Significance of Charge Regulation in the Analysis of Protein Charge Ladders

Irina Gitlin, Michael Mayer, and George M. Whitesides*

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

Received: September 25, 2002; In Final Form: December 4, 2002

Analysis of protein charge ladders using capillary electrophoresis (CE) provides a method of determining charges of proteins. This method has disregarded the effect of charge compensation—a response of the protein and its environment to a change in electrostatic potential on the surface of the protein. This work examines the difference in charge, ΔZ, between the first two rungs of the ladder of bovine carbonic anhydrase II (BCAII) as a function of pH and ionic strength using CE. These data were analyzed in three ways: using models based on Hückel theory and on charge regulation, and using linear regression. These analyses were in only qualitative agreement, and the differences between them suggest that simple theoretical models for the behavior of colloidal particles cannot establish the value of ΔZ accurately in proteins. Linear regression of mobilities of the rungs of charge ladders—a method proposed in earlier work—continues to be a computationally convenient method of estimating the charge Z_0 of native proteins, but the accuracy of this method depends on the value of ΔZ. The absolute value of ΔZ cannot presently be established accurately. In the case of BCAII, we suggest $\Delta Z = -0.93$ for the difference in charge between the first two rungs of the charge ladder at pH = 8.4 and 10 mM ionic strength. An estimate of the uncertainty in this value for BCAII due to uncertainties in the values of pK_a of amino acids and of the hydrodynamic radius is ±0.02. Other uncertainties not considered in this analysis will make this value larger.

Introduction

A protein charge ladder is a collection of derivatives of a protein generated by converting its charged groups (most commonly lysine ϵ-NH_3^+ but also aspartate or glutamate $-CO_2^-$) into electrically neutral ones (ϵ-$NHCOCH_3$ or $-CO_2CH_3$).[1] In free solution capillary electrophoresis (CE), these modified proteins separate into distinct peaks or "rungs"; each rung contains regioisomers with approximately the same charge: that is, at least nominally, the same number of modified groups. Charge ladders provide a self-calibrating tool for estimating certain basic physical parameters of proteins such as charge,[2,3] hydrodynamic radius,[4] and electrostatic contributions to the free energy of binding of ligands[5,6] and of protein folding.[7]

The simplest analysis of charge ladders assumes that the charge difference (ΔZ) between the consecutive rungs of a charge ladder generated by acetylation of lysine ϵ-NH_3^+ groups is a full unit of charge ($\Delta Z = -1$). Menon and Zydney have made the point that ΔZ may have a value different from -1 if the change in charge of the ϵ-amino group on acetylation (which is, in fact, -1, provided that this group is completely protonated before acetylation) is offset by a compensating change in charge elsewhere in the protein.[8] This concept—charge compensation—is well-developed and extensively modeled in colloid chemistry.[9,10] Menon and Zydney suggested that it also applies to proteins and proposed a model to analyze the adjustment of charge of the macromolecule and the surrounding buffer upon annihilation of a positive charge from the surface of the protein.

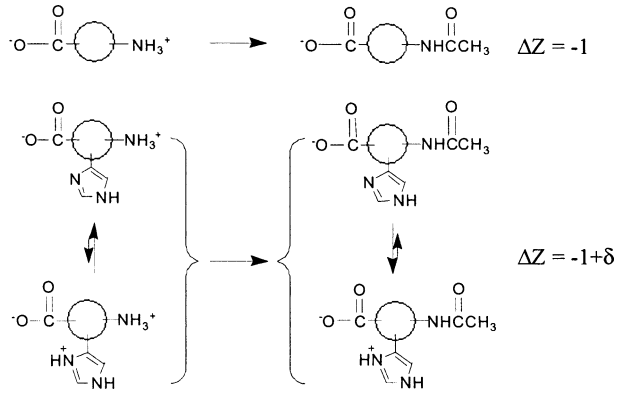
The analysis of the response of the charge of a protein to the acetylation of a lysine ϵ-NH_3^+ can be divided into three levels. First, there is a physical effect—charge compensation, in which the charge on the rest of the protein adjusts to the neutralization of charge on the ϵ-NH_3^+ in a way that reduces the total charge on the protein. Second, there are mechanisms for this charge compensation, which are discussed below. Third, there are theoretical models that allow the extent of charge compensation to be estimated based on assumptions about the character of electrostatic interactions in these systems.

Possible mechanisms that might make the value of $|\Delta Z|$ smaller than 1 on eliminating the positive charge from the ϵ-NH_3^+ group, and thus reducing the electrostatic potential on the surface of the protein, include (i) an increase in the local concentration of protons, (ii) perturbations of the pK_a values of other ionizable amino acids, and/or (iii) an increase in the local concentration of buffer ions that would effectively screen the increased negative charge. Changes in protein conformation, in solvation, or in the character of the dielectric medium in which the charges interact may also be important, albeit impractical to model. The lysine ϵ-NH_2 group might, in principle, also be incompletely protonated, so that the charge on this group before acylation would be less than +1; we assume, however, that this protonation is complete at values of pH commonly used for electrophoretic analysis of proteins. The extent of charge compensation resulting from any combination of i−iii would be predicted to be most significant in buffers of low ionic strength and at values of pH close to the pK_a values of ionizable amino acids. In working with proteins at values of pH commonly used (pH 7.0−8.5), the most plausible amino acid to show changes in the extent of protonation with changes of Z would be histidine (pK_a = 6.0−7.0).[11] Scheme 1 suggests the physical basis of mechanism (i) of charge compensation, using histidine as an example. The value of ΔZ to be used at a

* To whom correspondence should be addressed. Tel: (617)495-9430. Fax: (617)495-9857. E-mail: gwhitesides@gmwgroup.harvard.edu.

