Chapter 6
Continuum Electrostatic Analysis of Proteins

G. Matthias Ullmann and Elisa Bombarda

6.1 Introduction

Electrostatic interactions play an important role in many biochemical systems especially because of their long range. Electrostatic interactions often guide the association of binding partners, but they also tune catalytic properties of the active site of enzymes. For instance, the protonation and redox behavior of residues and prostetic groups in protein is heavily influenced by the electrostatics of the surrounding. Moreover, regulation of biochemical processes is often mediated by electrostatic modifications of proteins such as for instance by phosphorylation of serine, threonine or tyrosine residues or by acetylations of lysine residues. However, the electrostatics of a protein is not only influenced by the distribution of charged and polar aminoacid residues in the protein but also by the surrounding solvent and the ions that are dissolved therein. The solvent may screen charge-charge interactions and stabilizes the structure of protein by solvating charged aminoacids. Another example of such effect are membrane potentials, that can influence the conformation of proteins.

There are different ways to describe solvent effects theoretically [1]. In molecular dynamics simulations, the solvent is described explicitly in form of individual solvent molecules, normally water molecules [2]. In such simulations, ions are

G.M. Ullmann (✉)
Structural Biology/Bioinformatics, University of Bayreuth, Universitätsstr. 30, NW I, 95447 Bayreuth, Germany
e-mail: ullmann@uni-bayreuth.de

E. Bombarda
Experimental Physics IV, University of Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany

G. Náray-Szabó (ed.), *Protein Modelling*, DOI 10.1007/978-3-319-09976-7_6

represented explicitly as well. These simulations require a considerable amount of computing time mainly to simulate the water and the ions solvating the protein. For this reason, molecular mechanics cannot be applied when information about longer time scales (more than micro-seconds) or many different states of a molecule are required. In such cases, continuum electrostatics, which relies on Maxwell's equations, is the approach of choice [3]. In continuum electrostatics, the protein and the surrounding solvent are described as dielectric continua. Since electrostatic interactions play a major role in biomolecular systems, continuum electrostatics has a broad range of applications in biomolecular modeling. Moreover in combination with a master equation approach, continuum electrostatics can even be used to describe the reaction kinetics of complex systems.

In this article, we describe some methods that are based on continuum electrostatic calculations. In the beginning, we will introduce the electrostatic model based on the Poisson-Boltzmann equation. In this part, we try to make the underlying theory understandable in order to give a feeling for its strength but also for its limitations. Moreover, we explain how the linearized Poisson-Boltzmann equation can be solved numerically. Afterwards, we discuss how the continuum electrostatic model was used to analyze the association of biomolecules and the thermodynamics of biochemical reactions. The major purpose of this review is to give a general overview of methods that rely on continuum electrostatics and discuss their physical basis.

6.2 The Continuum Electrostatic Model Based on the Poisson-Boltzmann Equation

6.2.1 The Physical Basis of the Poisson-Boltzmann Equation

The conceptual idea of modeling proteins using continuum electrostatics is relatively simple (Fig. 6.1). The protein is assumed to have a fixed structure defining a region of low polarizability which is embedded in a region with high polarizability representing the solvent. The polarizability is related to the relativedielectric constant of the medium, also called relative permittivity, the higher the dielectric constant the higher the polarizability. This model is mathematically represented by the Poisson equation with a spatially varying dielectric constant or actually better said dielectric permittivity (Eq. 6.1).

$$\nabla[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -4\pi\rho(\mathbf{r}) \tag{6.1}$$

where $\varepsilon(\mathbf{r})$ is the permittivity of the medium which varies spatially (inside and outside of the protein), ∇ is the differential operator, $\phi(\mathbf{r})$ is the electrostatic potential, $\rho(\mathbf{r})$ is the charge distribution within the protein and the solvent. The low dielectric region of the protein is delimited by assigning atomic radii to each atom and determining the solvent excluded volume by rolling a sphere over the protein [4]. Physically, the polarizability of a medium depends mainly on the mobility and the polarity of the molecules or molecular groups of the medium. Thus, solvents

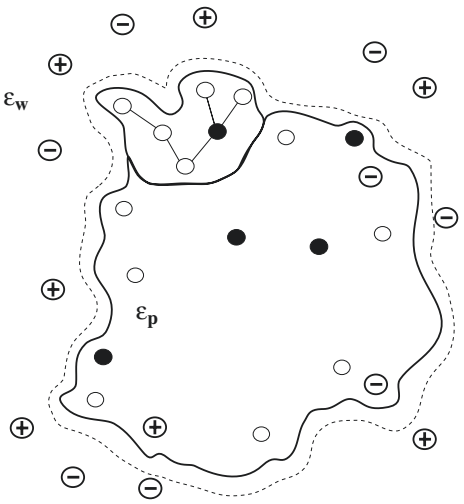

Fig. 6.1 Conceptual model of the continuum electrostatic approach. The protein is modeled as a dielectric continuum of low permittivity ε_p with fixed point charges. The protein is embedded in an environment with a high permittivity ε_w representing the solvent. In the continuum with a high permittivity, a charge density represents the ions dissolved in the aqueous solution. The *dotted line* marks the so-called Stern layer or ion-exclusion layer. Mobile ions are not allowed inside the protein volume and the Stern layer

with freely movable molecules and large dipoles have a high dielectric constant. A charged solute in a solvent induces a so-called reaction field (see Fig. 6.2). This reaction field is caused by the orientation of solvent molecules towards the charge, the so-called orientational polarization, and by the electronic polarization of the solvent molecules, i.e. by the deformation of the electron clouds. Apolar solvents such as octane have a relative dielectric constant of about 2, since the shielding due to orientational polarization is negligible while polar solvents such as water possess a large molecular dipole and have consequently a high dielectric constant.

There are two types of charges in this model, spatially fixed charges representing the charge distribution within the protein and mobile charges representing the ions in the solvent. The ions dissolved in the solvent are excluded from the volume of the protein. The charges of the protein are usually represented by point charges at the position of the nucleus of the atoms. These point charges allow to represent charged aminoacids in the protein such as for instance aspartate and glutamate residues but also dipoles like for instance in the protein backbone or in the side chains of uncharged aminoacids. Mathematically, the charge distribution is represented as

$$\rho(\mathbf{r}) = \rho_f(\mathbf{r}) + \rho_{ion}(\mathbf{r}) \tag{6.2}$$

where ρ_f represents the charge distribution due to the point charges in the solute, i.e. the protein, and ρ_{ion} represents the charge distribution of the ions dissolved

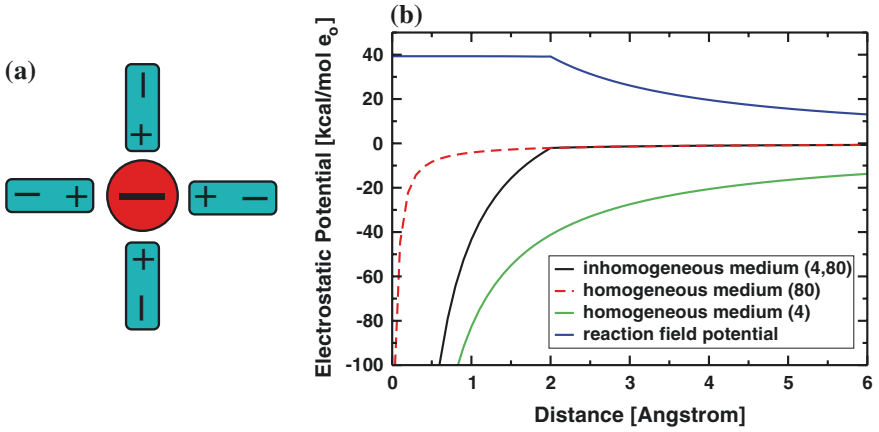

Fig. 6.2 Electrostatic potential of an anion in solution. **a** The solvent (*blue*) generates a solvation shell around the solvated anion (*red*). This solvation shell gives rise to the reaction field which counter-acts the electrostatic potential of the anion. **b** Electrostatic potentials of an anion with a radius of 2 Å in solution calculated with software APBS [5]. The *black line* shows the electrostatic potential for an inhomogeneous dielectric medium with a dielectric constant of 4 inside the ion and a dielectric constant of 80 in the solvent. The *green line* describes the electrostatic potential in a homogeneous medium with a dielectric constant of 4. The reaction field potential (*blue line*) is obtained as the difference between the *green line* and the *black line*. The *red dashed line* describes the electrostatic potential in a homogeneous medium with a dielectric constant of 80

in the solvent. These ions are represented by a charge density that adopts a Boltzmann distribution. This distribution can be approximated by

$$\rho_{ion}(\mathbf{r}) = \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_o \exp\left(\frac{-Z_i e_o \phi(\mathbf{r})}{RT}\right) \tag{6.3}$$

assuming that there is no correlation between the ions in the solution. In Eq. 6.3, Z_i is the charge number of the ion of type i, K is the number of the different ion types in the solution, e_o is the elementary charge, i.e. the charge of a proton, and c_i^{bulk} is the bulk concentration of the ion, i.e., the concentration where the protein electrostatic potential vanishes. Substituting Eqs. 6.2 and 6.3 in 6.1, the Poisson-Boltzmann equation for a medium with a spatially varying permittivity assumes the following form

$$\nabla[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -4\pi\left(\rho_f(\mathbf{r}) + \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_o \exp\left(\frac{-Z_i e_o \phi(\mathbf{r})}{RT}\right)\right) \tag{6.4}$$

This equation is a non-linear partial differential equation since the potential $\phi(\mathrm{r})$ occurs not only on the left side of the equation but also in a non-linear term, namely in the exponential, on the right side of the equation, which describes the ion distribution around the protein. Generally, non-linear partial differential

equations are difficult to solve even numerically. However, by approximating the exponential as

$$\sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_o \exp\left(\frac{-Z_i e_o \phi(\mathbf{r})}{RT}\right) \approx \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i e_o - \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i^2 e_o^2 \frac{\phi(\mathbf{r})}{RT} \quad (6.5)$$

and realizing that the first term on the right side is zero because of charge balance, one obtains the linearized Poisson-Boltzmann equation, i.e. the potential occurs on the right side of the equation only in a linear term.

$$\nabla[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -4\pi\left(\rho_{prot}(\mathbf{r}) - \sum_{i=1}^{K} c_i^{\text{bulk}} Z_i^2 e_o^2 \frac{\phi(\mathbf{r})}{RT}\right) \quad (6.6)$$

With the common definitions of the ionic strength $I = \frac{1}{2}\sum_{i=1}^{K} c_i^{\text{bulk}} Z_i^2$ and a modified inverse Debye length $\bar{\kappa} = \sqrt{\frac{8\pi N_A e_o^2 I}{k_B T}}$ the linearized Poisson-Boltzmann equation assumes the form that is found in some biophysics text books (Eq. 6.7).

$$\nabla[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -4\pi\rho_{prot}(\mathbf{r}) + \bar{\kappa}^2(\mathbf{r})\phi(\mathbf{r}) \quad (6.7)$$

As can be seen from Eq. 6.4, the Poisson-Boltzmann equation depends explicitly on temperature. However, this temperature dependence describes only the temperature dependence of the ion distribution. The temperature dependence of the dielectric constant is not explicitly included. Therefore, varying only the temperature in the Poisson-Boltzmann equation is not physically meaningful. Normally, room temperature is assumed in these kind of calculations and when another temperature is chosen, the dielectric constants should be adapted.

