Correlation between the Charge of Proteins in Solution and in the Gas Phase Investigated by Protein Charge Ladders, Capillary Electrophoresis, and Electrospray Ionization Mass Spectrometry

Jeffrey D. Carbeck,[†] **Joanne C. Severs,**[‡] **Jinming Gao,**[†] **Qinyuan Wu,**[‡] **Richard D. Smith,**[*,‡] **and George M. Whitesides**[*,†]

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, and Environmental Molecular Science Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99352

Received: January 13, 1998; In Final Form: July 13, 1998

Charge ladders of bovine carbonic anhydrase II, hen egg-white lysozyme, and bovine pancreatic trypsin inhibitor, prepared by partial acetylation of primary amino groups on the surface of the protein, have been analyzed by capillary electrophoresis (CE) and on-line electrospray ionization mass spectrometry (ESIMS) using solution conditions that maintain the native structure of the protein. CE was used to separate the proteins that constitute the charge ladder into individual "rungs"—protein derivatives that have the same number of acetylated amino groups and approximately the same net charge in solution. ESI was used to produce ions in the gas phase of the proteins that constitute each rung of the charge ladder; the mass spectra of these ions were obtained and analyzed. The distributions in charge states observed in the gas phase for the groups of proteins comprising each rung of the charge ladders were narrow, consistent with the retention of a compact structure of the proteins in the gas phase, and substantially independent of the number of acetylated amino groups. The ions observed in the gas phase had surface charge densities in a relatively narrow range of ~0.9−1.5 units of charge per 10^3 Å^2 of surface area (as estimated from crystallographic structures). These results demonstrate that the distribution of charge states for proteins produced in the gas phase by ESI do not necessarily reflect the net charge of the protein in solution or the number of amino groups on the protein.

Introduction

Electrospray ionization mass spectrometry (ESIMS) has rapidly become an important tool in the characterization of biopolymers. The ability to produce intact multiply charged ions of molecules with high molecular weights makes this technique applicable to large proteins and DNA segments.[1] A distinctive feature of ESI is that the conditions used to produce ions can be controlled to generate ions of proteins in the gas phase that appear to retain aspects of their solution-phase structure and noncovalent associations.

The ESI process produces ions of proteins—and other macromolecules—in the gas phase with a distribution in net charge typically due to variation in the number of protons associated with the macromolecule. This distribution in net charge results in a series of peaks when separated by MS. Each peak has a unique value of m/z, where m is the mass of the ion and z is its charge. We refer to the individual values of z observed in the gas phase as a charge state and the collection of values of z observed in the gas phase as the distribution of charge states.[2]

The physical and chemical origins of the distribution of charge states of proteins produced in the gas phase by ESI are unknown; several trends have, however, been observed. Both the magnitude of charging and the breadth of the distribution of charge states of ions of proteins produced in the gas phase by ESI are sensitive to the conformation of the protein in solution (i.e., native or denatured) from which the proteins are electrosprayed.[3,4] Early ESIMS studies of proteins typically employed solution conditions (e.g., 1−5% acetic acid) that resulted in the denaturation of the proteins. These conditions produced positively charged ions in the gas phase with a broad distribution of charge states and a maximum charge state that correlated well with the total number of sites having the largest proton affinities: normally, for proteins, Arg, Lys, His, and the N-terminal α-amino group.[4,5] Recent work indicated that gas-phase reactions, such as proton transfer, may also influence the distribution of charge states observed in the gas phase.[6] Several reports suggested a correlation between the net charge of proteins in solution and the maximum charge state of ions of proteins produced in the gas phase by ESI.[7−12] Significant deviations from these trends have been noted, and the origin of the distribution of charge states remains a subject of debate.[7−14]

Experimental studies designed to establish the relationship between the distribution of charge states of proteins in the gas phase and the location and nature of their ionizable sites have been complicated by the need to use different proteins with obvious structural differences. Possible approaches to reduce variations in the structure of model proteins for such studies—for example, the use of site-directed mutagenesis to produce individual protein variants having essentially the same structure but with different numbers of ionizable sites or values of net charge in solution—would constitute substantial efforts, and such systematic studies have not yet been conducted.

---

* Authors to whom correspondence should be addressed. Richard D. Smith, Phone: (509) 376-0723, Fax: (509) 376-5824, E-mail: rd_smith@pnl.gov; and George M. Whitesides, Phone: (617) 495-9430, Fax: (617) 495-9857, E-mail: gwhitesides@gmwgroup.harvard.edu.
† Harvard University.
‡ Pacific Northwest National Laboratory.

