Charge Effects on Folded and Unfolded Proteins†

Dirk Stigter and Ken A. Dill*
Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143
Received March 24, 1989; Revised Manuscript Received September 25, 1989

ABSTRACT: We develop a theory for the effects of charge on the stabilization of globular proteins. The folding process is modeled as occurring through a fictitious intermediate state along a two-part thermodynamic pathway in which the molecule (i) increases its density and then (ii) rearranges its ionic groups to the protein surface. The equilibrium for the binding of protons in salt solutions is assumed to be driven by the electrical potential due to the charge distribution, in addition to the intrinsic binding affinity and bulk proton concentration. The potential is calculated for inside and outside a porous sphere model of the protein using the Poisson–Boltzmann relation, wherein the interior dielectric constant is taken to be a linear function of the chain density. The model predicts the slope of the titration curves for native myoglobin in agreement with experiments by Breslow and Gurd (1962). From the similar experiments on the unfolded state, and from the experiments of Privalov et al. (1986) on the intrinsic viscosity of the unfolded molecules, the theory shows that the unfolded state has a much higher density than a chain in a Θ solvent and that the density increases with ionic strength. In addition, from the free energy of proton binding to the protein, we also calculate the electrostatic contributions to protein stability, a major contribution deriving from changes in ionization. We consider the example of the stability of myoglobin as a function of pH, ionic strength, and ionic groups buried in the native protein structure. We show that although maximum stability of most proteins should occur at their isoelectric point, the burial of nontitratable groups should lead to maximum stabilities at pH values other than the isoelectric point.

The stability of a protein is generally dependent upon pH. The denaturation temperature of a protein is often observed to have a broad maximum around the isoelectric pH of the molecule (Acampora & Hermans, 1967; Privalov & Khechinashvili, 1974; Privalov et al., 1986). This suggests that at the pH extremes where there is net charge on the molecule, the electrical free energy of the protein can be reduced by unfolding, a process that increases the average separation of like charges. The electrical driving force for unfolding can also be affected by ionic strength, since small ions are able to shield intrachain charge repulsions. Our purpose in this paper is to develop a theory for these electrical contributions to the stability of globular proteins.

The effects of charge are treated by using the following model. We consider the unfolded molecule to have charges randomly distributed throughout a "porous" sphere, which is otherwise filled by the solvent, a solution containing monovalent salt. The folding process is considered to follow a two-part fictitious thermodynamic pathway, chosen for mathematical convenience and previously described for modeling the thermal stabilities of proteins (Dill, 1985; Dill et al. 1989). (i) The protein molecule undergoes a density increase, whereby the ionizable groups become more concentrated and the salt solution is expelled, until the chain reaches a maximum compactness. Within the sphere, the ionizable groups remain distributed randomly. Along the radius, the degrees of ionization vary with the electrical driving forces. (ii) Then the molecule rearranges its residues to put its charges on the surface of the protein, a sphere of low dielectric constant. Both the compact spherical native protein and the porous spherical unfolded molecule are taken to be in ionization equilibrium with an aqueous solution of given pH and ionic strength, with the potential distribution given by the Poisson–Boltzmann equation. The model and the potential calculation are described in detail here.

†Supported by NIH.

Our models for the electrostatic free energies of porous and nonporous spheres are variants of models used earlier for the titration of proteins (Lindstrøm-Lang, 1924) and of polyelectrolytes (Hermans & Overbeek, 1948). The problem of the titration equilibrium is closely related to that of how electrostatic interactions affect stability (Hermans & Scheraga, 1961).[1] The present approach is developed to treat this combined problem. Various aspects of the theory are illustrated by using experimental titration data on both the folded and unfolded states of myoglobin, a well-studied example of a one-domain protein.

Characterization of the Protein Model

The unfolded protein is represented by a sphere in which protein residues and solution are mixed randomly, in ionization equilibrium with the surrounding solution of given pH and ionic strength. Along the contraction path, the radius R of this sphere decreases from R^* in the unfolded state to R_p in the randomly condensed state, as indicated in Figure 1. At the same time the protein density of the sphere, $\rho = (R_p/R)^3$, increases from the initial value $\rho^* = (R_p/R^*)^3$ to $\rho = 1$ for the fully condensed state which excludes any solution. In keeping with the earlier lattice treatment (Dill, 1985; Dill et al., 1989), the chain is divided into n segments which occupy n lattice sites, each with volume $v = (4\pi/3n)R_p^3$. In general,

[1] In this paper the electrical free energy g_{el} is derived in terms of the net charge number Z and the interaction factor w, taken empirically from titration curves as $g_{el} = kTwZ^2$ for both native and denatured protein. This approach implies that w is independent of the protein charge number Z, in agreement with the solid sphere model of the globular protein in eq 27. We note that the factor w is not exactly constant for our model of the unfolded protein. It follows from eq 30 that for the porous sphere model w depends on Z because the change of the α_i with Z depends on r and, hence, the shape of the charge distribution changes with Z. However, the resulting variations of w with Z are small for low densities, and for the ρ^* values in Figure 5, they are well within the errors inherent in the determination of w from experimental titration curves.

0006-2960/90/0429-1262$02.50/0

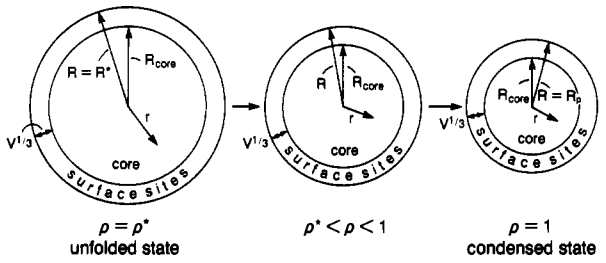

FIGURE 1

then, a sphere with density ρ has n/ρ lattice sites, n of which are filled with protein and the rest, $n(1/\rho - 1)$ sites, with salt solution. With an arrangement of such sites in concentric shells, when r is the radial coordinate, all sites in the outer shell, between $r = R$ and $r = R_{core} = R - v^{1/3}$, are counted as exterior or surface sites. This constitutes a fraction $f_e = 1 - (R_{core}/R)^3$ of the total number n/ρ of sites in the sphere. In the final state, which is globular and compact ($\rho = 1$), the $(1 - f_e)n$ interior or core sites are occupied predominantly by a fraction Φ of the hydrophobic residues in the chain.

The ionic charge on the protein derives from t different types of acid and basic groups, n_i groups of type $i = 1, 2, ..., t$ per protein molecule, each with intrinsic ionization constant k_i for proton binding. A fraction f_s of each type of these ionic groups is in surface sites, in the outer shell between $r = R_{core}$ and $r = R$; the charge deriving from these groups is treated as a uniform surface charge at $r = R$. The remaining ionizable groups, the fraction $1 - f_s$, are distributed uniformly in the core, in the volume $r < R_{core}$. Along the condensation path, the random distribution always gives $f_s = f_e$; hence, f_s is a function of the variable density ρ. Along the reconfiguration path, where $\rho = 1$, f_s is a function of the composition, Φ, and of the reorder variable θ, which is defined as the fraction of surface sites occupied by hydrophobic segments (Dill, 1985; Dill et al., 1989). The relation is

$$f_s = [(1 - \theta)/(1 - \Phi)]f_e \qquad \text{for } \rho = 1 \tag{1}$$

This completes the description of the folding pathway. We now turn to the electrostatic problem.

Self-Consistent Charge-Potential Treatment

The electrostatic free energy of a system is determined by the potential field experienced by each charge, but this potential, in turn, is generated by the charge distribution. Hence, the charge distribution and the potential field must be determined self-consistently. In the present model the potential field, ψ, is spherically symmetric and thus depends only on the radial distance r from the center of the molecule. If this field $\psi(r)$ is known, then the degree of ionization, α_i, for each type i of ionic group can be calculated as a function of the bulk proton concentration, $H = 10^{-pH}$, by using

$$q_i \ln \frac{\alpha_i(r)}{1 - \alpha_i(r)} = \ln H + \ln k_i - \frac{e\psi(r)}{kT} \qquad i = 1, 2, ..., t \tag{2}$$

where $q_i = +1$ for basic and $q_i = -1$ for acid groups, e is the protonic charge, and kT is Boltzmann's constant multiplied by absolute temperature. The first two terms on the right-hand side of eq 2 simply describe the ordinary binding equilibrium for ligands with intrinsic binding constant k_i, with independent binding sites, i.e., Langmuir-type binding. Electrostatic interactions between the ionized groups bias the binding and are taken into account in the last term on the right-hand side of eq 2, through the potential field $\psi(r)$ as noted above. This correction has been commonly used to treat the titration of proteins [for a review, see Steinhardt and Beychok (1964)] and polyelectrolytes (Katchalsky & Gillis, 1949). An alternative way to look at this potential correction is to recognize that whereas H is the bulk proton concentration, $H \exp(-e\psi/kT)$ is the proton concentration at the local binding site in the presence of the field of the other charges.