The linearity of Eq. 6.7 implies that the potentials of two charge distributions $\rho_1(r)$ and $\rho_2(r)$ are additive as long as the spatial distribution of the dielectric permittivity does not change, i.e., for the charge distribution $\rho(r) = \rho_1(r) + \rho_2(r)$ one can obtain the total potential as sum of the partial potentials $\phi(r) = \phi_1(r) + \phi_2(r)$ as long as the spatial distribution of the dielectric permittivity ε(r) stays the same. This property has important consequences for the various applications. For instance the calculation of relative binding constants, which is a typical application of continuum electrostatic calculations, relies on this property. Namely, the calculation of pH titration curves is only possible because of this linearity as will be explained in a later section.

6.2.2 *Solving the Linearized Poisson-Boltzmann Equation Numerically*

For a few simple geometries analytical solutions of the linearized Poisson-Boltzmann equation exist [6, 7]. For irregular geometries, this equation can be solved by numerical methods. The most popular methods to solve Poisson-Boltzmann equation rely

on regular finite difference methods [3, 5, 8–10], but also adaptive-grid methods [11], multi-grid-level based methods [12–15], boundary element methods [16], or finite element methods [17] can be used.

The principle idea of finite difference methods is to replace the differential $\nabla f(x)$ by a quotient of finite differences $\frac{f(x+h)-f(x)}{h}$, where h is a discretization coarse graining. The approximation approaches the exact result, when h goes to zero. In finite differences methods, every linear differential equation becomes a system of linear equations, which can be solved by numeric algorithms. The space in which the potential should be determined is discretized and the potential is calculated for each volume element. In order to obtain a numerical approach to solve the linearized Poisson-Boltzmann equation, we rearrange it

$$\nabla[\varepsilon(\mathbf{r})\nabla\phi(\mathrm{r})] - \bar{\kappa}^2(\mathrm{r})\phi(\mathrm{r}) + 4\pi\rho_{prot}(\mathrm{r}) = 0 \tag{6.8}$$

and integrate it over the descritezed volume elements

$$\int \nabla[\varepsilon(\mathrm{r})\nabla\phi(\mathrm{r})]\,\mathrm{dr} - \int \bar{\kappa}^2(\mathrm{r})\phi(\mathrm{r})\mathrm{dr} + 4\pi \int \rho_{prot}(\mathrm{r})\,\mathrm{dr} = 0 \tag{6.9}$$

One the basis of Gauss's theorem, the first integral can be transformed into a surface integral that can be approximated with a finite difference expression.

$$\int \nabla[\varepsilon(\mathrm{r})\nabla\phi(\mathrm{r})]\mathrm{dr} = \int [\varepsilon(\mathrm{r})\nabla\phi(\mathrm{r})]\mathrm{dA} \tag{6.10}$$

$$= \sum_{i=1}^{6} \frac{h^2\varepsilon_i(\phi_i - \phi_\mathrm{o})}{h} \tag{6.11}$$

$$= \sum_{i=1}^{6} h\varepsilon_i(\phi_i - \phi_\mathrm{o}) \tag{6.12}$$

The volume integrals on Eq. 6.9 can be written as

$$\sum_{i=1}^{6} h\varepsilon_i(\phi_i - \phi_\mathrm{o}) - h^3\bar{\kappa}_\mathrm{o}^2\phi_\mathrm{o} + 4\pi q_\mathrm{o} = 0$$

where $\bar{\kappa}_\mathrm{o}^2$ is the modified inverse Debye length (related to the ionic strength) that is associated with this grid point. Rearranging this equation gives

$$\phi_\mathrm{o} = \frac{\left(\sum_{i=1}^{6} h\varepsilon_i\phi_i\right) + 4\pi q_\mathrm{o}}{\left(\sum_{i=1}^{6} h\varepsilon_i\right) + h^3\bar{\kappa}_\mathrm{o}^2} \tag{6.13}$$

Equation 6.13 is the numeric solution of the linearized Poisson-Boltzmann equation. It says that the potential ϕ_o in the grid cell depends on the electrostatic

potential ϕ_i of the six surrounding grid cells, the dielectric constant ε_i between the present and neighboring grid cells, the charge q_o and the ionic strength parameter $\bar{\kappa}_o^2$ assigned to the grid cell. Such an equation exists for almost all grid cells in the lattice, except for those at the boundary of the box. For the grid points at the boundary of the box, a good initial value needs to be determined for instance from an analytical approximation or a from numerical solution that was obtained with a coarser grid resolution. The set of equations like Eq. 6.13 form a system of linear equations, which can be solved. A common way is to obtain the potential iteratively. First a value of the potential is assigned to each grid cell, for instance from an analytical approximation. Then the potential is iteratively calculated, i.e., the potential of the present iteration is calculated from the potential of the previous iteration. The iteration is continued until the potential is sufficiently accurate. In practice, the iteration is stopped when the difference between the electrostatic potentials that were determined in two subsequent iteration steps is sufficiently small.

To solve the Poisson-Boltzmann equation for molecular systems practically, a flowchart of the type represented in Fig. 6.3 is followed. Different implementations of Poisson-Boltzmann solvers may vary in details. Here, the description focuses on the standard finite difference methods.

As a first step, the parameters of the molecule are read in. In particular, the parameters are the coordinates, the radii and the partial charges of the atoms. Also the dielectric constant of the solvent and of the solute needs to be defined as well as the ionic strength and the probe sphere radii for defining the solvent accessible surface and the ion exclusion layer. The temperature, which influences only the ionic distribution in the solution (see the discussion above), can also be defined. Moreover, the parameters for the numeric solvers needs to be read such as the number of grid points, the position of the grid, and the grid spacing.

In the next step, the boundaries of the dielectric regions are calculated. On the basis of the coordinates and radii of atoms, molecular surfaces are calculated for each dielectric regions. Usually surface is defined by rolling balls over the atoms of the molecule, which are represented as spheres of a defined coordinates and radii [18] (Fig. 6.4). The rolling ball represents a solvent molecule. For water, a radius of 1.4 Å is generally assumed as radius of the molecule. The step of defining the boundaries between the high and low dielectric regions is crucial, since it defines at which positions the reaction field is formed. A too small solvent radius may lead to unrealistically small cavities inside the protein that are filled with high dielectric continuum. A too large solvent radius may cause that internal cavities and surface clefts of the protein, that are actually filled with water, are not filled with a high dielectric continuum and thus no reaction field can arise from these cavities.

To solve the Poisson-Boltzmann equation numerically, all physical properties of the system (charge, electrostatic potential, electrical permittivity and ion accessibility) have to be mapped onto a grid. An easy way to map charges to the grid is a linear interpolation scheme. In this approach, a charge q_p at a given position $\mathbf{r}_p$ is fractioned to the eight surrounding grid points at the positions $\mathbf{r}_a (a = 1 \cdots 8)$ as follows: $q_p = q_a \left(1 - \frac{r_{ax} - r_{px}}{h}\right)\left(1 - \frac{r_{ay} - r_{py}}{h}\right)\left(1 - \frac{r_{az} - r_{pz}}{h}\right)$, where r_{ax} is the x

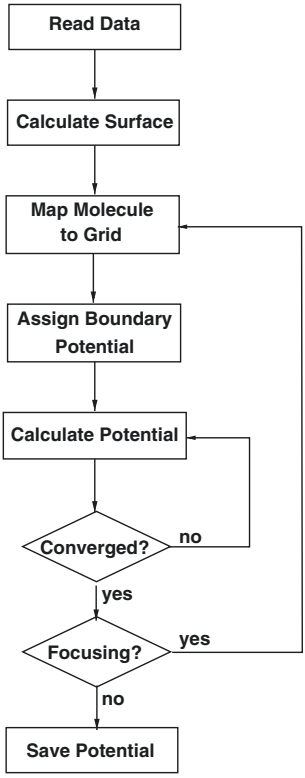

Fig. 6.3 Solving the Poisson-Boltzmann equation by a finite difference methods. First, all parameters for the calculation are read (coordinates, charges, radii, dielectric constants and grid definitions). Then the surface of the molecule is calculated as dielectric boundary. According to the boundaries, dielectric constants are mapped to the grid. The atomic charges are distributed over the surrounding grid points. Electrostatic potentials are assigned to the grid points based on an initial guess. It is important, that the initial potentials at the boundaries of the grid are very good approximations, because they remain constant during the calculation. The finite difference formulation of the Poisson-Boltzmann equation is solved by an iterative scheme, until the electrostatic potential does not change significantly anymore between two subsequent iterations. The computations require on one hand a large initial grid to minimize the error due to the approximated boundary potential and on the other hand a fine final grid to minimize the error due to the finite difference approximation. Both requirements can usually not be fulfilled directly due to a limited amount of memory. Therefore, consecutively smaller and finer grids are calculated using the previous grid to define the potentials at the boundaries. This method is called focusing. Finally, electrostatic energies are calculated as product of charge and electrostatic potential

-component of the vector a and h the grid spacing (mesh size). Analogously, the other neighboring grid points get the remaining fraction of the charge assigned according to their distance. The spatial dependent dielectric permittivity is defined on a grid, which is shifted by half a grid unit compared to the charge grid (ε_1 to ε_6 in Fig. 6.5). The surface of the molecule is used to assign the dielectric constant of the region, if the point is inside the surface. In a similar way, also the ionic

