Rates of Proton Transfer from Carboxylic Acids to Dianions, $CO_2(CH_2)_pCO_2^{2-}$, and Their Significance to Observed Negative Charge States of Proteins in the Gas Phase†

Arthur Blades, Michael Peschke, Udo Verkerk, and Paul Kebarle*

Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

Received: May 31, 2002; In Final Form: August 9, 2002

The carboxylic acid dianions, $CO_2(CH_2)_pCO_2^{2-}$, are the simplest model for two deprotonated acidic side chains, such as Glu or Asp, which are on opposite sides of a nondenatured globular protein. Rate constant determinations of the charge reducing reaction, $CO_2(CH_2)_pCO_2^{2-} + AH = HCO_2(CH_2)_pCO_2^- + A^-$, involving dianions with C_n where n ranges from 7 to 16 ($n = p + 2$) with a variety of oxygen acids AH including acetic acid, show that charge reduction (loss) occurs at collision rates for all of the above reagents. This is in contrast with results for the positively charged proteins. Charge loss at collision rates in the model reaction (for two lysine side chains), $NH_3 + H_3N(CH_2)_pNH_3^{2+} = NH_3(CH_2)_pNH_2^+ + NH_4^+$, occurs only for C_n when $n < 7$ ($n = p$). These results provide an explanation for the lower charged states of nondenatured proteins in the negative ion mode, relative to the positive ion mode, observed in the literature when the proteins are sprayed from aqueous solution with ammonium acetate buffer. According to the charge residue model (CRM), if an ammonium acetate buffer is used, charging of the protein will occur via NH_4^+ in the positive ion mode and $CH_3CO_2^-$ in the negative ion mode. The much lower tolerance for proximity of another charge in proteins in the negative ion mode, revealed by the rate measurements of the dianions reacting with acetic acid, is due to the different effects of alkyl substitution on the intrinsic basicities in the positive ion mode and on the intrinsic acidities in the negative ion mode.

Introduction

The positive and negative charged states of proteins observed in electrospray ionization mass spectrometry (ESIMS) are of great interest. Thus denatured proteins lead to higher charged states than those observed for the native protein, and this provides a qualitative measure of the degree of denaturing that has occurred.[1] The charge state is found also to depend on other electrolytes present in the solution, such as the chemical nature of the buffer salts, even when the pH is the same. For example, use of a buffer that uses alkylammonium ions such as triethylammonium leads to much lower charged states than use of an ammonium ion containing buffer.[2] Such manipulation allows a certain degree of control of the charge state.[2c] It is also observed that the charge states are lower in the negative ion mode relative to the charge states observed in the positive mode for the same protein in the same solution.[3]

An understanding of the above phenomena can be obtained only on the basis of the mechanism that leads to charging of the proteins in ESIMS. We will argue (see Results and Discussion) that the charging of the proteins is due to the small ions that are at the surface of the small droplets containing the protein and that there is a limit to the charges that can be provided to the protein by this mechanism.

Another property that is of obvious importance in determining the charge states is the (maximum) number of charges (protons) that the protein can hold when it enters the gas phase. This quality depends on the apparent basicity and number of the basic side chains (positive ion mode) that are near the surface of the protein. A protein that has $Z - 1$ protonated basic side chains could hold one more charge due to protons, if the apparent basicity, GB^{app}, defined below, of at least one of the not-yet-charged basic side chains is higher than a given value. Thus, when the charging is due to an NH_4^+ ion, the GB^{app} should be higher than the gas-phase basicity of ammonia, $GB(NH_3)$, assuming that the charging reaction occurs in a gas-phase-like medium. The same condition will hold if the protein has already Z protons and there is an NH_3 molecule that is hydrogen-bonded to the protonated side chain.

The diprotonated alkyl diamines[4−6] are a simple two charge model that is very useful for treatments dealing with the determination of the GB^{app} of polyprotonated proteins.

$$NH_3 + H_3N(CH_2)_pNH_3^{2+} = (NH_3(CH_2)_pNH_3 \cdot NH_3)^{2+} = NH_3(CH_2)_pNH_2^+ + NH_4^+ \quad (1)$$

The diamine can be considered as the simplest model for two lysine side chains on opposite sides of a protein. On heating, the complex with the ammonia molecule can either lose the ammonia molecule (desolvation, DS) or lose an ammonium ion (deprotonation, DP). Loss of charge, deprotonation, is favored at short distances, low p. At some given distance between the two protonated amino groups, the activation free energies for DP and DS can become equal, and when this condition is met, eq 2 will hold:

$$GB^{app}(NH_3(CH_2)_pNH_2^+) = GB(NH_3) \quad (2)$$

Recent work in this laboratory[6] involving kinetic rate measurements of the reaction

$$B + H_3N(CH_2)_pNH_3^{2+} = BH^+ + NH_3(CH_2)_pNH_2^+ \quad (3)$$

at 400 K with a series of bases B led to a rate constant that was approximately equal to half of the collision rate constant when

† Part of the special issue "Jack Beauchamp Festschrift".

10.1021/jp021326s CCC: $22.00
Published on Web 08/31/2002

$p = 7$ and B = NH_3. Bases with GB higher than that of NH_3 led to deprotonation rates at the collision rate. This means that eq 2 holds for $p = 7$. The rate of deprotonation by ammonia for diprotonated diamines with $p > 7$ was found to decrease rapidly with increasing values of p.