10.1021/jp027066w CCC: $25.00
Published on Web 01/15/2003

SCHEME 1



particular value of pH and buffer and the choice of pH and buffer for a particular investigation are important in understanding what kinds of information, extracted from analysis of charge ladders, can be considered quantitative and what semiquantitative or qualitative.

In this work, we have used bovine carbonic anhydrase (BCAII) as the model protein and CE as the analytical system to obtain experimental values of the mean electrophoretic mobilities of the rungs of the charge ladder derived from BCAII by acetylation. We have determined these mobilities as a function of ionic strength and pH. We have analyzed the data to infer the effective charge of the protein (in native form and acylated) in three ways. (i) We assume the value of ΔZ (the difference in charge between adjacent rungs of the charge ladder) to be $\Delta Z = Z_{n+1} - Z_n = -1.0$ (where n is the number of acetyl groups). This model is the simplest that can be drawn of the electrostatic interactions in a protein; it assumes that all of the changes of the protein are independent of one another. (ii) We calculate the value of Z_n (and ΔZ) from the values of mobility using Hückel theory (HL).[12] This theory is highly simplified; it is based on a model that treats the protein as a uniformly charged sphere and the solution as a continuum of uniform dielectric constant with mobile ions distributed according to the linearized Poisson−Boltzmann equation. (iii) We calculate the value of Z_n (and ΔZ) using a more elaborate, but still very simplified, electrostatic theory (the "charge regulation" (CR) model outlined by Menon and Zydney), which allows a change in the charge of the protein to modify its environment. This model is an adaptation of a model proposed by Linderstrøm-Lang to describe proton binding in hydrogen ion titration curves.[13,14] The model of Linderstrøm-Lang is based on the assumption of spherical protein, with evenly distributed charge and impenetrable to the solvent.

We consider the limitations of these models and conclude that the numerous assumptions and simplifications underlying them all make it impractical to use them to evaluate precisely the difference in charge of the rungs in a charge ladder. Although, as a result, we cannot make quantitative statements about the extent and significance of charge compensation in the analysis of charge ladders, we conclude—as suggested by Menon and Zydney—that ΔZ is probably significantly less than -1, rather than -1, and estimate plausible values of ΔZ as a function of pH. We conclude, in general agreement with Menon and Zydney,[8] that a value of $\Delta Z \cong -0.9$ is a plausible estimate of this parameter at pH $\sim$ 8.5 for BCAII.

Background

In CE, the mobility (μ, m^2 V^{-1} s^{-1}) of a protein is assumed to be directly proportional to the electrostatic force on it (that is, to its effective charge Z) and inversely proportional to its hydrodynamic drag. The drag, in turn, is related imprecisely to the shape and size of the molecule by assuming an empirical relationship. Within this approximate framework, the electrophoretic mobility can be expressed by eq 1

$$\mu = C_p \frac{Z}{MW^{\alpha}} \tag{1}$$

where MW is the molecular weight of the molecule, α is an empirical parameter that depends on the shape of the molecule, and C_p is a constant that depends on the properties of the solution and relates the ratio Z/MW^{α} to the mobility.[15] Experimentally, the mobility of an analyte can be calculated by comparison of its migration time (t_a, s) with the migration time of a neutral marker (t_{nm}, s) (eq 2); the latter accounts for the electroosmotic flow generated by the charged wall of a capillary.[15] In this equation, the symbols L_d (m) and L_t (m) are the distances from the inlet to the detector and from the inlet to the outlet of the capillary; V (V) is the voltage applied between the two ends of the capillary.

$$\mu = \frac{L_d L_t}{V}\left(\frac{1}{t_a} - \frac{1}{t_{nm}}\right) \tag{2}$$

A method to determine protein charge developed in our laboratory[16] involves generation of a charge ladder by acetylation of lysine groups, followed by determination of electrophoretic mobility using CE. Each rung of the resulting charge ladder differs from the preceding (differing by one $NHCOCH_3$ group) by ΔZ. For carbonic anhydrase (and most other proteins studied), the electrophoretic mobility of the first five rungs of the ladder correlates linearly with the number (n) of acetyl groups introduced.[16] The value of ΔZ depends on the pK_a of the modified group and the pH of the solution. In the case of acetylated BCAII, for which the average pK_a of lysine is 10.3 and the pH of the CE buffer is 8.4, we originally assumed ΔZ to be -1.[17] The x-intercept of the linear regression (LR) line through the values of mobility μ of the first five rungs plotted against $n\Delta Z$ gives the charge of the native protein Z_0 (Figure 1).

This method is attractive for computational and conceptual simplicity but depends for its accuracy in estimation of Z_0 on the value of ΔZ. If the value of $|\Delta Z| < 1$, then the value of Z_0 will be too large. The question is then, "What is ΔZ on acetylation of one fully protonated lysine ϵ-NH_3^+ group in a protein?"