Our purpose is to calculate the α_i values. They are determined by the self-consistent theory described below. Then, when the α_i values are known, the surface charge density of the sphere is given in terms of them by

$$\sigma = \frac{f_s e}{4\pi R^2} \sum_{i=1}^{t} q_i n_i \alpha_i(R) \qquad \text{at } r = R \tag{3}$$

and the fixed charge density in the core as a function of r is

$$\rho_{el}(r) = \frac{(1 - f_s)e}{(4\pi/3)R^3_{core}} \sum_{i=1}^{t} q_i n_i \alpha_i(r) \qquad \text{for } 0 < r < R_{core} \tag{4}$$

with eq 1 for f_s.

The fixed charges of the protein are neutralized by mobile small ions surrounding the spherical folded molecule and, along the condensation path, also inside the porous spherical unfolded states of the molecule. The distributions of these charges and of the electrostatic potential are related by the Poisson–Boltzmann (PB) equation. Although this equation is not exact and has a controversial history, Monte Carlo computations have shown that it is a good approximation for distributions of small monovalent ions (Torrie & Valleau, 1979; Snook & van Megen, 1981; Linse et al., 1982; Le Bret & Zimm, 1984; Mills et al., 1984). For fields generated by typical proteins in aqueous salt solutions, the linearized form of the PB equation has been found to be a good approximation (Stigter & Dill, 1989). Outside the sphere we thus have

$$\frac{d^2\psi}{dr^2} + \frac{2}{r}\frac{d\psi}{dr} = \kappa_w^2 \psi \qquad r > R \tag{5}$$

The Debye length outside the sphere, $1/\kappa_w$, is given by

$$\kappa_w^2 = \frac{e^2}{\epsilon_0 \epsilon_w kT} \sum_s n_s z_s^2 \qquad r > R \tag{6}$$

n_s is the concentration (number density) of ions of type s with valency z_s in the bulk solution, ϵ_0 is the permittivity of vacuum, and ϵ_w is the relative permittivity or dielectric constant of the solvent around the sphere.

We make the usual assumption that the potential vanishes in the bulk solution, that is

$$\psi = 0 \qquad \text{at } r = \infty \tag{7}$$

As long as the distribution of the charge fixed to the sphere has spherical symmetry, its details play no role in determining the field in the outside solution. Only the total net charge Z_{tot} of the sphere, which includes the mobile small ions inside it, is relevant. Then the outside field, satisfying eqs 5 and 7, is the well-known solution (Debye & Hückel, 1923)

$$\psi(r) = \psi(R)\frac{Re^{-\kappa_w(r-R)}}{r} \qquad r > R \tag{8}$$

where $\psi(R)$ is the surface potential of the sphere

$$\psi(R) = \frac{Z_{tot}e}{4\pi\epsilon_0\epsilon_w R(1 + \kappa_w R)} \tag{9}$$

The potential field inside the porous sphere arises from an averaging of the interactions among the many fixed and mobile charges. Therefore, we take for the dielectric constant inside the sphere, ϵ_{sphere}, the volume average of the solution value ϵ_w and the protein value ϵ_p

$$\epsilon_{sphere} = (1 - \rho)\epsilon_w + \rho\epsilon_p \qquad r < R \tag{10}$$

Similarly, we use for the ion concentration inside the sphere the average values $(1 - \rho)n_s$. Therefore, the Debye length inside the sphere, $1/\kappa_{sphere}$, is given by

$$\kappa^2_{sphere} = \frac{(1 - \rho)\epsilon_w}{(1 - \rho)\epsilon_w + \rho\epsilon_p}\kappa_w{}^2 \tag{11}$$

which follows from replacement of n_s/ϵ_w by $(1 - \rho)n_s/\epsilon_{sphere}$ in eq 6. In the outer shell the fixed ionic charges, treated as a surface charge, are taken into account through the boundary conditions below. Therefore, the PB equation for the outer shell of the sphere is with eq 11 for κ^2_{sphere}

$$\frac{d^2\psi}{dr^2} + \frac{2}{r}\frac{d\psi}{dr} = \kappa^2_{sphere}\psi \qquad R_{core} < r < R \tag{12}$$

For the core this equation is modified by the fixed charge density $\rho_{el}(r)$ of eq 4

$$\frac{d^2\psi}{dr^2} + \frac{2}{r}\frac{d\psi}{dr} = \kappa^2_{sphere}\psi - \frac{\rho_{el}}{\epsilon_0\epsilon_{sphere}} \qquad 0 < r < R_{core} \tag{13}$$

We now discuss the boundary conditions and the solution of eqs 12 and 13. Equation 8 gives for the potential gradient at the outside surface of the sphere

$$\left(\frac{d\psi}{dr}\right)_R = -\frac{1 + \kappa_w R}{R}\psi(R) \qquad r = R^+ \tag{14}$$

At $r = R$, ψ is continuous. Gauss' law combined with eq 14, with the surface charge from eq 3, and with the change of dielectric constant from ϵ_w to ϵ_{sphere} from eq 10 yields the gradient at the surface inside the sphere

$$\left(\frac{d\psi}{dr}\right)_R = -\frac{\epsilon_w(1 + \kappa_w R)}{\epsilon_{sphere}R}\psi(R) + \frac{\sigma}{\epsilon_0\epsilon_{sphere}} \qquad r = R^- \tag{15}$$

Since at $r = R_{core}$ there is no surface charge and the dielectric constant is uniform, ψ and $d\psi/dr$ are continuous at this boundary. Finally, without a point charge at $r = 0$, symmetry requires that

$$d\psi/dr = 0 \qquad \text{for } r \to 0 \tag{16}$$

Equation 8 gives the outside field, except for the initially unknown surface potential $\psi(R)$. The potential field inside the sphere is obtained by the successive stepwise numerical integration of eqs 12 and 13 by means of a fourth-order Runge–Kutta integration routine (Rice, 1983). This computer routine is designed for the simultaneous integration of a set of coupled first-order ordinary differential equations such as

$$dy_1/dr = f_1(r,y_1,y_2) \qquad dy_2/dr = f_2(r,y_1,y_2) \tag{17}$$

The technique is used for our second-order differential equations by defining $y_1 = \psi$ and $y_2 = d\psi/dr$. This converts eq 12 into the following set of coupled equations

$$dy_1/dr = y_2 \qquad \text{for } R_{core} < r < R \tag{18}$$

$$dy_2/dr = \kappa^2_{sphere}y_1 - (2/r)y_2$$

and eq 13 becomes

$$\frac{dy_1}{dr} = y_2 \qquad \text{for } r < R_{core} \tag{19}$$

$$\frac{dy_2}{dr} = \kappa^2_{sphere}y_1 - \frac{\rho_{el}(y_1)}{\epsilon_0\epsilon_{sphere}} - \frac{2}{r}y_2$$

For each given potential $y_1 = \psi$, the charge density ρ_{el} in eq 19 is obtained from eqs 2 and 4.