The result that eq 2 holds for $p = 7$ is in very good agreement with earlier work by Gronert[5] based on ab initio calculations of the energy surface of reaction eq 1, $p = 7$, when a small correction involving the entropies of activation for DS and DP is introduced.[6] Such a correction was needed because the ab initio calculations involved the internal energies and not the free energies.

The diprotonated diamine model can be used for the development of equations with which the GB^{app} of basic sites of proteins can be predicted approximately. This approach was first used by Williams and co-workers[4] in a series of papers that preceded the work of both Gronert[5] and Peschke et al.[6] These authors[4] proposed the relationships

$$GB_{sp}^{app} \approx GB_{int} - \frac{q^2}{4\pi\epsilon_o\epsilon_r r} \quad (4a)$$

$$GB_{sp,j}^{app} \approx GB_{int,j} - \sum_{i=1}^{i=n} \frac{q^2}{4\pi\epsilon_o\epsilon_r r_{i,j}} \quad (4b)$$

where sp stands for the site of protonation, q is the elementary charge, ϵ_o is the permittivity in a vacuum, and ϵ_r is the relative permittivity of the medium in which the charges interact. The distance between the charges is r. Equation 4a is for two charges, while 4b applies to multiple charges with distances $r_{i,j}$ between the sites of protonation. GB_{int} stands for the (intrinsic) basicities of the site of protonation in the absence of all charges. Values of GB_{int} for given side chains are estimated on the basis of experimentally determined GB of the corresponding amino acids. Equation 4a is based on a combination of electrostatics, expressing the Coulomb energy due to the repulsion between charges, and the gas-phase basicity GB_{int} based on experiment, while eq 4b is a summation of the pair-like repulsions.

Equation 4a can be subjected to experimental verification, see eqs 1−3, while experimental verification of the apparent basicities of side chains on multiply protonated proteins, eq 4b, is much more difficult because the identity of the given site that was deprotonated is not known.[6] For the positive ion charge states, eq 4a was shown to predict too low apparent basicities by both Gronert's ab initio calculations[5a] and the experimental rate measurements,[6] see eqs 1−3. Improved equations, also based on electrostatics and on some insights provided by Gronert's work,[5] were developed:

$$GB_{sp}^{app} \approx GB_{int} - \frac{q^2}{4\pi\epsilon_o r} + \frac{q^2}{4\pi\epsilon_o(R_{NN}(\max) + r)} + E_R + T\Delta\Delta S \quad (5a)$$

$$GB_{sp}^{app} = GB_{int} - \sum_{i=1}^{z} \frac{q^2}{4\pi\epsilon_o r_{i,j}} + \sum_{i=1}^{z} \frac{q^2}{4\pi\epsilon_o(r_{i,j} + R_{NN,j}(\max))} + E_R + T\Delta\Delta S \quad (5b)$$

where the third term in eqs 5a and 5b refers to the electrostatic repulsion of the transition state for deprotonation when the charge has been moved by a distance of $R_{NN}(\max)$ away from the other charge(s) (the reverse activation barrier), E_R refers to dipole, polarization, and charge delocalization terms, and $T\Delta\Delta S$ refers to the entropy differences in the transition states for desolvation and deprotonation. Because the polarization effects are included explicitly and implicitly in the intrinsic basicity, it is not necessary to include the relative permittivity of the medium. For more details of the terms, see Peschke et al.[6] These were then used to evaluate GB^{app} of basic side chains in cytochrome *c*, carbonic anhydrase, and pepsin, which also led to predictions of the number of charges (protons) that these proteins can hold.[6]

In the present work, we provide results from experimental rate measurement of the protonation of the dianions, $CO_2(CH_2)_pCO_2^{2-}$, by a variety of acids, AH, and particularly aliphatic carboxylic acids such as acetic acid, of which the conjugate base, the acetate anion, is often present in the buffers used to spray native (nondenatured) proteins. Protonation by AH leads to charge loss:

$$CO_2(CH_2)_pCO_2^{2-} + AH = HCO_2(CH_2)_pCO_2^{-} + A^- \quad (6)$$

Thus, eq 6, where AH = CH_3CO_2H, is the negative ion mode counterpart of the model eq 3 when the buffer is ammonium acetate. All of the experimental determinations of nondenatured proteins that provide a comparison between the positive and negative charged states have used this buffer.[3]

The dicarboxylic acids can be viewed as models for the acidic side chains glutamate and aspartate. The rate measurement results provide a very telling explanation as to why the charged states of proteins in the negative ion mode are generally significantly lower than those in the positive ion mode.

II. Experimental Section

The experimental measurements relating to reaction eq 6 were performed with a reaction chamber sampled by a quadrupole mass spectrometer that has been described.[7] The same apparatus was used previously for determination[6] of the proton-transfer rates involving the positive diprotonated diamines (eq 3). The reagent ions, $CO_2(CH_2)_pCO_2^{2-}$ (C_6, C_7, C_8, C_{10}, C_{12}, C_{14}, and C_{16}, where C_n corresponds to the diacid with $n = p + 2$), were produced by electrospray of solutions of the sodium salts, and the rate measurements were obtained by introducing these ions into a reaction chamber, which contained also 10 Torr of N_2 as bath gas and known low partial pressures (1−200 μTorr) of the acids, AH. The intensities of the ionic reactants were determined with a quadrupole mass spectrometer.

