

11-1-2021

Using Covalent Modifications to Distinguish Protein Electrospray Mechanisms: Charged Residue Model (CRM) vs. Chain Ejection Model (CEM)

Lars Konermann
konermann, konerman@uwo.ca

Douglas J. D. Pimlott
Western University

Follow this and additional works at: <https://ir.lib.uwo.ca/chempub>

 Part of the [Chemistry Commons](#)

Citation of this paper:

Konermann, Lars and Pimlott, Douglas J. D., "Using Covalent Modifications to Distinguish Protein Electrospray Mechanisms: Charged Residue Model (CRM) vs. Chain Ejection Model (CEM)" (2021). *Chemistry Publications*. 233.
<https://ir.lib.uwo.ca/chempub/233>

Invited Contribution for “Paul Kebarle” Virtual Special Issue of IJMS

**Using Covalent Modifications to Distinguish Protein
Electrospray Mechanisms: Charged Residue Model (CRM)
vs. Chain Ejection Model (CEM)**

Douglas J. D. Pimlott and Lars Konermann*

Department of Chemistry, The University of Western Ontario, London, Ontario,

N6A 5B7, Canada.

* corresponding author

E-mail address of the corresponding author: konerman@uwo.ca

Funding was provided by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-04243).

ABSTRACT

Different mechanisms have been proposed for the formation of gaseous protein ions during electrospray ionization (ESI). In the charged residue model (CRM) ions are produced upon nanodroplet evaporation to dryness. This mechanism is thought to dominate in native ESI, where proteins retain compact conformations, with charge states close to the Rayleigh charge of protein-sized aqueous droplets. Much higher charge states are generated from proteins that are unfolded in solution. The chain ejection model (CEM) has been proposed for ESI under such denaturing conditions. In the CEM proteins are gradually expelled, while mobile H^+ equilibrate between the droplet and its protruding tail. Providing clear-cut evidence for these scenarios remains difficult, because electrosprayed ions do not usually retain any features that reveal their formation mechanism. In this work we propose that the stepwise elimination of basic sites can serve to distinguish between the CRM and CEM. Using cytochrome *c* as a model system, we studied proteins that had between zero and 19 Lys blocked by acetylation. In native ESI (pH 7) the same low charge states were observed regardless of acetylation. This behavior is consistent with the CRM, where charge states are governed by protein size, rather than protein surface chemistry. Denaturing (pH 2) conditions resulted in much higher ESI charge states. Intriguingly, spectra acquired under these pH 2 conditions gradually shifted to lower charge states when the number of acetylated Lys was increased. This charge reduction is attributed to the fact that lowering the number of basic sites compromises the ability of the protein to compete with the droplet for mobile H^+ during the CEM. In conclusion, we illustrate that simple covalent modifications can help distinguish between protein ion formation via the CRM or the CEM.

1. Introduction

The inception of electrospray ionization (ESI) has opened up tremendous analytical opportunities [1]. Paul Kebarle (1926 - 2019) was a key contributor to this area, and his seminal reviews remain essential reading for any ESI practitioner [2,3]. In addition to using ESI as an ionization technique for mass spectrometry (MS) [4-6], electrosprayed ions can be interrogated by ion mobility spectrometry (IMS) [7-9] and complementary techniques [10-15]. Protein research is an area where ESI-MS has had a particularly large impact [5-8]. One of the most intriguing approaches in this context is “native” ESI, where experiments are designed to retain solution-like protein conformations and interactions in the gas phase [5,6,16-19].

The mechanisms whereby desolvated analyte ions are formed from nanometer-sized solvent droplets in the ESI plume have attracted considerable attention [20-32] and continue to be controversial [33-35]. In the commonly used positive ion mode, ESI nanodroplets are charged close to the Rayleigh limit z_R by excess charge carriers such as H^+ and NH_4^+ [3,36,37]

$$z_R = 8\pi/e \times (\epsilon_0 \gamma r_{droplet}^3)^{1/2} \quad (1)$$

where $r_{droplet}$ = droplet radius, γ = surface tension of water, ϵ_0 = vacuum permittivity, e = elementary charge [3]. Globular $[M + zH]^{z+}$ protein ions produced by native ESI are characterized by $z \approx z_R$ (with $r_{droplet} = r_{protein}$), suggesting that they were formed by nanodroplet evaporation to dryness in accordance with the charged residue model (CRM, Figure 1A) [3,23,27,37].