Charge of Proteins in Solution and in the Gas Phase

J. Phys. Chem. B, Vol. 102, No. 51, 1998 **10597**

Charge ladders[15] of proteins provide a simple means of obtaining a large number of derivatives of proteins that differ in the number of ionizable sites, but that have similar tertiary structures. The modification of a distribution of the ionizable groups on a protein in a way that annihilates their charge results in the formation of a distribution of protein derivatives differing in values of charge in solution, but having similar coefficients of friction.[16] The separation of this distribution of proteins in the solution phase by capillary electrophoresis (CE) results in families of protein derivatives that group, in terms of their electrophoretic mobility, into the "rungs" of a charge ladder. Each rung of the ladder consists of proteins with the same number of modified ionizable groups. Charge ladders can be conveniently generated by the treatment of proteins with acetic anhydride;[16] the resulting acetylation reactions convert positively charged Lys ϵ-NH$_3^+$ groups and the N-terminal α-NH$_3^+$ group to neutral *N*-acyl derivatives. Modification of other ionizable groups on a protein (e.g., $-N\equiv C(NH_2)_2^+$ on Arg, $-COOH$ on Glu and Asp) using other classes of reagents can also lead to protein charge ladders.[17] Examination of charge ladders using CE makes it possible to estimate the net charge of the native protein in solution, as well as charges of the protein derivatives that make up the charge ladder.[16,18]

In this work, we use the charge ladders of three proteins—bovine carbonic anhydrase II (BCA II, E. C. 4.2.1.1), bovine pancreatic trypsin inhibitor (BPTI) and hen egg-white lysozyme (HEL, E. C. 3.2.1.17)—produced by partial acetylation of primary amino groups with acetic anhydride to study the relationship between the number of amino groups on the proteins, the net charge of the proteins in solution, and the distribution of charge states of ions of the proteins produced in the gas phase by ESI. We use CE to separate the proteins that constitute a charge ladder into its individual rungs in the solution phase, on-line ESI to produce ions of these proteins in the gas phase, and MS to measure the mass and charge of these ions. CE–ESIMS was performed using an electrophoresis buffer and "gentle" ESI interface conditions that were intended to maintain the native structure of the protein.[3] The distributions of charge states observed in the gas phase were compared directly with the net charges of the proteins in solution, estimated by CE.[16] Our analyses confirm that each rung of the charge ladder consists of proteins with the same number of acetylated amino groups and give evidence that the compact structure of the native proteins in solution is preserved across the rungs of the charge ladders. This work also shows that the distribution of charge states of protein ions in the gas phase does not necessarily reflect either the net charge of the native protein in solution or the number of available amino groups.

Experimental Section

Formation of Protein Charge Ladders. All proteins were purchased from Sigma (St. Louis, MO) and used without additional purification. The initial total protein concentrations used for formation of the ladders were BCA II, 15 mg/mL; BPTI, 20 mg/mL; and HEL, 15 mg/mL in deionized water adjusted to pH 12 by the addition of 10 vol. % of 0.1 N NaOH. Acetic anhydride (5–20 equiv; 100 μM in dioxane) was added to the solutions of proteins, and the reactants were mixed by vortexing. Each rung in the final charge ladders was controlled to have relatively the same intensity by mixing the products from several acetylations done with different equivalents of acetic anhydride. The charge ladders were purified on NICK spin columns (Pharmacia Biotech) before analysis by CE–ESIMS.

Capillary Electrophoresis–Electrospray Ionization Mass Spectrometry. A Thermo Crystal 310 capillary electrophoresis system (Thermo CE, Franklin, MA) interfaced to a Finnigan TSQ 7000 triple quadrupole mass spectrometer equipped with an electrospray ionization interface (Finnigan MAT, San Jose, CA) was used for these studies. A CE–MS interface based on electrical contact across a micro-dialysis junction connecting the CE capillary and ESI emitter allowed operation under nondenaturing solution conditions.[19] The 200 μm i.d. polysulfone dialysis tubing (nominal molecular weight cutoff of 10 000) used for the microdialysis junction was obtained from A/G Technology Corporation (Needham, MA) and the fused silica capillary of dimensions 192 μm o.d. $\times$ 30 μm i.d. $\times$ 65 cm was from Polymicro Technologies Inc. (Phoenix, AZ). One end of a 2 cm segment of capillary was etched in 40% hydrofluoric acid for 30 min to produce a sharply tapered ESI emitter, and then coupled to the separation capillary by the microdialysis junction. An electrophoresis buffer of 10 mM ammonium acetate buffer was prepared from ammonium hydroxide and glacial acetic acid (Sigma, St. Louis, MO) using deionized distilled water from a Nanopure II water system (Barnstead, Dubuque, IA). The pH of the buffer was adjusted to 8.3 for the separation of the charge ladders of BPTI and HEL, and to 9.0 for the separation of the charge ladder of BCA II. The resolution of the rungs of the charge ladder of BCA II by CE decreases below a pH of 8.5 in ammonium acetate buffer. Samples were injected into the capillary, held at 27 °C, with 100 or 200 mbar pressure for 6 s. An ESI voltage of 1.6 kV was applied to the interface junction after the CE voltage had been applied. The ESI inlet capillary was set at 160 °C, corresponding to relatively gentle conditions compared to inlet temperature of 200 °C conventionally used for ESIMS with this instrument.