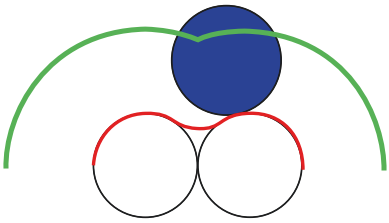

Fig. 6.4 Calculation of the solvent accessible surface. The solvent accessible surface (*red line*) of the two atoms (*white circle*) is calculated by a 'rolling ball' (*blue circle*). The surface of the Stern layer is shown by the *green line*

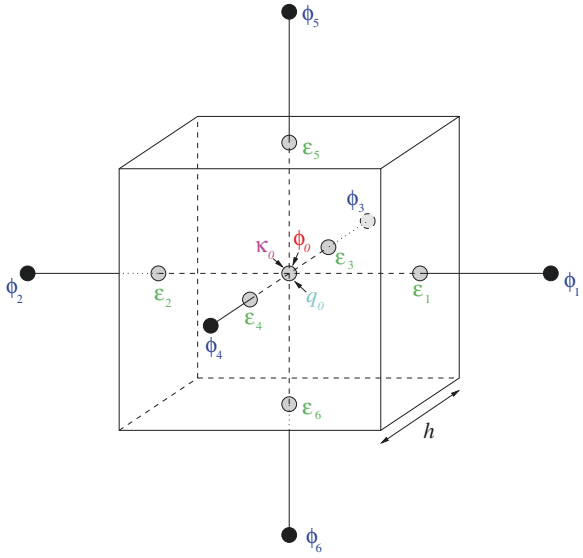

Fig. 6.5 Representation of one grid cell for solving the linearized Poisson-Boltzmann equation by a finite difference method. To the point in the center of the box, the electrostatic potential ϕ_o, the inverse Debye length κ, and the charge q_o are assigned. The *filled points* represent the centers of the neighboring grid cells to which the potentials ϕ_i are assigned. The dielectric constants ε_i are assigned to the lines connecting two neighboring grid points

accessibility is defined on a grid with the only difference that a larger probe sphere radius (usually of 2 Å) is used to define the surface.

In order to actually calculate the potential numerically, the electrostatic potential grid needs to be initialized with an initial guess. A reasonable starting point is the Debye-Hückel expression. It is important that the initial guess of the potential at the outer boundaries of the grid is given accurately enough, since these potential values will not change during the calculation. For this reason, it is important that the distance between the protein and the outer boundary of the grid is large enough (at least 10 Å). In order to obtain a sufficient precision and numerical stability of the calculated potentials, the grids used in the calculation should have a resolution of at least 0.25 Å. Even for relatively small proteins, the number of grid points needed at this resolution to cover the protein and an adequate part of the solvent would require a huge amount of memory. Therefore, the grid is refined in several steps. One starts with a grid that is large enough to hold the whole protein and has a distance of at

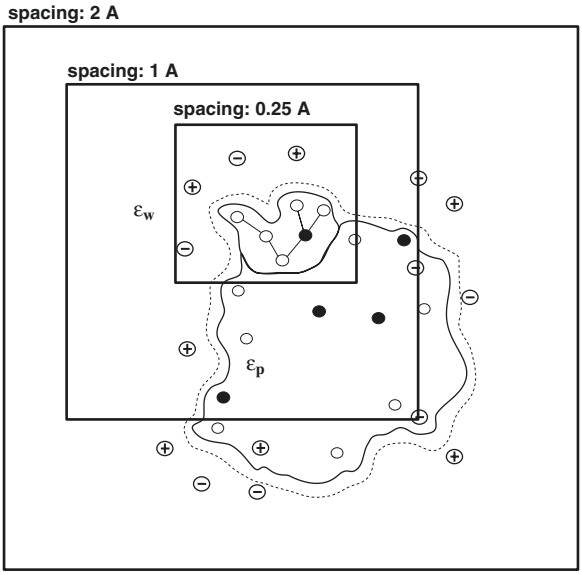

Fig. 6.6 Focusing of the grid for the calculation of the electrostatic potential using the Poisson-Boltzmann Equation. In order to calculate the electrostatic potential numerically, the protein needs to be mapped on a grid. To get a reasonable solution, the grid needs to be large enough to initialize the outer boundary of the grid with a reasonable analytic approximation. However, to get the good electrostatic potential, the grid needs to have a fine resolution. This fine resolution is obtained by first solving the Poisson-Boltzmann equation on a coarse grid. This solution is then used to initialize a smaller grid with a better resolution and so on until the desired resolution is reached. The outer grid is usually centered on the geometric center of the protein. Instead, the finest grid is usually centered on the geometric center of the group at which the exact electrostatic potential is desired

least 10 Å between the protein and the outer boundary of the grid. The center of the grid is usually chosen as the geometric center of the protein. This grid will be rather coarse (for instance 2.0 Å). Once the potential of this grid is converged, it can be used to initialize a finer grid (for instance 1.0 Å), which is embedded in the coarser grid. This procedure, called focusing [19], is repeated until a sufficiently fine grid can be used (Fig. 6.6). The finer grids are centered on the center of interest, for instance the set of atoms that form the site. Once the calculation is converged, the potential is stored and can be used for subsequent visual analysis using molecular visualization software or for calculations of electrostatic energies.

6.2.3 Electrostatic Potentials and Electrostatic Energies

The solution of the Poisson-Boltzmann equation is the electrostatic potential $\phi(\mathbf{r})$ which can be expressed as a potential that is composed of two parts.

$$\phi(\mathbf{r}) = \sum_{i=1}^{M} \frac{q_i}{4\pi\varepsilon_p|\mathbf{r} - \mathbf{r}_i'|} + \phi_{\mathrm{rf}}(\mathbf{r}) \tag{6.14}$$

The first term describes the Coulomb electrostatic potential at the position $\mathbf{r}$ caused by M point charges q_i at positions $\mathbf{r}_i'$ in a medium with a permittivity ε_p, the term $\phi_{\mathrm{rf}}(\mathbf{r})$ describes the reaction fieldpotential originating from the charge distribution and the dielectric boundary between the protein and the solvent as well as from the distribution of ions in the solution. The reaction field is always oriented opposite to the field of the solute and therefore shields the field of the solute. This reaction field is of great importance for understanding structural and functional properties of proteins. For instance, in aqueous solution the dipole of a peptide α-helixis counteracted by the reaction field, which drastically reduces the strength of the helix dipole compared to its value in vacuum [20]. Moreover, reaction field effects can explain the orientation of helices in membrane proteins [21].

The electrostatic potential that is obtained by solving the Poisson-Boltzmann equation has already a great value by itself. Visualization of this potential can give first insights into the interaction between molecules as shown for instance in Fig. 6.7 where different representations of the electrostatic potential of Cytochrome c Peroxidase are shown. By convention, red shows negative electrostatic potentials and blue shows positive electrostatic potentials.

Probably more important than visualizing the electrostatic potential is using it to calculate electrostatic energies. Such calculations can give quantitative insights into biochemical mechanisms. Two different kind of electrostatic energies can be distinguished: interaction energies and reaction field energies.

The interaction energy G_{inter} energy is obtained by charging charge set ρ_2 in the presence of the electrostatic potential caused by charge ρ_1. Assuming that the charge set ρ_2 consists of a single charge q_f, the interaction energy becomes

$$G_{\mathrm{inter}} = \int_0^{qf} \phi(\rho_1, \mathbf{r}_q) dq = \phi(\rho_1, \mathbf{r}_q) q_f \tag{6.15}$$

Since the potential $\phi(\rho_1, \mathbf{r}_q)$ at the position $\mathbf{r}_q$ of the charge q_f is totally independent of the charge q_f itself, the integration in Eq. 6.15 reduces to a simple multiplication. Equation 6.15 can be generalized to the interaction between two disjunct sets of charges $\{q\}$ and $\{p\}$, which is given by

$$G_{\mathrm{inter}} = \sum_{i=1}^{N_q} q_i \phi\big(\{p\}, \mathbf{r}_{q_i}\big) = \sum_{i=1}^{N_p} p_i \phi\big(\{q\}, \mathbf{r}_{p_i}\big) \tag{6.16}$$

where N_q and N_p are the number of charges in the charge sets $\{q\}$ and $\{p\}$, respectively, $\phi\big(\{p\}, \mathbf{r}_{q_i}\big)$ is the potential caused by the charge set $\{p\}$ at the position of the charge q_i and $\phi\big(\{p\}, \mathbf{r}_{q_i}\big)$ is the potential caused by the charge set $\{q\}$ at the position of the charge p_i. As can be seen from Eq. 6.16, this interaction energy is symmetric.

The reaction field energy in continuum electrostatics, which is also sometimes called self-energy, is the interaction energy of the charge set $\{q\}$ with its own reaction field potential ϕ_{rf}. To obtain this energy, one imagines the charging of a

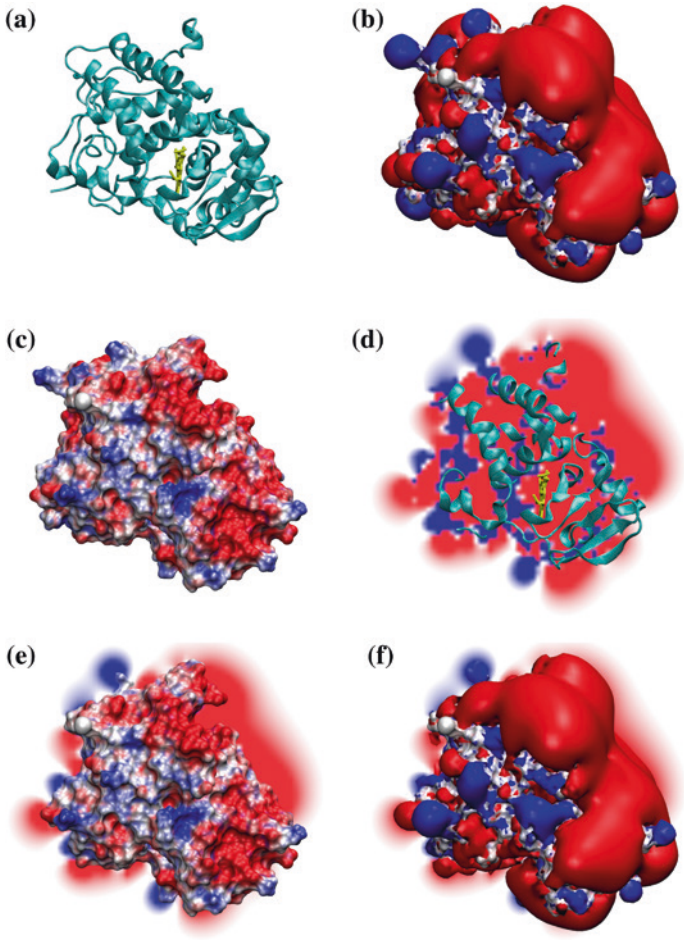

Fig. 6.7 Different visualizations of the electrostatic potential of cytochrome *c* peroxidase (CcP). The protein is shown in the same orientation in all pictures. **a** Cartoon representation of CcP showing the orientation of the protein. The heme is shown in a stick representation. **b** Isosurfaces of the electrostatic potential. *Blue* represent positive potentials, *red* negative potentials. The *red* and *blue surfaces* show where the potential has the value of $-1\ k_BT/e_\circ$ and $1\ k_BT/e_\circ$, respectively. **c** Electrostatic potential mapped to the molecular surface of the protein. The potential on the surface shows values between $-3\ k_BT/e_\circ$ (*red*) and $3\ k_BT/e_\circ$ (*blue*). **d** Slice thought the electrostatic potential. The potential on the slice is scaled between $-1\ k_BT/e_\circ$ (*red*) and $1\ k_BT/e_\circ$ (*blue*). **e** A combination of the representation shown in **c** and **d** gives an impression how the electrostatic potential fill the space but also allows to see more details on the molecular surface. **f** A combination of the representation shown in **c** and **b** gives a better impression how the electrostatic potential fill the space but lacks details on the molecular surface. The potentials were calculated using APBS [5] and visualized with VMD [22]