Much more highly charged ions are generated from proteins that are unfolded in solution. Unfolding can be triggered by supplementing samples with acid or base, organic cosolvents, heating, disulfide cleavage, or cofactor removal [27,38-43]. The empirical relationship between protein conformation and ESI charge states allows ESI-MS to be used for probing protein structural changes in solution [27,38-43].

The increased ESI charge states seen for unfolded proteins do *not* reflect the protein titration behavior in solution [44,45]. Instead, the high z values of these ions are related to the ESI mechanism. It has been proposed that unfolded proteins follow the chain ejection model (CEM, Figure 1B) [6,26,27,46,47]. In the CEM, solvent-exposed hydrophobic residues cause the unfolded protein to migrate to the droplet surface. The chain is then gradually expelled, proceeding through various stages where a steadily growing polypeptide tail protrudes from the droplet. Ejection is driven by electrostatic repulsion between the droplet and the protein tail. Excess H^+ that are highly mobile in water [48-51] and in proteins [52-55] undergo charge equilibration between droplet and the protruding protein, driven by the tendency of the system to lower its Coulomb energy by maximizing the distance between charges. Because of its stretched-out conformation, the protruding chain can accommodate many of these mobile H^+ . Charge equilibration takes place until droplet and protein separate, causing the protein to depart as a highly charged $[M + zH]^{z+}$ ion (Figure 1B) [27]. These CEM events are analogous to the ejection of highly charged $[M + zH]^{z+}$ monomers from gaseous multi-subunit proteins after collisional activation [56-60]. Additional support for the CEM comes from proton transfer experiments [47] and supercharging studies [26,27].

A third ESI scenario is the ion evaporation model (IEM), where compact, pre-charged species desorb from the droplet. The IEM is prevalent for small ions such as Na^+ [20,21,24,27,29]. Even for proteins, the IEM can be viable. However, this is the case only for proteins that are tightly folded, carry a high solution charge, and reside in relatively large droplets [61]. The CRM and CEM appear to be much more likely mechanisms for intact proteins [6,27,46,61]. Of note, the IEM ejection of small charge carriers such as H^+ and NH_4^+ plays an ancillary role during protein CRM by keeping the shrinking droplets close to z_R (Figure 1A) [25,27].

The ongoing discussions related to ESI mechanisms reflect the fact that it is difficult to provide experimental evidence for specific ion formation pathways. Electrosprayed ions do not

usually retain any features that reveal whether they were formed by the CRM, CEM, IEM, or by other mechanisms. Some efforts have been made to identify mechanism-specific attributes in experimental spectra, e.g., the tendency of proteins to form nonspecific adducts with nonvolatile solutes in native ESI is consistent with droplet evaporation to dryness during the CRM [3,16,61].

In the current work, we aim to provide additional experimental support for specific protein ESI mechanisms. We hypothesize the following: (1) The experimentally observed charge state of a CRM-generated globular protein ion should depend only on the protein radius (eq. 1). “Chemical details” such as the number of basic sites on the protein should not affect the charge of CRM-products [37,45]. (2) A key feature of the CEM is the equilibration of mobile H^+ between the droplet and its protein tail (Figure 1B) [27]. Protein-bound H^+ are known to reside mainly on Lys, Arg, and His side chains [59,62]. The H^+ binding capability of basic sites can be abrogated by covalent modifications [63]. Such covalent blockage can be expected to lower the charge states of CEM-generated protein ions, because charge equilibration should place fewer H^+ on the protein tail if some of the basic sites are eliminated. Our experiments on the model protein cytochrome *c* (cyt *c*) confirm these expectations, suggesting that covalent modifications can indeed help distinguish between the CRM and CEM.