A model that is often used in relating the electrophoretic mobilities of colloids and proteins to charge is that of Hückel.[12] As applied to a protein, it starts by assuming the protein is a uniformly charged hard sphere. It first relates the electrostatic potential ψ_s (V) on the surface of a sphere to the charge of the sphere Z through the solution to the linearized Poisson−Boltzmann equation (eq 3) (also known as the Debye−Hückel equation).

$$\psi_s = \frac{eZ}{4\pi\epsilon_o\epsilon R(1 + \kappa R)} \tag{3}$$

Here, e (C) is the electron charge, ϵ (unitless) is the dielectric constant of the buffer, ϵ_o (C V^{-1} m^{-1}) is the permittivity of vacuum, R (m) is the radius of the sphere, and κ (m^{-1}) is the inverse Debye length.[12] The Debye length is a quantity that depends on ionic strength; it suggests the average thickness of the double layer[18] that is adjacent to the charged particle. The

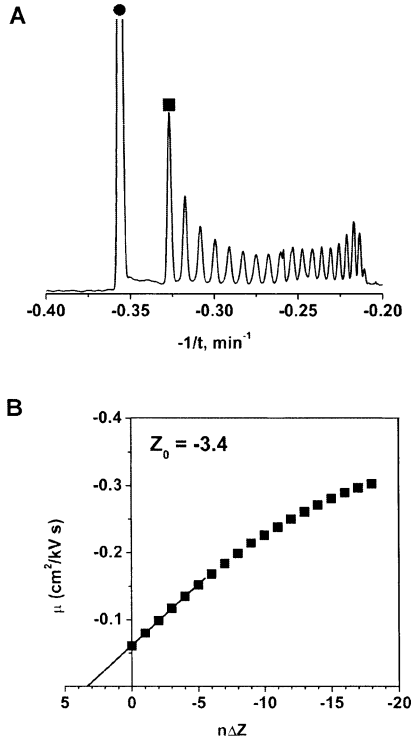


Figure 1. (A) Charge ladder of BCAII formed by partial acetylation of lysine ϵ-NH_3^+ groups and observed by CE. The peak marked by (●) is the neutral marker *p*-methoxybenzyl alcohol used to monitor electroosmotic flow. The peak marked by (■) is the native BCAII. Each following peak represents a mixture of proteins having the same number of acylated amino groups. The experiment is conducted in 25 mM Tris−192 mM Gly buffer, pH = 8.4. (B) Plot of charge ladder mobility calculated by eq 2 as a function of $n\Delta Z$ assuming a ΔZ value of −1. A best fit line is constructed through the first six members of the charge ladder. Extrapolation of the best fit line to the intercept with the abscissa yields an estimate of the charge of the native protein.

electrostatic potential ψ_s can be related to the particle mobility via the Hückel equation (eq 4)

$$\mu = \frac{2}{3}\frac{\epsilon_o \epsilon \psi_s}{\eta} \tag{4}$$

where η (kg m^{-1} s^{-1}) is the viscosity of the solution.[19] Equation 4 is valid for thick ($\kappa R < 1$) electrical double layers. Henry modified eq 4 to include electrophoretic retardation effects; this modification made the theory applicable to a wider range of electrical double layers[19] (eq 5), where eq 6 describes the function $f(\kappa R)$.[6]

$$\mu = \frac{2}{3}\frac{\epsilon_o \epsilon \psi_s}{\eta} f(\kappa R) \tag{5}$$

$$f(\kappa R) = 1 + \frac{(\kappa R)^2}{16} - \frac{5(\kappa R)^3}{48} - \frac{(\kappa R)^4}{96} + \frac{(\kappa R)^5}{96} + \left(\frac{(\kappa R)^4}{8} - \frac{(\kappa R)^6}{96}\right)\exp(\kappa R)\int_{\infty}^{\kappa R}\frac{\exp(-x)}{x}\,dx \tag{6}$$

The combination of eq 3 and 5 makes it possible to determine the charge of a protein directly from its mobility by eq 7, provided that the parameters in this equation (R, κ, η) are known.

$$\mu = \frac{eZ}{6\pi\eta R(1+\kappa R)} f(\kappa R) \tag{7}$$

When applied to proteins, the HL, described in eq 3−7, has several shortcomings. (i) The nonspherical shape of the proteins and the resulting approximation of its hydrodynamic radius as that of an equivalent sphere introduce uncertainties into the calculations. While the calculation of the surface potential from mobility data depends only logarithmically on the value of the radius (eq 5), the charge is directly related to the value of the radius.[21] Small variations in the radius can thus introduce large variations in charge. (ii) The model cannot account for heterogeneities in surface charge and the dielectric constant of the protein. (iii) The model is valid only for surface potentials below 25 mV, the limit at which linearization of the Poisson−Boltzmann equation applies.[22]

Menon and Zydney[8] proposed a theoretical method of determining the charge of proteins. They suggest that the local concentration of H^+ ions near the surface of a protein correlates with the electrostatic potential of its surface ψ_s via the Boltzmann distribution (eq 8), where e is the electronic charge, k (J K^{-1}) is Boltzmann's constant, and T (K) is temperature. The linearized Poisson−Boltzmann equation (eq 3) relates the electrostatic potential ψ_s to charge Z. The values of Z in turn can be found by summing up all the charged residues r_i of titratable species i at a particular local pH.

$$[H^+_{local}] = [H^+_{bulk}]\exp(-e\psi_s/kT) \tag{8}$$

For acidic residues where n_i is the total number of titratable species i and K_{ai} is the intrinsic dissociation constant of that species—that is, the dissociation constant in the absence of other charges—r_i is given by eq 9. For basic residues, eq 10 is a similar expression.

$$r_i(\text{acidic}) = \frac{n_i K_{ai}}{K_{ai} + [H^+_{bulk}]\exp(-e\psi_s/kT)} \tag{9}$$

$$r_i(\text{basic}) = n_i - \frac{n_i K_{ai}}{K_{ai} + [H^+_{bulk}]\exp(-e\psi_s/kT)} \tag{10}$$