particle in a dielectric medium, and asks what is the energy of this charging process. Analogous to Eq. 6.15, we can write

$$G_{\mathrm{rf}} = \int_0^{q_f} \phi_{rf}(q, \mathbf{r}_q)\mathrm{d}q \tag{6.17}$$

in contrast to before, the reaction field potential ϕ_{rf} depends on the charge of the particle. For simplicity, one assumes a linear response, i.e. $\phi_{rf} = Cq_f$. Thus, from Eq. 6.17, we obtain

$$G_{\mathrm{rf}} = \int_0^{q_f} Cq\mathrm{d}q = \frac{1}{2}Cq_f^2 = \frac{1}{2}\phi_{rf}(q_f, \mathbf{r}_q)q_f \tag{6.18}$$

The last term $\frac{1}{2}\,\phi_{rf}\left(\mathrm{q}_f, \mathbf{r}_q\right)q_f$ is obtained by using $\phi_{rf} = Cq_f$. Equation 6.18 can be generalized to obtain the reaction field energy of a charge set $\{q\}$

$$G_{\mathrm{rf}} = \frac{1}{2}\sum_{i=1}^{N_q} q_i\phi_{\mathrm{rf}}\left(\{q\}, \mathbf{r}_{q_i}\right) \tag{6.19}$$

As shown above, the factor $\frac{1}{2}$ in this equation is a consequence of the linear response ansatz.

Although simple at a first sight, the continuum electrostatic model is surprisingly successful in describing the properties and processes that are connected to the electrostatics of biomolecules. Such properties are for instance the association of proteins or redox and pH titration behavior. Its success relies probably on the fact that solvent degrees of freedom, which are normally difficult to sample in molecular dynamics simulations, are averaged from the beginning by assuming that the medium can be described by a dielectric continuum. To use the full power of continuum electrostatics, it needs to be combined with other techniques from statistical thermodynamics such as free energy calculations and Monte Carlo techniques as will be detailed below.

6.3 Electrostatic Association of Proteins

Since electrostatic force is long range, it plays a particular important role in the interaction of proteins. Especially in the case of electron transfer proteins, electrostatics plays a major role in the association process. In order to ensure a fast

turnover, electron transfer proteins often associate transiently. This feature implies that electron transfer complexes are often relatively loose and dynamic [23]. In these complexes, electrostatics helps to get the right balance between specificity and flexibility. The structural flexibility has been observed for electron transfer complexes and studied extensively especially for the complex of cytochrome *c* peroxidase and cytochrome *c* or the complex of plastocyanin and cytochrome *f* [24–28].

6.3.1 Electrostatic Docking of Proteins

The docking of proteins can be simulated by Metropolis Monte Carlo [24]. For this purpose, the electrostatic potential of one protein, usually the larger protein, is mapped to a three dimensional grid surrounding the protein. The second molecule is placed in a random orientation at a certain distance from molecule one, i.e., randomly on the surface of a sphere which surrounds molecule one. The sphere should be large enough that the electrostatic potential of protein one on the surface of this sphere is zero or at least equipotential, i.e., everywhere the same on the surface. Then the second protein is randomly rotated and translated in the electrostatic potential of the first protein. The new configurations are accepted according to the Metropolis criterion [29]. That means that configurations with energies that are lower than that of the previous configuration are always accepted and configurations with a higher energy are accepted with a probability that is proportional to $\exp(-\Delta E/RT)$, where ΔE is the interaction energy difference between the new and the old configuration, R is the gas constant and T is the absolute temperature. The interaction energy is calculated by multiplying the charge distribution of the second molecule with the potential of the first molecule according to Eq. 6.16. If the second molecule moves too far away from the first molecule, the simulation is restarted on the sphere mentioned above. The Metropolis Monte Carlo procedure generates a Boltzmann ensemble which describes the equilibrium distribution of molecule two around molecule one. Such an ensemble is experimentally accessible by paramagnetic relaxation enhancement (PRE) NMR spectroscopy using spin labels attached at defined positions of the protein [23].

We recently applied this Monte Carlo docking procedure to the complex of Cytochrome *c* and Cytochrome *c* Peroxidase [30, 31]. As can be seen in Fig. 6.8a, b, the electrostatics of the docking surface of the two proteins is complementary. While Cytochrome *c* Peroxidase shows a clear negative potential (Fig. 6.8a), Cytochrome *c* shows a negative potential (Fig. 6.8b). Figure 6.8c, d shows the density of the center of mass of Cytochrome *c* generated in the Monte Carlo simulation. The large extension of this density shows that the complex of Cytochrome *c* and Cytochrome *c* Peroxidase allows a larger flexibility. This flexibility is often observed in the case of electron transfer proteins, since it allows a fast turnover because of a larger dissociation rate [32].

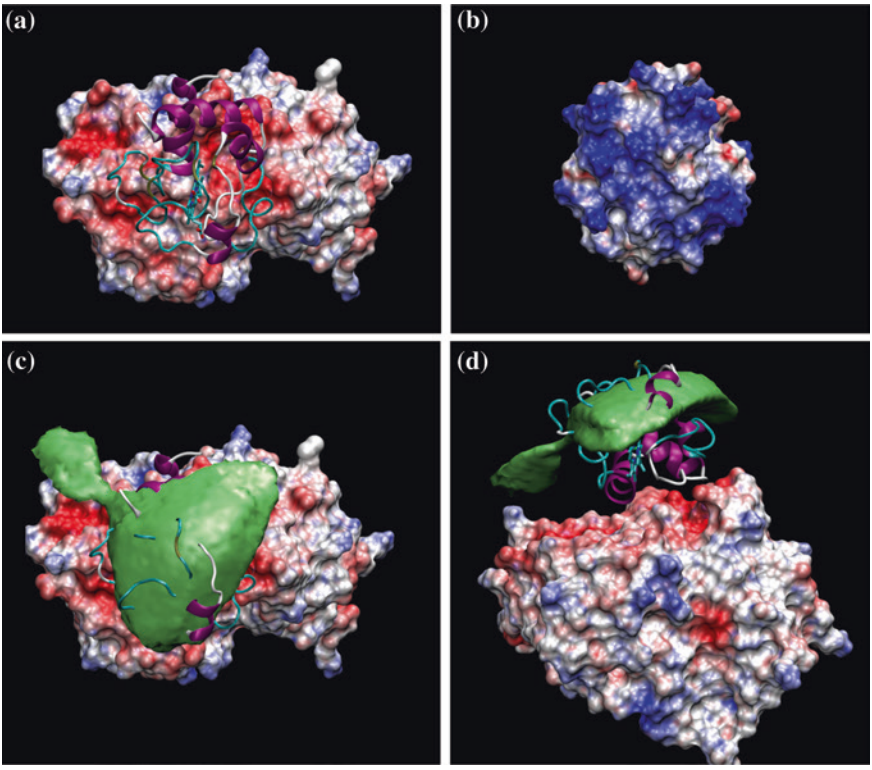

Fig. 6.8 Complex of cytochrome *c* (Cc) and Cytochrome *c* Peroxidase (CcP). The interaction of these two proteins is governed by electrostatic interactions. **a** Complex of Cc and CcP. Cc is shown as a cartoon. CcP is shown in a representation in which the electrostatic potential is mapped to its surface. **b** Cc with its electrostatic potential mapped to its surface. Cc is rotated by 180° around the y-axis (axis from *left* to *right* within the plain) compared to **a**. It can be seen that the electrostatic potential of Cc is positive and thus complementary to the electrostatic potential of CcP. **c** Simulated docking of Cc and CcP. The green density shows the region in which Cc can be found with a high probability. The density represents the frequency with which the center of mass of Cc is found in this certain volume area. The extension of the *green volume* indicates that the complex is not well defined but shows a high flexibility. This behavior was also found for other electron transfer complexes and is in agreement with experimental findings. **d** Simulated docking of Cc and CcP as in **c** but rotated by 90° around the x-axis

Brownian dynamics simulations are in some respect similar to the Monte Carlo approach [33–37]. In Brownian dynamics simulations, the association of two proteins is simulated using Newton's equations of motion combined with additional random and friction terms and the interaction is calculated based on electrostatic potentials obtained from the Poisson-Boltzmann equation. Brownian dynamics simulations enable to determine relative association rate constants and thus

to study for instance the influence of mutations or of ionic strength on association rates. These simulations have been also applied to a variety of electron transfer complexes and allowed to interpret experimental findings on the association between electron transfer partners [38–41].

6.3.2 Similarity of Electrostatic Potentials of Proteins

Proteins with similar function show often similar structures. A special case, which at the first sight seems to be an exception to this rule, can be found for the electron transfer between cytochrome b_6f and photosystem I in photosynthesis. This transfer is usually mediated by the blue copper protein plastocyanin. Under copper deficiency however, plastocyanin is replaced by the heme protein cytochrome c_6. Although plastocyanin and cytochrome c_6 differ considerably in composition and structure, they perform the same function in the photosynthetic electron-transport chain.

