# Salt Effects on Protein Titration and Binding

**Yaoqi Zhou**†

*Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138*

*Received: June 8, 1998; In Final Form: August 27, 1998*

Salt effects on protein titration and binding are investigated using a cavity-function theory for chemical association. Analytical expressions are obtained for the electrostatic interaction factor in protein titration and for the p*K* shift of the calcium-binding constant in calbindin $D_{9k}$. Although the structural details of proteins are neglected in the model, the theoretical results are in excellent agreement with experimental data.

## Introduction

The binding of a ligand to an enzyme is essential for its biological function.[1] An important binding class involves inorganic ions such as hydrogen, sodium, and calcium. The interactions of these small ions are primarily electrostatic. Thus, many charged groups on the protein can contribute to binding. The resulting complexity suggests that it would be of interest to introduce simplified models for ion binding. The simplified models are expected to be useful particularly for understanding more global effects (e.g., the effect of ionic strength on ion binding) rather than the specific interactions involved in binding at a particular site.

The purpose of this paper is to develop a simple and yet accurate analytic theory for salt effects on protein titration and binding. Understanding salt effects is of great importance because biomolecules in living systems are in solutions with ionic strengths of about 0.1 mol/L. Changing the salt concentration can cause the binding constant to change by several orders of magnitude.[2]

The first theoretical treatment of ion binding in native proteins was the work of Linderstrøm-Lang,[3] who modeled the native protein as a uniformly charged hard sphere and treated the electrostatic interaction by the Debye−Hückel (DH) approximation. More refined spherical models which introduce discrete charge distributions were initiated by Tanford and Kirkwood[4] and extended by others.[5−7] More recently, theories for atomistic models of proteins have been introduced based on the Poisson−Boltzmann (PB) equation,[8−13] a self-consistent free-energy approximation,[14] and site−site integral equations.[2,15,16] A theory for optimizing ligand−protein electrostatic interactions has also been developed.[17,18] The role of electrostatic interactions in protein stability and protein folding was discussed in several recent reviews.[19−21]

In this paper, we apply a recently developed cavity-function theory for chemical association[22] to determine the effect of solution conditions on ion binding in globular proteins. The cavity-function theory is a statistical thermodynamic theory. The zeroth-order theory combined with a "linear" approximation for the *n*th-order cavity function[23,24] yields thermodynamic results identical to those of Wertheim's first-order thermodynamic perturbation theory (TPT).[25] For ternary associations ($n = 3$), the zeroth-order approximation is found to be more accurate than first-order TPT and is comparable in accuracy to second-order TPT.[26] First-order TPT (or the zeroth-order theory combined with a "linear" approximation) is found to be reasonably accurate in many applications including chemical association of simple fluids, the equations of state for polymer chains,[22,27−30] and ionic associations[22,31] if the association occurs outside the strongly repulsive (hard or soft) cores of model reactants. Recently, the cavity-function theory has been extended to both a system in which chemical association occurs inside the hard core and systems of weak electrolytes.[32] The main advantage of cavity-function theory over other approaches, such as the site−site integral equation,[2,15] is that it is simple and analytical if an analytical approximation such as the exponential (EXP) approximation[33−35] is used.

In the present paper, the cavity function obtained with the EXP approximation is applied to calculate the effect of ionic strength on ligand binding. The method treats the coupling between the hard-core interaction and the electrostatic interaction consistently. In contrast, the DH-like approximation and the PB equation neglect the correlation between mobile ions with proteins and allow ions to overlap.[36] Quantitatively, the new analytic theory more accurately predicts the electrostatic interaction factors measured in protein-titration experiments[37] than the DH theory. It also accurately predicts the p*K* shifts in the binding of two calcium ions in the presence of potassium chloride.

## The Cavity-Function Theory

The cavity-function theory of chemical associations[22] considers an *n*-particle association (or binding) reaction

$$A + B + C + ... \rightleftharpoons ABC... \tag{1}$$

In chemical equilibrium, the chemical potential $\mu$ of the product ABC... is equal to the sum of the chemical potentials of the components A, B, C, ..., that is

$$\mu_{ABC...} = \mu_A + \mu_B + \mu_C + ... \tag{2}$$

When the chemical potentials are divided into their ideal and nonideal parts, we have

† E-mail: zhou@tammy.harvard.edu.

10.1021/jp982542x CCC: $15.00 
Published on Web 11/18/1998

$$\mu_{ABC...} = (\mu^{\circ}_{ABC...} + k_B T \ln \rho_{ABC...}) + k_B T \ln \gamma_{ABC...} \quad (3)$$

$$\mu_A = (\mu^{\circ}_A + k_B T \ln \rho_A) + k_B T \ln \gamma_A \quad (4)$$

$$\mu_B = (\mu^{\circ}_B + k_B T \ln \rho_B) + k_B T \ln \gamma_B \quad (5)$$

$$\mu_C = (\mu^{\circ}_C + k_B T \ln \rho_C) + k_B T \ln \gamma_C \quad (6)$$

and so on. The superscript ° denotes the standard-state chemical potential, $k_B$ is the Boltzmann constant, $T$ is the temperature, and $\rho_l$ and $\gamma_l$ are the number density and the activity coefficient, respectively, for species $l$. Equation 3 can be rewritten as[38]

$$K^{app} \equiv \frac{\rho_{ABC...}}{\rho_A \rho_B \rho_C \cdots} = K^{\circ} \frac{\gamma_A \gamma_B \gamma_C \cdots}{\gamma_{ABC...}} \quad (7)$$

where $K^{app}$ is the apparent equilibrium association constant and $K^{\circ}$ is the standard-state equilibrium constant which is defined by the following equation:

$$-k_B T \ln K^{\circ} \equiv \mu^{\circ}_{ABC...} - \mu^{\circ}_A - \mu^{\circ}_B - \mu^{\circ}_C - ... \quad (8)$$

If the resulting molecule ABC... is rigid, $K^{app}/K^{\circ}$ can be expressed in terms of the $n$-particle cavity function $y^{(n)}_{ABC...}(L_A, L_B, ...)$[22] as

$$K^{app}/K^{\circ} = y^{(n)}_{ABC...}(L_A, L_B, ...) \equiv \frac{\gamma_A \gamma_B \gamma_C \cdots}{\gamma_{ABC...}} \quad (9)$$

where the $n$ -particle cavity function $y^{(n)}_{ABC...}$ is evaluated at the positions $L_A$, $L_B$, ..., of A, B, ... in the product ABC... Eq 9 is the $n$-particle analogue of the equation for binary association ($n = 2$) obtained by Chandler and Pratt.[39] It provides an alternative method for calculating the apparent equilibrium constant $K^{app}$ given the value of $K^{\circ}$; i.e., the approach is not useful for the specific equilibrium constant $K^{\circ}$ but is useful for nonideal corrections. Equation 9 focuses on calculating the value of the $n$-particle cavity function instead of calculating the individual activity coefficients as in eq 7. This avoids the difficulty often encountered in calculating the activity coefficient for the molecular species ABC...

