

22.3. ELECTROSTATIC INTERACTIONS IN PROTEINS

for example. Rather, it is one of the quantities that one aims to extract from theoretical models to compare with an experiment.

22.3.2.4. Calculation of energies and forces

Once the electrostatic potential distribution has been obtained, calculation of experimental properties usually requires evaluation of the electrostatic energy or force. For a linear system (where the dielectric and ionic responses are linear) the electrostatic free energy is given by

$$\Delta G^{\text{el}} = 1/2 \sum_i \varphi_i q_i, \quad (22.3.2.6)$$

where φ_i is the potential at an atom with charge q_i . The most common source of nonlinearity is the Boltzmann term in the PB equation (22.3.2.4) for highly charged molecules such as nucleic acids. The total electrostatic energy in this case is (Reiner & Radke, 1990; Sharp & Honig, 1990; Zhou, 1994)

$$\Delta G^{\text{el}} = \int_V \{ \rho^e \varphi - (\epsilon E^2 / 8\pi) - kT \sum_i c_i^0 [\exp(-z_i e \varphi / kT) - 1] \} \text{d}\mathbf{r}, \quad (22.3.2.7)$$

where the integration is now over all space.

The general expression for the electrostatic force on a charge q is given by the gradient of the total free energy with respect to that charge's position,

$$\mathbf{f}_q = -\nabla_{\mathbf{r}_q}(G^{\text{el}}). \quad (22.3.2.8)$$

If the movement of that charge does not affect the potential distribution due to the other charges and dipoles, then equation (22.3.2.8) can be evaluated using the 'test charge' approach, in which case the force depends only on the gradient of the potential or the field at the charge:

$$\mathbf{f} = q\mathbf{E}. \quad (22.3.2.9)$$

However, in a system like a macromolecule in water, which has a non-homogeneous dielectric, forces arise between a charge and any dielectric boundary due to image charge (reaction potential) effects. A similar effect to the 'dielectric pressure' force arises from solvent-ion pressure at the solute-solvent boundary. This results in a force acting to increase the solvent exposure of charged and polar atoms. An expression for the force that includes these effects has been derived within the PB model (Gilson *et al.*, 1993):

$$\mathbf{f} = \rho^e \mathbf{E} - (1/2)E^2 \nabla \epsilon - kT \sum_i c_i^0 [\exp(-z_i e \varphi / kT) - 1] \nabla A, \quad (22.3.2.10)$$

where A is a function describing the accessibility to solvent ions, which is 0 inside the protein, and 1 in the solvent, and whose gradient is nonzero only at the solute-solvent surface. Similarly, in a two-dielectric model (solvent plus molecule) the gradient of ϵ is nonzero only at the molecular surface. The first term accounts for the force acting on a charge due to a field, as in equation (22.3.2.9), while the second and third terms account for the dielectric surface pressure and ionic atmosphere pressure terms respectively. Equation (22.3.2.10) has been used to combine the PB equation and molecular mechanics (Gilson *et al.*, 1995).

22.3.2.5. Numerical methods

A variety of numerical methods exist for calculating electrostatic potentials of macromolecules. These include numerical solution of self-consistent field electrostatic equations, which has been used in conjunction with the protein dipole-Langevin dipole method (Lee *et al.*, 1993). Numerical solution of the Poisson-Boltzmann

equation requires the solution of a three-dimensional partial differential equation, which can be nonlinear. Many numerical techniques, some developed in engineering fields to solve differential equations, have been applied to the PB equation. These include finite-difference methods (Brucoleri *et al.*, 1996; Gilson *et al.*, 1988; Nicholls & Honig, 1991; Warwicker & Watson, 1982), finite-element methods (Rashin, 1990; Yoon & Lenhoff, 1992; Zauhar & Morgan, 1985), multigridding (Holst & Saied, 1993; Oberoi & Allewell, 1993), conjugate-gradient methods (Davis & McCammon, 1989) and fast multipole methods (Bharadwaj *et al.*, 1994; Davis, 1994). Methods for treating the nonlinear PB equation include under-relaxation (Jayaram, Sharp & Honig, 1989) and powerful inexact Newton methods (Holst *et al.*, 1994). The nonlinear PB equation can also be solved *via* a self-consistent field approach, in which one calculates the potential using equation (22.3.2.5), then the mobile charge density is calculated using equation (22.3.2.3), and the procedure is repeated until convergence is reached (Pack & Klein, 1984; Pack *et al.*, 1986). The method allows one to include more elaborate models for the ion distribution, for example incorporating the finite size of the ions (Pack *et al.*, 1993). Approximate methods based on spherical approximations (Born-type models) have also been used (Schaeffer & Frommel, 1990; Still *et al.*, 1990). Considerable numerical progress has been made in finite methods, and accurate rapid algorithms are available. The reader is referred to the original references for numerical details.