The functional equivalence of the two proteins can be understood on the basis of similarity of their electrostatic potentials [42, 43]. The similarity of the electrostatic potentials is defined on the basis of the integral-based Hodgkin index H_{ab}^{elec} [44].

$$H_{ab}^{elec} = \frac{2 \int \phi_a \phi_b \, \mathrm{d}V}{\int \phi_a^2 \, \mathrm{d}V + \int \phi_b^2 \, \mathrm{d}V} \tag{6.20}$$

The potentials ϕ of the structurally different molecules a and b are integrated over the whole volume V. The numerator quantifies the spatial overlap of the electrostatic potentials ϕ, while the denominator normalizes this value such that the resulting similarity index H_{ab}^{elec} falls in the interval between -1 and $+1$. The value $+1$ corresponds to molecules with identical potentials, whereas -1 corresponds to electrostatic complementarity, i.e., potentials of the same magnitude but opposite sign. In order to optimize the superposition of the two molecules, the Coulomb potentials are approximated by Gaussian potentials and Eq. 6.20 is minimized with respect to the relative orientation of the two molecules [43]. The structural superposition, which was obtained by minimizing Eq. 6.20, was used to identify functionally equivalent residues in plastocyanin and cytochrome c_6 [42]. Interestingly, it can be seen that functional analogous aminoacids enable the specific recognition for the two isofunctional proteins. Figure 6.9 shows the electrostatic potential and the cartoon representations of plastocyanin and cytochrome c_6 of *Chlamydomonas rheinhardii* in the orientation that corresponds to the superposition of the two proteins.

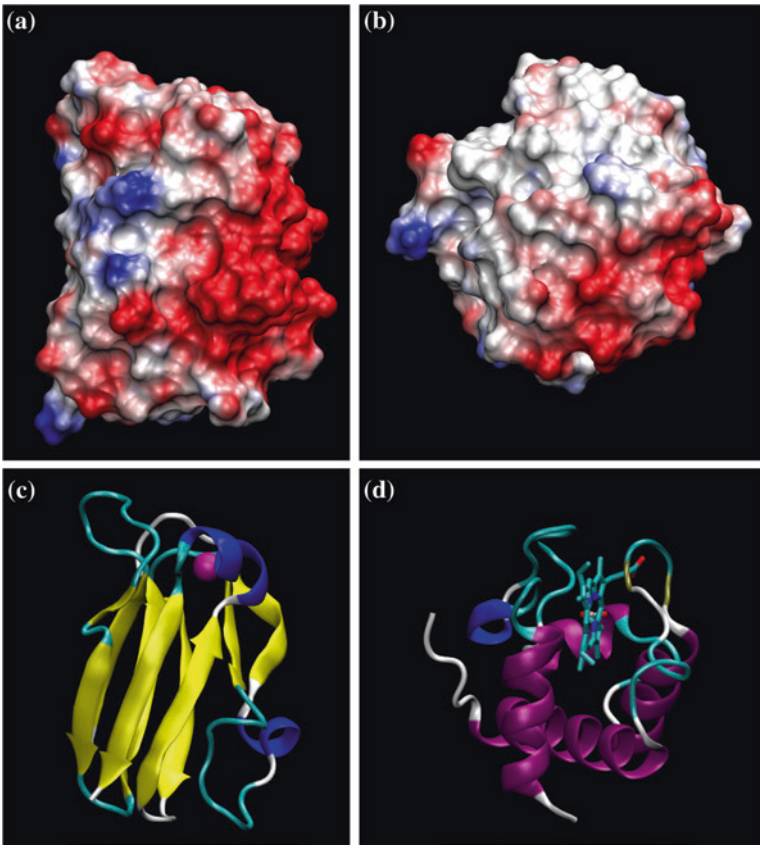

Fig. 6.9 Comparison of the electrostatic potentials of the isofunctional proteins plastocyanin and cytochrome c_6. **a** Electrostatic potential of plastocyanin mapped to its surface. **b** Electrostatic potential of cytochrome c_6 mapped to its surface. **c** Cartoon representation of plastocyanin. The copper ion is shown in a space-filled model. **d** Cartoon representation of cytochrome c_6. The heme is shown in a stick model

6.4 Titration Behavior of Proteins

The electrostatics of a protein is determined by the charge states of its protonatable and redox-active groups. However, often the description of proton binding and redox equilibria in proteins is considerably more complicated than that of small molecules because of the mutual interaction of the many protonatable and redox-active groups in one protein. Here we outline the methods how to describe such equilibria.

6.4.1 Shifted Titration Curves in Proteins

The titration curve of aminoacids in proteins can be shifted compared to the titration curve in aqueous solution. This shift has two main causes. First, the charges and dipoles of the protein may stabilize or destabilize a charge state in the protein. Second, the desolvation when a charged group is brought from the highly polar aqueous solution into the usually apolar interior of the protein destabilizes charged states. The actual direction of the shift depends on the balance of the different effects. Generally, it can be said that buried charges in the protein are unfavorable and thus the aminoacids tend to be uncharged under such circumstances. However, this rule of thumb may not apply if ion pairs are buried.

One way to obtain the protonation energy of a site within a protein is to calculate the difference between the pK_a value of an appropriate model compound in aqueous solution and the pK_a value of the protonatable group in the protein. The pK_a value associated with this protonation energy is called intrinsic pK_a value. In proteins with many interacting titratable residues, the intrinsic pK_a value is the pK_a value the titratable group would have if all other titratable groups in the protein are in a defined reference protonation state. To compute intrinsic pK_a values, pK_a values of a model compounds are required, such as for instance an aminoacid with blocked terminal amino- and carboxyl group. These values can be obtained either from experiment or from quantum chemical calculations. Sometimes only quantum chemical calculations are able to obtain the pK_a values for model compounds of prosthetic groups in proteins, because appropriate model compounds cannot be synthesized.

If only electrostatic contributions cause the difference between the protonation energies of a titratable group in a protein and in aqueous solution, the Poisson-Boltzmann equation provides a reasonable approximation of this energy difference. Transferring the protonatable group i with a given pK_a value $\mathrm{p}K_{a,i}^{\mathrm{model}}$ from aqueous solution into a protein causes an energy shift. This energy shift can be separated into two contributions. The first energy contribution $\Delta\Delta G_{\mathrm{Born}}$ is a Born-energy-like term (Eq. 6.21), which arises from the interaction of the charges of the protonatable group with its reaction field.

$$\begin{aligned}\Delta\Delta G_{\mathrm{Born,i}}^{\mathrm{prot}} = &\frac{1}{2}\sum_{a=1}^{N_{Q,i}} Q_{a,i}^{h}[\phi_p(\mathbf{r}_a; Q_i^h) - \phi_m(\mathbf{r}_a; Q_i^h)] \\ &- \frac{1}{2}\sum_{a=1}^{N_{Q,i}} Q_{a,i}^{d}[\phi_p(\mathbf{r}_a; Q_i^d) - \phi_m(\mathbf{r}_a; Q_i^d)]\end{aligned} \tag{6.21}$$

lThe second energy contribution $\Delta\Delta G_{\mathrm{back}}$ arises from the interaction of the charges of the protonatable group with non-titrating background charges (Eq. 6.22).

$$\Delta\Delta G_{\mathrm{back,i}}^{\mathrm{prot}} = \sum_{a=1}^{N_p} q_a[\phi_p(\mathbf{r}_a; Q_i^h) - \phi_p(\mathbf{r}_a; Q_i^d)] - \sum_{a=1}^{N_m} q_a[\phi_m(\mathbf{r}_a; Q_i^h) - \phi_m(\mathbf{r}_a; Q_i^d)] \tag{6.22}$$

The summations in Eq. 6.21 run over the $N_{Q,i}$ atoms of group μ that have different charges in the protonated (h) ($Q^h_{a,i}$) and in the deprotonated (d) ($Q^d_{a,i}$) form. The first summation in Eq. 6.22 runs over the N_p charges of the protein that belong to atoms in non-titratable groups or to atoms of titratable groups (not i) in their uncharged protonation form. The second summation in Eq. 6.22 runs over the N_m charges of atoms of the model compound that do not have different charges in the different protonation forms. The terms $\phi_m\left(\mathbf{r}_a, Q^h_i\right)$, $\phi_m\left(\mathbf{r}_a, Q^d_i\right)$, $\phi_p\left(\mathbf{r}_a, Q^h_i\right)$, and $\phi_p\left(\mathbf{r}_a, Q^d_i\right)$ denote the values of the electrostatic potential at the position $\mathbf{r}$ of the atom a. The electrostatic potential was obtained by solving the Poisson-Boltzmann equation numerically using the shape of either the protein (subscript p) or the model compound (subscript m) as dielectric boundary and assigning the charges of the titratable group μ in either the protonated (Q^h_i) or the deprotonated (Q^d_i) form to the respective atoms. These two energy contributions and the pK_a value of the model compound $\mathrm{p}K^{\mathrm{model}}_{a,\mu}$ are combined to obtain the so-called intrinsic pK_a value $\mathrm{p}K^{\mathrm{intr}}_{\mu}$ (Eq. 6.23) of the residue.

$$\mathrm{p}K^{\mathrm{intr}}_i = \mathrm{p}K^{\mathrm{model}}_{a,i} - \frac{1}{RT\ln 10}\left(\Delta\Delta G^{\mathrm{prot}}_{\mathrm{Born,i}} + \Delta\Delta G^{\mathrm{prot}}_{\mathrm{back,i}}\right) \tag{6.23}$$

The intrinsic pK_a value is the pK_a value that this group would have, if all other protonatable groups are in their reference protonation form.

6.4.2 Microstate Model

An additional complication in proteins is that proteins usually contain more than one titratable group. Also the interaction W_{ij} between the two groups i and j in their charged form can be calculated using the electrostatic potential obtained from the PBE (Eq. 6.24).

$$W_{ij} = \sum_{a=1}^{N_{Q,i}} \left[Q^h_{i,a} - Q^d_{i,a}\right]\left[\phi_p\left(\mathbf{r}_a, Q^h_j\right) - \phi_p\left(\mathbf{r}_a, Q^d_j\right)\right] \tag{6.24}$$

The interaction between the titratable groups can lead to titration curves that do not show a standard sigmoidal shape that can be fitted with the Henderson-Hasselbalch equation [45]. The difficulties can be resolved if the problem is formulated in terms of well-defined microstates of the protein which have a certain probability, instead of considering the protein as a system of groups with a certain protonation probability, as will be outlined now. Since the formalism can be easily extended to treat not only protonation equilibria but also redox equilibria [43], we explain it here in a more generalized form.