For given values of R, R_{core}, f_s, n_i, q_i, H, and k_i, the self-consistent surface potential $\psi(R)$ is determined in an iteration procedure aimed at satisfying the final boundary condition given by eq 16. An integration is started at the surface of the sphere with a choice of $y_1 = \psi(R)$ and the corresponding value $y_2 = (d\psi/dr)_R$ from eq 15, using eqs 3 and 2 for σ. With such initial values $[y_1(R), y_2(R)]$ the integration of eq 18 now proceeds stepwise inward from $r = R$ to $r = R_{core}$. At this point the current values $[y_1(R_{core}), y_2(R_{core})]$ are used to initialize the integration of eq 19 which is terminated near $r = 0$ with a test of eq 16. The latter boundary condition is satisfied only for the exact, but initially unknown, starting potential $\psi(R)$. The final slope, $y_2 = (d\psi/dr)_{r\to 0}$, is positive or negative, depending on whether the integration was started with a surface potential which is too high or too low. This is used as a criterion in a bisection iteration method (Rice, 1983) to approach the value of $\psi(R)$ that satisfies eq 16 and thus determine the self-consistent potential and charge distributions in the sphere.

We describe below tests of the model against experimental data. Then in the section following, we develop a theory for the effects of charge on protein stability.

Protein Titrations

Globular Proteins. Linderstrøm-Lang (1924) was the first to apply the Debye–Hückel theory (1923) to protein titrations. The Linderstrøm-Lang theory assumes that the folded protein is a nonporous spherical Z-valent ion with radius R which excludes (the center of) small ions to a radius R_{ex}. The treatment leads to the prediction of activity coefficients. However, it also predicts an interaction factor w which more readily can be determined from experiments. The factor w is related to the electrical free energy g_{el} for charging the native protein [see, e.g., Tanford and Kirkwood (1957)] and is defined by

$$\text{pH} + q_i \log\frac{\alpha_i}{1 - \alpha_i} = -\text{p}K_i - 0.868wZ = -\text{p}K_i - \frac{0.434}{kT}\frac{\partial g_{el}}{\partial Z} \tag{20}$$

where $\text{p}K_i = -\log k_i$ and the change from natural log introduces $\log e = 0.434$. For the Linderstrøm-Lang model we have

$$g_{el} = \int_0^Z e\psi(R,Z')\,dZ' \tag{21}$$

The Debye–Hückel theory (Debye & Hückel, 1923) gives for the surface potential of the protein ion with uniform surface charge Ze

$$\psi(R,Z) = \frac{Ze}{4\pi\epsilon_0\epsilon_w R} - \frac{Ze}{4\pi\epsilon_0\epsilon_w}\frac{\kappa_w}{1 + \kappa_w R_{ex}} \tag{22}$$

The first term in eq 22 is the Coulomb potential of the charge Ze at the surface of the protein ion. The second term is the potential of the ionic atmosphere with charge $-Ze$. Substitution of eq 22 into eqs 21 and 20 yields the well-known expression for w

$$w = \frac{1}{2ZkT}\frac{\partial}{\partial Z}\int_0^Z e\psi(R,Z')\,dZ' = \frac{e\psi(R,Z)}{2ZkT} = \frac{e^2}{8\pi\epsilon_0\epsilon_w kT}\left(\frac{1}{R} - \frac{\kappa_w}{1 + \kappa_w R_{ex}}\right) \tag{23}$$

Predictions of eq 23, with $R_{ex} = R + 2.5$ Å, are often some-

Table I: Comparison of Experimental Interaction Factors w_{exp} with Theoretical Results from Equation 23 with $R_p = R_{ex} = R_0$ and, in Parentheses, with $R_p = R_{ex} = R_0 + 3$ Å

protein	R_0 (Å)	M_{salt}	w_{exp}	w_{calc}
myoglobin[a]	17.3	0.06	0.085	0.086 (0.067)
		0.16	0.050	0.063 (0.048)
bovine serum albumin[b]	26.8	0.01	0.054	0.071 (0.061)
		0.03	0.036	0.053 (0.044)
		0.08	0.028	0.038 (0.032)
		0.15	0.024	0.030 (0.025)
ribonuclease[c]	15.8	0.01	0.112	0.149 (0.117)
		0.03	0.093	0.119 (0.092)
		0.15	0.061	0.075 (0.056)
conalbumin[d]	28.1	0.01	0.046	0.066 (0.057)
		0.03	0.035	0.049 (0.041)
		0.10	0.025	0.032 (0.027)

[a] Breslow and Gurd (1962). [b] Tanford et al. (1956a). [c] Tanford et al. (1956b). [d] Wishnia et al. (1961).

Table II: Ionizable Groups of Metmyoglobin following Breslow and Gurd (1962) at 25 °C

group	n_i	pK_i	group	n_i	pK_i
ϵ-amino	19	10.60	histidine (unfolded)	12	6.48
α-amino	1	7.80	hemic acid	1	8.90
histidine (globular)	6	6.62	carboxyl	23	4.40

what too high; nevertheless, agreement with experiments is better than 20% in a number of cases (Steinhardt & Beychok, 1964; Breslow & Gurd, 1962; Wishnia et al., 1961; Tanford et al., 1956a,b).

Our model for the globular state of proteins differs from the Linderstrøm-Lang model in one principal respect. Our model is simpler because we assume $R_{ex} = R$; i.e., we do not assume a charge-free shell of thickness $R_{ex} - R$ that is impenetrable by small ions. Since, as a consequence, in our model the countercharge may approach the protein up to the surface of the fixed charges, the predicted surface potentials and w factors are somewhat lower than the predictions of the Linderstrøm-Lang model. We believe our approach is better justified since the surfaces of globular proteins are not smooth. The charged surface groups may extend somewhat further into the solution to reduce their Born energy, allowing small counterions to penetrate into the high-potential regions between the fixed charges, similar to the situation at the surface of detergent micelles (Stigter & Mysels, 1955). Therefore, our simpler model should be a better approximation for globular proteins than the Linderstrøm-Lang model. In any case, the differences in w are small.

Comparison of theoretical predictions of w with experiments requires choice of the radius $R_p = R = R_{ex}$ of the globular sphere in eq 23. In Table I we compare experimental interaction values, w_{exp}, for several globular proteins with two model results obtained as follows. Assuming a specific volume of 0.73 mL/g for globular protein with molecular weight M, a radius R_0 is calculated from $(4/3)\pi R_0^3 = 0.73M/N_{av}$. Table I shows values of w for model spheres with $R_p = R_0$ and, in parentheses, for $R_p = R_0 + 3$ Å. The data show that there is satisfactory agreement between theory and experiment for an assumed radius R_p of the model sphere which is not much greater than R_0.

Tables II and III show the experimental results on metmyoglobin reported by Breslow and Gurd (1962). The experimental w values in Table III were derived from the titration of histidines in the globular protein and from the titration of carboxyls in the unfolded state. For the globular protein w_{exp} agrees quite well with the model calculation, with $R = R_{ex} =$ 18 Å in eq 23, as shown by the comparison in Table III. The results in this section show that the model gives satisfactory agreement with experiments on the titration of representative one-domain globular proteins.

Table III: Titrations of Metmyoglobin

	globular		unfolded	
M_{salt}	w_{exp}	w_{calc}	w_{exp}	ρ^*
0.16	0.050	0.059	0.034	0.51
0.06	0.085	0.081	0.044	0.39

Denatured Protein. In this section, we apply the present theory to the prediction of w for unfolded proteins as a function of the molecular density ρ of the porous sphere. This is then compared with titration experiments on unfolded myoglobin. For globular proteins the application of eq 20 is simple because all charged groups are at the surface and, in our model, at the same potential $\psi(R)$. For the globular molecules the interaction correction $2wZ$ in eq 20 can be evaluated directly as the potential $e\psi/kT$ at the charged groups, as indicated in eqs 2 and 23. This is not the case for the unfolded protein where the experimentally determined factor w refers to an average value for the whole protein molecule, with its uneven distribution of charge and potential. Now w is derived most conveniently from the charge and the electrical free energy as follows.