For binary association ($n = 2$), the two-particle cavity function can be obtained from integral-equation theories for radial-distribution functions.[40] The two-particle cavity function $y_{AB}(r)$ is by definition related to the radial-distribution function $g_{AB}(r)$ by

$$y_{AB}(r) = e^{\beta u_{AB}(r)} g_{AB}(r) \quad (10)$$

where $u_{AB}(r)$ is the pair interaction potential between particles A and B and $r$ is the distance between the two particles. For binding that involves more than two particles ($n > 2$), the $n$-particle cavity function is usually unknown. Thus, one has to make a further approximation. In this paper, we use the linear approximation which expresses the $n$-particle cavity function in terms of the product of two-particle cavity functions:[23,24]

$$y^{(n)}_{ABC...}(L_A, L_B, ...) = y_{AB}(L_{AB})\, y_{BC}(L_{BC}) ... \quad (11)$$

where $L_{AB} = |L_A - L_B|$, $L_{BC} = |L_B - L_C|$, etc. are the bond lengths and A and B, B and C, and so on are nearest neighbors. The linear approximation is essentially exact if the bound state is a linear molecule.[24,26] For the binding of several small ligands to a protein, the linear approximation is expected to be accurate when the binding sites are far from each other and the interactions among them are weak. This is often a reasonable assumption for global effects, although sometimes binding sites may have strong interactions (e.g., for pH titration involving neighboring residues[41]).

### Protein Titration

The general equations for the cavity-function theory are described in the previous section. Here we show how to apply the theory to the problem of protein titration and specifically to determine the effect of external salt on the titration behavior.

We consider a protein in its native conformation immersed in a salt solution at a constant pH. Native proteins normally consist of many acidic and basic groups that release or bind hydrogen ions depending on pH.[37] Because the equations for acidic and basic groups are essentially the same, we focus on the equation for acidic sites. For a site i on the protein surface, we have the equilibrium

$$P_i + H^+ \rightleftharpoons [HP_i]^+ \quad (12)$$

If the fraction of unbound site i, $P_i$, is denoted as $\alpha_i$, the fraction of site i bound to hydrogen, $HP_i$, is $1 - \alpha_i$. Through the use of the cavity-function theory, we have

$$K^{app}_i = \frac{C_{HP_i}}{C_{P_i} C_{H^+}} = \frac{1 - \alpha_i}{\alpha_i C_{H^+}} = K^{\circ}_i y_{HP_i}(L_{HP}) \quad (13)$$

where $C_{P_i}$ and $C_{HP_i}$ are the concentrations of unbound site i and bound site i, respectively, $C_{H^+}$ (the concentration of hydrogen ions) is related to pH via the equality $C_{H^+} = 10^{-pH}$, the bond-length $L_{HP}$ is the distance from the center of the bonded hydrated hydrogen to the center of the protein, $K^{app}_i$ is the apparent association constant, and $K^{\circ}_i$ is the association constant under ideal conditions.

The association constant $K^{\circ}_i$ can be expressed as the product of the nonelectrostatic (intrinsic) part $K^{intr}_i$ (ref 4) and the electrostatic part $K^{ele}_i$ which is equal to $\exp[-\beta u^{ele}_{HP}(L_{HP})]$.[40] Thus, we have

$$K^{\circ}_i = K^{intr}_i e^{-\beta u_{HP}{}^{ele}(L_{HP})} \quad (14)$$

where $u^{ele}_{HP}(L_{HP})$ is the electrostatic interaction potential between the hydrogen ion and the protein in the absence of salts at the bond length $L_{HP}$. If one assumes that the protein charge $Z_P e$ is uniformly distributed over the spherical protein surface, we have $u^{ele}_{HP}(L_{HP}) = Z_H Z_P e^2/(\epsilon L_{HP})$, where $Z_H$ (=+1) is the charge number of hydrogen ion, $e$ is the electronic charge, and $\epsilon$ is the dielectric constant. This assumption, which is quite drastic, is the same one used in Linderstrøm-Lang theory.[3] Similar to earlier work on charged-sphere models for proteins,[3,42−44] we also assume that the dielectric constant of water can be used to describe the protein-hydrogen-ion interaction because all charges are assumed to be at the surface of the protein, thus the value of the dielectric constant of the protein interior has no effect on the results. Besides, there is no experimental data for the dielectric constant in the vicinity of the protein binding sites.

Because $K^{intr}_i$ is an association constant rather than a dissociation constant, the experimental value of $pK^{intr}_i$ is equal to $\log K^{intr}_i$. Thus, eq 13 can be rewritten as follows:

$$pH - \log \frac{\alpha_i}{1 - \alpha_i} = pK^{intr}_i - 0.4343 \beta u^{ele}_{HP}(L_{HP}) + \log y_{HP}(L_{HP}) \quad (15)$$