22.3.3. Applications

An exhaustive list of applications of classical electrostatic modelling to macromolecules is beyond the scope of this chapter. Three general areas of application are discussed.

22.3.3.1. Electrostatic potential distributions

Graphical analysis of electrostatic potential distributions often reveals features about the structure that complement analysis of the atomic coordinates. For example, Fig. 22.3.3.1(a) shows the distribution of charged residues in the binding site of the proteolytic enzyme thrombin. Fig. 22.3.3.1(b) shows the resulting electrostatic potential distribution on the protein surface. The basic (positive) region in the fibrinogen binding site, which could be inferred from close inspection of the distribution of charged residues in Fig. 22.3.3.1(a), is clearly more apparent in the potential distribution. Fig. 22.3.3.1(c) shows the effect of increasing ionic strength on the potential distribution, shrinking the regions of strong potential. Fig. 22.3.3.1(d) is calculated assuming the same dielectric for the solvent and protein. The more uniform potential distribution compared to Fig. 22.3.3.1(b) shows the focusing effect that the low dielectric interior has on the field emanating from charges in active sites and other cleft regions.

22.3.3.2. Charge-transfer equilibria

Charge-transfer processes are important in protein catalysis, binding, conformational changes and many other functions. The primary examples are acid-base equilibria, electron transfer and ion binding, in which the transferred species is a proton, an electron or a salt ion, respectively. The theory of the dependence of these three equilibria within the classical electrostatic framework can be treated in an identical manner, and will be illustrated with acid-base equilibria. A titratable group will have an intrinsic ionization equilibrium, expressed in terms of a known intrinsic $\text{p}K_a^0$, where $\text{p}K_a^0 = -\log_{10}(K_a^0)$, K_a^0 is the dissociation constant for the reaction $\text{H}^+ \text{A} = \text{H}^+ + \text{A}$ and A can be an acid or a base. The $\text{p}K_a^0$ is determined by all the quantum-chemical, electrostatic and environ-

22. MOLECULAR GEOMETRY AND FEATURES

mental effects operating on that group in some reference state. For example, a reference state for the aspartic acid side-chain ionization might be the isolated amino acid in water, for which $\text{p}K_a^0 = 3.85$. In the environment of the protein, the $\text{p}K_a$ will be altered by three electrostatic effects. The first occurs because the group is positioned in a protein environment with a different polarizability, the second is due to interaction with permanent dipoles in the protein, the third is due to charged, perhaps titratable, groups. The effective $\text{p}K_a$ is given by

$$\text{p}K_a = \text{p}K_a^0 + (\Delta\Delta G^{\text{rf}} + \Delta\Delta G^{\text{perm}} + \Delta\Delta G^{\text{tit}})/2.303kT, \quad (22.3.3.1)$$

where the factor of $1/2.303kT$ converts units of energy to units of $\text{p}K_a$. The first contribution, $\Delta\Delta G^{\text{rf}}$, arises because the completely solvated group induces a strong favourable reaction field (see Section 22.3.2.3) in the high dielectric water, which stabilizes the charged form of the group. (The neutral form is also stabilized by the solvent reaction field induced by any dipolar groups, but to a lesser extent.) Desolvating the group to any degree by moving it into a less polarizable environment will preferentially destabilize the charged form of that group, shifting the $\text{p}K_a$ by an amount

$$\Delta\Delta G^{\text{rf}} = (1/2)\sum_i (q_i^d \Delta\varphi_i^{\text{rf},d} - q_i^p \Delta\varphi_i^{\text{rf},p}), \quad (22.3.3.2)$$

where q_i^p and q_i^d are the charge distributions on the group, $\Delta\varphi_i^{\text{rf},p}$ and $\Delta\varphi_i^{\text{rf},d}$ are the changes in the group's reaction potential upon moving it from its reference state into the protein, in the protonated (superscript p) and deprotonated (superscript d) forms, respectively, and the sum is over the group's charges. The contribution of the permanent dipoles is given by