To avoid sample degradation, evident from changes in mass in the mass spectra as well as degradation of the quality of separation by CE, the ladders were analyzed within a day of production and the samples were stored at 4 °C between runs. At pH 9.0, BCA II is negatively charged and bare fused silica capillaries were used with the standard arrangement of the cathode for CE at the outlet of the capillary. Interactions between the proteins that are positively charged at pH 8.3—BPTI and HEL—and the negatively charged walls of the silica capillary were avoided by the noncovalent coating of the interior of the capillary with a polycationic polymer, Polybrene (Aldrich, Milwaukee, WI).[20] This coating reversed the charge at the wall of the capillary, and consequently the cathode for CE was located at the inlet of the capillary to maintain the electroosmotic flow toward the ESI–mass spectrometer interface.

Results and Discussion

CE–ESIMS of BCA II. We performed CE–ESIMS analysis of the charge ladder of BCA II (MW 29 kD, comprising two isozymes with values of pI of 5.4 and 5.9) using an electrophoresis buffer of 10 mM ammonium acetate (pH 9.0) and a bare, fused silica capillary. Native BCA II has 18 Lys, 9 Arg, and 11 His residues;[18] the N-terminal α-amino group of native BCA II is acetylated. The separation of the proteins that constitute the charge ladder of BCA II by CE results in 19 observed peaks corresponding to the native BCA II, and the partially acetylated derivatives having from one to 18 of the Lys ϵ-amino groups acetylated (Figure 1a). The positive-ion ESI mass spectra obtained for the protein derivatives that constitute each rung of the charge ladder of BCA II show that
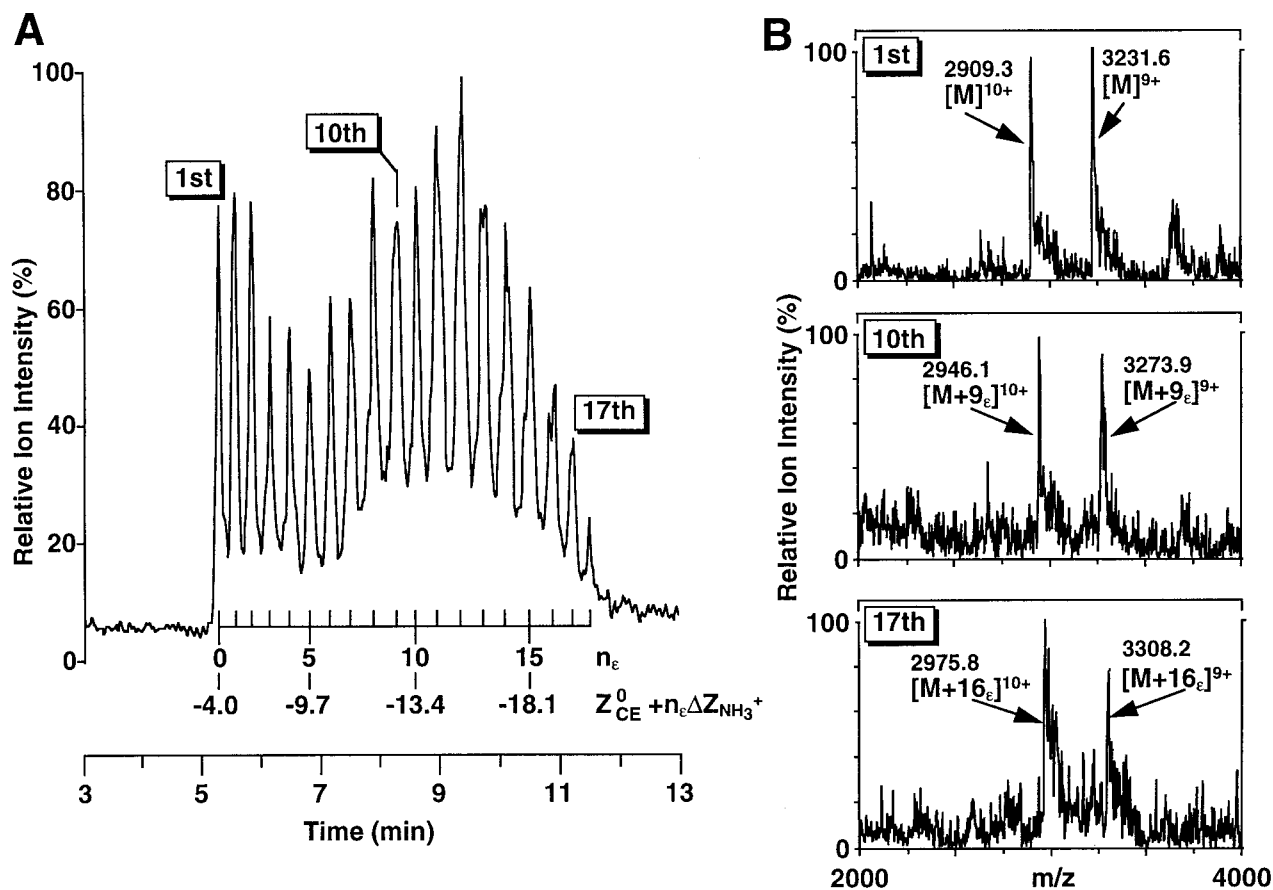
Figure 1. (A) CE−ESIMS electropherogram of a charge ladder of BCA II formed by partial acetylation of Lys ϵ-amino groups with acetic anhydride. The number of acetylated Lys ϵ-amino groups, n_ϵ, as well as the estimated charge, $Z_{CE}^0 + n_\epsilon\Delta Z_{NH3}^+$, of the proteins that make up each of the rungs of the ladder is indicated below the electropherogram, where Z_{CE}^0 is the charge of the native protein estimated by CE, and ΔZ_{NH3}^+ is the change in charge due to the acetylation of a Lys ϵ-amino group (assuming a pK_a of 10.2, $\Delta Z_{NH3}^+ = -0.94$ at pH 9.0). Separations were performed at 25 °C in 10 mM ammonium acetate buffer (pH 9.0) with an applied field of 437 V/cm for 3 min, reduced to 206 V/cm for the remainder of the separation. The emerging proteins were detected using ESIMS by scanning m/z from 2900 to 3390 at 1 s/scan. (B) ESI mass spectra obtained from the accumulation of scans for the first, 10th, and 17th peaks to emerge after separation by CE. Ten scans from m/z 2000 to 4000 were summed for each peak. Each peak is labeled with its value of m/z and identified using the notation $[M + n_\epsilon]^{\pm z}$. The quantity in brackets represents the net mass of the ion, where M is the mass of the native protein, and n_ϵ is the number of acetylated Lys ϵ-amino groups; the superscript $\pm z$ represents the charge of the ion in the gas phase.