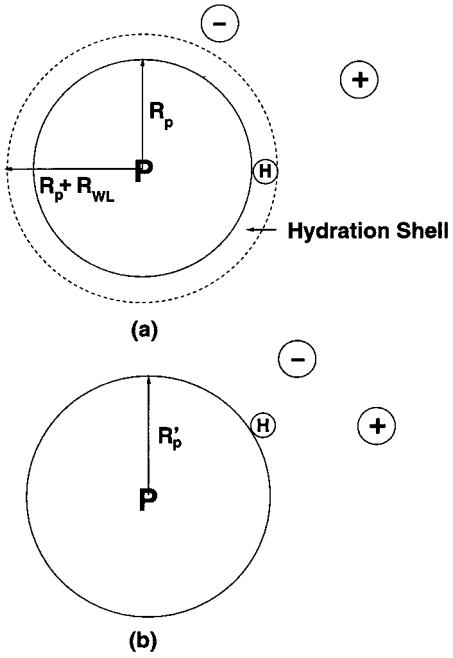


**Figure 1.** Schematic plot of the models for the globular protein (P), hydrogen ions (H), salt cations (+), and anions (−). (a) The radius of the hydrated protein is equal to the radius of the protein plus the thickness of the water layer, $R_{WL}$. Hydrogen ion binding in proteins can be thought of as replacing a water molecule in the hydration shell by a hydrated hydrogen ion ($H_3O^+$). (b) An "effectively" equivalent model. See text for details.

where $\log(e) = 0.4343$ and the sub-subscript i on the cavity function is dropped because we have assumed that charged binding sites are uniformly distributed over the protein surface. In our theory, the discrete effect of charged sites is taken into account through the intrinsic association constant $K_i^{intr}$.

Equation 15 can be further simplified by separating the hard-core contribution to the cavity function from the remainder. We shall label the remainder as the electrostatic part because this term of the cavity function is different from unity only after the electrostatic interaction is initiated. Thus, $y_{HP}^{ele}(L_{HP})$ is defined as $y_{HP}(L_{HP})/y_{HP}^{core}(L_{HP})$. On the basis of eq 10; one can also write $y_{HP}^{ele}(L_{HP}) = \exp[\beta u_{HP}^{ele}(L_{HP})]g_{HP}^{ele}(L_{HP})$, where the electrostatic part of the radial distribution-function $g_{HP}^{ele}(L_{HP})$ is similarly defined as $[g_{HP}^{ele}(L_{HP}) = g_{HP}(L_{HP})/g_{HP}^{core}(L_{HP})]$. Thus, eq 15 can be rewritten as

$$\mathrm{pH} - \log\frac{\alpha_i}{1-\alpha_i} = \mathrm{p}K_i^{intr} + \log y_{HP}^{core}(L_{HP}) + \log g_{HP}^{ele}(L_{HP}) \tag{16}$$

To perform a calculation, we have to set up a model for the protein and the hydrogen ion. However, hydrogen binding should not be thought of as a hydrogen ion in contact with the hard surface of a protein because the binding-reaction product is still the same protein but with a different charge number. To be consistent, the hydrogen ion must bind inside the protein so that the size of the newly produced protein is unchanged. The actual binding process may be described as follows. A native protein is normally enclosed by one or more layers of water as shown in Figure 1a. The binding of one hydrated hydrogen ion to the protein can be thought of as "replacing" one water molecule in the protein's first hydration shell by the hydrated hydrogen ion. Thus, one can model the hydrated protein as a charged hard sphere with a hard-sphere radius that is equal to the protein radius, $R_P$, plus the thickness of the first hydration shell, $R_{WL}$. The hydrated hydrogen ion is modeled as a charged hard sphere with a radius of a water molecule, $R_H$ ($=R_W$). As a result, the binding of hydrogen to protein occurs inside the hydrated protein "hard" core (see Figure 1a). This "internal" binding requires the calculation of the cavity function for which the hydrated hydrogen ion is inside the *hydrated* protein hard core because $L_{HP} < R_P + R_{WL}$. However, there is no accurate method to evaluate the cavity function for two ionic particles at a distance inside ionic hard cores although an approximate method has been suggested for the case of equal-sized overlapping ionic cores.[32] Here, we shall borrow an idea of Amos and Jackson, who replace the overlapping hard-sphere model with two "equivalent" tangent hard spheres with effective diameters.[45] We assume that $y_{HP}^{core}(L_{HP})\, g_{HP}^{ele}(L_{HP})$ for the overlapping hydrogen and protein model (Figure 1a) is equivalent to $y_{HP}^{core}(R'_P + R'_H)\, g_{HP}^{ele}(R'_P + R'_H)$ for a protein and a hydrogen at contact with effective radii $R'_P$ and $R'_H$, respectively (Figure 1b). The accuracy of this assumption depends on the recipe that relates $R'_P$ and $R'_H$ to $R_P$ and $R_H$. In this paper, we do not try to find such a recipe because there is no unique method for determining the radii $R_P$ and $R_H$. Rather, we take the effective hydrated protein radius $R'_P$ to be the radius of a sphere which contains one protein molecule and 20% water by weight. This choice of radius is commonly used by protein chemists in protein titration.[37] The radius takes into account approximately the first hydration shell. For convenience, we also assume that the difference between the effective radius of a hydrated hydrogen ion $R'_H$ and the radius of water is negligible; i.e., $R'_H = R_W$. As we have mentioned, we are unable to assess directly the loss in accuracy because of the replacement of two overlapping ionic spheres with two ionic spheres at contact. The overall accuracy of the theory will be judged by comparing the predictions of our theory with experimental data and by testing the sensitivity of our results on the choice of size parameters.

For the configuration of the effective-size hydrated hydrogen and the effective-size hydrated protein at contact, $\sigma_{HP} = R'_P + R'_H$ and $y_{HP}^{core}(\sigma_{HP}) = g_{HP}^{core}(\sigma_{HP})$. The latter equation is due to $\beta u_{HP}^{core}(r) = 0$ for $r \geq R'_P + R'_H$. Thus, eq 16 can be rewritten as

$$\mathrm{pH} - \log\frac{\alpha_i}{1-\alpha_i} = \mathrm{p}K_i^{intr} + \log g_{HP}(\sigma_{HP}) \tag{17}$$

where the radial-distribution function $g_{HP}(\sigma_{HP}) = g_{HP}^{core}(\sigma_{HP})\, g_{HP}^{ele}(\sigma_{HP})$.