From eqs 3 and 4, the number of fixed charges Z of the porous sphere is

$$Z = \sum_{i=1}^{t}\left[f_s q_i n_i \alpha_i(R) + \frac{(1-f_s)q_i n_i}{(4\pi/3)R^3_{core}}\int_0^{R_{core}} \alpha_i(r) 4\pi r^2\, dr\right] \quad (24)$$

where the first term in brackets gives the fixed surface charge and the second term gives the fixed charge distributed in the core. The electrical free energy g_{el} is evaluated as the work done reversibly to charge up the porous sphere to Z charges. This work does not depend on the particular choice of charging process, but only on the final state, that is, on the distribution of the fixed charge as given in eq 24, subject to the self-consistent potential distribution. In the Debye–Hückel linear approximation the work of charging equals the sum, or integral, of $(1/2)\psi\, d(Ze)$ over all fixed charge Ze, where at the surface $\psi = \psi(R)$, and in the core $\psi = \psi(r)$ varies with r. In view of eq 24 the electrical free energy of the porous sphere becomes

$$g_{el} = \sum_{i=1}^{t}\left[\frac{f_s q_i n_i \alpha_i(R) e\psi(R)}{2} + \frac{(1-f_s)q_i n_i}{(4\pi/3)R^3_{core}}\int_0^{R_{core}} \frac{\alpha_i(r) e\psi(r)}{2} 4\pi r^2\, dr\right] \quad (25)$$

After the determination of the self-consistent charge-potential distribution of the model sphere for the desired parameters by the method outlined in the previous sections, the information then is available for the evaluation of eqs 24 and 25. Since the protein charge varies with pH, we may compute w using the difference equation

$$w = \frac{1}{2ZkT}\frac{\partial g_{el}}{\partial Z} = \frac{1}{2ZkT}\frac{\Delta g_{el}/\Delta pH}{\Delta Z/\Delta pH} \quad (26)$$

Using the above procedure, with $\Delta pH = 0.02$ in eq 26, we have evaluated the interaction factor w for the myoglobin model along the condensation path, with the parameters defined in Table II, and taking for the globular molecule radius $R_p = 18$ Å, $n = 110$ lattice sites (corresponding to 153 amino

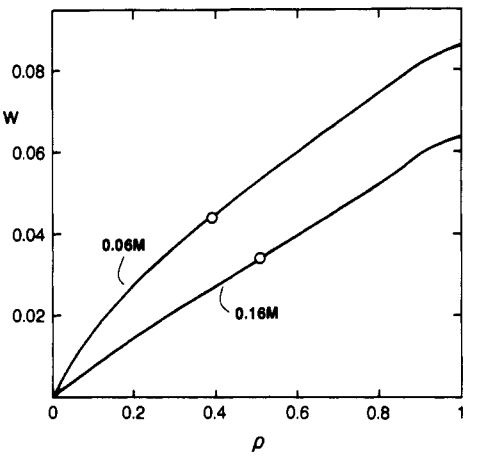

FIGURE 2: Electrostatic interaction factor w from eq 26 for titration of carboxyl groups of metmyoglobin along hypothetical contracting path, as function of density ρ of model sphere, at 25 °C and pH 3.5, ionic strength as indicated in figure. Circles indicate experimental w values from Table III.

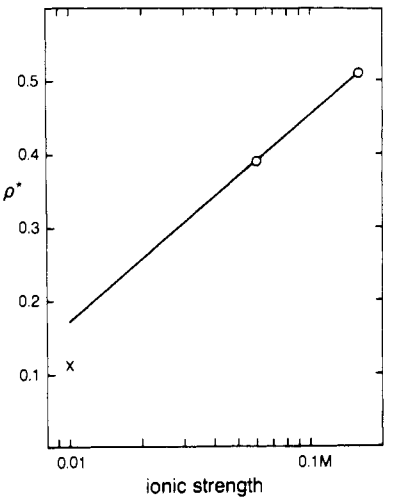

FIGURE 3: Density of unfolded metmyoglobin at 25 °C as a function of ionic strength. (Circles) From titration data, w_{exp} in Figure 2. (Cross) From viscosity data by Privalov et al. (1986).

acids), and fraction of hydrophobic residues $\Phi = 0.45$. The dielectric constant of globular proteins is not well-known. On the basis of recent theoretical work by Gilson and Honig (1986) we choose $\epsilon_p = 3.5$. The results for w as a function of the chain density ρ are plotted in Figure 2 for the two values of the ionic strength in Table III. The curves show how sensitive w is to the salt concentration and to the density ρ.

The circles in Figure 2 indicate the experimental values for the unfolded state, w_{exp} in Table III. When these experimental results for w and the known salt concentrations are substituted in the theory, Figure 2 shows that the predicted density of the unfolded state is $\rho^* = 0.51$ and $\rho^* = 0.39$ in solutions of 0.16 and 0.06 M ionic strength, respectively. This prediction, that the unfolded state has extremely high density, is consistent with predictions from a very different theoretical treatment (Dill, 1985; Dill et al., 1989). This predicted density for the unfolded state is much greater than that a polypeptide chain would have in a Θ solvent, in which the chain would obey random flight statistics. For a random flight chain, we may calculate the density as follows. The length of the fully stretched myoglobin chain of 153 residues is about $L = 153 \times 3.8 = 581$ Å. With a persistence length $P = 19$ Å for polypeptide chains (Flory, 1969), one obtains a radius of gyration of the coil $s = (LP/3)^{1/2} = 61$ Å, and an equivalent sphere radius $R = s(5/3)^{1/2} = 78$ Å, giving a protein density $\rho = (R_p/R)^3 = (18/78)^3 = 0.012$ in the sphere. This is much less than the values for the unfolded state, $\rho^* = 0.39$ and $\rho^* = 0.51$, obtained above.

Moreover, the experimental interaction factors, w_{exp}, are considerably larger than those which would be predicted for the titration of a polyelectrolyte molecule in solution, further supporting the view that the unfolded protein is relatively dense. Titration curves of synthetic linear polyelectrolytes such as poly(acrylic acid) and poly(methacrylic acid) are well explained with a model of a uniformly charged cylinder (Kotin & Nagasawa, 1962; Sugai & Nitta, 1973). Neglecting end effects for a cylinder with length L, radius a, and linear charge density Ze/L, the Debye–Hückel approximation for the interaction factor w of such a model is, from eq 23

$$w = \frac{1}{2Z}\frac{e\psi_s}{kT} = \frac{e^2}{4\pi\epsilon_0\epsilon_w kTL}\frac{K_0(\kappa_w a)}{\kappa_w a K_1(\kappa_w a)} \qquad (27)$$

where ψ_s is the surface potential of the cylinder (Stigter, 1975) and K_0 and K_1 are modified Bessel functions. With a radius of $a = 3.65$ Å for the unfolded myoglobin cylinder and $L = 581$ Å as before, eq 27 predicts $w = 0.014$ in 0.16 M and $w = 0.018$ in 0.06 M ionic strength solutions. These interaction factors, predicted for a polypeptide in a good solvent, are much less, by a factor of nearly 3, than the experimental values, w_{exp} in Table III, for the unfolded state. We conclude that the values of w_{exp} obtained by experiments on unfolded myoglobin are consistent only with a high unfolded-state density, similar in some regards to polymers in poor solvents.

A further test of the theory can be made by comparison with viscosity measurements of the unfolded state density. In figure 3, we make a linear extrapolation to predict the density the unfolded molecules would have in a solution at 0.01 M ionic strength, to compare with viscosity measurements which have been made at that ionic strength by Privalov et al. (1986). They have reported intrinsic viscosity data for unfolded metmyoglobin in 10 mM sodium acetate buffer solutions at various temperature and pH values. In the pH range where metmyoglobin unfolds at 25 °C, the intrinsic viscosity increases from $[\eta] = 3.2$ cm^3/g at pH 4.80 to $[\eta] = 20.5$ cm^3/g at pH 3.40. Neglecting drainage of the unfolded coil, a high estimate of its density is given by the ratio of $[\eta]$ values, $\rho^* = 3.2/20.5 = 0.156$. Partial drainage may be introduced by polymer viscosity theory, on the basis of the model of a porous sphere (Debye & Bueche, 1948; Brinkman, 1948). This lowers the density of unfolded metmyoglobin with $[\eta] = 20.5$ cm^3/g to $\rho^* = 0.125$ if we use a high estimate of the friction factor of the fully drained sphere and to $\rho^* = 0.096$ if we use a low estimate of this fraction factor, which is an uncertain parameter of the hydrodynamic model. We take the average, $\rho^* = 0.11$, as the experimental density of the unfolded state in 0.01 M salt solutions, indicated by a cross in Figure 3.