the charges of the ions generated in the gas phase from the protein derivatives that make up the charge ladder of this protein do not vary with the number of acetylated Lys ϵ-amino groups; the +9 and +10 charge states predominate in the distribution (Figure 1b).

We used CE to estimate the charge of native BCA II, $Z_{CE}^0 = -4.0$ in ammonium acetate buffer at pH 9.0.[16] We assume the annihilation of charge associated with Lys ϵ-NH$_3^+$ groups increases the negative charge of the protein in solution and thereby increases its electrophoretic mobility relative to the native protein. Previously, using this assumption, we assigned the first peak to emerge by CE to native BCA II and the subsequent 18 peaks to the partially acetylated derivatives with the number of acetylated Lys ϵ-amino groups, n_ϵ, increasing from one to 18.[18,21] ESIMS analysis of the protein derivatives that make up each rung of the ladder confirms this assignment. The mass spectra in Figure 1 show that each rung of the ladder separated by CE in solution corresponds to a set of regio-isomeric derivatives of BCA II that have the same value of n_ϵ, and that n_ϵ increases in units of one from 0 to 18 across the rungs of the charge ladder.

The charge of the proteins that constitute each rung of the charge ladder in solution is estimated from the charge of the native protein, Z_{CE}^0, and the change in charge due to the acetylation of n_ϵ Lys ϵ-amino groups, $n_\epsilon\Delta Z_{NH3}^+$. Using a value of 10.2 for the pK_a of the ϵ-amino group of Lys, ΔZ_{NH3}^+ is estimated to be -0.94 at pH 9.0. The charge of the proteins in solution that make up the charge ladder of BCA II estimated in this way vary from -4.0 to ~-21. From the mass spectra we observe that the charge of the ions generated in the gas phase from the protein derivatives that make up the charge ladder of BCA II do *not* correlate with the charge of the protein derivatives in solution.