The radial-distribution function $g_{HP}(\sigma_{HP})$ can be calculated via various integral-equation approximations such as the hypernetted chain approximation.[40] Here, we shall use the EXP approximation[33−35] as an illustrative example. The EXP approximation exponentiates the electrostatic part of the pair-correlation function obtained in the mean-spherical approximation (MSA).[46,47] It is known that the EXP approximation gives considerably better structural information for the charged hard-sphere system than does the MSA and the contact value of the radial-distribution function in the EXP approximation is very accurate at least for monovalent salts modeled by charged hard spheres.[35] In the EXP approximation, the contact value of the radial-distribution function satisfies

$$g_{HP}(\sigma_{HP}) = [g_{HP}(\sigma_{HP})]^{core} e^{[g_{HP}^{MSA}(\sigma_{HP})]^{ele}} \tag{18}$$

where the superscripts "core" and "ele" denote the hard core and electrostatic contributions to the contact value of the pair-correlation function $g_{HP}(\sigma_{HP})$, respectively. Substituting eq 18

into eq 17, we have (cf. (10)

$$\mathrm{pH} - \log\frac{\alpha_i}{1-\alpha_i} = \mathrm{p}K_i^{\mathrm{intr}} + \log[g_{\mathrm{HP}}(\sigma_{\mathrm{HP}})]^{\mathrm{core}} + 0.4343[g_{\mathrm{HP}}^{\mathrm{MSA}}(\sigma_{\mathrm{HP}})]^{\mathrm{ele}} \quad (19)$$

The hard-core contact value $[g_{\mathrm{HP}}(\sigma_{\mathrm{HP}})]^{\mathrm{core}}$ for unequal-size hard-sphere mixtures can be obtained from the MCSL equation of state:[48]

$$[g_{\mathrm{HP}}(\sigma_{\mathrm{HP}})]^{\mathrm{core}} = \frac{1}{(1-\eta)} + \frac{3\xi_2\sigma_{\mathrm{H}}\sigma_{\mathrm{P}}}{(1-\eta)^2(\sigma_{\mathrm{H}}+\sigma_{\mathrm{P}})} + \frac{2\xi_2{}^2\sigma_{\mathrm{H}}{}^2\sigma_{\mathrm{P}}{}^2}{(1-\eta)^3(\sigma_{\mathrm{H}}+\sigma_{\mathrm{P}})^2} \quad (20)$$

where $\sigma_l$ is the diameter of the charge hard-sphere species $l$, $\sigma_{\mathrm{H}} = 2R'_{\mathrm{H}}$, $\sigma_{\mathrm{P}} = 2R'_{\mathrm{P}}$, $\eta = (\pi/6)\sum\rho_l\sigma_l^3$ and $\xi_2 = (\pi/6)\sum\rho_l\sigma_l^2$. Results obtained from eq 20 are in excellent agreement with computer simulation results of hard-sphere mixtures and are more accurate than the contact values obtained in the Percus−Yevick approximation.[48] In eq 20, the units of length are angstroms (Å) and the units for number densities are $\text{Å}^{-3}$. The conversion from number density $\rho_l$ (in $\text{Å}^{-3}$) to concentration $C_l$ (in moles/liter) is $\rho_l = 6.022 \times 10^{-4}C_l$.

The expression of $[g_{HP}^{\mathrm{MSA}}(\sigma_{HP})]^{\mathrm{ele}}$ for an unequal-sized charged hard-sphere mixture is also well-known.[49,50] To simplify the expressions, we assume that the diameters of the salt cations and those of the salt anions are the same (i.e., $\sigma_+ = \sigma_- = \sigma_{\mathrm{S}}$) and that the concentration of hydrogen ions is much less than that of the salt concentration. The latter is usually true if the pH value is greater than 2. Thus, we have[49,50]

$$[g_{\mathrm{HP}}^{\mathrm{MSA}}(\sigma_{\mathrm{HP}})]^{\mathrm{ele}} = -\frac{\lambda}{\sigma_{\mathrm{HP}}}\frac{Z_{\mathrm{H}}Z_{\mathrm{P}}}{(1+\Gamma\sigma_{\mathrm{H}})(1+\Gamma\sigma_{\mathrm{P}})} \quad (21)$$

where $\lambda = \beta e^2/\epsilon$, $\Gamma = (\sqrt{1+2\kappa\sigma_{\mathrm{S}}} - 1)/2\sigma_{\mathrm{S}}$, and $\kappa$ is the inverse Debye screening length $[=\sqrt{4\pi\lambda(\rho_+Z_+{}^2+\rho_-Z_-{}^2)}]$.

Experimentally, one plots $\mathrm{pH} - \log[\alpha_i/(1-\alpha_i)]$ versus $Z_{\mathrm{p}}$ to determine the effective $[\mathrm{p}K_i]^{\mathrm{eff}}$ and the electrostatic interaction factor $\omega$[37]

$$\mathrm{pH} - \log\frac{\alpha_i}{1-\alpha_i} = [\mathrm{p}K_i]^{\mathrm{eff}} - 0.8686\omega Z_{\mathrm{p}} \quad (22)$$

From eqs 19 and 21, we have

$$[\mathrm{p}K_i]^{\mathrm{eff}} = \mathrm{p}K_i^{\mathrm{intr}} + \log[g_{\mathrm{HP}}(\sigma_{\mathrm{HP}})]^{\mathrm{core}} \quad (23)$$

and

$$\omega = \frac{\lambda}{2(R'_{\mathrm{H}}+R'_{\mathrm{P}})(1+2\Gamma R'_{\mathrm{H}})(1+2\Gamma R'_{\mathrm{P}})} \quad (24)$$

Because the hard-core contact value $[g_{\mathrm{HP}}(\sigma_{\mathrm{HP}})]^{\mathrm{core}}$ increases as the salt concentration increases,[48] eq 23 predicts a larger effective $[\mathrm{p}K_i]^{\mathrm{eff}}$ for a higher salt concentration. Such changes are normally smaller than experimental error because the salt concentration in experiments is typically less than 0.1 mol/L (or $6.022 \times 10^{-5}$, $\text{Å}^{-3}$). For higher salt concentrations, however, the hard-core contribution can be significant. For example, 1.2 mol/L KCl solution (modeled by charged hard spheres of 5 Å in diameter, see below) is enough to increase the $\mathrm{p}K_i$ by 0.1 units.

