive comparison of GB energies with FDPB, leading to a parameterization of the GB model for use with the analytical solvation radii of Hawkins *et al.* [27] and the AMBER force field. Structures, energies and conformational changes (including A↔B transitions) in the GB simulations agree well with explicit solvent simulations.

35. • David L, Luo R, Gilson M: Comparison of generalized Born and Poisson models: energetics and dynamics of HIV protease. *J Comput Chem* 2000, 21:295-309.
The first MD simulation of a protein using FDPB to model the effect of solvent. This paper also describes the testing and reparameterization of GB with the solvation radii given in [26] and the CHARMM22 force field. FDPB yields stable simulations close to the crystal structure (~1.5 Å rms deviation). GB yields a slightly larger rms deviation (~2 Å).

36. Calimet N, Schaefer M, Simonson T: Protein molecular dynamics with the Generalized Born/ACE solvent model. *Proteins* 2001, in press.

37. Zou X, Sun Y, Kuntz I: Inclusion of solvation in ligand binding free energy calculations using the generalized Born model. *J Am Chem Soc* 1999, 121:8033-8043.

38. Dominy B, Brooks C III: Methodology for protein–ligand binding studies: application to a model for drug resistance, the HIV/FIV protease system. *Proteins* 1999, 36:318-331.

39. Majeux N, Scarsi M, Caflisch A: Efficient electrostatic solvation model for protein–fragment docking. *Proteins* 2001, 42:256-268.

40. Arora N, Bashford D: Solvation energy density occlusion approximation for evaluation of desolvation penalties in biomolecular interactions. *Proteins* 2001, 43:12-27.

41. Kangas E, Tidor B: Electrostatic specificity in ligand design. *J Chem Phys* 2000, 112:9120-9131.

42. You T, Bashford D: Conformation and hydrogen ion titration of proteins: a continuum electrostatic model with conformational flexibility. *Biophys J* 1995, 69:1721-1733.

43. • Alexov EG, Gunner MR: Calculated protein and proton motions coupled to electron transfer: electron transfer from QA– to QB in bacterial photosynthetic reaction centers. *Biochemistry* 1999, 38:8253-8270.
Electron transfer in a photosynthetic reaction center is studied with a multiple-conformer FDPB model, in which sidechains close to the chromophores and approximately 100 crystal waters are allowed several possible conformations. Explicit conformational transitions are observed in response to proton binding and electron transfer. These provide a microscopic model of dielectric relaxation in the immediate vicinity of the chromophores. Experimental pH-dependent redox properties are well reproduced.

44. Alexov EG, Gunner MR: Modeling the effects of mutations on the free energy of the first electron transfer from QA– to QB in photosynthetic reaction centers. *Biochemistry* 2000, 39:5940-5952.

45. Spassov V, Bashford D: Multiple-site ligand binding to flexible macromolecules: separation of global and local conformational change and an iterative mobile clustering approach. *Protein Sci* 1999, 20:1091-1111.

46. •• Havranek J, Harbury P: Tanford-Kirkwood electrostatics for protein modeling. *Proc Natl Acad Sci USA* 1999, 96:11145-11150.
A novel reincarnation of the image charge method for proteins. This general method is one of the main alternatives to GB-like models and would appear better suited to the globular geometry of soluble proteins. Here, numerical advances and some approximations borrowed from GB lead to high efficiency. Along with explicit treatment of multiple sidechain conformations, the model gives accurate results for BPTI titration.

47. Tanford C, Kirkwood J: Theory of protein titration curves. General equations for impenetrable spheres. *J Am Chem Soc* 1957, 79:5333-5339.

48. Schaefer M, Bartels C, Karplus M: Solution conformations and thermodynamics of structured peptides: molecular dynamics simulation with an implicit solvation model. *J Mol Biol* 1998, 284:835-847.

49. Luo R, David L, Hung H, Devaney J, Gilson M: Strength of solvent-exposed salt bridges. *J Phys Chem B* 2000, 103:727-736.

50. Landau L, Lifschitz E: *Electrodynamics of Continuous Media.* New York: Pergamon Press; 1980.

51. Gilson M, Davis M, Luty B, McCammon J: Computation of electrostatic forces on solvated molecules using the Poisson–Boltzmann equation. *J Phys Chem* 1993, 97:3591-3600.

52. Marchi M, Ballone P, Levy N, Borgis D: A dielectric continuum molecular dynamics method. *J Chem Phys* 2001, in press.

53. York D, Karplus M: A smooth solvation potential based on the conductor-like screening model. *J Phys Chem A* 1999, 103:11060-11079.