In the apparatus used, the reactant ion, when inside the reaction chamber, is exposed to a weak drift field by applying a small voltage, $V_d = 5$ V, between the ion entrance orifice (IN) and the ion exit orifice (OR) (8 mm apart, see Figure 1 in Blades et al.[7]). This drift field and the pressure of the bath gas N_2 control the drift velocity and thus also the reaction time t. The value of the drift voltage, V_d, is low so that the drifting ions have thermal internal energies. The drift times of the ions in the reaction chamber are in the 100−1000 μs range depending on the value of the drift field. In the present experiments, the drift field was the same as that used in the previous work.[6] A rough estimate of the drift time, $t = 270$ μs, for the $NH_3(CH_2)_7NH_3^{2+}$ ion was obtained.[6] The drift time for the present $CO_2(CH_2)_pCO_2^{2-}$ ions should be similar, but slightly higher, because of the greater chain length of the dianions used.

For the proton-transfer reaction eq 6,

$$CO_2(CH_2)_pCO_2^{2-} + AH = (CO_2(CH_2)_pCO_2H)^- + A^-$$

$$\ln([CO_2(CH_2)_pCO_2^{2-}]_{AH}/[CO_2(CH_2)_pCO_2^{2-}]) = kt[AH] \quad (7)$$

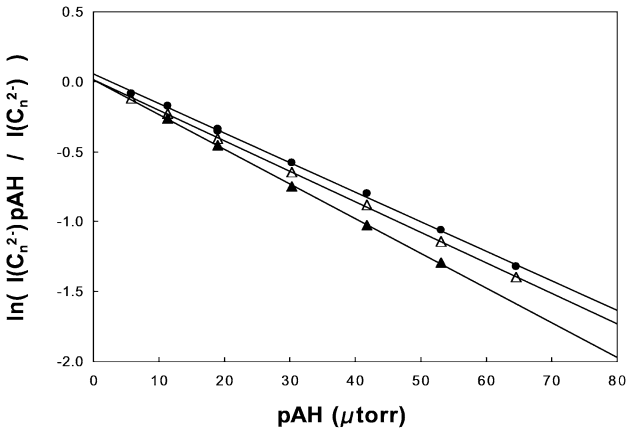

Figure 1. Plot corresponding to eq 7, in which C_n^{2-} is $CO_2(CH_2)_pCO_2^{2-}$ and n is the number of carbon atoms: $n = p + 2$. The ion intensities, $I(C_n^{2-})_{p_{AH}}$ and $I(C_n^{2-})$, are observed intensities with a partial pressure p of AH and without AH in the reaction chamber. Data are for C_{10}^{2-} and AH = (●) acetic acid, (△) pivalic acid (*tert*-$C_4H_9CO_2H$), and (▲) *para*-cresol. Slopes of plots after pressure is converted to concentration lead to the product between rate constant k and reaction time t. These are given in Table 1.

TABLE 1: Results from Proton Transfer Rate Measurements for Reaction $CO_2(CH_2)_pCO_2^{2-}$ + AH = $HCO_2(CH_2)_pCO_2^{-}$ + A^-

		$kt \times 10^{13}$ (cm^3 molecule^{-1})[b]	
AH	GB(A^-)[a]	$n = 10$[c]	$n = 16$[d]
CH_3CO_2H (acetic)	0	9.2	10
$CH_3CH_2CO_2H$ (propionic)	−1.1		11
$CH_3(CH_2)_3CO_2H$ (valeric)	−2.3		11.4
$(CH_3)_3CCO_2H$ (pivalic)	−3.9	9.8	11.5
p-cresol	+2.7	11.0	

[a] Gas-phase basicity of anion A^-, GB(A^-), relative to GB($CH_3CO_2^-$) = 0. Values taken from Cumming and Kebarle[12a] and Caldwell et al.[12b] [b] Product of rate constant k and reagent dianion drift time t through ion source from plots in Figures 1 and 2. The time t is constant when same reagent dianion is involved. [c] Results for dianion C_{10} ($p = 8$). Data from Figure 1. Rough estimates of the rate constant k can be obtained with a reaction time $t = 300$ μs. This value leads to $k = 3 \times 10^{-9}$ cm^3 molecule^{-1} s^{-1} for CH_3CO_2H. The value $t = 300$ μs is based on estimates[6] of the ion drift times through the reaction chamber, see Experimental Section. [d] Results for dianion with $n = 16$ carbon atoms. Data from Figure 2.

eq 7 is used to obtain a value for the product kt. $[CO_2(CH_2)_pCO_2^{2-}]_{AH}$ is the ion concentration after time t, that is, at the exit of the reaction chamber, when a constant concentration [AH] is present in the reactor. $[CO_2(CH_2)_pCO_2^{2-}]$ is the ion concentration after time t, when [AH] = 0. The ion concentration ratio in eq 7 is replaced with the corresponding ion intensity ratio observed with the mass spectrometer.