As the salt concentration approaches zero, $\Gamma \to \kappa/2$. Neglecting the $\kappa^2$ term in eq 24, we have

$$\omega = \frac{\lambda}{2}\left[\frac{1}{(R'_{\mathrm{H}}+R'_{\mathrm{P}})} - \frac{\kappa}{1+\kappa(R'_{\mathrm{H}}+\mathrm{R}'_{\mathrm{P}})}\right] \quad (25)$$

This equation is similar but not equivalent to the well-known expression obtained in the Debye−Hückel approximation[3,37]

$$\omega^{\mathrm{DH}} = \frac{\lambda}{2}\left[\frac{1}{\mathrm{R}'_{\mathrm{P}}} - \frac{\kappa}{1+\kappa(R_{\mathrm{S}}+R'_{\mathrm{P}})}\right] \quad (26)$$

where $R_{\mathrm{S}}$ $(=\sigma_{\mathrm{S}}/2)$ is the radius of the salt cations or anions. In other words, the low-concentration limit of eq 24 does not recover the equation obtained in the DH approximation.

Stigter and Dill[43] suggested an empirical equation by letting $R_{\mathrm{S}} = 0$ in eq 26,

$$\omega^{\mathrm{SD}} = \frac{\lambda}{2}\left[\frac{1}{\mathrm{R}'_{\mathrm{P}}} - \frac{\kappa}{1+\kappa\mathrm{R}'_{\mathrm{P}}}\right] \quad (27)$$

They obtained this equation by arguing that ions can penetrate further inside proteins because of imperfection in the protein surface. In fact, their equation can be obtained naturally as the $R'_{\mathrm{H}} \to 0$ limit of eq 25. Thus, the improvement of eq 27 over the DH equation (eq 26) can be explained as a consequence of more consistent handling of the coupling of hard-core and electrostatic interactions inherited from the EXP approximation. However, an obvious defect in eq 27 is the implication that the electrostatic interaction factor $\omega$ is independent of the size of salt. This is the direct consequence of taking the low-concentration limit to yield eq 25.

Table 1 compares the theoretical electrostatic interaction factors $\omega$ with the experimentally measured ones for several globular proteins. The radius of each hydrated globular protein, $R'_{\mathrm{P}}$, is estimated from the volume occupied by a single protein based on its molecular weight and specific volume plus the volume occupied by water, assuming that the model sphere consists of 20% water by weight.[51−54] The radius of a hydrogen ion is 1.5 Å, the approximate radius of a water molecule.[55] Similar values for the radius of a hydrogen ion are employed in earlier works.[56−58] The radii of $K^+$ and $Cl^-$ are taken to be 2.5 Å, the same as those used in earlier protein-titration studies[37] to facilitate comparison. The dielectric constant of water is 78.7 at $T = 25$ °C. As has been found earlier,[37] the prediction of the DH approximation for $\omega$ (eq 26) is somewhat too high. Stigter and Dill's empirical recipe[43] does improve the agreement between the theoretical and experimental results whereas our results are in the best overall agreement with the experimental data. The agreement, however, is slightly worse as the salt concentration decreases. This is probably because the structureless charged hard-sphere model for proteins is less accurate at low salt concentrations. At low salt concentrations, the interaction between protein and hydrogen ion is stronger because of less shielding from salts. As a result, the actual locations of binding sites and the actual shape of a protein become more important.

To verify the accuracy of estimated hydrated protein sizes, we also calculate the radius of gyration based on X-ray crystal structures published in the protein data bank. In Table 1, we present the geometric radius of gyration defined as

$$R_{\mathrm{g}}{}^2 = \frac{1}{n}\sum_{i=1}^{n}[(x_i-x_{\mathrm{c}})^2 + (y_i-y_{\mathrm{c}})^2 + (z_i-z_{\mathrm{c}})^2] \quad (28)$$

**TABLE 1: Comparison between Theoretical and Experimental Electrostatic Interaction Factors $\omega$[a]**

| protein | $R'_P$ (Å)[b] | $R'_g$ (Å)[c] | KCl concn (M) | $\omega$ DH[d] | $\omega$ SD[e] | this work[f] | expt[g] |
|---|---|---|---|---|---|---|---|
| myoglobin | 18.4 | 15.2 | 0.06 | 0.087 | 0.078 | 0.072 | 0.085 |
| | | | 0.16 | 0.069 | 0.057 | 0.053 | 0.050 |
| bovine serum albumin | 32.0 | | 0.01 | 0.056 | 0.054 | 0.051 | 0.054 |
| | | | 0.03 | 0.043 | 0.039 | 0.038 | 0.036 |
| | | | 0.08 | 0.033 | 0.028 | 0.027 | 0.028 |
| | | | 0.15 | 0.027 | 0.022 | 0.022 | 0.024 |
| ribonuclease | 17.1 | 14.3 | 0.01 | 0.137 | 0.133 | 0.120 | 0.112 |
| | | | 0.03 | 0.112 | 0.106 | 0.096 | 0.093 |
| | | | 0.15 | 0.079 | 0.066 | 0.061 | 0.061 |
| conalbumin | 30.0 | 29.5 | 0.01 | 0.062 | 0.060 | 0.056 (0.050[h]) | 0.046 |
| | | | 0.03 | 0.048 | 0.044 | 0.042 (0.037[h]) | 0.035 |
| | | | 0.10 | 0.034 | 0.029 | 0.028 (0.025[h]) | 0.025 |

[a] The diameters of $K^+$ and $Cl^-$ are 5 Å,[37] the diameter of a hydrogen ion is 3 Å (the approximate size of a water molecule); and the dielectric constant of water is 78.7 at $T = 25$ °C. [b] Estimated hydrated protein radius $R_P'$. [c] The radius of gyration based on crystal structures. [d] Equation 26. [e] Equation 27. Different radii are used in the original Stigter and Dill paper. The radii used here are consistent with the radius of gyration calculations. [f] Equation 24. [g] Myoglobin,[51] bovine serum albumin,[53] ribonuclease,[52] and conalbumin.[54] [h] Calculated with $R_P' = R_g + 2R_W = 32.5$ Å.