CE−ESIMS of HEL and BPTI. We performed CE−ESIMS analysis of the charge ladder of HEL (MW 14.3 kD; pI 11.1) using an electrophoresis buffer of 10 mM ammonium acetate (pH 8.3) and a fused silica capillary coated noncovalently with the cationic polymer, Polybrene. Native HEL has 11 Arg, 6 Lys, and 1 His residues. The separation of the charge ladder of HEL by CE produced seven peaks corresponding to the native HEL, and the six partially acetylated derivatives (Figure 2a). The positive-ion ESI mass spectra obtained for the protein derivatives that constitute each rung of the charge ladder of HEL, again, show that the charge of the ions generated in the gas phase from the protein derivatives that make up the charge ladder of HEL do not vary with the number of acetylated Lys ϵ-amino groups; the +7 and +8 charge states predominate the distribution (Figure 2b).
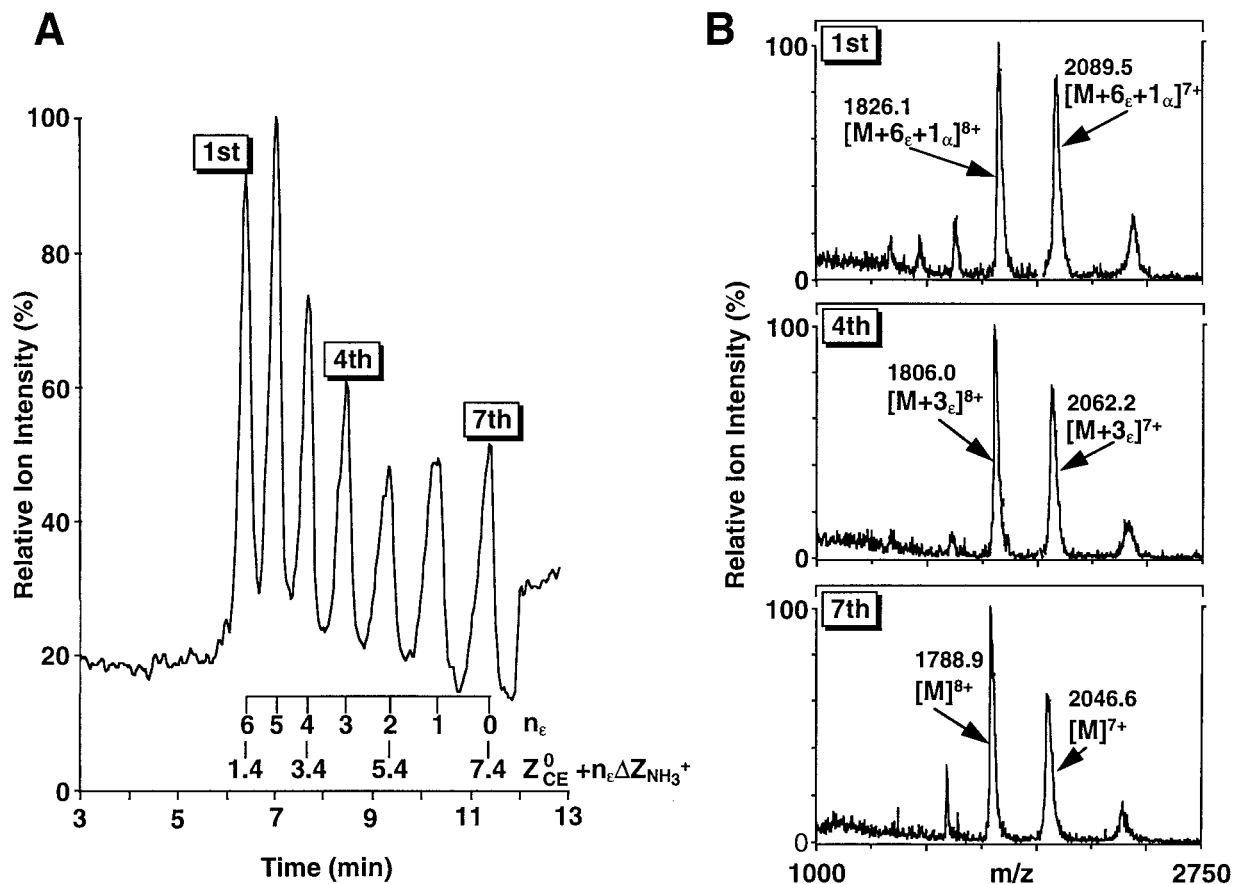
Charge of Proteins in Solution and in the Gas Phase

J. Phys. Chem. B, Vol. 102, No. 51, 1998 **10599**

Figure 2. (A) CE–ESIMS electropherogram of a charge ladder of HEL formed by partial acetylation of Lys ϵ-amino groups with acetic anhydride. The number of acetylated Lys ϵ-amino groups, n_ϵ, as well as the estimated charge, $Z_{CE}^0 + n_\epsilon \Delta Z_{NH3}^+$, of the proteins that make up each of the rungs of the ladder are indicated below the electropherogram. ΔZ_{NH3}^+ is assumed equal to -1.0 at pH 8.3. Separations were performed at 25 °C in a silica capillary coated with Polybrene using 10 mM ammonium acetate buffer (pH 8.3) with an applied field of 437 V/cm. The emerging proteins were detected using ESIMS by scanning m/z from 500 to 2500 at 2 s/scan. (B) ESI mass spectra shown were obtained from the accumulation of scans for the first, fourth, and seventh peaks to emerge after separation by CE. Ten scans from m/z 1000 to 2750 were summed for each peak. Each peak is labeled with its value of m/z and identified using the notation $[M + n_\epsilon + n_\alpha]^{\pm z}$. The quantity in brackets represents the net mass of the ion, where M is the mass of the native protein, and n_ϵ and n_α are the number of acetylated Lys ϵ-amino groups and N-terminal α-amino groups, respectively; the superscript $\pm z$ represents the charge of the ion in the gas phase.

Native HEL has a net charge of $+7.4$ at pH 8.4, as estimated by CE.[16] In contrast to negatively charged BCA II, the annihilation of charge associated with Lys ϵ-NH_3^+ groups *decreases* the positive charge of the protein and thereby *decreases* its electrophoretic mobility relative to the native protein. Previously, we assigned the first peak to emerge by CE to the per-acetylated HEL, and the subsequent six peaks to the partially acetylated derivatives of HEL with the number of acetylated Lys ϵ-amino groups, n_ϵ, decreasing from 6 to 0, with the last peak to emerge assigned to native HEL.[20] The mass spectra in Figure 2b show that each rung of the ladder separated by CE in solution corresponds to a set of regio-isomeric derivatives of HEL that have the same value of n_ϵ, and that n_ϵ decreases in units of one from 6 to 0 across the rungs of the charge ladder.