Determinations with a range of constant concentrations [AH] at constant drift field, that is, constant t, lead to plots such as that shown in Figure 1. The plots are linear as expected from eq 7, and the slopes give the values for kt for each acid, AH; t is constant because the same reagent ion is involved. Therefore, the slopes kt provide the relative values for the proton-transfer rate constants; kt values obtained for several different dianions and different acids AH are tabulated in the Results and Discussion section, see Table 1. Additional information concerning the validity of eq 7 is given in Peschke et al.[6]

Results and Discussion

(a) Results and Significance of Rate Measurements of Reactions $CO_2(CH_2)_pCO_2^{2-}$ + AH = $CO_2(CH_2)_pCO_2H^-$ + A^-. The negative values of the slopes, equal to kt for reaction eq 6, with the C_{10} dianion, $CO_2(CH_2)_8CO_2^{2-}$ (C_{10}^{2-}), exhibit very slowly increasing values in the order: acetic acid, pivalic acid, and *p*-cresol, see Figure 1 and Table 1. The reaction time t is expected to be the same because the same reagent ion is involved. There are two possibilities: (a) The rates could be increasing because reactions (eq 6) are endoergic. In that case, the reaction rates will be below collision rates and increase as GB(A^-) decreases. This is clearly not the case, because cresol, of which the anion has the highest basicity, leads to the highest rate (see Figure 1 and Table 1). (b) The rates are at the collision limit. In that case, they should increase in proportion to the Langevin rate constant coefficients, $(\alpha/\mu)^{0.5}$. This is found to be the case. Thus using polarizabilities for *para*-cresol and acetic acid, $\alpha = 13.5$ and 5.7 Å^3, estimated with the additivity rules[8] and the corresponding reduced mass μ, one obtains a prediction of the ratio of the collision rates, k_c(cresol)/k_c(acetic acid) = 1.25, which is very close to the observed ratio $kt = 1.2$, Table 1.

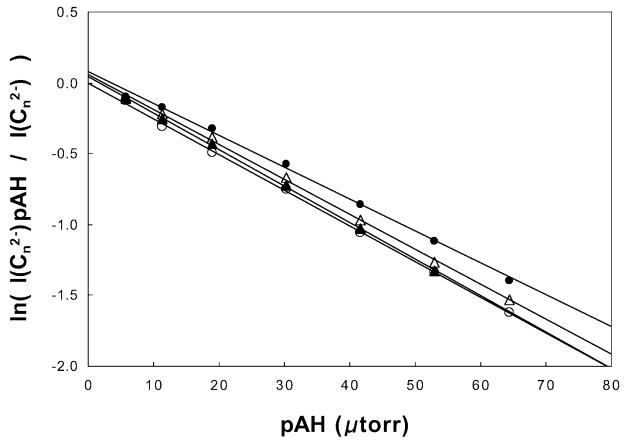

Figure 2. Plot corresponding to eq 7, in which C_n^{2-} is $CO_2(CH_2)_pCO_2^{2-}$ and n is the number of carbon atoms: $n = p + 2$. The ion intensities, $I(C_n^{2-})_{p_{AH}}$ and $I(C_n^{2-})$, are observed intensities with a partial pressure p of AH and without AH in the reaction chamber. Data are for C_{16}^{2-} and and AH = (●) acetic acid, (△) propionic acid, (▲) valeric acid (*n*-$C_4H_9CO_2H$), and (○) pivalic acid (*tert*-$C_4H_9CO_2H$). Values of kt are given in Table 1.

The kt values for the C_{16}^{2-}, Figure 2 and Table 1, are seen to be slightly higher than those for C_{10}^{2-}. We attribute this shift to an increase of the ion drift time for the longer C_{16}^{2-} ions. Apart from this difference, the rates for the two dianions are seen to be essentially the same and should correspond to collision rates for both systems.

An estimated drift time t, see Table 1, leads to rate constants of a magnitude near $k \approx 3 \times 10^{-9}$ cm^3 molecule^{-1} s^{-1} for acetic acid and C_{10}^{2-}. This is a value close to expected collision rates and represents additional evidence that the disappearance of the dianions occurs at collision rates.

The proton-transfer reaction, eq 6, is by far the major process involved in the disappearance of the dianions. Very minor amounts of the adducts $(C_n \cdot HA)^{2-}$ are also observed (see Figure 3). The adducts, which are the intermediates in the reaction of eq 6, are of course the most stable product. However, when formed, they are internally excited and require collisional stabilization by the bath gas. At the temperature used, 400 K, the lifetime of the excited adducts is too short and the majority decomposes via the proton-transfer reaction, eq 6. The small amount of adducts, Figure 3, is seen to increase in the order of increasing basicity of A^-. This is expected because the activation energy for proton transfer increases with increasing basicity of

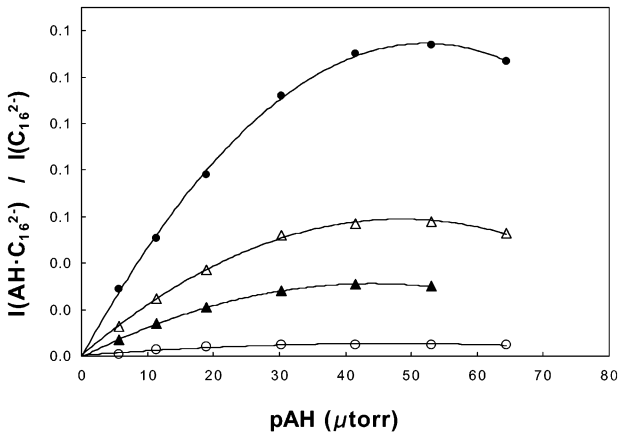

Figure 3. Fraction of *A*, observed ion intensity for adduct ions formed by reaction $C_{16}^{2-} + AH = (C_{16}{\cdot}AH)^{2-}$, relative to *B*, total decrease of intensity of C_{16}^{2-} due to adduct formation and proton transfer from AH. Results show that adduct formation is only a minor product that decreases with decreasing basicity of A^-. Symbols represent (●) acetic acid, (△) propionic acid, (▲) valeric acid, and (○) pivalic acid.