Let us consider a system that possesses N protonatable sites and K redox-active sites. Such a system can adopt $M = 2^{N+K}$ states assuming that each sites can exist in two forms. The interaction between them can be modeled purely electrostatically, i.e. the electronic coupling is negligible. Each state of the system can be

written as an $N + M$-dimensional vector $\vec{x} = (x_1, \ldots, x_{N+K})$, where x_i is 0 or 1 if site i is deprotonated (reduced) or protonated (oxidized), respectively. Each state of the system has a well-defined energy which depends on the energetics of the individual sites and the interaction between sites. The energy of a state $\vec{x}_\nu$ is given by [46–50]:

$$\begin{aligned} G(\vec{x}_\nu) = &\sum_{i=1}^{N} \left(x_{\nu,i} - x_i^\circ\right) RT \ln 10 \left(\mathrm{pH} - \mathrm{p}K_{a,i}^{\mathrm{intr}}\right) \\ &- \sum_{i=1}^{K} \left(x_{\nu,i} - x_i^\circ\right) F \left(E - E_i^{\mathrm{intr}}\right) \\ &+ \frac{1}{2} \sum_{i=1}^{N+K} \sum_{j=1}^{N+K} \left(x_{\nu,i} - x_i^\circ\right) \left(x_{\nu,j} - x_j^\circ\right) W_{ij} \end{aligned} \tag{6.25}$$

where R is the gas constant; T is the absolute temperature; F is the Faraday constant; $x_{\nu,i}$ denotes the protonation or redox form of the site i in state $\vec{x}_\nu$, x_i° is the reference form of site i; $\mathrm{p}K_{a,i}^{\mathrm{intr}}$ and E_i^{intr} are the $\mathrm{p}K_a$ value and redox potential, respectively, that site i would have if all other sites are in their reference form (intrinsic $\mathrm{p}K_a$ value and intrinsic redox potential); E is the reduction potential of the solution; pH is the pH value of the solution; W_{ij} represents the interaction of site i with site j.

Equilibrium properties of a physical system are completely determined by the energies of its states. To keep the notation concise, states will be numbered by Greek indices, i.e., for state energies we write G_ν instead of $G(\vec{x}_\nu)$. For site indices, the roman letters i and j will be used. The equilibrium probability of a single state $\vec{x}_\nu$ is given by

$$P_\nu^{eq} = \frac{e^{-\beta G_\nu}}{Z} \tag{6.26}$$

with $\beta = 1/RT$ and Z being the partition function of the system.

$$Z = \sum_{\nu=1}^{M} e^{-\beta G_\nu} \tag{6.27}$$

The sum runs over all M possible states. Properties of single sites can be obtained from Eq. 6.26 by summing up the individual contributions of all states. For example, the probability of site i being oxidized is given by

$$\langle x_i \rangle = \sum_{\nu}^{M} x_{\nu,i} P_\nu^{eq} \tag{6.28}$$

where $x_{\nu,i}$ denotes the protonation or redox form of site i in the charge state $\vec{x}_\nu$. For small systems, this sum can be evaluated explicitly. For larger systems, Monte-Carlo techniques can be used to determine these probabilities [51, 52].

For a system of interacting sites, the protonation or reduction probabilities $\langle x_i \rangle$ can show a complex shape, thus rendering the assignment of pK_a values or midpoint potentials to individual sites difficult or even meaningless [45, 53–55]. The energy differences between microstates, however, remain well defined and thus form a convenient basis to describe the system. For individual sites in such a complex system, one can however define pH-dependent pK_a values and solution redox potential dependent midpoint potentials [56].

$$\mathrm{p}K_i = \mathrm{pH} + \frac{1}{\ln 10}\left(\ln \frac{\langle x_i \rangle}{1 - \langle x_i \rangle}\right) \tag{6.29}$$

$$E_i^{\mathrm{o}} = E - \frac{RT}{F}\left(\ln \frac{\langle x_i \rangle}{1 - \langle x_i \rangle}\right) \tag{6.30}$$

These values define properties that are directly related to a free energy difference [56] and are relevant for understanding enzymatic mechanisms.

6.4.3 An Illustrative Example

Titration curves with a non-standard sigmoidal shape can be seen in many protein titration studies, but their interpretation is often very complicated. Sometimes also small molecules show a complex titration behavior. One example is diethylenetriamine-pentaacetate [57–59]. Here, we discuss a fictitious molecule with three groups and each of them can bind a proton. The intrinsic pK_a values and the interaction energies are given in Fig. 6.10a. The resulting microscopic equilibria are given in Fig. 6.10b. The individual titration are given in Fig. 6.10c. Given all the different energy terms, the population of the microscopic and macroscopic states in dependence of pH can be calculated (Fig. 6.10d). Figure 6.10e shows how the individual titration curves and the population of the different microstates contribute to the titration curves of the individual sites.

The titration curve of the central group is unusual because of its non-monotonic shape. In particular between pH 5 and pH 9, the protonation probability increases with increasing pH (i.e. with decreasing proton concentration). The population of the microstates give a physical rational for the unusual, irregular titration behavior of the central group. At high pH (low proton concentration), the protons bind preferable to the central group. Binding the second proton to one of the terminal groups while the central group stays protonated is unfavorable, because these two proton binding sites repel each other. Therefore, when the second proton binds, it is more favorable to deprotonate the central group and to protonate both terminal groups. When the two terminal amines are protonated, the two protons are at a greater distance from each other, and thus they repel each other less as compared to protonating one terminal and the central amine. Finally at very low pH (high proton concentration), all three sites will bind a proton.

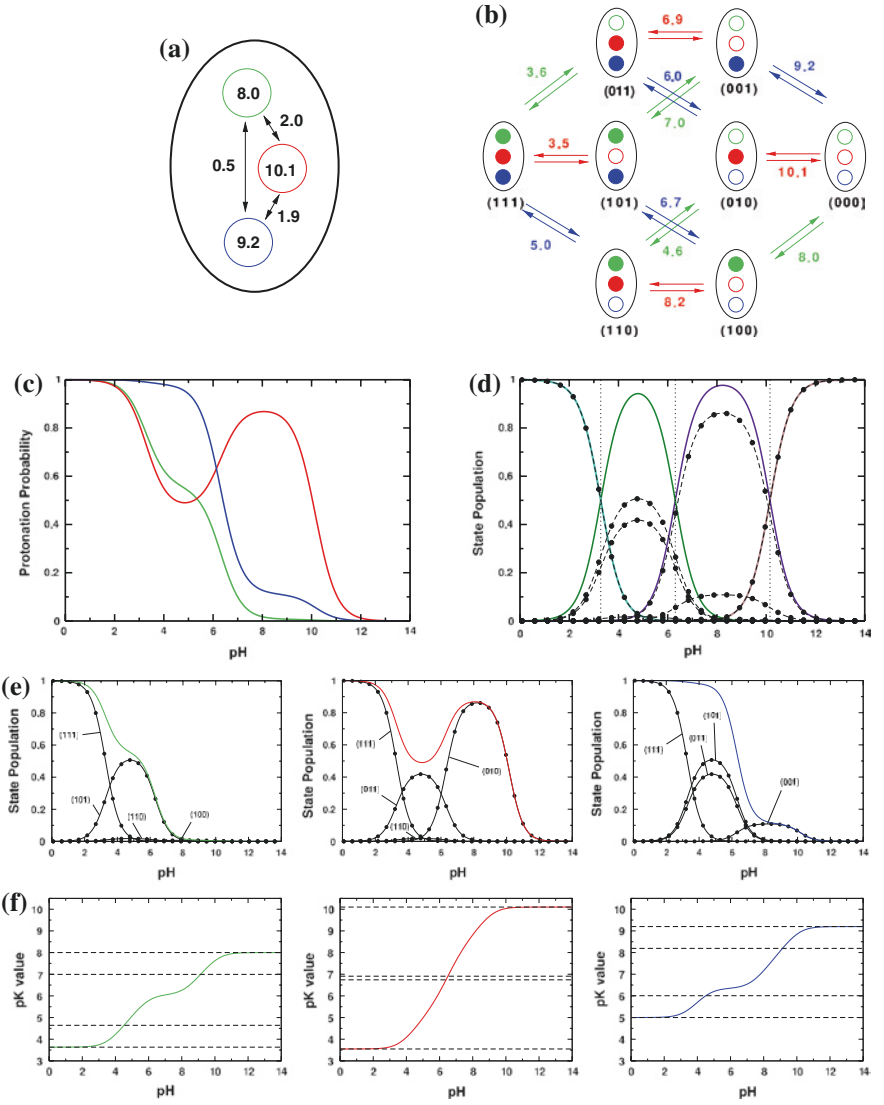

Fig. 6.10 Titration behavior of a fictitious molecule with three titratable groups. **a** Schematic drawing of a fictitious molecule with three titratable groups. The *numbers in the circles* are the intrinsic pK_a values, the numbers at the arrows are the respective interaction energies. All numbers are given in pK_a units. **b** Microscopic protonation equilibria. The numbers at the arrows indicate the microscopic pK_a values. **c** Individual titration curve of the three sites. **d** Populations of the macroscopic (*solid lines*) and microscopic (*lines with circles*) states of the system. The *dotted lines* mark the macroscopic pK_a values. **e** Contributions of the different microstates to the titration curves of the three individual sites. **f** Effective (*solid lines*) and microscopic (*dashed lines*) pK_a values of the three sites

Obviously, it is not easily possible to assign pK_a values to the individual sites. All groups are associated with four microscopic pK_a values (see Fig. 6.10f). However, using Eq. 6.29, it is possible to assign an effective pK_a value to each site (see Fig. 6.10e). It can be seen that the effective pK_a value assumes values that vary within the limits of the microscopic pK_a value.

Even more interesting than the pH-dependence of the pK_a values is the pH-dependence of the protonation energy. The protonation free energy ΔG_{prot} is given by

$$\Delta G_{\text{prot}} = RT\ln 10(\text{pH} - \text{p}K_a) \tag{6.31}$$

This equation leads to a linear dependence of the protonation energy for pH-independent pK_a values that are found in the case of isolated titratable groups. However, if the pK_a value is pH-dependent, the pH-dependence of the protonation free energy becomes non-linear.