Reaction of HEL with acetic anhydride may also result in the acetylation of the N-terminal α-amino group. As a result, each rung of the charge ladder of HEL may be composed of two components: proteins with and without the N-terminal α-amino group acetylated. Using conventional CE with longer capillaries and electrophoresis buffers that are not compatible with ESIMS, the two components of each peak can be partially resolved.[20] The peaks in the mass spectra of Figure 2b are identified on the basis of the most abundant ion species. We observe that the first rung of the charge ladder to emerge by CE is comprised primarily of proteins with the N-terminal

α-amino group acetylated. In contrast, the fourth and seventh rungs of the charge ladder to emerge by CE contain little protein with the N-terminal α-amino group acetylated. From these measurements we infer that the probability of acetylation of the N-terminal α-amino group of HEL increases with the number of acetylated Lys ϵ-amino groups. Due to the relatively low mass resolution obtained with a quadrupole mass spectrometer and the possibilities of adduction of acetate ions to the protein, quantitative assignment of the two components was not practical. It is possible that all the rungs of the ladder actually include some proteins with the N-terminal α-amino group acetylated.

The charge ladder of BPTI (MW 6.5 kDa; pI 10.5) was also analyzed under the same conditions as those employed for the charge ladder of HEL. Native BPTI has 6 Arg and 4 Lys residues and 3 disulfide bridges that contribute to its compact tertiary structure;[22] the N-terminal α-amino group of native BPTI is acetylated. Native BPTI has a net charge of $+5.5$ at pH 8.4, as estimated by CE.[16] Again, the mass spectra show that the charge of the ions generated in the gas phase from the protein derivatives that make up the charge ladder of BPTI do not vary with the number of acetylated Lys ϵ-amino groups; the $+4$ and $+5$ charge states are the predominant species. The temperature of the heated ESI inlet capillary was reduced from 160 to 110 °C to see if the distribution of charge states varied on altering

TABLE 1: Correlation of Charge States of Proteins in the Gas Phase with the Charge of the Protein in Solution Estimated by CE, Z^0_{CE}, the Number of Basic Residues, the Molecular Surface Area, and the Charge per Molecular Surface Area

protein	Z^0_{CE} [a]	no. of basic residues	molecular surface area ($\times 10^{-3}$ Å^2)[b]	charge states observed in the gas phase	charge per molecular surface area ($\times 10^3$ Å^{-2})
BPTI	+5.5	Arg(6) Lys(4) His(0)	3.46	+4, +5	1.2, 1.4
HEL	+7.4	Arg(11) Lys(6) His(1)	5.50	+7, +8	1.3, 1.5
BCA II	−4.0	Arg(9) Lys(18) His(11)	10.1[c]	+9, +10	0.89, 0.99

[a] The charge of BPTI and HEL were estimated in 25 mM Tris−192 mM Gly buffer, pH 8.4, and the charge of BCA II was estimated in 10 mM ammonium acetate buffer, pH 9.0, using a combination of CE and charge ladders.[16] [b] The molecular surface area for each of the native proteins was calculated from its crystal structure obtained from the Brookhaven Protein Data Base using Quanta and the CHARMM parameter set. [c] The molecular surface area of BCA II was estimated from the crystal structure of the human CA II isozyme. This isozyme has a distribution of charge states in the gas produced by ESI that are similar to those of BCA II (i.e., +9, +10).

the conditions of ion desolvation. The same distribution of charge states was obtained at both temperatures.

Distribution of Charge States of Protein Charge Ladders Produced by ESI. Previously, we observed that the binding affinities of the proteins that constitute the charge ladder of BCA II for neutral inhibitors are independent of the number of acetylated Lys ϵ-amino groups, and inferred that acetylation of the Lys ϵ-amino groups of BCA II does not profoundly alter the native structure of its active site in solution.[21] In this work, we observe that the distributions of charge states of proteins that make up the charge ladder of BPTI, BCA II, and HEL produced in the gas phase by ESI were relatively narrow: all three proteins gave just two predominant charge states in the gas phase (an observation that has previously been correlated with solution conditions that maintain the native conformation of proteins[3]). We infer that acetylation of the primary amino groups does not induce major changes in the structure of these three proteins in solution under the conditions used.