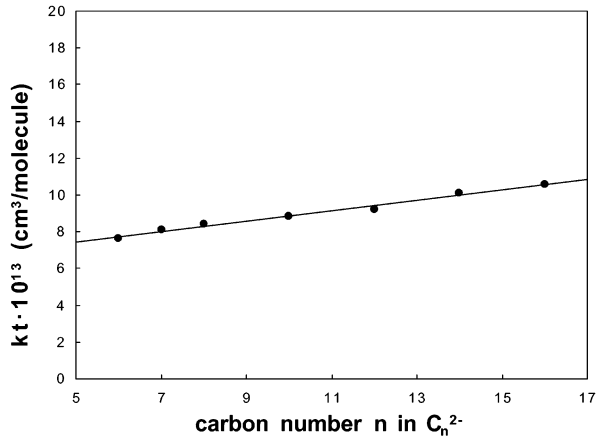

Figure 4. Values of *kt* determined for reaction $C_n^{2-} + CH_3CO_2H = C_nH^- + CH_3CO_2^-$. The observed small increase of *kt* with *n* is attributed to increase of ion drift time *t* with increasing *n* and essentially constant rate constants *k* at the collision limit.

A^-, vide infra, and this increases the lifetime and thus also the collisional stabilization of the complex.

Shown in Figure 4 are the *kt* values determined for the reaction of eq 6, where in a series of experiments dianions, C_n^{2-}, with increasing *n* were used while the acid, AH, was the same (acetic acid). The *kt* values are seen to increase from $kt = 7.6 \times 10^{-13}$ cm^3 molecule^{-1} to $kt = 10.6 \times 10^{-13}$ cm^3 molecule^{-1} for an increase of $n = 6$ to $n = 16$. The small increase in *kt* with *n* is attributed to an increase of the ion drift time *t*, while the rate constants remain essentially constant at collision rates. To our knowledge, no suitable expressions for the mobilities of doubly charged ions of the present type are available in the literature. The mobilities are not expected to change very much because the collision cross sections for these still relatively small ions are expected to be determined by the Langevin forces between the two single charges and the collision gas (nitrogen) molecules. This expectation is in line with the *kt* values in Table 1 and the results in Figure 4, which show very small changes of *kt* with *n*.

The results above for acetic acid show a remarkable difference when compared with the rate results for the positive di-ion model $H_3N(CH_2)_pNH_3^{2+}$. For the positive ion case, charge loss by proton transfer to NH_3 was observed[6] only for $p < 7$. On the other hand, for the negative ion case, one observes charge loss by proton transfer from acetic acid to C_n^{2-} for *n* values right up to $n = 16$ and, probably, also considerably higher *n*. (see section b). The distance between the two N atoms is 10 Å in the $p = 7$ diamine.[5a] The distance between the two oxygens in the C_{16}^{2-} dianion can be estimated to be about 25 Å. Many of the acidic, respectively, basic, side chains near the surface of midsize proteins such as carbonic anhydrase are at distances that are much less than 25 Å. This means that with ammonium acetate as buffer and assuming that the charging in the negative ion mode is due to the small ions, that is, acetate anions at the surface of the droplet containing the protein, many of the acidic side chains will remain not charged due to the presence of charges on side chains as far away as 25 Å. In the positive ion mode, charge loss will occur only when the other charged basic sites are at much shorter distances. The findings based on the reaction rates, thus, provide an experimental answer to the question, why are the observed charge states in the negative ion mode lower than those in the positive ion mode?

(b) Mechanism of Protein Charging and Reasons for Lower Charge States of Proteins in the Negative Ion Mode. As already mentioned in the Introduction, it is an experimental fact[2] that in the positive ion mode the average charge state observed with ammonium acetate can be changed to progressively lower charge states by changing the cations from NH_4^+ to alkylammonium ions, such as $MeNH_3^+$, $(Me)_2NH_2^+$, $(Me)_3NH^+$, where the conjugate bases have increasing gas-phase basicities. The question can be asked: what is the cause of this phenomenon? One possibility, considered by Smith and coworkers,[3c] was the creation of vapor-phase bases NH_3, $MeNH_2$, etc. from the buffer salts on evaporation of the droplets formed by electrospray, followed by deprotonation of the proteins in the gas phase by these bases. The authors rejected this possibility estimating that the pressure of the bases produced in this manner would be too low.