The total charge on the protein is then the sum of the charge due to negatively charged (acidic) residues and positively charged (basic) residues (eq 11). The charge Z is calculated iteratively from equations 8, 3, and 11. Following Menon and Zydney, we will refer to these calculations as the CR model.

$$Z = \sum_i r_i(\text{basic}) - \sum_i r_i(\text{acidic}) \tag{11}$$

The CR model makes a number of approximations; three are especially important. (i) It assumes a spherical shape for a protein and uses an approximate hydrodynamic radius. (ii) It assumes that all titratable residues of the same kind of group (e.g., all lysine ϵ-NH_3^+ groups) have the same value of pK_a. In a protein, the equilibrium between the charged and the neutral states of ionizable residues is influenced by interactions with permanent dipoles of the proteins, by other titratable groups, and by restricted interactions with water.[23] These interactions can shift the values of pK_a of amino acids substantially from their average values: theoretical[24] and experimental[25] studies of individual residues have demonstrated variations of more than two units in pK_a values for residues of the same type. (iii) It assumes that the effect of a change in the surface potential of the protein will be felt primarily in a change in the local concentration of protons in solution (that is, in the local pH). The change in surface potential may also be influenced by

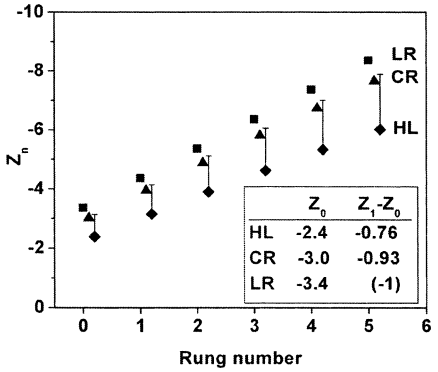


Figure 2. Values of charge on the first six rungs of the BCAII charge ladder as calculated by different methods discussed in the text. The squares refer to the charge calculated by the graphical LR, triangles refer to CR (eqs 8, 3, 11), and diamonds refer to the Hückel model (HL; eq 7). The calculations assumed the BCAII radius to be 2.05 nm. The inset shows the values of Z_0 and $\Delta Z = Z_1 - Z_0$ determined by the three methods (or assumed in the case of LR). The CR and HL data points are slightly displaced off the integral values of rung numbers for clarity of the error bars. The error bars on Hückel data show the sensitivity of the HL calculations to the numerical value of the hydrodynamic radius of a protein. The radius is varied from 2.05 to 2.5 nm.

interactions with buffer ions—either as loose association or tight binding.[26] The fact that buffer ions are not treated explicitly in this physical picture is, however, not important, since the core of the proposal by Menon and Zydney,[8] and of earlier discussions by Carbeck,[1] is that the extent of protonation of other ionizable groups in the protein (especially the imidazole groups in histidine) will change with changes in Z_0. A change in the extent of protonation can be considered equivalently to be the result of a change in pK_a at constant pH or of a change in pH at constant pK_a.

In summary, the empirical LR model is based on experimental mobility data and allows us to find the charge of a native protein at the conditions in which the experiment is conducted. The HL also utilizes experimental mobility data and relates it to the surface potential of a protein and to charge. Thus, the charge of individual rungs of the charge ladder can be calculated. The CR theory is not based on the experimental data; it is able to generate the charge of the native protein and of subsequent rungs based on the assigned conditions (pH, ionic strength) and on the physical parameters of a protein (size, amino acid composition, values of pK_a of the amino acids).

Results and Discussion

Charge as a Function of Rung Number Calculated by Different Methods. We collected experimental electrophoretic mobilities for the BCAII charge ladder using Tris−Gly buffer (25 mM Tris−192 mM Gly, 10 mM ionic strength,[27] pH 8.4, 25 °C). We then calculated the charge corresponding to the first six rungs based on HL (eq 7) and based on LR with $\Delta Z = -1$. We also calculated the charge of those rungs by CR theory for conditions used in CE experiments (eqs 3, 8−11). Figure 2 summarizes those calculations. The values of Z_0—the charge of the native protein—were estimated to be −3.4 by LR, −3.0 by CR, and −2.4 by HL. The values of Z_n estimated from the experimental mobilities by the LR method are consistently more negative than those calculated using the CR method or the Hückel model. The estimate of Z_n by LR (assuming $\Delta Z = -1.0$) is consistently closer to the theoretical values calculated by the CR than to those estimated from experimental mobilities using the Hückel model.

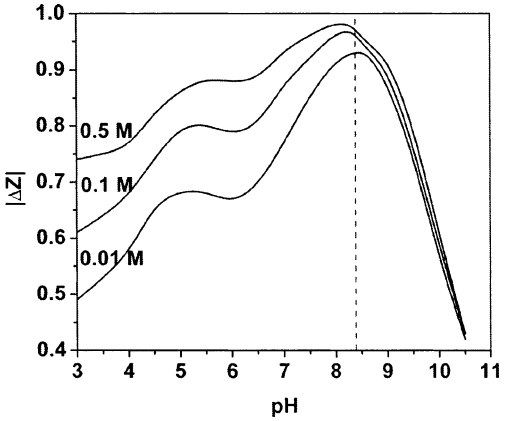


Figure 3. Plot of $|\Delta Z|$ as a function of pH at three different ionic strengths, calculated by CR theory for BCAII. The maximum ΔZ occurs at pH ∼ 8.5 and approaches unity as the ionic strength increases. In the solution of 0.01 M ionic strength, pH 8.4 (shown as dotted line), ΔZ is calculated to be 0.93. The value of $R = 2.05$ nm is used in the calculations.