Even if this example seems to be an extreme example, similarly complicated protonation equilibria caused by charge-charge interactions occur frequently in proteins. A sign of such complications are irregular titration curves. Moreover, even if the titration curves in proteins show apparently a standard sigmoidal shape, the interactions between titratable groups may lead to pH-dependent pK_a values [56]. Such pH-dependent pK_a values may lead to nearly pH-independent protonation energies in a certain pH range and thus may explain why some particular residue can function as proton donor or acceptor over a large pH range allowing catalysis under different pH conditions. Probably for this reason, there are often more protonatable residues in the active site of enzymes than the specific function would require.

6.5 Other Applications of Continuum Electrostatic Calculations

The continuum electrostatic calculations have a great potential for studying the mechanism of enzymes. A step forward in the analysis of enzymatic mechanisms is the combination of continuum electrostatics with quantum chemical calculations as inCOSMO [60] orCPCM [61]. However, it is relatively difficult to model the protein environment appropriately with such methods. An attractive alternative is the self-consistent reaction field method (SCRF) which was developed by Tomasi and coworkers [62] and applied to protein systems by Bashford, Noodlemann and coworkers [63–67]. This method combines quantum chemical calculations with Poisson-Boltzmann calculations and allows to account for the charge distribution within the protein and solvent polarization effects in quantum chemical calculations.

To explore possible mechanisms, it is often required to examine many different possibilities. Sometimes, even many mechanism may be possible at the same time and a single answer may not exist. Such complex reaction schemes can be explored with the help of the microstate model introduced in Sect. 6.4.2.

The kinetics of such reactions can be simulated by a master equation approach. The rate constants which are required for such simulations can be calculated using electrostatic methods [68–71]. Thus, combined with a master equation approach, continuum electrostatics offers also a possibility to access the non-equilibrium behavior of biomolecular systems. In the microstate formalism given by Eqs. 6.25–6.28, charge transfer events are described as transitions between well-defined microstates of a system. The time dependence of the population of each microstate can be simulated using a master equation

$$\frac{d}{dt}P_\nu(t) = \sum_{\mu=1}^{M} k_{\nu\mu}P_\mu(t) - \sum_{\mu=1}^{M} k_{\mu\nu}P_\nu(t) \tag{6.32}$$

where $P_\nu(t)$ denotes the probability that the system is in charge state ν at time t, $k_{\nu\mu}$ denotes the probability per unit time that the system will change its state from μ to ν. In Eq. 6.32, the first sum includes all the reactions that generate state ν, the second sum includes all the reactions that destroy state ν. The summations run over all possible states μ. In order to restrict the number of states and only consider states that are accessible in a certain energy range, methods like extended Dead End Elimination [72] can be used. Simulating charge transfer by Eq. 6.32 assumes that these processes can be described as a Markovian stochastic process. This assumption implies that the probability of a given charge transfer only depends on the current state of the system and not on the way in which the system has reached this state. The system given by Eq. 6.32 is a system of coupled linear differential equations with constant coefficients, for which an analytical solution exists [69, 70]. Equation 6.32 describes the time evolution of the probability distribution of microstates of the system. For these microstates, energies G_ν and transition probabilities $k_{\nu\mu}$ can be assigned unambiguously. The time-dependent probability of finding a single site in a particular form can be obtained by summing up individual contributions from the time-dependent probabilities $P_\nu(t)$.

$$\langle x_i \rangle(t) = \sum_{\nu}^{M} x_{\nu,i}P_\nu(t) \tag{6.33}$$

The application of the method outlined about to electron transfer reactions is particularly attractive, since their rates can be estimated using the rate law developed by Moser and Dutton [73, 74] which relies on the Marcus theory [75] and agrees well with experimental data. Mainly three factors govern the rate constants of biological electron transfer reactions: the energy difference between the donor state and the acceptor state, the environmental polarization (reorganization energy), and the electronic coupling between the redox sites. The energy barrier for the transfer process is given in the framework of Marcus theory as

$$\Delta G^{\neq} = \frac{(\Delta G^\circ + \lambda)^2}{4\lambda} \tag{6.34}$$

where ΔG° is the energy difference between donor and acceptor state and λ is the reorganization energy. The electronic coupling between the redox sites is accounted for by a distance-dependent exponential function $A\exp(-\beta(R-R_\circ))$, where R is the edge-to-edge distance between the electron transfer centers, $R_\circ$ represents a van der Waals contact distance and A represents an optimal rate.

The free energy ΔG° for a transition between two states ν and μ can be calculated within the electrostatic model using Eq. 6.25. The reorganization energy λ contains two contributions, $\lambda = \lambda_o + \lambda_i$, where λ_o is the solvent reorganization energy and λ_i is the inner sphere reorganization energy. λ_o was shown to be accessible to calculations using electrostatic potentials obtained from the solution of the Poisson-Boltzmann equation [75, 76]. The inner sphere reorganization energy λ_i can be estimated by quantum chemical calculations and it is often found to be significantly smaller than the solvent reorganization energy [77–80].

For analyzing a complex charge transfer system, it is of particular interest to follow the flow of charges through the system, i.e., the charge flux. The flux from state ν to state μ is determined by the population of state ν times the probability per unit time that state ν will change into state μ, i.e., by $k_{\mu\nu}P_\nu(t)$. The net flux between states μ and ν is thus given by

$$J_{\nu\mu}(t) = k_{\nu\mu}P_\mu(t) - k_{\mu\nu}P_\nu(t) \tag{6.35}$$

The net flux (Eq. 6.35) is positive if there is a net flux from state μ to state ν. This flux analysis allows to deduct the reaction mechanism from even very complex reaction schemes [70].

In cases when the number of possible microstates get too large, the differential equation can not be solved analytically anymore. Thus, approximations and simulations need to be applied. One attractive simulation method is the dynamical Monte Carlo Simulation scheme [81], which allows to simulate very complex reaction mechanism such as proton transfer through a protein matrix [82]. Again, the reaction parameter can be obtained from continuum electrostatic calculations. Each simulation trajectory describes one particular reaction path through the possible states of the system. A reaction mechanism can then be inferred from the analysis of many such trajectories. Up to now, this method was only applied to relatively simple systems [82]. However, future application to enzymes which involve chemical transformation, proton and electron transfer as well as conformational changes seem possible making dynamical Monte Carlo simulations a promising future road to analyze enzyme function.

6.6 Conclusion

Electrostatic interactions play a major role in biomolecular systems. In particular, the mechanism of enzyme cannot be understood without the correct evaluation of the effect of all charges involved in the enzyme or in the environment. The methods based on continuum electrostatics are extremely valuable in this task, since they allow

the analysis of the electrostatic interactions involving the macromolecular partners in their environment at meaningful time scales. In this review, we have shown how continuum electrostatic methods provide essential information both on the thermodynamics and the kinetics of biological mechanisms. These methods model essential biophysical aspects correctly and allow a computationally efficient calculation of biochemical reactions. We believe that continuum electrostatics has a broad range of applications in biomolecular modeling and the value of the method will become even more obvious when more and larger protein machines will be investigated.

Acknowledgments This work was supported by the DFG Grants UL 174/8 and BO 3578/1.

References

1. Feig M (ed) (2010) Modeling solvent environments. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
2. van Gunsteren WF, Bakowies D, Baron R, Chandrasekhar I, Christen M et al (2001) Biomolecular modeling: goals, problems, perspectives. Angew Chem Int Ed Engl 45:4064–4092
3. Honig B, Nicholls A (1995) Classical electrostatics in biology and chemistry. Science 268:1144–1149
4. Richards FM (1977) Areas, volumes, packing and protein structure. Annu Rev Biophys Bioeng 6:151–176
5. 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
6. Kirkwood JG (1934) Theory of solutions of molecules containing widely separated charges with special application to zwitterions. J Chem Phys 2:351–361
7. Daune M (1999) Molecular Biophysics. University Press, Oxford
8. Warwicker J, Watson HC (1982) Calculation of the electrostatic potential in the active site cleft due the α-helix dipols. J Mol Biol 186:671–679
9. Bashford D (1997) An object-oriented programming suite for electrostatic effects in biological molecules. In Yutaka I, Rodney RO, John VWR, Marydell T (eds) Scientific computing in object-oriented parallel environments. Springer, Berlin, pp 233–240
10. Im W, Beglov D, Roux B (1998) Continuum solvation model: electrostatic forces from numerical solutions to the Poisson-Boltzmann equation. Comp Phys Comm 111:59–75
11. Boschitsch AH, Fenley MO (2011) A fast and robust Poisson-Boltzmann solver based on adaptive cartesian grids. J Chem Theor Comput 7:1524–1540
12. Holst MJ, Saied F (1993) Multigrid solution of the Poisson-Boltzmann equation. J Comput Chem 14:105–113
13. Holst MJ, Saied F (1995) Numerical-solution of the nonlinear Poisson-Boltzmann equation: developing more robust and efficient methods. J Comput Chem 16:337–364
14. Holst MJ (2001) Adaptive numerical treatment of elliptic systems on manifolds. Adv Comp Math 15:139–191
15. 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
16. Sklenar H, Eisenhaber F, Poncin M, Lavery R (1990) Including solvent and counterion effects in the force fields of macromolecular mechanics: the field integrated electrostatic approach (FIESTA). In: David LB, Richard L (eds) Theoretical biochemistry and molecular biophysics, pp 317–335