A much more likely alternative is based on a consideration of the electrospray mechanism by which the proteins are transferred to the gas phase. Recent research and reviews[9] on the mechanism of the generation of gas-phase ions by the ESI method indicate that the small ions (such as inorganic ions, Na^+, NH_4^+ etc., or organic ions, such as protonated organic bases BH^+, or negative ions, such as Cl^- and $CH_3CO_2^-$ in the negative ion mode) are produced by the ion evaporation model (IEM).[9a−d,f] Large macro-ions and typically the nondenatured globular (native) proteins are produced by the charge residue model (CRM).[9a,c,e,f]

The most significant evidence that the multiply charged native proteins are produced by CRM was provided by de la Mora.[9e] He showed that the experimentally observed number of charges, Z_{obs}, reported in the ESIMS literature was approximately equal to the charge at the surface of the precursor water droplet containing the protein, when the evaporating droplet has just reached the size of the protein. This charge, which we call Z_{CRM}, can be evaluated with the Rayleigh equation[8] and a value for the radius *R* of the protein (see Figure 1 in de la Mora[9e]).

de la Mora[9e] did not discuss the chemistry of protonation, that is, how exactly is the charge at the surface of the disappearing droplets converted to charge of the proteins. We assume[6,10] that as the last water evaporates most of the ionized basic side chains can be expected to become neutralized by reacting with nearby counterions. Such a process will be fostered by prior ion pairing, caused by the increasing electrolyte concentration in the evaporating droplets. In the last stage, the excess (unpaired) ions at the surface of the droplet, which provide the charge of the droplet, will end up on the protein

and can react with side chains at the surface of the protein. Recent work from this laboratory[6,10] has provided evidence that when the protein is sprayed from aqueous solution containing ammonium acetate as the major electrolyte, the charging is due to NH_4^+ ions in the positive ion mode.[10] Similarly, one can expect charging by $CH_3CO_2^-$ ions in the negative ion mode.

The lowest-energy products will be the proton-bridged adducts:

$$\text{protein}\cdots(CH_2)_2CO_2{-}H{-}O_2CCH_3^- \quad (8)$$

The side chain models aspartate or glutamate. A complete proton transfer leading to charging by deprotonation of the side chain occurs later in the "desolvation" stage in the sampling system, either in the heated (100−200 °C) sampling capillary leading to the mass spectrometer or in the CID stage because of ion acceleration by the electric field applied between sampling capillary skimmer electrodes.

$$\text{protein}\cdots(CH_2)_2CO_2{-}H{-}O_2CCH_3^- = \text{protein}\cdots(CH_2)_2CO_2^- + CH_3CO_2H \quad (9)$$

The experimental results in section a demonstrated that the adduct of the model compound, which has one more negative charge located some 25 Å away from the hydrogen bond, will on heating not acquire a second charge, but lose it, by proton-transfer back to the "side chain":

$$^-O_2C(CH_2)_{14}CO_2{-}H{-}O_2CCH_3^- = {}^-O_2C(CH_2)_{14}CO_2H + CH_3CO_2^- \quad (10)$$

This result is in stark contrast with the positive ion mode, where NH_4^+ was the charging reagent. In that case, as predicted by the model compound $NH_3(CH_2)_pNH_3^{2+}$, loss of the second charge occurred only after the two charge sites were much closer, 10 Å. This difference must be responsible for the lower charge states observed for proteins sprayed with ammonium acetate buffer, and it is of interest to elucidate the causes of the difference.

Charge loss is promoted by a proximity of the charges, while charge retention is favored by the difference of the gas-phase basicities of the sites competing for the proton, see eq 4 from Williams[4] and see also improved equations by Gronert[5] and eq 5. In the positive ion mode, the basicities involved are the intrinsic basicity, GB_{int}, of the side chain and the basicity of the conjugate base, NH_3. In the negative ion mode, it is the intrinsic basicity of the deprotonated side chain and the basicity of the acetate anion. There is a fundamental difference in the way the basicities change with alkyl substitution in these two cases. The basicities of the amines[11] increase rapidly with alkyl substitution (relative values in kcal/mol): $GB(NH_3) = 0$; $GB(CH_3NH_2) = 10.9$; $GB(CH_3CH_2NH_2) = 14.1$; $GB(CH_3{-}CH_2)_7NH_2) \approx 18$.

On the other hand the relative basicities of the anions A^- (equal to the gas-phase acidities of AH)[12] change very much more slowly. There is actually a decrease of acid strength (increase of $GB(A^-)$) from formic to acetic acid by 3.2 kcal/mol[12a] and then very small increases of acid strength from acetic to propionic, butyric, valeric, and so on, see Table 1 and acidity determinations.[12]

This contrasting alkyl substituent effect for the positive and negative systems is well understood. For ammonia and the alkylamine bases, substitution by alkyl groups with increasing size leads to stabilization of the protonated base by both increased σ electron donation from the carbon adjacent to the N atom and by the increasing polarizability of the alkyl group. On the other hand, the deprotonated carboxylic group is destabilized by σ electron donation from the adjacent carbon, and stabilization is provided only by the polarizability of the substituent.[12b,13]

Konermann and Douglas[3a] observed a shift of average charge state from +10 to −6 for lysozyme, from +8 to −5 for cytochrome *c*, and from +6 to −5 for ubiquitin. Similarly, Smith and co-workers[3c] observed a shift from +8 to −5 for cytochrome *c* when the above proteins were sprayed in aqueous solutions with ammonium acetate as buffer. The observed higher charges for the positive ion states are in line with the discussed small differences of the basicities of carboxylate anions, which do not lead to a strong chemical preference for retention of the charge in the negative ion state.