where $n$ is the total number of atoms, $x_c = (\sum_{i=1}^{n} x_i)/n$, $y_c = (\sum_{i=1}^{n} y_i)/n$, $z_c = (\sum_{i=1}^{n} z_i)/n$, and $(x_i, y_i, z_i)$ are the coordinates for atom $i$. In eq 28, we sum over all atoms except water that are identified in the protein data bank. We have checked the effect of including hydrogen atoms on the geometric radius of gyration. This is accomplished by using the H-Build facility in CHARMM.[59] We find that adding hydrogen atoms does not significantly change the radius of gyration. For example, the geometric radius of gyration for myoglobin after including the hydrogen atoms is 15.5 Å, which is only slightly larger than the 15.2 Å obtained without hydrogen atoms. It is found that $R_g = 15.2$ Å for myoglobin (PDB ref no.:, 1MBS) and 14.3 Å for ribonuclease (PDB ref no., 3RN3). These radii of gyration, $R_g$, plus the diameter of water, 3 Å, are very close to the estimated hydrated protein radius, $R'_p$. However, the radius of gyration for conalbumin (PDB ref no.: 1OVT) is 29.5 Å, which is about the same as $R'_p$ (30 Å). Thus, the size of hydrated conalbumin is underestimated. As Table 1 shows, using $R'_p = R_g + 2R_W = 32.5$ Å for conalbumin improves the already good agreement between theoretical predictions and experimental data. Thus, $R_g + 2R_W$ is a more accurate way to estimate the size of a hydrated protein if $R_g$ is known.

We have shown that the theory yields accurate predictions of electrostatic interaction factors $\omega$ when the effective protein hard-core diameter is taken as the size of a protein plus that of a water molecule and the effective hydrated hydrogen hard-core diameter is taken to be the size of a water molecule. The error on the size of a hydrated hydrogen is expected to be small because the hydrogen ion itself can be regarded as a point ion and the value agrees with various estimates.[56] On the other hand, the predicted value of electrostatic interaction factors roughly scales with the estimated protein size. For example, increasing or decreasing the conalbumin's size (30 Å) by 3 Å, the diameter of a water molecule, will decrease or increase the electrostatic interaction factor $\omega$ by approximately 10%, which is about the experimental error.[53]

Before closing this section, we shall emphasize that, unlike the recently developed theories based on the PB equation,[9−13] the cavity-function theory cannot be used to predict the site-specific p$K_i$. The cavity-function theory only predicts changes in the p$K_i$ due to the presence of salts.

### Calcium Binding

We investigate the binding of two calcium ions in calbindin $D_{9k}$ as an example of an association event involving more than two particles. The p$K$ shift of calcium binding in calbindin $D_{9k}$ has been investigated both experimentally and theoretically.[2,60] The chemical equilibrium of the binding of two $Ca^{2+}$ ions can be expressed as follows:

$$P + 2Ca^{2+} \rightleftharpoons [PCa_2]^{4+} \tag{29}$$

The apparent equilibrium constant $K^{app}$ is associated with the actual equilibrium constant by (cf. eq 9)

$$K^{app}/K^\circ = y^{(3)}_{CaPCa} \approx [y_{PCa}(L_{PCa})]^2 \tag{30}$$

where the linear approximation (eq 11) for the three-particle cavity function has been used. Equation 30 can be rewritten in terms of the p$K$ shift:

$$pK \equiv -\log K^{app} = pK_0 - 2\log y^{core}_{PCa}(L_{PCa}) - 2\log y^{ele}_{PCa}(L_{PCa}) \tag{31}$$

where $pK_0 = -\log K^\circ$ (ref 2).

In Svensson and Woodward's model,[2] calbindin $D_{9k}$ is represented by a large hard sphere with a 11.6-Å radius and 48 charged hard-sphere sites which are located at the X-ray crystal coordinates occupied by the carboxylic oxygens and the lysine nitrogens. A partial charge of −0.5 is assigned to all carboxylic oxygens, and a charge of + 1 is assigned to all lysine nitrogens. In addition, $Ca^{2+}$, $K^+$, and $Cl^-$ are modeled as charged hard spheres, each of which has a diameter of 4 Å. In this paper, we shall use the much simpler charged hard-sphere model for the protein. Following Svensson and Woodward, the radius of the charged hard sphere for protein is chosen to be 11.6 Å. [For reference, our calculated geometric radius of gyration for calbindin $D_{9k}$ (PDB ref no.: 3ICB) is 11.5 Å.] Here, we will not add a layer of hydration to the protein to facilitate the comparison between our results and those of Svensson and Woodward. (Results with a hydrated protein as the core will be discussed, however.) For the same reason, we use their models for $Ca^{2+}$, $K^+$, and $Cl^-$ although the 4 Å diameter for $K^+$ or $Cl^-$ is smaller than the 5 Å diameter for $K^+$ or $Cl^-$ in the previous section (this diameter difference leads to a 3% difference in final results). Furthermore, we assume that the binding between $Ca^{2+}$ and calbindin $D_{9k}$ occurs at the surfaces of their hard cores (see Figure 2).

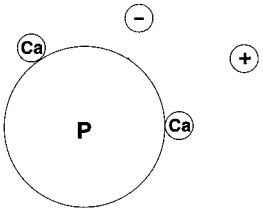


**Figure 2.** Schematic plot of charged hard-sphere models for calbindin $D_{9k}$ (P), calcium ions ($Ca^{2+}$), salt cations (+), and salt anions (−). The binding between calbindin $D_{9k}$ and two calcium ions is assumed to occur at their hard-core surfaces.