2. Materials and Methods

Horse heart cyt *c* (12360 Da) and acetic anhydride were purchased from Millipore Sigma (St. Louis, MO, USA). All other chemicals were supplied by Fisher Scientific (Nepean, ON) or Caledon Laboratories (Georgetown, ON). Samples contained 5 μ M cyt *c* in aqueous solution in the presence of 10 mM ammonium acetate (pH 7). For Lys acetylation, acetic anhydride was added in 250- to

20000-fold molar excess relative to the protein. The reaction was allowed to proceed for 10 minutes at 0 °C, followed by solvent exchange using centrifuge filters (Amicon Ultra 0.5, 10 kDa MWCO) at $13000 \times g$ for 15 minutes to remove unreacted acetic anhydride and acetic acid produced by acetic anhydride hydrolysis. Solvent exchange was repeated 3 times to ensure the sample returned to pH 7 in 10 mM ammonium acetate. For native ESI-MS these samples were directly infused into the ESI source. For denaturing experiments the solutions were supplemented with 0.12 % formic acid (pH 2). In other words, both types of samples experienced acetylation under the same conditions at pH 7. Mass spectra were acquired on a SYNAPT G2-Si instrument in positive ion mode (Waters, Milford, MA, USA). Aqueous solutions were infused at $5 \mu\text{L min}^{-1}$ at an ESI voltage of +2.8 kV, without organic cosolvents. Spectra were acquired at a cone voltage of 20 V, with source and desolvation temperatures of 80 and 250 °C respectively. Optical experiments were conducted on samples identical to those used for ESI-MS, except that pH 2 samples contained HCl instead of formic acid. Circular dichroism spectra were acquired on a J-810 spectropolarimeter (JASCO, Easton, MD), and fluorescence data were recorded on a PTI Fluorolog QM-7/2005 instrument. All experiments were performed in triplicate, with independent samples for each replicate.

3. Results and Discussion

Choice of Model Protein. Cyt *c* in neutral aqueous solution has a compact native fold (Figure 2A). Acidification to pH 2 causes unfolding, with near-complete breakdown of tertiary and secondary structure [64]. The protein possesses a large number (24) of basic side chains, giving it a pI around 9.6 [65]. Almost all of these basic sites are solvent-exposed in the native crystal structure, the majority of them (19/24) being Lys (Figure 2B). Lys ϵ -NH₂ groups can be acetylated using acetic anhydride ($\text{R-NH}_2 \rightarrow \text{R-NH-CO-CH}_3$, $\Delta M = +42$ Da) [63,66-68]. This modification abrogates the

capability of Lys to undergo protonation in solution, i.e., the positive solution charge and the pI of cyt *c* drop in accordance with the degree of acetylation [66]. Similarly, acetylation greatly reduces the capability of Lys to accommodate protons in the gas phase, as this amine \rightarrow amide conversion lowers the gas phase basicity of Lys by as much as ~ 70 kJ mol⁻¹ [47]. Previous reports suggest that acetylated cyt *c* retains its native structure in neutral solution [66]. The combination of these attributes makes cyt *c* a promising model system for the experiments of this work. Incidentally, cyt *c* biosynthesis involves acetylation of the N-terminus, such that this site is not basic [69].

Protein Acetylation. Incubation of cyt *c* with acetic anhydride generated abundant acetylation. Some of the resulting mass distributions are exemplified in Figure 3A. The number of acetyl tags increased with the molar excess of labeling agent, evident from the progressive shift of the peak envelopes to higher mass as a result of multiple +42 Da modifications. The highest number of acetyl tags observed was 19, consistent with the number of Lys in cyt *c* [69] and with the fact that acetic anhydride is an amino group-specific labeling agent [63,66-68]. High anhydride concentrations can sometimes cause modifications at other side chains [63,70], however, the agreement between expected and observed maximum acetylation in Figure 3A indicates that such over-acetylation is not prevalent under the conditions used here (the two Cys in cyt *c* are protected as thioethers [69]). Increasing acetylation resulted in some Na⁺ adducts in the spectra (asterisks in Figure 3A), suggesting that the acetic anhydride contained Na⁺ as a trace impurity. Such adducts are common in ESI-MS [71]. Also, acetylated Lys may be more effective at solvating Na⁺ ions in the gas phase. The quality of our data was high enough to discern Na⁺ adducts from “clean” multiply acetylated $[M + zH]^{z+}$ ions (Figure 3A). Adducted signals were thus excluded from the discussion below.