Influence of the Hydrodynamic Radius. Estimation of the radius of 2.05 nm for BCAII used in previous work[3] was based on the average value of the partial specific volume for the proteins (0.72 mL/g)[28] and the assumption of spherical shape. We searched for other methods, experimental or theoretical, to get a better estimate for the hydrodynamic radius of a protein. In the case of lysozyme—the best characterized globular protein—the estimate of the radius of gyration as measured by light scattering experiments ranged from 1.41[29] to 2.2 nm,[30] albeit in different conditions. An estimate of the radius of lysozyme based on the average volume gave a value of $R = 1.59$ nm.[3] The relative standard deviation of these values is ±19%. Allowing the radius of BCAII to deviate from the previously used value of 2.05 by the same percentage yielded the following results by the Hückel calculations: $Z_0 = -1.9$ (at $R = 1.7$) and −3.1 (at $R = 2.5$). The sensitivity of the Hückel theory to the value of R is demonstrated in Figure 2 by the error bars generated as R is changed from 2.05 to 2.50 nm.

The CR model is less sensitive to the value of the radius than the Hückel model because it is not based on the motion of a charged sphere in an electric field and thus does not calculate hydrodynamic drag. An increase in radius from 2.05 to 2.50 nm results in the change in Z_0 from −3.0 to −3.1.

***Z_0* and *ΔZ* as a Function of Ionic Strength.** We studied the behavior of charge ladders of BCAII in solutions of varying ionic strength and evaluated the charge of the native protein under those conditions. As the ionic strength, and thus the shielding, of the buffer increases, we expect the charge of the protein to become more negative. With increasing shielding, we also expect to reduce the electrostatic interactions between the charged residues (e.g., between Lys and His as depicted in Scheme 1) and thus to bring the value of $\Delta Z = Z_1 - Z_0$ closer to a full unit of charge.[31] This trend is evident in the theoretical CR calculations shown in Figure 3.[32] The value of $|\Delta Z|$ is maximal at pH ∼ 8.5 and approaches unity at that pH at very high (>0.5 M) ionic strengths.

The values of Z_0, as estimated by the three methods, follow the predicted trend with increasing ionic strength: they become more negative at higher ionic strength (Figure 4). The values of ΔZ found using Hückel calculations, however, do not approach unity as the ionic strength increases. The origin of the observed values is not immediately clear but may be due to the assumption of an incorrect value for the hydrodynamic radius in the calculations. Indeed, using a higher value of R (e.g., values

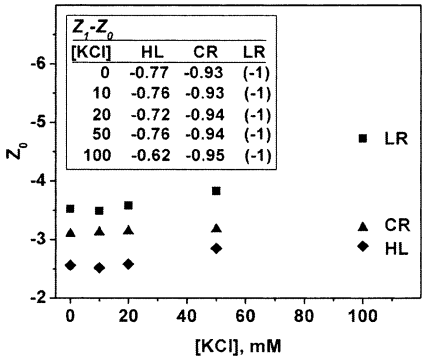


Figure 4. Charge of native BCAII as a function of increasing concentration of KCl in 25mM Tris−192 mM Gly buffer, pH = 8.4, 25 °C. The squares refer to the charge estimated by the LR method, triangles refer to CR theory, and diamonds refer to HL. The inset provides the values of $\Delta Z = Z_1 - Z_0$ as calculated by the three methods. The value of $R = 2.05$ nm is used in the CR and HL calculations.

TABLE 1: Charge of the Native BCAII and the Charge Difference between the Native and the First Rung of the Charge Ladder as a Function of pH in 25 mM Tris−192 mM Gly Buffer (A) and in 20 mM Phosphate Buffer (B)[a]

A

pH	I (mM)	ψ_s^{HL} (mV)	Z_0^{HL}	Z_0^{CR}	Z_0^{LR}	$Z_1 - Z_0^{CR}$
8.0	9	−10.74	−2.0	−2.7	−3.2	0.91
8.4	10	−12.65	−2.4	−3.0	−3.4	0.93
9.0	21	−11.44	−2.5	−3.7	−3.7	0.87
9.4	36	−10.82	−2.8	−4.6	−3.9	0.78
8.4 + 50 mM KCl	60	−9.84	−3.5	−3.1	−3.8	0.95

B

pH	I (mM)	ψ_s^{HL} (mV)	Z_0^{HL}	Z_0^{CR}	Z_0^{LR}	$Z_0 - Z_1^{CR}$
7.4	51	−10.15	−2.9	−2.1	−3.7	0.89
8.0	57	−10.83	−3.2	−2.8	−3.9	0.95
8.4	59	−11.05	−3.3	−3.1	−3.9	0.95
9.0	60	−11.71	−3.5	−3.8	−4.3	0.88
9.4	61	−12.97	−3.9	−4.7	−4.9	0.78
10.0	62	−17.06	−5.2	−7.4	−7.7	0.58

[a] Ionic strength I is shown explicitly at each pH value. The value of $R = 2.05$ nm is used in CR and HL calculations.

closer to the upper limit of 2.5 nm) results in values of ΔZ^{HL} closer to −1 (calculations not shown). Another contribution to the value of ΔZ at high ionic strengths may be variations in hydrodynamic radius with ionic strength. Grigsby et al.[33] estimated the hydrodynamic radius of lysozyme using dynamic light scattering while changing the ionic strength of the buffer and found no clear trend. The observed changes in the value of the radius were attributed to the changes in hydration layer and to the binding of ions at the surface of the protein. These effects are dependent on the type of ion present and on its chaotropic or kosmotropic nature.[33] The complex dependence of the hydrodynamic radius, and thus of the drag on ionic strength, for which the Hückel model cannot account, may be the source of error in the values of Z_0 and ΔZ.