**TABLE 2: Comparison between Theoretical and Experimental p*K* Shifts in the Binding of Two $Ca^{2+}$ Ions to Calbindin $D_{9k}$[a]**

| | | KCl | $\Delta pK$[b] | | | |
|---|---|---|---|---|---|---|
| protein | $Z_P$ | concn (M) | RISM-MSA[c] | MC[c] | this work[d] | expt[c] |
| wild type | −8 | 0.05 | 2.2 | 2.0 | 2.4 (2.7)[e] | 2.9 |
| | | 0.10 | 2.8 | 2.6 | 3.0 (3.4)[e] | 3.6 |
| | | 0.15 | 3.2 | 2.9 | 3.4 (3.8)[e] | 4.0 |
| E17Q | −7 | 0.05 | 1.9 | 1.8 | 2.1 | 2.2 |
| | | 0.10 | 2.5 | 2.3 | 2.6 | 2.8 |
| | | 0.15 | 2.8 | 2.6 | 2.9 | 3.0 |
| E17Q+D19N | −6 | 0.05 | 1.6 | 1.5 | 1.8 | 1.7 |
| | | 0.10 | 2.1 | 2.0 | 2.3 | 2.2 |
| | | 0.15 | 2.4 | 2.2 | 2.5 | 2.5 |
| E17Q+D19N+E26Q | −5 | 0.05 | 1.4 | 1.3 | 1.5 | 1.4 |
| | | 0.10 | 1.8 | 1.7 | 1.9 | 1.9 |
| | | 0.15 | 2.1 | 1.9 | 2.1 | 2.0 |

[a] The diameter of calbindin $D_{9k}$ is 23.2 Å, the diameters of $Ca^{2+}$, $K^+$, and $Cl^-$ are 4 Å, the dielectric constant of water is 78.7, and the temperature is 25 °C. [b] $\Delta\ pK = pK - pK|_{C_{KCl}=0.002\ M}$. [c] A much more complicated interaction-site model for protein was used for the RISM-MSA and MC studies.[2] [d] Equation 32. [e] Calculated by assuming $Z_P = -9$.

For charged hard-sphere models, eq 31 becomes (cf. eq 21)

$$pK = pK_0 - 2\log[g_{PCa}(\sigma_{PCa})]^{core} + \frac{0.8686\lambda\, Z_{Ca}Z_P}{\sigma_{PCa}}\left[\frac{1}{(1+\Gamma\sigma_{Ca})(1+\Gamma\sigma_P)} - 1\right] \quad (32)$$

where the EXP approximation has been used in the derivation and $Z_{Ca} = +2$. The hard-core term in eq 32 is negligible if the salt concentration is less than 0.2 mol/L. The hard-core term will be significant only if the salt concentration is higher than 1 mol/L.

Table 2 compares the p*K* shifts ($\Delta pK = pK - pK|_{C_{KCl}=0.002\ M}$) calculated by the cavity-function theory to results obtained by the site−site integral-equation theory of the reference interaction-site model (RISM), by MC simulations, and by experiments at various KCl concentrations for the wild-type calbindin $D_{9k}$ and several mutants.[2] The mutants are obtained by replacing the negatively charged amino acids glutamate (E) and aspartate (D) with the neutral glutamine (Q) and asparagine (N), respectively. The nomenclature E17Q stands for the replacement of the glutamate (E) at position 17 by glutamine (Q). Each replacement increases the protein charge number by 1 from −8 with no changes in the size of the protein.[2] Table 2 shows that our results are in excellent agreement with the experimental results for the three mutants whereas they are slightly lower than the experimental values for the wild type. Our predictions for the values of the p*K* shift are even closer to experimental results in all cases than the values predicted in the MC simulations and the RISM-MSA integral-equation theory, despite the fact that our model is much simpler. This strongly suggests that at least for calbindin $D_{9k}$ the simple charged hard-sphere model is accurate enough to predict the salt effects on the p*K* shift of the protein. The close agreement between our theoretical results and those of RISM-MSA indicates that structural details do not play important roles in the salt effects on the p*K* shift even at a concentration as low as 0.05 mol/L.

As we have indicated in Table 2, an excellent agreement between theoretical predictions and experimental results is found for the three mutants but not for the wild type of calbindin $D_{9k}$. It is interesting that the agreement for the wild type protein will be significantly improved if a negative charge of −9 rather than −8 is assigned to the protein. However, this improvement is likely a coincidence. More work is needed to understand why all theoretical predictions (including MC simulations) fail for the wild type.

If the diameter of protein calbindin $D_{9k}$ is taken as 26.2 Å (thus including one layer of water molecules) the theoretical p*K* shift decreases 0.1−0.2 p*K* units. As a result, the agreement between the theoretical results and the experimental data becomes worse. In the previous section, we have shown that excellent agreement between theoretical predictions and experimental data for electrostatic interaction factors is achieved when the first hydration shell is included as part of the hard core of the protein. Thus, calcium ions experience a smaller protein hard core than hydrogen ions. This is likely a consequence of the assumption that the size of a protein is unchanged during protonation. It is also physically reasonable because calcium ions are much heavier than hydrogen ions, and as a result, the calcium ions do not "feel" the hydration shell as the hard part of a protein.

## Summary

In this paper, we demonstrated that a cavity-function theory of chemical association is quantitatively accurate in predicting the salt effects on protein titration and binding. One attractive feature of the theory is that its analytical expression in the EXP approximation is as simple as but is more accurate than the expression in the DH approximation. The theory can be easily applied to the case in which the size of salt cations is different from the size of anions. If needed, the accuracy of the theory can also be systematically improved by using increasingly accurate integral-equation theories for cavity functions.

As we have shown, the cavity-function theory is very accurate in predicting salt effects on protein titration and binding. In particular, our simple structureless charged hard-sphere model provides predictions for the p*K* shifts as accurate as a more complicated multisite protein model based on the X-ray crystal structure.[2,60] Thus, at least for the cases studied here, the structural details do not play important roles in global effects on protein titration and binding. An extension of the theory to investigate localized binding effects using an atomistic model is in progress.

**Acknowledgment.** I am grateful to Aaron Dinner and Professor Martin Karplus for helpful discussions and comments. This material is based upon work supported by a postdoctoral fellowship from the Program in Mathematics and Molecular Biology at the University of Berkeley, which is supported by the National Science Foundation under Grant No. DMS-9406348. The government has certain rights in this material. This work is also supported in part by a National Institutes of Health postdoctoral fellowship.