17. Cortis CM, Friesner RA (1997) Numerical solution of the Poisson-Boltzmann equation using tetrahedral finite-element meshes. J Comp Chem 18:1591–1608
18. Lee B, Richards FM (1971) The interpretation of protein structures: estimation of static accessibility. J Mol Biol 55:379–380
19. Klapper I, Hagstrom R, Fine R, Sharp K, Honig B (1986) Focusing of electric fields in the active site of Cu–Zn superoxide dismutase: effects of ionic strength and amino-acid modification. Proteins 1:47–59
20. Sengupta D, Behera RN, Smith JC, Ullmann GM (2005) The α-helix dipole: screened out? Structure 13:849–855
21. Sengupta D, Meinhold L, Langosch D, Ullmann GM, Smith JC (2005) Energetics of helical-peptide orientations in membranes. Proteins 58:913–922
22. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14:33–38
23. Bashir Q, Scanu S, Ubbink M (2011) Dynamics in electron transfer protein complexes. FEBS J 278:1391–1400
24. Ullmann GM, Knapp EW, Kostic NM (1997) Computational simulation and analysis of the dynamic association between plastocyanin and cytochrome *f*. Consequences for the electron-transfer reaction. J Am Chem Soc 119:42–52
25. Ubbink M, Ejdebäck M, Karlson BG, Bendall DS (1998) The structure of the complex of plastocyanin and cytochrome *f*, determined by paramagnetic NMR and restrained rigid body molecular dynamics. Structure 6:323–335
26. Qin L, Kostic NM (1993) Importance of protein rearrangement in the electron-transfer reaction between the physiological partners cytochrome *f* and plastocyanin. Biochemistry 32:6073–6080
27. Pearson DC, Gross EL, David E (1996) Electrostatic properties of cytochrome *f*: implications for docking with plastocyanin. Biophys J 71:64–76
28. De Rienzo F, Gabdoulline RR, Menziani MC, De Benedetti PG, Wade RC (2001) Electrostatic analysis and Brownian dynamics simulation of the association of plastocyanin and cytochrome *f*. Biophys J 81:3090–3104
29. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH (1953) Equation of state calculation by fast computing machines. J Chem Phys 21:1087–1092
30. Bashir Q, Volkov AN, Ullmann GM, Ubbink M (2010) Visualization of the encounter ensemble of the transient electron transfer complex of cytochrome c and cytochrome c peroxidase. J Am Chem Soc 132:241–247
31. Volkov AN, Bashir Q, Worrall JAR, Ullmann GM, Ubbink M (2010) Shifting the equilibrium between the encounter state and the specific form of a protein complex by interfacial point mutations. J Am Chem Soc 132:11487–11495
32. Kostić NM (1996) Dynamic aspects of electron-transfer reactions in metalloprotein complexes. In: Metal-containing polymeric materials. Plenum Press, New York, pp 491–500
33. Madura JD, Davis ME, Gilson MK, Wade RC, Luty BA et al (1994) Biological applications of electrostatic calculations and Brownian dynamics. Rev Comp Chem 5:229–267
34. Gabdoulline RR, Wade RC (2002) Biomolecular diffusional association. Curr Opin Struct Biol 12:204–213
35. Andrew SM, Thomasson KA, Northrup SH (1993) Simulation of electron-transfer self-exchange in cytochromes *c* and b_5. J Am Chem Soc 115:5516–5521
36. Northrup SH, Allison SA, McCammon JA (1984) Brownian dynamics simulation of diffusion-influenced bimolecular reactions. J Chem Phys 80:1517–1524
37. Northrup SH (1994) Hydrodynamic motions of large molecules. Curr Opin Struct Biol 4:269–274
38. Pearson DC, Gross EL (1995) The docking of cytochrome *f* with plastocyanin: three possible complexes. In: Mathis P (ed) Photosynthesis: from light to biosphere, vol II. Kluwer Academic Publishers, New York, pp 729–732
39. Haddadian EJ, Gross EL (2006) A Brownian dynamics study of the effects of cytochrome *f* structure and deletion of its small domain in interactions with cytochrome c_6 and plastocyanin in *Chlamydomonas reinhardtii*. Biophys J 90:566–577

40. Gross EL, Rosenberg I (2006) A Brownian dynamics study of the interaction of phormidium cytochrome f with various cyanobacterial plastocyanins. Biophys J 90:366–380
41. Haddadian EJ, Gross EL (2005) Brownian dynamics study of cytochrome f interactions with cytochrome c(6) and plastocyanin in *Chlamydomonas reinhardtii* plastocyanin, and cytochrome c(6) mutants. Biophys J 88:2323–2339
42. Ullmann GM, Hauswald M, Jensen A, Kostic NM, Knapp EW (1997) Comparison of the physiologically-equivalent proteins cytochrome c_6 and plastocyanin on the basis of their electrostatic potentials. Tryptophane 63 in cytochrome c_6 may be isofunctional with tyrosine 83 in plastocyanin. Biochemistry 36:16187–16196
43. Ullmann GM, Hauswald M, Jensen A, Knapp EW (2000) Superposition of ferredoxin and flavodoxin using their electrostatic potentials. Implications for their interactions with photosystem I and ferredoxin: NADP reductase. Proteins 38:301–309
44. Hodgkin E, Richards W (1987) Molecular similarity based on electrostatic potential and electric field. Int J Quant Chem Quant Biol Symp 14:105–110
45. Klingen AR, Ullmann GM (2006) Theoretical investigation of the behavior of titratable groups in proteins. Photochem Photobiol Sci 5:588–596
46. Bashford D, Karplus M (1990) pK_as of ionizable groups in proteins: atomic detail from a continuum electrostatic model. Biochemistry 29:10219–10225
47. Ullmann GM, Knapp EW (1999) Electrostatic computations of protonation and redox equilibria in proteins. Eur Biophys J 28:533–551
48. Ullmann GM (2000) The coupling of protonation and reduction in proteins with multiple redox centers: theory, computational method, and application to cytochrome c_3. J Phys Chem B 104:6293–6301
49. Gunner MR, Mao J, Song Y, Kim J (2006) Factors influencing the energetics of electron and proton transfers in proteins. What can be learned from calculations. Biochim Biophys Acta 1757:942–968
50. Nielsen JE, McCammon JA (2003) Calculating pKa values in enzyme active sites. Protein Sci 12:1894–1901
51. 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. Proc Natl Acad Sci USA 88:5804–5808
52. Ullmann RT, Ullmann GM (2012) GMCT: a Monte Carlo simulation package for macromolecular receptors. J Comp Chem 33:887–900
53. Ullmann GM (2003) Relations between protonation constants and titration curves in polyprotic acids: a critical view. J Phys Chem B 107:6293–6301
54. Onufriev A, Case DA, Ullmann GM (2001) A novel view on the ph titration of biomolecules. Biochemistry 40:3413–3419
55. Onufriev A, Ullmann GM (2004) Decomposing complex ligand binding into simple components: connections between microscopic and macroscopic models. J Phys Chem B 108:11157–11169
56. Bombarda E, Ullmann GM (2010) pH-dependent pk_a values in proteins-a theoretical analysis of protonation energies with practical consequences for enzymatic reactions. J Phys Chem B 114:1994–2003
57. Kula R, Sawyer D (1964) Protonation studies of anion of diethylenetriaminepentaacetic acid by nuclear magnetic resonance. Inorg Chem 3:458
58. Sudmeier JL, Reilley CN (1964) Nuclear magnetic resonance studies of protonation of polyamine and aminocarboxylate compounds in aqueous solution. Analyt Chem 36:1698–1706
59. Letkeman P (1979) An NMR protonation study of metal diethylenetriaminepentaacetic acid complexes. J Chem Ed 56:348–351
60. Klamt A, Schuurmann G (1993) Cosmo: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. J Chem Soc Perkin Trans 2:799–805
61. Cossi M, Rega N, Scalmani G, Barone V (2003) Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. J Comput Chem 24:669–681

62. Miertus S, Scrocco E, Tomas J (1981) Electrostatic interaction of a solute with a continuum. A direct utilization of Ab initio molecular potentials for the prevision of solvent effects. Chem Phys 55:117–129
63. Chen JL, Noodleman L, Case D, Bashford D (1994) Incorporating solvation effects into density functional electronic structure calculations. J Phys Chem 98:11059–11068
64. Li J, Fischer CL, Chen JL, Bashford D, Noodleman L (1996) Calculation of redox potentials and pK_a values of hydrated transition metal cations by a combined density functional and continuum dielectric theory. J Phys Chem 96:2855–2866
65. Richardson WH, Peng C, Bashford D, Noodleman L, Case DA (1997) Incorporating solvation effects into density functional theory: calculation of absolute acidities. Int J Quant Chem 61:207–217
66. Li J, Nelson MR, Peng CY, Bashford D, Noodleman L (1998) Incorporating protein environments in density functional theory: a self-consistent reaction field calculation of redox potentials of [2Fe_2S] clusters in ferredoxin and phthalate dioxygenase reductase. J Phys Chem A 102:6311–6324
67. Liu T, Han WG, Himo F, Ullmann GM, Bashford D et al (2004) Density functional vertical self-consistent reaction field theory for solvatochromism studies of solvent-sensitive dyes. J Phys Chem B 108:11157–11169
68. Sham YY, Muegge I, Warshel A (1999) Simulating proton translocations in proteins: probing proton transfer pathways in the *Rhodobacter sphaeroides* reaction center. Proteins 36:484–500
69. Ferreira A, Bashford D (2006) Model for proton transport coupled to protein conformational change: application to proton pumping in the bacteriorhodopsin photocycle. J Am Chem Soc 128:16778–16790
70. Becker T, Ullmann RT, Ullmann GM (2007) Simulation of the electron transfer between the tetraheme-subunit and the special pair of the photosynthetic reaction center using a microstate description. J Phys Chem B 111:2957–2968
71. Bombarda E, Ullmann GM (2011) Continuum electrostatic investigations of charge transfer processes in biological molecules using a microstate description. Faraday Discuss 148:173–193
72. Kloppmann E, Ullmann GM, Becker T (2007) An extended dead-end elimination algorithm to determine gap-free lists of low energy states. J Comp Chem 28:2325–2335
73. Moser CC, Keske JM, Warncke K, Farid RS, Dutton PL (1992) Nature of biological electron transfer. Nature 355:796–802
74. Page CC, Moser CC, Chen X, Dutton PL (1999) Natural engineering principles of electron tunneling in biological oxidation–reduction. Nature 402:47–52
75. Marcus RA (1963) Free energy of nonequilibrium polarization systems. II. Homogeneous and electrode systems. J Chem Phys 38:1858–1862
76. Sharp KE (1998) Calculation of electron transfer reorganization energies using the finite difference Poisson Boltzmann model. Biophys J 73:1241–1250
77. Marcus RA, Sutin N (1985) Electron transfer in chemistry and biology. Biochim Biophys Acta 811:265–322
78. Williams RJP (1999) Electron transfer and proton coupling in proteins. J Solid State Chem 145:488–495
79. Olsson MHM, Ryde U, Roos BO (1998) Quantum chemical calculation of the reorganization energy of blue copper proteins. Prot Sci 81:6554–6558
80. Ryde U, Olsson MHM (2001) Structure, strain and reorganization energy of blue copper models in the protein. Int J Quant Chem 81:335–347
81. Gillespie DT (2001) Approximate accelerated stochastic simulation of chemically reacting systems. J Chem Phys 115:1716–1733
82. Till MS, Becker T, Essigke T, Ullmann GM (2008) Simulating the proton transfer in gramicidin A by a sequential dynamical Monte Carlo method. J Phys Chem B 112:13401–13410