Figure 3B displays the dependence of average acetylation level N_{avg} on the molar excess of acetic anhydride, calculated as

$$N_{avg} = \frac{\sum(N_i \times I_i)}{\sum(I_i)} \quad (2)$$

where I_i is the signal intensity of cyt c with a given number of acetyl tags (N_i). This analysis was conducted for the 8+ charge state at pH 7 because (as noted below) these solutions provide the most unbiased view of the acetylation behavior. The number of acetyl tags increased sharply at low concentrations of labeling agent. Beyond 5000-fold molar excess N_{avg} started to level off, reaching a value of 16.7 ± 1 for samples that had been exposed to 20000-fold molar excess (Figure 3B).

Acetylation Dependence of ESI Charge State Distributions. In an initial series of experiments, we acquired cyt c mass spectra using native ESI, i.e., in aqueous solution at pH 7. Non-acetylated controls showed a narrow distribution of low charge states, comprising only 8+ and 7+ ions (Figure 4A), consistent with earlier reports [72,73]. ESI measurements on acetylated samples at pH 7 produced very similar results, with 8+ and 7+ as the dominant charge states, even for cyt c that had undergone near-complete acetylation in 20000-fold excess of acetic anhydride (Figure 4B-D).

Acidification of the samples to pH 2 resulted in a very different behavior. ESI mass spectra of unlabeled cyt c at pH 2 showed very high charge states, with a unimodal distribution that peaked at 16+ and extending to 21+ (Figure 4E). The observation of these highly charged ions is consistent with the fact that cyt c is extensively unfolded at pH 2 [64]. Similar spectra have previously been reported for cyt c and many other proteins under denaturing solvent conditions [27,38-43]. Interestingly, the ESI charge state distributions at pH 2 were highly sensitive to acetylation. 250-fold excess of acetic anhydride caused a bimodal charge state distribution; the main maximum had shifted from 16+ to 14+, and intense 9+/8+ signals appeared (Figure 4F). Upon acetylating the protein further, both parts of this bimodal distribution shifted to even lower charge states, until

ultimately the spectra were dominated by 8+ and 7+ ions for 20000-fold excess of acetic anhydride (Figure 4G, H).

Complementary to Figure 4, we compiled the ESI-MS data in a way that highlights how the intensity of each charge state (summed over all acetylation signals with the same z) depends on the labeling level. The resulting plots reaffirm that ESI charge state distributions acquired at pH 7 were virtually unaffected by changes in the concentration of labeling reagent (Figure 5A). In contrast, at pH 2 the presence of acetyl groups caused major shifts to lower charge states (Figure 5B). To highlight the strikingly different behavior at pH 7 and pH 2 even more clearly, we determined average charge states z_{avg} according to

$$z_{avg} = \frac{\sum(z_i \times I_i)}{\sum(I_i)} \quad (3)$$

where I_i is the added intensity of all acetylation peaks that correspond to a given charge state z_i . When including all of the peaks in the pH 2 spectra, z_{avg} dropped from 15.7 to 9.3 between zero and 20000-fold excess of acetic anhydride (Figure 5C, red). When considering only the high charge states of the bimodal pH 2 data (21+ to 10+), the drop in z_{avg} was still very significant, from 15.7 to 11.6 (Figure 5C, magenta). This is in contrast to the pH 7 data, which showed hardly any change with $z_{avg} = 7.5$ and $z_{avg} = 7.4$ for unlabeled and fully labeled samples, respectively (Figure 5C, black).

Yet another way of dissecting the data is by determining the charge state distributions of individual acetylation levels, i.e., for all protein ions carrying a specific number of acetyl groups. Once again, the data obtained in this way show that the charge state distributions acquired at pH 7 remained virtually unchanged for proteins with different numbers of acetylated Lys (Figure 5D). This is in striking difference to pH 2, where an increasing number of acetyl tags shifted the spectra to lower charge states (Figure 5E). At pH 2 the value of z_{avg} decreased almost linearly with $\Delta z_{avg} \approx -0.3$ per acetylated Lys when considering all the peaks (Figure 5F, red). When considering only the

high charge states in the pH 2 spectra, the slope was somewhat lower, $\Delta z_{avg} \approx -0.2$ per acetylated Lys (Figure 5F, magenta). At pH 7 z_{avg} remained constant (Figure 5F, black).