***Z*₀ and Δ*Z* as a Function of pH.** We evaluated the charge of native BCAII as a function of pH in two buffer systems, Tris−Gly (25 mM Tris−192 mM Gly) and phosphate (20 mM). We compared the values of charge on the native BCAII as estimated experimentally from the LR method, as calculated from the experimental mobility using the Hückel model, and as calculated using the theoretical CR model (Table 1). As expected, the charge of the native protein becomes more negative as the pH of the buffered solution increases. The largest discrepancy between the values of ΔZ^{CR} and the values of ΔZ^{LR} is observed at pH = 7.4, pointing at the increasing effect of charge compensation as the pH of the buffer approaches the value of pK_a of His residues. The same value of pH, but of different ionic strength, in two buffer systems results in substantially different values of charge calculated by LR, reaffirming the strong effect of the ionic strength on the charge of a protein. An experiment in which the ionic strength of Tris−Gly buffer (pH 8.4) was matched to that of phosphate (pH 8.4) by addition of KCl resulted in $Z_0 = -3.8$ in Tris−Gly (last line in Table 1A) and $Z_0 = -3.9$ in phosphate (third line in Table 1B), as estimated using LR. These values are within the experimental error of each other. We therefore conclude that the two buffer systems analyzed, Tris−Gly and phosphate, do not differ in their influence of Z_0 of BCAII and infer that the structure of the buffer may have a smaller effect than the pH on the value of ΔZ. In general, however, the charge of proteins with affinities for particular ions is affected by the buffer composition and the valence of the buffer ions.[34] LR analysis is capable of capturing these effects, whereas CR theory cannot explicitly account for an influence of the structure of the ions on the mobility of the protein.

It is interesting to note that as the pH of the Tris−Gly buffer system increases from 8.4 to 9.4 (by addition of NaOH), the surface potential becomes less negative (Table 1A). When, however, Z_0 is calculated from the surface potential via the Poisson−Boltzmann equation (eq 3), the calculation results in increased net charge. We attribute this result to the significant increase in ionic strength—and therefore, shielding—in the Tris−Gly buffer system with higher pH values. Thus, the effect of screening at higher ionic strength is indeed qualitatively captured by the Hückel theory. In the phosphate buffer system, the ionic strength is kept almost constant, and the Hückel calculations result in monotonically decreasing surface potential and charge with increasing pH. The LR method captures the trend of screening as well: the charge decreases monotonically as the pH and ionic strength increase in the Tris−Gly buffer system.

Estimating the Uncertainty of the Value of ΔZ^{CR}. To estimate the uncertainty in the variables that play a role in the CR model (in particular the error in the values of the protein radius R and the dissociation constants pK_a of ionizable residues), we used experimentally measured values of R and pK_a in other, better characterized, protein systems. As discussed earlier, the estimated uncertainty of the hydrodynamic radius R of BCAII is ±0.4 nm. This uncertainty translates into an error of ±0.02 in the value of ΔZ. To estimate the variations in the values of pK_a of ionizable amino acids, we looked at the values of pK_a of lysine residues in lysozyme, measured by hydrogen ion titrations.[25] The average pK_a of lysines in lysozyme was found to be 10.4 ± 0.2. We use the uncertainty of ±0.2 units in the pK_a for Lys, Tyr, and His residues (residues with pK_a closest to the working pH of 8.4) to estimate the uncertainty of ΔZ at pH of 8.4 for BCAII. This analysis yields an estimated uncertainty of ±0.01 introduced by the uncertainty of the pK_a values. Assuming that the uncertainties from radius and pK_a values are independent, we estimate the overall uncertainty in ΔZ^{CR} due to these terms to be ±0.02. Other factors not included in the analysis will increase this uncertainty.

Conclusion

We conclude (in agreement with Menon and Zydney[8]) that charge compensation influences the change in charge of a protein on acetylation in such a way that $|\Delta Z| < 1$. Our previous assumption[2,3,16] that $\Delta Z = -1.0$ is unjustified and probably wrong. HL and CR calculations give strong, if qualitative,

evidence for a smaller value of ΔZ ($\Delta Z = -0.93 \pm 0.02$ at pH = 8.4, $I = 10$ mM). The mechanism of charge compensation—that is, the allocation of compensating shifts in charge among changes in the values of pK_a of functional groups, changes in local pH, and changes in ionic screening—cannot be satisfactorily resolved with the experimental data or the theoretical tools that are available to us.

The approximations in both HL and CR approaches are sufficiently large that neither can be assumed to give quantitatively correct answers. The complexities of the subject—an irregularly shaped, hydrated protein with heterogeneous surface charge distribution and heterogeneous low dielectric core, moving in a nonideal ionic solution, under the influence of a high electric field—make it impractical for us to try to construct a full theoretical treatment. To do so would require solving the full Poisson−Boltzmann and Navier−Stokes equations and accounting for hydration and ion binding. Work in this area has not resulted in full agreement with experimental data at all conditions and pH ranges.[35,36] Simplified HL and CR models are not believable at the level of precision required to give a quantitative value to ΔZ.