We compare the latter value with the above results for ρ^* from titration data. In Figure 2 the present model predicts a lower ρ^*, that is, an increase in coil size, for decreasing ionic strength I, in agreement with polyelectrolyte theory (Hermans & Overbeek, 1948; Stigter, 1982, 1985). The theory (Stigter, 1985) suggests that a $\rho^* - \log I$ plot is slightly convex. Therefore, results of the linear extrapolation in Figure 3 of the ρ^* values in Figure 2 to $\rho^* \simeq 0.17$ in 0.01 M ionic strength solution may be somewhat too high. Additional experiments would be helpful. In summary, the titration and viscosity data yield three approximate ρ^* values of unfolded myoglobin in different ionic strength solutions, to be compared later with predictions of the protein folding theory.

The model for folded and unfolded states is quite simple, but its surprisingly good performance suggests that whatever

errors it has may compensate each other. Corrections can be made for (i) the self-energy of the smeared charges representing discrete ionized groups, (ii) a nonrandom distribution of ionic groups at the surface of the protein sphere, and (iii) a spread of intrinsic pK values in each class of ionizable groups. Some of these considerations have been addressed in more recent work reviewed, e.g., by Matthew (1985), including the relation between charge site geometry and dielectric value (Tanford & Kirkwood, 1957), and the relation between solvent accessibility and effective pK values (Matthew et al., 1978). For most proteins it is found that only a few percent of the ionizable groups exhibit abnormal pK values because they are inaccessible to solvent (Rashin & Honig, 1984). Of particular interest is the recent work by Gilson and Honig (1988), who, for arbitrary geometry and distributions of discrete fixed charges, solve the Poisson–Boltzmann equation by a finite difference method.

It follows from the comparison of theory and experiment above, for the titration behavior of folded and unfolded proteins, that the present approach should also provide a satisfactory foundation for treatment of charge effects on protein stability, undertaken below. The prediction of protein titration behavior is probably more demanding of an electrostatics theory than the prediction of protein stability. Titration theory can be tested on native and unfolded protein separately, but it is only differences in free energy that enter the theory of protein stability. These differences may be considerably averaged, because of the wide variation in the many different environments of the ionizable groups in the ensemble of unfolded states. This further suggests that the refinements indicated above may be of less importance for stability than for titrations.

Free Energy of Charged Proteins

In the remainder of the paper, we extend the theory above to treatment of the electrostatic contributions to protein stability, the free energy of the folded state minus the free energy of the unfolded state. In an earlier paper (Stigter & Dill, 1989) we compared two ways to evaluate the free energy of charged colloids, by a thermodynamic charging process and by using the binding polynomial. We developed both methods for colloidal particles on which all fixed charges are at the same electrostatic potential. This uniformity of the electrostatic potential applies to our present model of the folded protein, since it is characterized here as a nonporous sphere with a uniform surface charge. We found that the results of both methods, although in different analytical form, are numerically not significantly different. It is difficult to apply the binding polynomial method rigorously to systems in which the potential of the fixed charges is nonuniform. On the other hand, application of the thermodynamic charging method remains straightforward. Therefore, we use here the charging process for the porous sphere model along the folding pathway.

We consider a solution of colloid spheres which each carry t types of acid and basic groups, n_i groups of type i = 1, 2, ..., t. Using the charging process, we derived Δg_b, the proton binding free energy per molecular sphere, relative to its unionized state. When all ionic groups are at the surface, $r = R$, the earlier treatment (Stigter & Dill, 1989) yields

$$\Delta g_b = kT \sum_{i=1}^{t} n_i \ln (1 - \alpha_i) + g_{dl} \qquad (28)$$

where the ionization degrees α_i satisfy the relevant eqs 2 for $r = R$. The *total* free energy of the electrical double layer in eq 28, g_{dl}, and the *electrical* free energy of the electrical double layer in eq 20, g_{el}, are related by (Verwey & Overbeek, 1948)

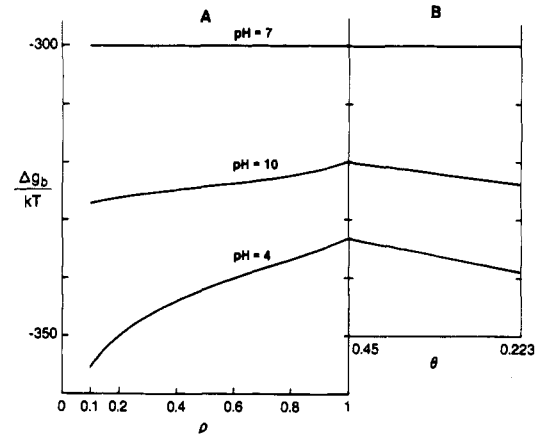

FIGURE 4: Proton binding free energy of metmyoglobin with unfolded ionic groups of Table II along hypothetical folding pathway in aqueous solutions of 0.01 M ionic strength at 25 °C for pH values as indicated. (A) Contracting path with variable density ρ. (B) Reconfiguration path with variable reorder parameter θ.

$$g_{dl} = -\int_0^{\psi(R)} Z(\psi') e d\psi'(R) = -Ze\psi(R,Z) + \int_0^Z e\psi(R,Z')\, dZ' = -Ze\psi(R,Z) + g_{el} \qquad (29)$$

In the Debye–Hückel approximation, when Z and $\psi(R)$ are linearly related, we have

$$g_{dl} = -g_{el} = -1/2 Ze\psi(Z,R) = -\frac{e\psi(R)}{2} \sum_{i=1}^{t} q_i n_i \alpha_i(R) \qquad (30)$$

In a study of the charge effects on the stability of ribonuclease, Hermans and Scheraga (1961) derived eq 28 for a single class i of ionizable groups on the protein.

We now turn to the porous sphere model, with a fraction f_s of the ionic groups at the surface and the remainder distributed in the core. Similar to eq 28, we have contributions from the surface and now also from each spherical shell with thickness dr in the core. Integrating over the core, we now obtain for all groups of the porous sphere

$$\Delta g_b = g_\alpha + g_{dl} \qquad (31)$$

where

$$g_\alpha = kT \sum_{i=1}^{t} \left[f_s n_i \ln (1 - \alpha_i(R)) + \frac{(1 - f_s) n_i}{(4\pi/3) R^3_{core}} \int_0^{R_{core}} \ln (1 - \alpha_i(r)) 4\pi r^2\, dr \right] \qquad (32)$$

which largely arises from the translational entropy of the protons on the ionizable groups and, using the Debye–Hückel approximation as in eq 30, $g_{dl} = -g_{el}$ is now given by eq 25 for the porous sphere.

In Figure 4 the proton binding free energy of metmyoglobin in solutions of 0.01 M ionic strength at 25 °C is plotted for three pH values along the hypothetical folding pathway. Panel A shows $\Delta g_b/kT$ along the condensation path, when the proton density of the model sphere increases from $\rho = 0.1$ to $\rho = 1$. In panel B, for the reconfiguration path at $\rho = 1$, the reorder parameter θ decreases from $\theta = \Phi = 0.45$ for random distribution to its minimum value, reached for $f_s = 1$ in eq 1, when all ionic groups are at the surface of the sphere.

Figure 4 shows that at pH 7, very near the isoelectric point of myoglobin, Δg_b is essentially identical for the folded and unfolded states. This is because $\psi \simeq 0$ in the model sphere

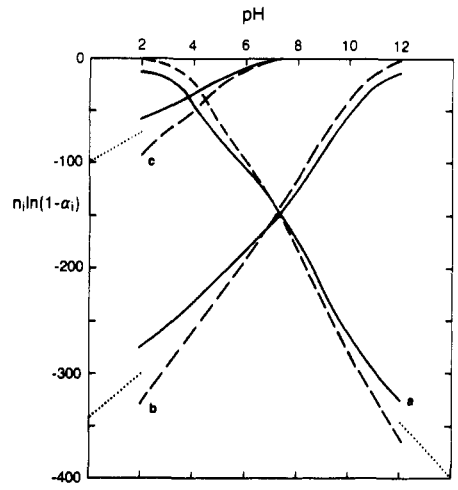

FIGURE 5: Proton binding entropy of various groups in metmyoglobin with unfolded groups of Table II as function of pH in aqueous solutions of 0.01 M ionic strength at 25 °C. (Solid curves) Globular state. (Dashed curves) Unfolded state with $\rho = 0.1$. (Dotted lines) Limiting slope at extreme pH. (Curves a) Carboxylic acid, $n_i = 23$. (Curves b) ϵ-amino, $n_i = 19$. (Curves c) Histidine, $n_i = 12$.

along the entire folding path and, hence, the α_i values in eqs 2, 25, and 32 are independent of configuration. Although there is no electrostatic interaction, $g_{dl} = -g_{el} = 0$, $\Delta g_b = g_\alpha$ has a large negative value relative to the hypothetical reference state in which α_i for 0 for all acid and basic groups. For pH 4 and 10, the free energy becomes large for the collapsed disordered "intermediate", $\rho = 1$ and $\theta = 0.45$. In this hypothetical intermediate state the average distance between the ionic charges is at a minimum; hence, their repulsion is maximal and, as a result, the Δg_b curves exhibit a maximum.