$$\Delta\Delta G^{\text{tit}} = \sum_i (q_i^d - q_i^p) \varphi_i^{\text{perm}}, \quad (22.3.3.3)$$

where φ_i^{perm} is the interaction potential at the i th charge due to all the permanent dipoles in the protein, including the effect of screening. It is observed that intrinsic $\text{p}K_a$'s of groups in proteins are rarely shifted by more than 1 $\text{p}K_a$ unit, indicating that the effects of desolvation are often compensated to a large degree by the $\Delta\Delta G^{\text{perm}}$ term (Antosiewicz *et al.*, 1994). The final term accounts for the contribution of all the other charged groups:

$$\Delta\Delta G^{\text{tit}} = \sum_i (q_i^d \langle \varphi_i \rangle_{\text{pH},c,\Delta V}^d - q_i^p \langle \varphi_i \rangle_{\text{pH},c,\Delta V}^p), \quad (22.3.3.4)$$

where $\langle \varphi_i \rangle$ is the mean potential at group charge i from all the other titratable groups. The charge states of the other groups in the protein depend in turn on their intrinsic ' $\text{p}K_a$'s', on the external pH if they are acid-base groups, the external redox potential, ΔV , if they are redox groups and the concentration of ions, c , if they are ion-binding sites, as indicated by the subscript to $\langle \varphi_i \rangle$. Moreover, the charge state of the group itself will affect the equilibrium at the other sites. Because of this linkage, exact determination of the complete charged state of a protein is a complex procedure. If there

are N such groups, the rigorous approach is to compute the titration-state partition function by evaluating the relative electrostatic free energies of all 2^N ionization states for a given set of pH, c , ΔV . From this one may calculate the mean ionization state of any group as a function of pH, ΔV *etc.* For large N this becomes impractical, but various approximate schemes work well, including a Monte Carlo procedure (Beroza *et al.*, 1991; Yang *et al.*, 1993) or partial evaluation of the titration partition function by clustering the groups into strongly interacting sub-domains (Bashford & Karplus, 1990; Gilson, 1993; Yang *et al.*, 1993).

Calculation of ion-binding and electron-transfer equilibria in proteins proceeds exactly as for calculation of acid-base equilibria, the results usually being expressed in terms of an association constant, K_a , or a redox midpoint potential E_m (defined as the external reducing potential at which the group is half oxidized and half reduced, usually at pH 7), respectively.

22.3.3.3. Electrostatic contributions to binding energy

The electrostatic contribution to the binding energy of two molecules is obtained by taking the difference in total electrostatic energies in the bound (AB) and unbound $A + B$ states. For the linear case,

$$\Delta\Delta G_{\text{bind}}^{\text{elec}} = (1/2)\sum_i^{N_A} q_i^A (\varphi_i^{AB} - \varphi_i^A) + (1/2)\sum_j^{N_B} q_j^B (\varphi_j^{AB} - \varphi_j^B), \quad (22.3.3.5)$$

where the first and second sums are over all charges in molecule A and B , respectively, and φ^x is the total potential produced by $x = A, B$, or AB . From equation (22.3.3.5), it should be noted that the electrostatic free energy change of each molecule has contributions from intermolecular charge-charge interactions, and from changes in the solvent reaction potential of the molecule itself when solvent is displaced by the other molecule. Equation (22.3.3.5) allows for the possibility that the conformation may change upon binding, since different charge distributions may be used for the complexed and uncomplexed forms of A , and similarly for B . However, other energetic terms, including those involved in any conformational change, have to be added to equation (22.3.3.5) to obtain net binding free energy changes. Nevertheless, changes in binding free energy due to charge modifications or changes in external factors such as pH and salt concentration may be estimated using equation (22.3.3.5) alone. For the latter, salt effects are usually only significant in highly charged molecules, for which the nonlinear form for the total electrostatic energy, equation (22.3.2.4), must be used. The salt dependence of binding of drugs and proteins to DNA has been studied using this approach (Misra, Hecht *et al.*, 1994; Misra, Sharp *et al.*, 1994; Sharp *et al.*, 1995), including the pH dependence of drug binding (Misra & Honig, 1995). Other applications include the binding of sulfate to the sulfate binding protein (Aqvist *et al.*, 1991) and antibody and antigen interactions (Lee *et al.*, 1992; Slagle *et al.*, 1994).