Estimation of ΔZ from CR analysis suggests a value of −0.93 per acylation for BCAII at pH = 8.4, 10 mM ionic strength. Carrying out the LR analysis with such ΔZ yields better agreement in Z_0 between all models than with $\Delta Z = -1.0$. One difference between $\Delta Z = -1.0$ and $\Delta Z = -0.93$ is plausibly a small shift in the pK_a of one or more histidines or other ionizable residues with pK_a values near the value of the working pH, such as Lys, Tyr, and Zn−OH. Analysis of estimated curves of ΔZ as a function of pH suggests that ΔZ will be at a local minimum at pH = 6.0. This minimum is due to histidines strongly participating in charge compensation near their pK_a value.

Elimination of the effect of charge compensation (i.e., to achieve $|\Delta Z| > 0.99$) would require that no value of pK_a lie within three units of the working pH, according to CR calculations for BCAII. It would be worthwhile to use site-directed mutagenesis to replace His groups with neutral amino acids, for example, to compare the behavior of a protein with and without contributions of His to charge compensation. We have not done these experiments, and the capability of CE to resolve these differences is yet to be established.

This analysis suggests (but does not prove) that values of charge (e.g., Z_0) extracted from charge ladders by the method of LR[1,16] and assuming $\Delta Z = -1.0$ will be too high in magnitude by approximately 10%. In any event, the method does not provide a quantitative estimation of Z_0. It seems unlikely that theoretical analysis will allow the extraction of a more accurate value; the problem is simply too complex. Until some method is developed to calibrate ΔZ experimentally, estimation of Z_0 must be considered semiquantitative rather than quantitative. We believe that estimation of Z_0 at pH ∼ 8.5, taking $\Delta Z \sim -0.90$, is most likely to yield minimum error in the analysis of BCAII. The exact value of ΔZ for other proteins is likely to depend on their particular three-dimensional structure and proximity of charged residues to the one being modified.

For most applications of charge ladders, the uncertainty that emerges from this analysis is not problematic. The change in mobility that occurs on acylation of a lysine ϵ-NH_3^+ group gives a measure of the change in the net charge that results in adding (subtracting) one full unit of charge at one localized functional group. The complex network of electrostatic interactions that is characteristic of proteins decreases this change in charge. Addition of a unit of charge by some other mechanism (ligand binding, chemical modification, backbone or side chain cleavage, binding of surfactants) will have the same effect. The shifts in mobility due to these processes are interpretable in terms of the unit of electrostatic response ($\Delta\mu$ and ΔZ) from acetylation.

Experimental Section

Materials. BCAII (pI 5.9, E. C. 4.2.1.1), acetic anhydride, dioxane, and *p*-methoxybenzyl alcohol were purchased from Sigma-Aldrich (St. Louis, MO). Fused silica capillaries were purchased from Polymicro technologies (Phoenix, AZ). NICK spin columns containing G-50 Sephadex gel were purchased from Pharmacia Biotech (Piscataway, NJ). The pH of Tris−Gly buffer (25 mM Tris−192 Gly) was adjusted with 1 M acetic acid or 1 N sodium hydroxide for pH studies. Phosphate buffers, 20 mM monobasic, 20 mM dibasic, and 20 mM tribasic, were mixed in necessary proportions to achieve the desired pH between 7.4 and 10.

Acetylation of BCAII. BCAII was dissolved in 0.01 N NaOH at a concentration of 0.1 mM. Solutions of acetic anhydride (20 and 50 mM) were made in dioxane. These reagents were prepared immediately before use. Two aliquots of the protein solution (100 μL each) were reacted with 5 μL of the acetic anhydride solutions, resulting in 10- and 25-fold excess of acetic anhydride per batch, thus favoring the earlier and the later rungs of the ladder. The reactions were allowed to go to completion (∼30 min), and modified BCAII was purified on NICK spin columns. The two BCAII samples were mixed together and diluted in electrophoresis buffer (∼10 μM) for analysis.

CE. CE experiments were conducted on a Beckman P/ACE 5500 system. Uncoated fused silica capillaries (47 cm total length, 40 cm to the detector, 50 μm ID) were used for analysis. Samples were pressure-injected at 0.5 psi for 4−6 s and separated at an applied voltage of 15 kV. For the experiments conducted at high ionic strength (100 mM KCl), the applied voltage was reduced to 5 kV to minimize Joule heating. An electrostatically neutral marker, *p*-methoxybenzyl alcohol (100 mM), was added to each sample for measurements of electroosmotic flow. Detection was done by UV absorbance at 214 nm.

Acknowledgment. This work was supported by the National Institute of Health Grant GM 51559. I.G. acknowledges the National Science Foundation Graduate Research Fellowship for support. M.M. acknowledges the Swiss National Science Foundation for support. We are grateful to Prof. Andrew Zydney (Pennsylvania State University) for raising the issue of charge regulation in proteins and for a helpful discussion during the progress of this work.