Acetylation Dependence of Protein Conformation in Solution. Earlier studies on cyt *c* and other proteins suggested that Lys acetylation only causes minor conformational changes in solution, at least for pH 7 samples [66,74,75]. Because conformation is a key factor for protein ESI behavior we re-examined this aspect by conducting solution-phase optical experiments on cyt *c*. CD spectroscopy reports on secondary structure [76]. CD data acquired at pH 7 were relatively similar for all acetylation levels, with minima at 208 nm and 222 nm which reflect a high α -helical content (Figure 6A) [69,76]. A ~20% increase in CD signal amplitude at 222 nm compared to the unmodified protein indicates that some regions become more helical after extensive acetylation. Trp fluorescence spectra report on protein compactness. In native cyt *c* the single Trp (residue 59) is quenched by the nearby heme [77]. All of the pH 7 samples were virtually non-fluorescent (Figure 6C), implying that Trp-59 maintains close contact with the heme even after maximal acetylation. These data support the view that cyt *c* remains tightly folded at pH 7, regardless of acetylation [66,74,75].

As expected, non-acetylated cyt *c* at pH 2 showed a dominant CD minimum at around 203 nm (Figure 6B) which reflects the fully unfolded structure of the acid-denatured protein [64,76]. Surprisingly, acetylation triggered the formation of some helical structure, evident from CD minima at 208 and 222 nm (Figure 6B). However, this acetylation-induced secondary structure appears to be non-native, because the pH 2 CD spectra (Figure 6B) are quite different from those in Figure 6A (the main minimum for native cyt *c* is at 222 nm, while the acetylated pH 2 samples have their main minima at 208 nm). The CD spectra of the acetylated pH 2 samples are reminiscent of those reported for acid/methanol molten-globules [78]. Figure 6D reveals that acetylated cyt *c* at pH 2 is dominated by conformers that are quite compact, as seen by the low Trp fluorescence of these samples (Figure

6D). Overall, the optical data of Figure 6C, D confirm that unmodified cyt *c* at pH 2 is fully unfolded, while after acetylation the samples are dominated by non-native, but relatively compact conformations. This compaction may be caused by the fact that acetylation reduces the extent of charge-charge reduction in solution at pH 2 compared to the unmodified protein [79].

Ion Production for Native Cyt c at pH 7: Evidence for CRM Behavior. Like many other proteins under native ESI conditions, cyt *c* ions retain a compact globular shape that can be characterized by an effective radius $r_{protein}$ [5,6,16-19,37]. According to the CRM, nanodroplets in the ESI plume shrink by solvent evaporation until $r_{droplet} = r_{protein}$ (Figure 1A). As the final solvent layer disappears, the remaining solution charges bind to the protein [3,23,27,37,45]. Eq. 1 dictates that the net charge of the shrinking droplets is close to z_R . This net charge includes *all* charge carriers, including the protein [27]. For example, a vanishing 8+ droplet might contain an 8+ protein and no additional dissolved ions. Alternatively, the 8+ droplet might contain a 10- protein and 18 excess NH_4^+ . In both cases, droplet evaporation to dryness will generate a $[\text{M} + 8\text{H}]^{8+}$ gaseous ion (protein-bound NH_4^+ will release ammonium, leaving behind H^+ [3]). The charge state of a CRM-generated protein ion should therefore only depend on $r_{protein}$, not on protein-specific parameters such as the number of basic sites or the protein charge in solution.