The charge of the protein ions produced in the gas phase is a measure of the extent of protonation of the basic groups of the protein, which, in turn, depends on the numbers and affinities of these groups for protons in the gas phase: that is, their values of gas-phase basicity, defined as the negative of the free energy of protonation. The total, or apparent gas-phase basicity is consider to be the sum of an intrinsic basicity and the contributions from the electrostatic interactions of the other charged sites on the molecule. The side chains with the largest values of intrinsic gas-phase basicity are those of Arg, Lys, and His. The side chains of Arg clearly have the largest value of intrinsic gas-phase basicity; there are conflicting reports of the relative values of intrinsic gas-phase basicity of the side chains of Lys and His.[6,23] Previous calculations of the electrostatic interactions between charged groups on protein ions in the gas phase show that the electrostatic interactions can contribute to more than half the value of the total gas-phase basicity of these side chains (on the order of 100 kcal/mol).[6,23]

The distributions of charge states produced by ESI were centered around +7 and +8 for HEL, +4 and +5 for BPTI, and +9 and +10 for BCA II. In Table 1, we compare the charge states observed in the gas phase with the charge in solution, the number of basic residues, and the molecular surface area for the native proteins used in this study. The protein ions produced in the gas phase by ESI using solution conditions that maintain the native structure of the protein and a "gentle" ESI interface have charge state distributions that do not depend on the net charge of the protein in solution. The magnitude of the charge states produced in the gas phase *increase* with the molecular surface area of the native protein; the charge per unit surface area of these three proteins is similar.
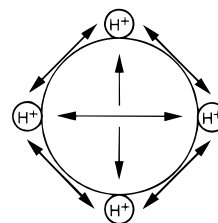
These results indicate that the values of gas-phase basicity, and hence the magnitude and distributions of charge states of protein ions produced in the gas phase, are probably determined

SCHEME 1: A Schematic Illustration Indicating the Mechanisms That Determine the Net Charge of a Protein in the Solution Phase and in the Gas Phase[a]
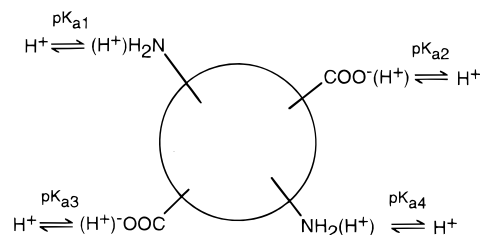
[a] In the solution phase, the net charge of a protein is determined largely by its *chemical* properties: the number and kind of ionizable groups, the ionization constant of these groups (as described by their values of pK_a), and the value of pH of the solution. The values of pK_a of these ionizable groups *are*, to some extent affected by the values of the local electrostatic potential arising from the other charged groups on the protein. In the gas phase, the values of the apparent gas-phase basicity of basic side chains and hence the net charge of ions of a protein produced by ESI are determined largely by its *physical* properties; that is, the molecular surface area of the protein and the Coulombic repulsion between charged sites on its surface.

largely by structural factors such as the density of charged groups (i.e., the distances between charged groups) on the surface of the protein (Scheme 1). Solution conditions that denature a protein and thus change its conformation (and, on average, increase the distance between protonatable groups) may also change the number of sites available for charging;[7] there is some evidence to support this view. These observation are consistent with the suggestion that the charge of protein ions in a given gas-phase environment is substantially defined by structurally dependent charge−charge repulsion as well as the nature of available charge sites: that is, the number and intrinsic values of gas-phase basicity of the basic residues.[2,3] We have previously suggested that repulsive Coulombic interactions generally become the most important factor in determining the charge states of a protein in the low dielectric environment of the gas phase. This view is also consistent with the recent results of Williams and co-workers,[6,24−28] and of Karishikoff et al.,[23] and is compatible with the observation that nondena-

Charge of Proteins in Solution and in the Gas Phase

J. Phys. Chem. B, Vol. 102, No. 51, 1998 **10601**

turing ESI conditions produce ions of proteins in the gas phase with smaller values of charge and narrower distribution of charge states, as compared to ions of proteins produced using denaturing ESI conditions. These smaller values of charge and narrower distributions of charges states are presumably due to greater Coulombic interactions between charged groups on the surface of the protein in its compact, native state.

The effect of acetylation of the primary amino groups of proteins on their values of charge in the gas phase is expected to depend on the number of these groups and on their values of gas-phase basicity. From Table 1, we observe that the total number of basic side chains for the three proteins studied here is greater than the largest value of charge in the gas phase, even when all of the primary amino groups are acetylated. Indeed, the number of Arg residues alone is enough to account for all but the largest observed charge state of BCA II in the gas phase. We infer that there are at least two reasons why the charge states in the gas phase do not depend on the number of amino groups on the protein: (i) the Lys ϵ-amino groups are not protonated significantly in the gas phase; (ii) acetylation of the Lys ϵ-amino groups alters the values of gas-phase basicity of the remaining protonatable groups such that the *pattern* of protonation changes upon acetylation, while the *number* of protonated groups remains constant. It is possible that the modification of basic residues of proteins other than Lys ϵ-amino groups may result in significant changes in the observed distributions of charge states in the gas phase.