## References and Notes


(1) Fersht, A. *Enzyme Structure and Mechanism*; W. H. Freeman: New York, 1985.
(2) Svensson, B.; Woodward, C. E. *J. Phys. Chem.* **1995**, *99*, 1614.
(3) Linderstrøm-Lang, K. *C. R. Trav. Lab. Carlsberg* **1924**, *15*, 1.
(4) Tanford, C.; Kirkwood, J. G. *J. Am. Chem. Soc.* **1957**, *79*, 5333.
(5) Matthews, J. B.; Gurd, F. R. N. *Methods Enzymol.* **1986**, *130*, 413.
(6) States, D. J.; Karplus, M. *J. Mol. Biol.* **1987**, *197*, 122.
(7) Jayaram, B. *J. Phys. Chem.* **1994**, *98*, 5773.
(8) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *157*, 671.
(9) Gilson, M. K.; Honig, B. H. *Biopolymers* **1986**, *25*, 2097.
(10) Bashford, D.; Karplus, M. *J. Mol. Biol.* **1988**, *203*, 507.
(11) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219.
(12) Yang, A.-S.; Honig, B. *J. Mol. Biol.* **1994**, *237*, 602.
(13) Holst, M.; Kozack, R. E.; Saied, F.; Subramaniam, S. *Proteins* **1994**, *18*, 231.
(14) Mehler, E. L. *J. Phys. Chem.* **1996**, *100*, 16006.
(15) Kitao, A.; Hirata, F.; Go, N. *J. Phys. Chem.* **1993**, *97*, 10231.
(16) Ullner, M.; Woodward, C. E.; Jönsson, B. *J. Chem. Phys.* **1996**, *105*, 2056.
(17) Lee, L.-P.; Tidor, B. *J. Chem. Phys.* **1997**, *106*, 8681.
(18) Chong, L. T.; Dempster, S. E.; Hendsch, Z. S.; Lee, L.-P.; Tidor, B. *Protein Sci.* **1998**, *7*, 206.
(19) Ooi, T. *Adv. Biophys.* **1994**, *30*, 105.
(20) Honig, B.; Yang, A.-S. *Adv. Protein Chem.* **1995**, *46*, 27.
(21) Dill, K. A.; Stigter, D. *Adv. Protein Chem.* **1995**, *46*, 59.
(22) Zhou, Y.; Stell, G. *J. Chem. Phys.* **1992**, *96*, 1507.
(23) Stell, G. *The Equilibrium Theory of Classical Fluids*; Frisch, H. L.; Lebowitz, J. L., Eds.; Benjamin: New York, 1964.
(24) Uehara, Y.; Lee, Y. T.; Ree, T.; Ree, F. H. *J. Chem. Phys.* **1979**, *70*, 1884.
(25) Wertheim, M. S. *J. Chem. Phys.* **1987**, *87*, 7323.
(26) Müller, E. A.; Gubbins, K. E. *Mol. Phys.* **1993**, *80*, 957.
(27) Chapman, W. G.; Jackson, G.; Gubbins, K. E. *Mol. Phys.* **1988**, *65*, 1057.
(28) Banaszak, M.; Chiew, Y. C. *Phys. Rev. E* **1993**, *48*, 3760.
(29) Ghonasgi, D.; Chapman, W. G. *AIChE J.* **1994**, *40*, 878.
(30) Zhou, Y.; Hall, C. K. *Biopolymers* **1996**, *38*, 273.
(31) Zhou, Y.; Yeh, S.; Stell, G. *J. Chem. Phys.* **1995**, *102*, 5785.
(32) Zhou, Y.; Stell, G. *J. Chem. Phys.* **1995**, *102*, 8089.
(33) Andersen, H. C.; Chandler, D. *J. Chem. Phys.* **1972**, *57*, 1918.
(34) Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1972**, *57*, 1930.
(35) Andersen, H. C.; Chandler, D.; Weeks, J. D. *J. Chem. Phys.* **1972**, *57*, 2626.
(36) Triolo, R.; Grigera, J. R.; Blum, L. *J. Phys. Chem.* **1976**, *80*, 1858.
(37) Tanford, C. *Adv. Protein Chem.* **1962**, *17*, 70.
(38) Atkins, P. W. *Physical Chemistry*; W. H. Freeman: New York, 1986.
(39) Chandler, D.; Pratt, L. R. *J. Chem. Phys.* **1976**, *65*, 2925.
(40) Friedman, H. L. *A Course in Statistical Mechanics*; Prentice Hall: Englewood Cliffs, NJ, 1985.
(41) Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 9556.
(42) Tanford, C.; Buzzel, J. G.; Rands, D. G.; Swanson, S. A. *J. Am. Chem. Soc.* **1955**, *77*, 6421.
(43) Stigter, D.; Dill, K. A. *Biochemistry* **1990**, *29*, 1262.
(44) Bleil, R.; Mohanty, U. *J. Phys. Chem.* **1995**, *99*, 2915.
(45) Amos, M. D.; Jackson, G. *Mol. Phys.* **1991**, *74*, 191.
(46) Lebowitz, J. L.; Percus, J. K. *Phys. Rev.* **1966**, *144*, 251.
(47) Waisman, E.; Lebowitz, J. L. *J. Chem. Phys.* **1972**, *56*, 3086.
(48) Mansoori, G. A.; Carnahan, N. F.; Starling, K. E.; Leland, T. W. *J. Chem. Phys.* **1971**, *54*, 1523.
(49) Blum, L. *Mol. Phys.* **1975**, *30*, 1529.
(50) Blum, L.; Høye, J. S. *J. Phys. Chem.* **1977**, *81*, 1311.
(51) Breslow, E.; Gurd, F. R. N. *J. Biol. Chem.* **1962**, *237*, 371.
(52) Tanford, C.; Hauenstein, J. D.; Rands, D. G. *J. Am. Chem. Soc.* **1955**, *77*, 6409.
(53) Tanford, C.; Swanson, S. A.; Shore, W. S. *J. Am. Chem. Soc.* **1955**, *77*, 6414.
(54) Wishnia, A.; Weber, I.; Warner, R. C. *J. Am. Chem. Soc.* **1961**, *83*, 2071.
(55) Zhou, Y.; Stell, G.; Friedman, H. L. *J. Chem. Phys.* **1988**, *89*, 3836.
(56) Marcus, Y. *Chem. Rev.* **1988**, *88*, 1475.
(57) Tawa, G. I.; Pratt, L. R. *J. Am. Chem. Soc.* **1995**, *117*, 1625.
(58) Vilarino, T.; de Vicente, M. E. S. *J. Phys. Chem.* **1996**, *100*, 16378.
(59) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(60) Svensson, B.; Jösson, B.; Woodward, C. E.; Linse, S. *Biochemistry* **1991**, *30*, 5209.