The aforementioned expectation is confirmed by our pH 7 experiments, where z_{avg} remained virtually unchanged at $\sim 7.4+$ when the number of basic sites was lowered from 24 down to 5 by Lys acetylation (Figure 5C, E). Because the protein net charge exceeds the number of basic side chains, it can be assumed that some protonation also occurs at “non-traditional” sites such as Pro and Gln [45,80]. The invariability of ESI charge states is in stark contrast to the solution charge of cyt *c* at pH 7 which drops from 7.7+ down to 11.3- upon acetylation of 19 Lys residues (calculated using pK_a values from [79]). Eq. 1 therefore allows native ESI charge states to be predicted on the basis

of $r_{protein}$ alone, without requiring any other protein-specific parameters [3,23,27,37,45]. For the case considered here, eq. 1 predicts a CRM charge of 7.9+ ($r_{protein} = 1.7$ nm, $M = 12360$ Da, $d = 1$ g cm⁻³ [37], $\gamma = 0.0589$ N m⁻¹ [81]). This prediction is close to the measured value of $z_{avg} \approx 7.4+$ at pH 7. We conclude that the experimentally observed insensitivity of ESI charge states to the number of basic sites at pH 7 represents a hallmark of the CRM. These findings are consistent with ESI-MS data on other acetylated proteins [74]. Native ESI experiments on myoglobin variants previously found a slight dependence of charge states on the amino acid composition [82], but those effects might be attributable to conformational factors (discussed in more detail below). In contrast, under the conditions of the current work, CRM charge states at pH 7 were remarkably insensitive to changes in the number of basic sites.

The above CRM-related statements call for a slight qualifier. The high solution charge of unmodified cyt *c* (7.7+ at pH 7) implies that a fraction of protein ions may form via the IEM, while the majority follows the CRM [61]. However, acetylation of a few Lys reduces the cyt *c* solution charge sufficiently to eliminate protein IEM. Already at the lowest concentration of labeling agent (250-fold excess) the average number of acetylated Lys is 6.3 (Figure 3B) such that the cyt *c* solution charge drops close to zero at pH 7. Thus, protein IEM can be excluded for native ESI measurements on any of the labeled samples studied here [61].

High ESI Charge States of Unfolded Cyt c at pH 2: Evidence for CEM Behavior. As noted, the ESI mass spectra of acetylated cyt *c* at pH 2 showed bimodal charge state distributions (Figure 4F-H). Low charge states (9+ to 7+) can be attributed to compact (albeit nonnative) solution conformers, that dominate these samples as seen from the optical data of Figure 6C, D. These low charge states will be discussed in the subsequent section.

Here we focus on the highly charged ions (21+ to 10+) in the bimodal charge state distributions of Figure 4E-H. These highly charged ions reveal that, even after acetylation, all of the the pH 2 samples still contain a sub-population of solution-phase conformers that are extensively unfolded [27,38-43]. The concentration of these unfolded chains in solution appears to be quite low (Figure 6), but the presence of exposed hydrophobic residues dramatically enhances their ionization efficiency such that they are remain observable with relatively high signal intensity in Figure 4E-H [83,84]. A striking feature if these highly charged ions is their dramatic acetylation-induced charge state shift, with maxima that change from 16+ in the absence of acetylation down to 11+ for the fully acetylated protein (Figure 4E-H, see also the magenta profiles in Figure 5C, F). In the following discussion we make the case that this acetylation-induced charge state shift is a direct consequence of ion formation via the CEM.

A central element of the CEM is the occurrence of electrostatically driven charge equilibration (via migration of mobile H^+) between the droplet and its protein tail. This equilibration takes place until the protein separates from the droplet. The z value of the gaseous protein ion is therefore determined by the number of excess H^+ that reside on the chain at the point when separation occurs [6,26,27,46,47]. Under CEM conditions it is to be expected that reducing the number of basic sites will lower the capacity of the protein to accommodate H^+ , thereby decreasing the ESI charge state. The behavior of the highly charged *cyt c* ions in our pH 2 experiments (21+ to 10+, Figures 4, 5) confirms this CEM-based expectation.

An electrostatic toy model can be used to semi-quantitatively illustrate the effects of Lys acetylation on ESI charge states under CEM conditions. The model has previously been used in a CEM context [85]. We consider a stretched-out *cyt c* chain that is about to separate from the ESI droplet (Figure 7A). The droplet and all basic residues (Lys, Arg, His) are assumed to be spherical conductors. The net charge of the system q_{tot} (which represents the excess protons in the system)

can equilibrate freely between the droplet and the basic residues, driven by the tendency of the system to minimize its electrostatic energy. The positions of basic sites along the chain reflect the *cyt c* sequence. The droplet radius r_0 was chosen according to eq. 1, subject to the condition that the non-acetylated protein had to have a 15.7+ charge state (matching the experimental z_{avg} at pH 2 prior to acetylation, Figure 5C, F). Subsequent Lys acetylation was implemented by eliminating basic sites from the model. Ten different random acetylation patterns were tested for each acetylation level. Additional details are outlined in the SI.