We now turn to the dependence of Δg_b on pH, already discernible in Figure 4. In Figure 5 the separate contributions for the different values of i of $n_i \ln(1 - \alpha_i)$ to g_α/kT are shown as a function of pH for the three main types of ionic groups of myoglobin, ϵ-amino, imidazole, and carboxylic acid, as solid curves for the globular molecule and as dashed curves for the unfolded state with density $\rho = 0.1$, in solutions of 0.01 M ionic strength. The general shape of the curves is as expected from ionization behavior. For example, at low pH the acid groups tend to be un-ionized, $\alpha_i \to 0$, and the basic groups tend to full ionization, $\alpha_i \to 1$. Un-ionized groups do not contribute to g_α since $\ln(1 - \alpha_i) \to \ln 1 = 0$. For nearly full ionization, the protein charge is constant and, hence, the potential is also constant with further change in pH. In that case at the pH extremes $\alpha_i \to 1$ and

$$\ln(1 - \alpha_i) = \ln \alpha_i - q_i(\ln H + \ln k_i - (e\psi/kT)) = (q_i \ln 10)\text{pH} + \text{constant} \quad (33)$$

Equation 33 shows that at low pH $n_i \ln(1 - \alpha_i)$ for the basic groups, with $q_i = 1$, have a positive limiting slope $2.303n_i$ (see curves b and c, Figure 5). Similarly, curves a for the acid groups, with $q_i = -1$, have a negative limiting slope, $-2.303n_i$, at high pH. The limiting slopes, indicated in Figure 5 by dotted lines, are outside the pH range of the calculations for ionic strength 0.01 M.

The shape of the curves in Figure 5 is influenced by the interaction between the groups; this is measured by the electrostatic potential ψ in eq 2. At low pH, ψ is on average more positive for the globular than for the unfolded protein since there is higher charge density in the globular state. Therefore, at low pH, the acid groups are more ionized and the basic groups are less ionized in the globular than in the unfolded state. The opposite holds at high pH, where ψ is more negative

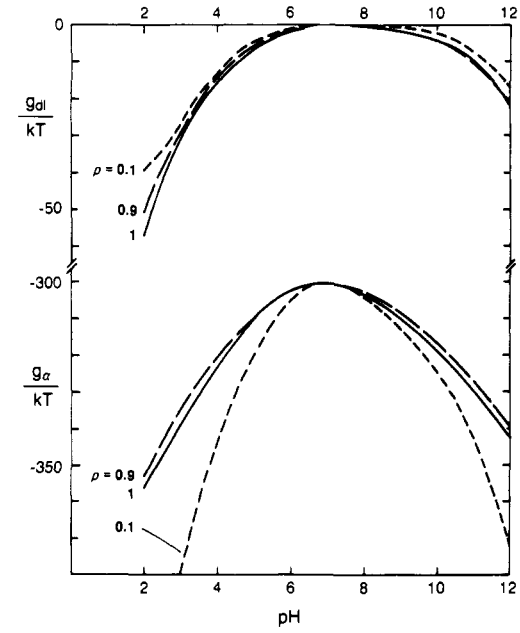

FIGURE 6: Proton binding entropy g_α and free energy of electrical double layer g_{dl} of metmyoglobin with unfolded charges of Table II as a function of pH in aqueous solutions of ionic strength 0.01 M at 25 °C. (Solid curves) Globular state. (Dashed curves) Unfolded states with density ρ as indicated.

for the folded than for the unfolded state. In between, at the IEP near pH 7, there are no interactions because $\psi = 0$ everywhere and, hence, the two sets of curves cross at the IEP.

Figure 6 shows g_α and g_{dl} as a function of pH, for the globular myoglobin as solid curves and as dashed curves along the unfolding pathway, for $\rho = 0.9$ and 0.1. The general shape of the curves is typical for mixed basic and acid polyelectrolytes, with a maximum at the IEP. At this maximum the free energy of the electrical double layer, g_{dl}, vanishes because the molecule has no net charge. For charged proteins, on either side of the IEP, g_{dl} is negative because formation of the electrical double layer is a spontaneous process (compare also eq 29). At the IEP the acid groups are still partly ionized, as are the basic groups. Therefore, at the IEP, the entropic contribution to the binding free energy, g_α, is large and negative with respect to the hypothetical standard state of the protein in which all acid and basic groups are un-ionized (compare Figure 5). The net change of g_α upon unfolding is small due to near cancellation of the large contributions. The same is true for the change of g_{dl} along the unfolding path. It is interesting that the variations with pH of g_α and g_{dl} are of the same order of magnitude, as are the changes of g_α and g_{dl} along the unfolding path, although here the changes may be in opposite directions, resulting in a smaller change of the sum Δg_b.

Electrostatic Contribution to Free Energy of Folding

Protein stability is a balance of various free energy contributions. If, in addition, ligand molecules are present which can bind to the protein, they will lead to another contribution to the free energy. The ligands of interest in this case are protons. If the free energy of binding ligand to one state of the protein differs from the free energy of binding to another state, then the ligand binding will affect the relative stabilities to the same degree. That is, the free energy difference in stability of the folded state relative to any particular unfolded state (ρ,θ) due to proton binding is

$$\delta g_r(\rho,\theta) = \Delta g_b(\text{folded}) - \Delta g_b(\rho,\theta) \quad (34)$$

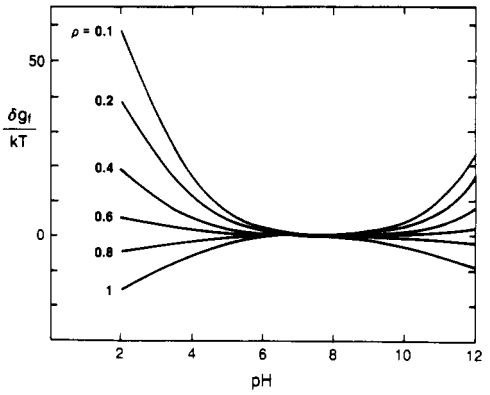

FIGURE 7: Proton binding contribution to folding free energy of metmyoglobin as function of pH along contracting path, with density ρ as indicated, in aqueous solutions of 0.01 M ionic strength at 25 °C. Unfolded ionic groups of Table II also for globular reference state.

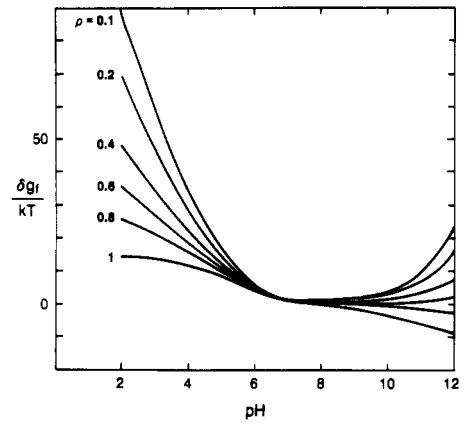

FIGURE 8: Proton binding contribution to folding free energy for systems of Figure 7 with 6, instead of 12, titratable histidines in globular reference state (see Table II).