54. David L, Field M: Basis set approach to solution of Poisson equation for small molecules immersed in solvent. *J Comput Chem* 1997, 18:343-350.

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

56. •• Baker N, Sept D, Holst M, McCammon J: The adaptive multilevel finite element solution of the Poisson-Boltzmann equation on massively parallel computers. *J Comput Chem* 2001, in press.
FDPB revisited: modern multigrid technology is brought to bear on an old problem. The PB equation is recast in a 'weak' integral form and discretized with the help of simple basis functions on an adaptive, multilevel grid. A 600,000 particle microtubule is modeled on a parallel computer.

57. •• Im W, Berneche S, Roux B: Generalized solvent boundary potential for computer simulations. *J Chem Phys* 2001, 114:2924-2937.
A new, elegant, analytical method to compute solvation forces with a continuum model. The solute charge distribution is approximated by a basis set expansion (analogous to a multipole expansion). Tests on a soluble protein and on a membrane protein show good accuracy and efficiency, indicating that the method will be useful for MD simulations with continuum solvent.

58. Simonson T: Electrostatic free energy calculations for macromolecules: a hybrid molecular dynamics continuum electrostatics approach. *J Phys Chem B* 2000, 104:6509-6513.

59. Holst M, Baker N, Wang F: Adaptive multilevel finite element solution of the Poisson–Boltzmann equation I: algorithms and examples. *J Comput Chem* 2000, 21:1319-1342.

60. Simonson T: Dielectric relaxation in proteins: macroscopic and microscopic models. *Int J Quantum Chem* 1999, 73:45-57.

61. Boresch S, Ringhofer S, Hochtl P, Steinhauser O: Towards a better understanding of biomolecular solvation. *Biophys Chem* 1999, 78:43-68.

62. Sharp K: Calculation of electron transfer reorganization energies using the finite-difference Poisson-Boltzmann model. *Biophys J* 1998, 73:1241-1250.

63. Simonson T, Perahia D: Microscopic dielectric properties of cytochrome c from molecular dynamics simulations in aqueous solution. *J Am Chem Soc* 1995, 117:7987-8000.

64. •• Simonson T, Archontis G, Karplus M: A Poisson-Boltzmann study of charge insertion in an enzyme active site: the effect of dielectric relaxation. *J Phys Chem B* 1999, 103:6142-6156.
A critical comparison between macroscopic and microscopic treatments of dielectric relaxation in an enzyme active site. A continuum model with a protein dielectric of 5–6 reproduces the MD relaxation free energy in response to charge insertion, indicating that the active site is only moderately polarizable. A different parameterization, with a protein dielectric of one, is required to reproduce the static electrostatic potential in the active site. This stems from the use of a molecular mechanics charge set in the FDPB calculations. It shows that, in current applications, the charge set may not always be consistent with the dielectric constants used.

65. Marcus R: Electron transfer reactions in chemistry: theory and experiment. In *Protein Electron Transfer.* Edited by Bendall D. Oxford: BIOS Scientific Publishers; 1996:249-272.

66. Sham Y, Chu Z, Warshel A: Consistent calculations of pKas of ionizable residues in proteins: semi-microscopic and microscopic approaches. *J Phys Chem B* 1997, 101:4458-4472.

67. Mertz E, Krishtalik L: Free energy of charge transfer and intraprotein electric field: method of calculation depends on the charge state of protein at a given structure. *Bioelectrochem Bioenerg* 1999, 48:397-405.

68. Mertz E, Krishtalik L: Low dielectric response in enzyme active site. *Proc Natl Acad Sci USA* 2000, 97:2081-2086.

69. Higo J, Kono H, Nakamura H, Sarai A: Solvent density and long-range dipole field around a DNA-binding protein studied by molecular dynamics. *Proteins* 2000, 40:193-206.

70. Gunner M, Saleh M, Cross E, ud-Doula A, Wise M: Backbone dipoles generate positive potentials in all proteins: origins and implications of the effect. *Biophys J* 2000, 78:1126-1144.

71. • Pitera J, Falta M, van Gunsteren W: Dielectric properties of proteins from simulations: the effects of solvent, ligands, pH, and temperature. *Biophys J* 2001, in press.
The so-called Frohlich–Kirkwood dielectric constant of four proteins was calculated from the dipole fluctuations observed in MD simulations with explicit solvent and accurate long-range electrostatics. Consistent with earlier studies, the proteins have large (15–40) average dielectric constants because of motions of their charged sidechains. When these are omitted from the analysis, a low dielectric of 2–4 is obtained. This illustrates the spatial heterogeneity of protein dielectric properties (although the low value is probably representative of the properties of the bulk of the molecule). The Frohlich–Kirkwood dielectric constant measures the average polarizability of the protein and should be used accordingly.