It is also notable that the charge states decrease from 10 (lysozyme) to 8 (cytochrome *c*) to 6 (ubiquitin) in the positive ion state, but a very small decrease, 6 to 5 to 5, is observed for the negative ion mode. This can be explained. In the positive mode, the capacity of the above proteins to hold charge is high[6] and the observed charge for the above proteins is determined by the amount of charge that the droplet can deliver according to the charge residue model.[6,9a,10] This charge decreases with the size of the proteins, that is, in the order lysozyme, cytochrome *c*, ubiquitin. In the negative mode, the charge is determined by the ability of the proteins to hold it, that is, by the apparent basicity of the carboxylate groups available. The apparent basicities will also depend on the size of the proteins, but because of the lower chemical ability of the carboxylate side chains to hold charge, these groups must be at much larger distances from each other. Because Coulombic energy changes become less sensitive to distance changes when the charges are further apart, it would be expected that in the negative ion mode size differences in proteins will show less effect on the observed charge state distribution.

We hope to be able to predict the expected negative charge states after having developed an equation for GB^{app} for the anions of the acidic side chains. This equation would be analogous to the equations developed for the basic side chains relevant for the positive ion mode[6] (see eq 5 above and also eqs 20 and 21 in ref 6). It was shown[6] that for the positive charge state the maximum number of the basic sites that can hold protons (equal to N_{SB} in the notation used[6] and calculated with eqs 20 and 21) can be equal to or larger than Z_{obs} (the charge state observed). This comes about because the positively charged proteins can tolerate a high density of charges, and this leads to N_{SB} values that are generally higher than Z_{CRM}, the charges provided by the small ions on the droplet that contains the protein. This means that it is Z_{CRM} that is charge limiting. Therefore, Z_{obs} in the positive ion mode is not a good test for the reliability of the calculated N_{SB}. On the other hand, for the negatively charged proteins, the charge state will be much lower, and therefore, the calculated maximum number of charges should be, in general, equal to Z_{obs}. The observed degree of agreement in the negative ion mode between Z_{obs} and the calculated number will thus provide a good test of the calculations and models used.

Negative charge states of proteins have proven much less useful than the positive states in applied biochemical mass spectrometry. The negative states are generally observed at lower intensities, and often, the interpretation of the results is more difficult.[3] However, there are special cases in which the negative states may be of advantage. The mass spectrometric study of

nondenatured protein−protein and protein−substrate complexes produced by electrospray ionization is at present a most active area of research,[14] which is making important contributions to biochemistry and biopharmacology. Some of these studies attempt to correlate the noncovalent binding energy of the complex in the biological environment with the binding energy determined in the gas phase by mass spectrometric techniques.

Examples of such recent work are experiments using the blackbody infrared radiative dissociation (BIRD) technique to thermally dissociate multiply protonated protein complexes in a Fourier transform ion cyclotron resonance (FTICR) mass spectrometer.[15,16] The multiple charges present on the protein were found to have a significant effect on the activation energies of the decomposition.[16] These results clearly demonstrate that it is very desirable to understand the origin of the charges on the proteins formed by ESI, know the positions on which the protons reside, and have an understanding of the ability of the protons to migrate to other basic groups, when the protein is heated to higher temperatures including temperatures that will lead to thermal decomposition. Such an understanding might allow the evaluation of the Coulombic energy terms due to the charges. Subtraction of the Coulombic energy terms from the observed activation energy might then lead to energy values that are representative of the bond energy of the complex due to noncovalent bonds that hold it together in solution.

In the positive ion mode, the expected migration of the protons during the decomposition of the complex presents a big hurdle in attempts to unravel the Coulombic contribution to the activation energy.[16] This migration is expected to be very facile particularly from protonated lysine side chains. The peptide carbonyl oxygens have relatively high gas-phase basicity and thus represent good "stepping stones" as proton acceptors in such a migration. In the negative ion mode, the proton migration is expected to be very slow. Only proton transfer from one acidic side chain to another can be expected, and the acidic side chains are generally too far from each other to allow such proton transfer. This can be expected to lead to a localization of the charges on the deprotonated side chains with the lowest GB^{app}. The positions of these charges can be predicted by calculation.[6] The contribution of the Coulombic repulsion to the observed activation energy could in such a case be evaluated by methods similar to those used in the derivation of eq 5.

Conclusions

(1) The charge residue model (CRM) for proteins according to which the small ions on the surface of the droplet, which contains one protein, provide the charges of the protein,[9e] combined with the gas-phase ion chemistry for the charging reactions, provides a consistent account of observed charged states of proteins. Thus, in the positive ion mode, the decrease of charge observed[2] with buffer alkylammonium ions relative to the charge state observed with the ammonium ion is caused by the higher basicities of the conjugate alkylamines relative to that of ammonia.

The observed decreased charge state in the negative ion mode compared to that in the positive ion mode, when ammonium acetate is used as buffer[3], can be also explained. In the positive ion mode, the intrinsic basicities of the basic side chains are much higher than the basicity of ammonia, and this causes charge retention. In the negative ion mode, the difference between the intrinsic acidity of the side chains and the acidity of acetic acid is very small. Therefore, even very remote negative charges can cause the loss of charge by proton transfer from acetic acid to the protein side chain.