Conclusions

CE−ESIMS is shown to be a useful tool for the study of charge ladders of proteins and for the examination of the relationship between the properties of proteins in the solution phase and the gas phase. CE−ESIMS confirms the previous assignment of the composition of the rungs of a charge ladder under similar solution conditions; each rung of the ladder is composed of regio-isomeric derivatives of proteins that have the same number of acetylated Lys ϵ-amino groups. The observation that only a few charge states are observed in the gas phase for the derivatives of the protein that constitute a charge ladder is consistent with retention of a compact structure of the protein derivatives resembling that of the unmodified protein.

We demonstrate that both the magnitude of charging and the breadth of the distribution of charge states of ions of proteins produced in the gas phase by ESI do not correlate with the charge of the proteins in solution or their number of amino groups. Instead, the magnitude of charge states appears to correlate with the available surface area of proteins. Ions of proteins produced in the gas phase by ESI had surface charge densities in a relatively narrow range: ∼0.9−1.5 units of charge per 10^3 Å^2 of surface area; these observations suggest that, in the gas phase, the values of gas-phase basicity of the basic residues, and, hence, the values of net charge of proteins are probably determined largely by electrostatic interactions between charged groups on the surface of proteins (Scheme 1).

References and Notes

(1) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F. *Mass Spectrom. Rev.* **1990**, *9*, 37−70.

(2) Smith, R. D.; Light-Wahl, K. J. *Biol. Mass Spectrom.* **1993**, *22*, 493−501.

(3) Light-Wahl, K. J.; Schwartz, B. L.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 5271−5278.

(4) Smith, R. D.; Loo, J. A.; Ogorzalek, R. R.; Busman, M.; Udseth, H. R. *Mass Spectrom. Rev.* **1991**, *10*, 359−452.

(5) Smith, R. D.; Loo, J. A.; Loo, R. R. O.; Busman, M.; Udseth, H. R. *Mass Spectrom. Rev.* **1992**, *11*, 431−443.

(6) Schnier, P. D.; Gross, D. S.; Williams, E. R. *J. Am. Chem. Soc.* **1995**, *117*, 6747−6757.

(7) Chowdhury, S. K.; Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1990**, *112*, 9012−9013.

(8) Kelly, M. A.; Vestling, M. M.; Fenselau, C. C.; Smith, P. B. *Org. Mass Spectrom.* **1992**, *27*, 1143−1147.

(9) Guevremont, R.; Siu, K. W. M.; LeBlanc, J. C. Y.; Berman, S. S. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 216−224.

(10) LeBlanc, J. C. Y.; Siu, K. W. M.; Guevremont, R. *Anal. Chem.* **1994**, *66*, 3289−3296.

(11) Anderegg, R. J.; Wagner, D. S. *Anal. Chem.* **1994**, *66*, 706−711.

(12) Ashton, D. S.; Beddell, C. R.; Cooper, D. J.; Green, B. N.; Oliver, R. W. A. *Org. Mass Spectrom.* **1993**, *28*, 721−728.

(13) LeBlanc, J. C. Y.; Wang, J.; Guevremont, R.; Siu, K. W. M. *Org. Mass Spectrom.* **1994**, *29*, 587−593.

(14) Wang, G.; Cole, R. B. *Org. Mass Spectrom.* **1994**, *29*, 419−427.

(15) Carbeck, J. D.; Colton, I. J.; Gao, J.; Whitesides, G. M. *Acc. Chem. Rev.* **1997**, *31*, 343−350.

(16) Colton, I. J.; Anderson, J.; Gao, J.; Chapman, R.; Isaacs, L.; Whitesides, G. M. *J. Am. Chem. Soc.* **1997**, *119*, 12701−12709.

(17) Anderson, J. R.; Colton, I. J.; Whitesides, G. M., unpublished.

(18) Gao, J.; Gomez, F. A.; Haerter, R.; Whitesides, G. M. *Proc. Natl. Acad. Sci.* **1994**, *91*, 12027−12030.

(19) Severs, J. C.; Harms, A. C.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1175−1178.

(20) Cordova, E.; Gao, J.; Whitesides, G. M. *Anal. Chem.* **1997**, *69*, 1370−1379.

(21) Gao, J.; Mammen, M.; Whitesides, G. M. *Science* **1996**, *272*, 535−537.

(22) Castro, M. J. M.; Anderson, S. *Biochemistry* **1996**, *3*, 11435−11446.

(23) Miteva, M.; Demirev, P. A.; Karshikoff, A. D. *J. Phys. Chem. B* **1997**, *101*, 9645−9650.

(24) Williams, E. R. *Trends Anal. Chem.* **1994**, *13*, 247−251.

(25) Gross, D. S.; Mills, P. D.; Tien, L. T.; Johnson, P. J.; Williams, E. R. *Proceedings of the 41st ASMS Conference on Mass Spectrometry and Allied Topics*, San Francisco, CA, 1993.

(26) Fang, L. L.; Zhang, R.; Williams, E. R.; Zare, R. N. *Anal. Chem.* **1994**, *66*, 3696−3701.

(27) Williams, E. R. *J. Mass. Spectrom.* **1996**, *31*, 831−842.

(28) Gross, D. S.; Williams, E. R. *J. Am. Chem. Soc.* **1995**, *117*, 883−890.