References and Notes

(1) Carbeck, J. D.; Colton, I. J.; Anderson, J. R.; Deutch, J. M.; Whitesides, G. M. *J. Am. Chem. Soc.* **1999**, *121*, 10671−10679.
(2) Gao, J.; Gomez, F. A.; Härter, R.; Whitesides, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12027−12030.
(3) Gao, J.; Whitesides, G. M. *Anal. Chem.* **1997**, *69*, 575−580.
(4) Carbeck, J. D.; Negin, R. S. *J. Am. Chem. Soc.* **2001**, *123*, 1252−1253.
(5) Gao, J.; Mammen, M.; Whitesides, G. M. *Science* **1996**, *272*, 535−537.
(6) Caravella, J. A.; Carbeck, J. D.; Duffy, D. C.; Whitesides, G. M.; Tidor, B. *J. Am. Chem. Soc.* **1999**, *121*, 4340−4347.
(7) Negin, R. S.; Carbeck, J. D. *J. Am. Chem. Soc.* **2002**, *124*, 2911−2916.
(8) Menon, M. K.; Zydney, A. L. *Anal. Chem.* **2000**, *72*, 5714−5717.
(9) Keh, H. J.; Ding, J. M. *Langmuir* **2002**, *18*, 4572−4583.

(10) Pujar, N. S.; Zydney, A. L. *J. Colloid Interface Sci.* **1997**, *192*, 338−349.
(11) Creighton, T. E. *Proteins. Structure and Molecular Properties*; W. H. Freeman and Co: New York, 1993.
(12) Hunter, R. J. *Foundations of Colloid Science*; Oxford University Press: New York, 1991; Vol. 1.
(13) Linderstrøm-Lang, K. C. *Compt. Rend. Trav. Lab. Carlsberg* **1924**, *15*, 1−29.
(14) Tanford, C. *Adv. Protein Chem.* **1962**, *17*, 69−161.
(15) Grossman, P. D. *Capillary Electrophoresis*: *Theory and Practice*; Academic Press: San Diego, 1992.
(16) Colton, I. J.; Anderson, J. R.; Gao, J.; Chapman, R. G.; Isaacs, L.; Whitesides, G. M. *J. Am. Chem. Soc.* **1997**, *119*, 12701−12709.
(17) At these values of pH and pK_a, the lysine NH_2 group is 98.8% protonated.
(18) The double layer is a diffuse cloud of ions carrying the charge opposite to that of the charged particle.
(19) Probstein, R. F. *Physicochemical Hydrodynamics*; John Wiley & Sons: New York, 1994.
(20) The variable *x* is a dummy variable used to define the integral.
(21) Beychok, S.; Warner, R. C. *J. Am. Chem. Soc.* **1959**, *81*, 1892−1897.
(22) At pH 8.4 in Tris−Gly buffer, the electrostatic potential of the first four rungs of the BCAII ladder is below |25 mV|. The electrostatic potential of the later rungs lies above |25 mV|. In general, in other experimental conditions presented here, the potential of first six rungs is below |25 mV|.
(23) Yang, A.; Gunner, M. R.; Sampogna, R.; Sharp, K.; Honig, B. *Proteins*: *Struct., Funct., Genet.* **1993**, *15*, 252−265.
(24) Beroza, P.; Case, D. A. *J. Phys. Chem.* **1996**, *100*, 20156−20163.
(25) Kuramitsu, S.; Hamaguchi, K. *J. Biochem.* **1980**, *87*, 1215−1219.
(26) Kuehner, D. E.; Engmann, J.; Fergg, F.; Wernick, M.; Blanch, H. W.; Prausnitz, J. M. *J. Phys. Chem. B* **1999**, *103*, 1368−1374.
(27) The ionic strength of a buffer is calculated by

$$I = 1/2 \sum_i c_i n_i^2$$

where c_i is the concentration of the ionic species *i* and n_i is the valency of the ion *i*. The fraction of ionized weak acids or bases at a particular pH is found from acid−base equilibrium and K_a values.
(28) Average partial specific volume is calculated from the values of partial specific volumes of 20 proteins given in Creighton, T. E. *Proteins. Structure and Molecular Properties*; W. H. Freeman and Co: New York, 1993; p 266.
(29) Svergun, D. I.; Richard, S.; Koch, M. H. J.; Sayers, Z.; Kuprin, S.; Zaccai, G. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2267−2272.
(30) Skouri, M.; Delsanti, M.; Munch, J.; Lorber, B.; Giege, R. *FEBS Lett.* **1991**, *295*, 84−88.
(31) The two charges, located on the protein, interact through the low dielectric core of the protein as well as through the high dielectric solvent medium. The screening of charge−charge interactions will, therefore, be less than for small, separate ions in solution.
(32) The following values were used for the ionizable residues (quantity and pK_a value) for BCAII: C-term: 1, 3.2; Asp: 19, 3.5; Glu: 11, 4.5; His 8, 6.2; Zn^{2+}−OH: 1, 7; Lys: 18, 10.3; Tyr: 7, 10.3; Arg: 9, 12.5 (Menon, M. K.; Zydney, A. L. *Anal. Chem.* **2000**, *72*, 5714−5717). The N-terminal amino group is acetylated in the native form of BCAII, and three of its 11 His groups are complexed with the Zn^{2+} ion. The slight discrepancy in the value of ΔZ_{max} between our and Zydney's calculations comes from the number of His and N-terminal acetylations used. We thank Prof. Zydney for helping us determine the reasons for this discrepancy.
(33) Grigsby, J. J.; Blanch, H. W.; Prausnitz, J. M. *J. Phys. Chem. B* **2000**, *104*, 3645−3650.
(34) Menon, M. K.; Zydney, A. L. *Anal. Chem.* **1998**, *70*, 1581−1584.
(35) Allison, S. A.; Tran, V. T. *Biophys. J.* **1995**, *68*, 2261−2270.
(36) Allison, S. A.; Potter, M.; McCammon, J. A. *Biophys. J.* **1997**, *73*, 133−140.