(2) A better understanding of the gas-phase ion chemistry of multiply charged states can be beneficial in studies of the binding energies of charged protein−protein and protein−substrate complexes in the gas phase, which seek for correlations with the noncovalent binding energies in solution.[14−16] Before examining the possible existence of such correlations, one must be able to extract the contribution of the Coulombic repulsions in the gas phase. The approach used in the evaluation of the charge states formed by CRM[6] is expected to be very similar to the approach that would have to be used to unravel the Coulombic contributions to the dissociation energy of multiply charged protein complexes. Use of the negative charge states can offer some advantages. Charge migration via proton transfer, which occurs in the thermal decomposition of positive charge states[16]and complicates the modeling, will probably not occur for negative charge states because the acidic side chains able to hold a charge are much fewer and the peptide backbone does not provide acidic sites that can facilitate the proton migration.

References and Notes

(1) (a) Konermann, L.; Silva, E. A.; Sogbein, O. F. *Anal. Chem.* **2001** *73*, 4836. (b) Ahmad, A.; Madhusudanan, K. P.; Bhakuni, V. *Biochim. Biophys. Acta* **2000**, *1480*, 201. (c) McLafferty, F. W.; Guan, Z.; Hampts, U.; Wood, T. D.; Kelleher, N. L. *J. Am. Chem. Soc.* **1998**, *120*, 4732. (d) Reitman, C. T.; Velazquez, I.; Tapia, O. *J. Phys. Chem. B* **1998**, *102*, 9344. (e) Konermann, L.; Douglas, D. J. *Biochemistry* **1997**, *36*, 12296 and references therein.

(2) (a) Le Blanc, J. C. Y.; Wang, J.; Guevremont, R.; Siu, M. *Org. Mass Spectrom.* **1994**, *29*, 587. (b) Iavarone, A. T.; Jurchen, J. C.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 976. (c) Lemaize, D.; Marie, G.; Serani, L.; Laprevote, O. *Anal. Chem.* **2001**, *73*, 1899.

(3) (a) Konermann, L.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1248. (b) Ogorzalek, R. R.; Smith, R. D. *J. Mass Spectrom.* **1995**, *30*, 339. (c) Wingel, B. E.; Light-Wahl, K. J.; Ogorzalek-Loo, R. R.; Udset, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 536.

(4) (a) Gross, D. S.; Rodriguez-Cruz, S. E.; Bock S.; Williams, E. R. *J. Phys. Chem.* **1995**, *99*, 4034. (b) Gross, D. S.; Williams, E. R. *J. Am. Chem. Soc.* **1995**, *117*, 883. (c) Schnier, P. D.; Gross, D. S.; Williams, E. R. *J. Am. Chem. Soc.* **1995**, *117*, 6747. (d) Schnier, P. D.; Gross, D. S.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1086. (e) Williams, E. R. *J. Mass Spectrom.* **1996**, *31*, 831.

(5) (a) Gronert, S. *J. Am. Chem. Soc.* **1996**, *118*, 3525. (b) Gronert, S. *J. Mass Spectrom.* **1999**, *34*, 787. (c) Gronert, S. *Int. J. Mass Spectrom.* **1999**, *185/186/187*, 351.

(6) Peschke, M.; Blades, A. T.; Kebarle, P. *J. Am. Chem. Soc.*, submitted for publication.

(7) Blades, A. T.; Klassen, J. S.; Kebarle, P. *J. Am. Chem. Soc.* **1996**, *118*, 12437.

(8) Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8543.

(9) (a) Kebarle, P.; Ho, Y. In *Electrospray Ionization Mass Spectrometry*; Cole, R. B., Ed.; John Wiley & Sons: New York, 1997; p 55. (b) Kebarle, P.; Peschke, M. *Anal. Chim. Acta* **2000**, *406*, 11. (c) Cole, R. B. *J. Mass Spectrom.* **2000**, *35*, 763. (d) Gamero-Castano, M.; de la Mora, F. J. *J. Mass Spectrom.* **2000**, *406*, 67. (e) de la Mora, F. J. *Anal. Chim. Acta* **2000**, *406*, 105. (f) Gamero-Catano, M.; de la Mora, F. J. *J. Mass Spectrom.* **2000**, *35*, 780. (g) Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804.

(10) Felitsyn, N.; Peschke, M.; Kebarle, P. *Int. J. Mass Spectrom.*, in press.

(11) Hunter, E. P.; Lias, S. G. *J. Phys. Chem. Ref. Data* **1998**, *27*, 3.

(12) (a) Cumming, G. B.; Kebarle, P. *Can. J. Chem.* **1978**, *56*, 1. (b) Caldwell, G.; Rennebogg, R.; Kebarle, P. *Can. J. Chem.* **1989**, *67*, 611.

(13) Brauman, J. I.; Blair, L. K. *J. Am. Chem. Soc.* **1971**, *93*, 4315.

(14) (a) Ganem, B.; Li, Y.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294. (b) Banem, B.; Li, Y. T.; Henion, J. *J. Am. Chem. Soc.* **1991**, *113*, 7818. (c) Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1991**, *113*, 8534. (d) Schwartz, B. S.; Light-Wahl, K. J.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 201. (e) Light-Wahl, K. J.; Schwartz, B. L.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 5277. (f) Lafitte, D. *Eur. J. Biochem.* **1999**, *261*, 337.

(15) (a) Price, W. D.; Schnier, P. D.; Jockbush, R. H.; Strittmacher, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640. (b) Gross, B. S.; Zhao, Y.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 519.

(16) Felitsyn, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2001**, *73*, 4647.