72. Hoefinger S, Simonson T: Dielectric relaxation in proteins: a continuum electrostatics model incorporating dielectric heterogeneity of the protein and time-dependent charges. *J Comput Chem* 2001, 22:290-305.

73. Laberge M, Kohler M, Vanderkooi J, Friedrich J: Sampling field heterogeneity at the heme of c-type cytochromes by spectral hole burning spectroscopy and electrostatic calculations. *Biophys J* 1999, 77:3293-3304.

74. Manas E, Wright W, Sharp K, Friedrich J, Vanderkooi J: The influence of protein environment on the low temperature electronic

spectroscopy of Zn-substituted cytochrome c. *J Phys Chem B* 2000, 104:6932-6941.

75. Song X, Chandler D: Dielectric solvation dynamics of molecules of arbitrary shape and charge distribution. *J Chem Phys* 1998, 108:2594-2600.

76. Jordanides X, Lang M, Song X, Fleming G: Solvation dynamics in protein environments studied by photon echo spectroscopy. *J Phys Chem B* 1999, 103:7995-8005.

77. Baptista A, Martel P, Soares C: Simulation of electron-proton coupling with a Monte-Carlo method: application to cytochrome c3 using continuum electrostatics. *Biophys J* 1999, 76:2978-2998.

78. Rabenstein B, Ullmann M, Knapp E: Electron transfer between the quinones in the photosynthetic reaction center and its coupling to conformational changes. *Biochemistry* 2000, 39:10487-10496.

79. Ullmann M: The coupling of protonation and reduction in proteins with multiple redox centers: theory, computational method and application to cytochrome c3. *J Phys Chem B* 2000, 104:6293-6301.

80. • Demchuk E, Genick U, Woo T, Getzoff E, Bashford D: Protonation states and pH titration in the photocycle of photoactive yellow protein. *Biochemistry* 2000, 39:1100-1113.

Proton binding in photoactive yellow protein (PYP) provides a severe test of continuum models. PYP has several unusual protonation states and large electrostatic interactions. This situation is well removed from the preferred domain of the null model (surface groups with tiny pK_a shifts). This careful study analyzes model dependency on several factors, including the protein dielectric constant. Although the dielectric constant can be adjusted to reproduce most of the experimental data using physically reasonable values (~2), the sensitivity of results to the exact dielectric is found to be enormous, with several relevant pK_as shifting by 10–30 units when the dielectric is changed from two to six. Thus calibrated, the model supports the view of the protein interior as a low dielectric medium (as opposed to models that use a high protein dielectric for pK_as) and provides structural insights.

81. Lamotte-Brasseur J, Lounnas V, Raquet X, Wade R: pKa calculations for class A beta-lactamases: influence of substrate binding. *Protein Sci* 1999, 8:404-409.

82. Lamotte-Brasseur J, Dubus A, Wade R: pKa calculations for class C beta-lactamases: the role of Tyr-150. *Proteins* 2000, 40:23-28.

83. Elcock A: Realistic modeling of the denatured states of proteins allow accurate calculations of the pH dependence of protein stability. *J Mol Biol* 1999, 294:1051-1062.

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

85. • Warwicker J: Simplified methods for pKa and acid pH-dependent stability estimation in proteins: removing dielectric and counterion boundaries. *Protein Sci* 1999, 8:418-425.

This study of proton binding in four proteins shows that a model with a protein dielectric of 80 (i.e. no protein–solvent dielectric boundary) performs as well on average as models that use a protein dielectric of 20 (or less) for the protein. Details about the location of the titrating sites considered are not given, but the success of the method suggests that they are largely solvent accessible. The null model was not tried.

86. • Blomberg N, Gabdoulline R, Nilges M, Wade R: Classification of protein sequences by homology modeling and quantitative analysis of electrostatic similarity. *Proteins* 1999, 37:379-387.

Electrostatic potentials computed with continuum models are used for comparison of protein structures within the Pleckstrin homology domain family. Some variants of the method are fast enough to be used for large-scale structure comparison, illustrating the possible use of electrostatic models in a genomics context.

87. Lee LP, Tidor B: Barstar is electrostatically optimized for tight binding to barnase. *Nat Struct Biol* 2001, 8:73-76.

88. • Lee MR, Baker D, Kollman P: 2.1 and 1.8 Å average Calpha RMSD structure predictions on two small proteins, HP-36 and S15. *J Am Chem Soc* 2001, in press.

An energy function combining molecular mechanics and dielectric continuum solvent provides an accurate ranking of model structures generated with the *ab initio* Rosetta approach.