Results of $\delta g_f/kT$ versus pH for the system under discussion are shown in Figure 7 for various ρ values along the contracting path. Whereas, at the extremes of pH, the predominant driving forces for binding are due to the binding entropies in g_α (as noted above and see Figure 6), the binding entropies nearly cancel in taking the difference to compute the relative stability (eq 34). From the thermodynamic relation [Hermans & Scheraga, 1961; see also eqs 4 and 6 of Stigter and Dill (1989)]

$$(\partial \Delta g_b/kT)/\ \partial \ln H = -Z \qquad (35)$$

it follows that the slope of the curves in Figure 7 is related to the charge difference δZ between the folded and the unfolded (ρ,θ) state according to

$$(\partial \delta g_f/kT)/\partial \mathrm{pH} = (\ln 10)\delta Z \qquad (36)$$

In a test of the self-consistency and accuracy of our computations we found that for the systems of Figure 7, where δZ ranges about from −7 to +7 protonic charges, eq 36 was always satisfied to better than 0.01 charge in δZ.

In many cases ionic groups are buried inside the structure of a native protein. For example, in metmyoglobin 6 of the 12 histidines are not titrated in the native form (see Table II). Consequently, these six un-ionized histidines do not contribute to the Δg_b of the folded state. This is readily taken into account in the reference state of δg_f. Figure 8 shows these results for δg_f, on the basis of a reference globular state in eq 34 with 6 instead of 12 histidines. Comparison with Figure 7 reveals that, as expected, the change of reference state does not modify δg_f at high pH where all histidines are un-ionized even in contact with the solution. However, the change at low pH is substantial. Here the increased proton binding capacity, relative to the folded state, increases δg_f by $2.303kT$ per pH unit for every extra proton per molecule bound by the unfolded protein [compare Stigter and Dill (1989)]. The overall influence of ionic groups buried in native proteins is that the stability versus pH curve of the protein becomes more asymmetric.

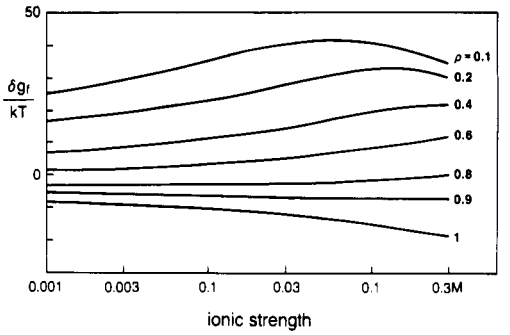

FIGURE 9: Proton binding contribution to folding free energy of metmyoglobin as function of ionic strength along contracting path, with density ρ as indicated, in aqueous solutions of pH 3 at 25 °C. Unfolded ionic groups of Table II also for globular reference state.

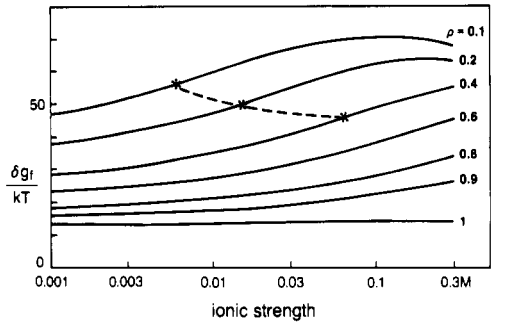

FIGURE 10: Proton binding contribution to folding free energy for systems of Figure 9 with 6, instead of 12, titratable histidines in globular reference state (see Table II). Stars indicate ρ^* values from Figure 3.

We now consider the dependence of the charge effects in protein folding on ionic strength for the systems of Figures 7 and 8. The influence of ionic strength on the binding free energy Δg_b along the folding path, and of the globular state, is quite large. However, the difference δg_f changes little with ionic strength. In Figure 9 δg_f is plotted versus ionic strength in solutions of pH 3, for various densities ρ of myoglobin with the unfolded set of ionic groups of Table II. A major factor in this dependence is the average distance between the protein charges in the two states, relative to the Debye length. We now account for the buried ionic groups in globular myoglobin. Figure 10 is for the same system as Figure 9, but now with 6, instead of 12, histidines in the folded reference state of δg_f in eq 34. The main effect is a shift of all curves to more positive values. In general, the δg_f curves in Figure 10 rise with increasing ionic strength. It is inappropriate to interpret this as implying that the folded protein becomes less stable upon the addition of salt. As discussed earlier, the equilibrium density ρ^* of the unfolded state increases when salt is added. Such ρ^* values, taken from the straight line in Figure 3 without consideration of the possible influence of pH differences, are marked as stars in Figure 10. The connecting curve decreases with increasing ionic strength, indicating that the folded state is favored by the addition of salt. This illustrates the importance of proper accounting for environmental changes on ρ^* and the unfolded states.

Born Energy and Other Factors

We assume, in eq 10, that our spherical protein has a lower dielectric constant than the surrounding aqueous solution. It is, therefore, appropriate to consider a correction which accounts for the fact that the Born energy of an ionized group in the core of the model sphere is increased relative to the energy of the same ionized group at the surface of the sphere or in the surrounding solution. Such a correction changes the intrinsic binding constant k_i in eq 2 for the ionic groups in the core. Therefore, we have repeated all of the above calculations with a model that now also includes an additional term to account approximately for the Born energy. We find that predictions for the folded state, and unfolded states with density $\rho^* \leq 0.85$, are little affected by this additional term. Born free energies are, of course, maximal for ionic groups buried in the native, globular state (Rashin & Honig, 1984). The consequence is that in most cases they are not ionized at all. Therefore, such buried groups can be dealt with satisfactorily by simply discounting them in the proton binding calculation of the globular state, as demonstrated in Figure 8. In summary, since we believe this approach is less arbitrary than using simplified models of the Born energy, we have preferred it for the protein stability calculations.

Consistent with the omission of the Born energy of the buried groups, we have used the same intrinsic pK values in the folded and unfolded state for the remaining ionic groups. This is not an essential feature of the model and can be readily modified.

The present model only attempts to treat the nonspecific effects of charge repulsion on protein stability. The smeared-charge treatment we have used does not address effects of specific interactions, such as salt bridges. It may, however, serve as a base-line model, to which electrostatic free energies of specific interactions might be added, by summing interactions of charges at specific sites in the folded protein [by use of methods described by Friend and Gurd (1979), for example]. A general type of interaction, neglected here, is the attraction between ions of opposite charge which will lead to a negative electrostatic contribution to the free energy. Calculations of Friend and Gurd (1979) have shown that these interactions might lead to considerable stabilization.

Conclusions

We have presented a model for the effects of proton binding equilibria on folded and unfolded states of globular protein molecules. Our purpose has been to provide a framework in which folded and unfolded states can be treated on equal footing. Although several methods currently exist for treating specific electrostatic interactions in folded proteins (Friend & Gurd, 1979; Weiner et al., 1984; Gilson & Honig, 1988), no corresponding method has been available to predict effects of electrostatics on the unfolded state. Therefore, theory has been unavailable for calculating electrostatic effects on protein stability, the difference of free energies between folded and unfolded states. The present treatment was developed to address that need. It is a first approximation which accounts only for the nonspecific effects of charge repulsions. Our folded-state model is a simplification of the 1924 model of Linderstrøm-Lang. The folded molecule is considered to be a sphere with a uniform surface charge surrounded by salt solution, as in the earlier model, but without an exclusion shell for the charge of the small ions which is now allowed up to the charged protein surface. The charge effects are modeled by using the Poisson–Boltzmann equation. The model gives somewhat better predictions than the Linderstrøm-Lang model for the pH titration curves of globular proteins. The unfolded state of the protein is modeled as a porous sphere of low dielectric constant, penetrated by salt solution. We study the range of possible densities of the unfolded molecules. Using this model for the unfolded state to interpret several different experimental results lead to the conclusion that the density of the unfolded state is highly dependent upon ionic strength and is often quite high. Finally, we consider the effects of the proton binding equilibria and ionic strength on protein stability. Because there is no net charge at the isoelectric point, the predicted electrostatic contribution to stability is negligible at that pH. However, at the extremes of pH, charge effects can lead to destabilization of the folded state, relative to the unfolded molecule, by several tens of kT. Moreover, when there are buried nontitratable ionic groups, then the electrostatic contribution to destabilization becomes asymmetric, and the maximum stability of the protein is not at the isoelectric pH.

Acknowledgments

We are indebted to the Department of Materials Science and Mineral Engineering, University of California, Berkeley, for hospitality to D.S.

Registry No. H^+, 12408-02-5.

References

Acampora, G., & Hermans, J., Jr. (1967) *J. Am. Chem. Soc. 89*, 1543–1547.

Breslow, E., & Gurd, F. R. N. (1962) *J. Biol. Chem. 237*, 371–381.

Brinkman, H. C. (1948) *Proc. Acad. Sci., Amsterdam 50*, 618, 821.

Debye, P., & Hückel, E. (1923) *Phys. Z. 24*, 185.

Debye, P., & Bueche, A. M. (1948) *J. Chem. Phys. 16*, 573–579.

Dill, K. A. (1985) *Biochemistry 24*, 1501–1509.

Dill, K. A., Alonso, D. O. V., & Hutchinson, K. (1989) *Biochemistry 28*, 5439–5449.

Flory, P. (1969) *Statistical Mechanics of Chain Molecules*, pp 42, 111, Interscience, New York.

Friend, S. H., & Gurd, F. R. N. (1979) *Biochemistry 18*, 4612–4619.

Gilson, M. K., & Honig, B. H. (1986) *Biopolymers 25*, 2097–2119.

Gilson, M. K., & Honig, B. (1988) *Proteins: Struct. Funct., Genetics 4*, 7–18.

Hermans, J. J., & Overbeek, J. Th. G. (1948) *Recl. Trav. Chim. Pays-Bas 67*, 761.

Hermans, J., Jr., & Scheraga, H. A. (1961) *J. Am. Chem. Soc. 83*, 3283.

Katchalsky, A., & Gillis, J. (1949) *Recl. Trav. Chim. Pays-Bas 68*, 879.

Kotin, L., & Nagasawa, N. (1962) *J. Chem. Phys. 36*, 873–879.

Le Bret, M., & Zimm, B. H. (1984) *Biopolymers 23*, 271–286.

Linderstrøm-Lang, K. (1924) *C. R. Trav. Lab. Carlsberg 15*, No. 7.

Linse, P., Gunnarson, G., & Jonsson, B. (1982) *J. Phys. Chem. 86*, 413–421.

Matthew, J. B. (1985) *Annu. Rev. Biophys. Chem. 14*, 387–417.

Matthew, J. B., Friend, S. H., Botelho, L. H., Lehman, L. D., Hanania, G. I. H., & Gurd, F. R. N. (1978) *Biochem. Biophys. Res. Commun. 81*, 416.

Mills, P., Anderson, C. F., & Record, M. T., Jr. (1985) *J. Phys. Chem. 89*, 3984–3994.

Privalov, P. L., & Khechinashvili, N. N. (1974) *J. Mol. Biol. 86*, 665-684.
Privalov, P. L., Griko, Yu. V., Venyaminov, S. Yu., & Kutyshenko, V. P. (1986) *J. Mol. Biol. 190*, 487-498.
Rashin, A., & Honig, B. (1984) *J. Mol. Biol. 173*, 515-521.
Rice, J. R. (1983) *Numerical Methods, Software and Analysis*, McGraw-Hill, New York.
Snook, I., & van Megen, W. (1981) *J. Chem. Phys. 75*, 4104.
Steinhardt, J., & Beychok, S. (1964) *Proteins (2nd Ed.)*, Chapter 8.
Stigter, D. (1975) *J. Colloid Interface Sci. 53*, 296-305.
Stigter, D. (1982) *Macromolecules 15*, 635-641.
Stigter, D. (1985) *Macromolecules 18*, 1619-1627.
Stigter, D., & Mysels, K. J. (1955) *J. Phys. Chem. 59*, 45-51.
Stigter, D., & Dill, K. A. (1989) *J. Phys. Chem. 93*, 6737-6743.
Sugai, S., & Nitta, K. (1973) *Biopolymers 12*, 1363-1376.
Tanford, C., & Kirkwood, J. G. (1957) *J. Am. Chem. Soc. 79*, 5333-5339.
Tanford, C., Swanson, S. A., & Shore, W. S. (1956a) *J. Am. Chem. Soc. 77*, 6414-6421.
Tanford, C., Hauenstein, J. D., & Rands, D. G. (1956b) *J. Am. Chem. Soc. 77*, 6409-6413.
Torrie, G. M., & Valleau, P. (1979) *Chem. Phys. Lett. 65*, 343-346.
Verwey, E. J. W., & Overbeek, J. Th. G. (1948) *Theory of the Stability of Lyophobic Colloids*, Elsevier, New York.
Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., & Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765-784.
Wishnia, A., Weber, I., & Warner, R. C. (1961) *J. Am. Chem. Soc. 83*, 2071-2081.

Tyrosine-96 as a Natural Spectroscopic Probe of the Cytochrome P-450$_{cam}$ Active Site

William M. Atkins and Stephen G. Sligar*
Departments of Biochemistry and Chemistry, University of Illinois, Urbana, Illinois 61801
Received May 10, 1989; Revised Manuscript Received August 30, 1989

ABSTRACT: The previously described correlation between the ferric spin equilibrium of cytochrome P-450$_{cam}$ and the environmental polarity of tyrosine residues (Fisher et al., 1986) has been further examined with the use of site-directed mutagenesis and active-site affinity reagents. Whereas the wild-type demonstrates an increase in environmental polarity of approximately one tyrosine residue, the mutant protein Y96F, in which Tyr-96 has been changed to Phe-96, demonstrates a lack of spin-state-dependent change in the second-derivative ultraviolet absorption spectrum. This suggests that the active-site Tyr-96 serves as a ultraviolet spectroscopic probe which can be utilized to determine the relative degree of water access to the active site for various substrate/protein complexes. The affinity reagent isobornyl mercaptan has been used to demonstrate the utility of this probe in determining the active-site polarity when substrate analogues are bound at the active site. In addition, the sensitivity of Tyr-96 to environmental polarity has been used to demonstrate that the product/enzyme complex, formed with 5-*exo*-hydroxycamphor, may be associated with increased water access to the heme iron. This may provide a means for turning off electron transfer when the product, instead of the substrate, is bound at the active site.

Cytochrome P-450$_{cam}$, isolated from *Pseudomonas putida*, catalyzes the regiospecific hydroxylation of the monoterpene camphor in the first committed step of the complex catabolism of this compound, thus allowing the organism to utilize camphor as a sole carbon source. Cytochrome P-450$_{cam}$ has served as the prototypical model for the chemical and physical characterization of the various oxidized, reduced, and oxygenated states of the heme center for the entire family of P-450 monooxygenases (Murray et al., 1986). This is largely due to its soluble nature, its availability in gram quantities, and the recently published set of X-ray crystal structures for the camphor-bound, substrate-free, and metyrapone-inhibited forms of the enzyme (Poulos et al., 1985, 1986, 1987, 1988).

It is well documented that camphor binding is associated with a change from the low-spin, aquo-liganded ferric enzyme to a high-spin, five-coordinate form. This change in spin state, which can be monitored spectrally by a change in the visible region of the absorbance spectrum, is accompanied by a change in reduction potential from −300 to −170 mV, allowing for electron transfer from the physiological electron-transfer partner putidaredoxin (Sligar, 1976). The observed change in reduction potential is of central importance in the regulation of catalysis by this enzyme, and structural differences between the low-spin and high-spin forms of the enzyme must be understood if the complete reaction cycle and mechanisms of substrate oxidation are to be completely documented.

The relationship between the active-site structure of P-450$_{cam}$ and the regulation of the spin state is being aided through the use of site-directed mutagenesis (Atkins & Sligar, 1988, 1989), rapid reaction kinetics (Fisher & Sligar, 1987) and equilibrium potential measurements (Fisher & Sligar, 1988), and insight gleaned from the available X-ray structures. Recently, second-derivative ultraviolet (UV)[1] spectroscopy was used as probe of the P-450$_{cam}$ active site, and it was demonstrated that the ferric spin equilibrium is coupled to the relative environmental polarity or degree of solvent exposure of approximately one tyrosine residue (Fisher et al., 1985). A linear correlation was observed between the fraction of high-spin enzyme ob-

* Address correspondence to this author at the Department of Biochemistry, University of Illinois.

[1] Abbreviations: Y96F, site-directed mutant in which Tyr-96 has been changed to Phe-96; IBM, isobornyl mercaptan; UV, ultraviolet.

0006-2960/90/0429-1271$02.50/0