The CEM toy model predicts a dramatic dependence of the protein ESI charge state on the number of acetylated Lys. Two scenarios were tested; ejection with the N-terminus first and with the C-terminus first. For both scenarios the modeled average protein charge state dropped from 15.7+ to 6.0+ as the number of acetylated Lys was raised from zero to 19 (blue and green symbols, Figure 7B). For proteins containing zero up to ~10 acetyl groups, the predicted trend is in remarkable agreement with experimental data for the highly charged *cyt c* ions at pH 2 (magenta profile in Figure 7B). This agreement supports the view that the experimentally observed drop in z_{avg} (from 16+ to 11+) with increasing Lys acetylation for highly charged *cyt c* ions at pH 2 represents a characteristic feature of the CEM.

For proteins carrying more than ~10 acetyl groups the model predictions diverge from the experimental data, suggesting that additional considerations are required to account for the behavior of these very highly acetylated proteins (such as allowing for conformations that are not fully stretched [86], or the possible protonation at sites other than Lys/Arg/His [47,80]). It is nonetheless remarkable that the extremely simple CEM model of Figure 7 can successfully describe the charge state shifts of highly charged *cyt c* ions at pH 2 for chains carrying up to ~10 acetyl groups.

Low ESI Charge States of Unfolded Cyt c at pH 2. Interpreting the behavior of the low charge states (9+ to 7+) in the pH 2 spectra is more difficult. The bimodal nature of these spectra suggests that there is an ESI mechanistic difference between high charge states (see previous section) and the low charge states (discussed here). It is intriguing that the low charge states shift from 9+/8+ (Figure 4F) to 8+/7+ with increasing acetylation (Figure 4H). On the one hand, it might be tempting to attribute this acetylation dependence to ion formation via the CEM, analogous to the discussion of highly charged ions in the preceding section. On the other hand, there are several arguments against the CEM for these low charge states. (i) Their charge states are very similar to those of the CRM ions formed at pH 7 (Figure 4A-D). (ii) The optical data of Figure 6C/D reveal that the low charge ions at pH 2 were formed from compact solution conformers. (iii) It is unlikely that the electrostatic repulsion of these low charge ions would be sufficient for CEM ejection from the droplet [27].

In our view, the most likely formation mechanism for the low charge states at pH 2 is the CRM. We attribute the slight charge state shift from 9+/8+ to 8+/7+ in Figure 4F-H to acetylation-induced compaction of the protein in solution. The net *cyt c* solution charge at pH 2 under the conditions of Figure 4F is 18+ (~7 acetyl groups), while the fully acetylated protein in Figure 4H (~19 acetyl groups) has a solution charge of 6+. The lower solution charge of the fully acetylated protein reduces intramolecular charge repulsion, thereby favoring a slightly more compact solution structure. From eq. 1 it can be estimated that a small change in effective protein radius (from 1.8 nm to 1.7 nm) can account for the experimentally observed charge state shift from 9+/8+ to 8+/7+. In summary, the CRM in combination with a subtle acetylation-induced conformational change appears to be what governs the formation of *cyt c* low charge state ions at pH 2.

4. Conclusions

Understanding how proteins are transferred from solution into the gas phase during ESI remains challenging [20-35]. Many recent advances in this area have been based on molecular dynamics simulations [24,27,28,30,31,87-89], but such computational data should be taken with a grain of salt until they are backed up by experiments. When detecting electrosprayed proteins by IMS/MS it is usually impossible to extract the ion “history”, i.e., the mechanism by which analytes were released from solution. The fact that that denaturing ESI produces higher protein charge states than native ESI strongly suggests that different mechanisms are at play in these two cases. However, charge states alone do not provide clear-cut answers when it comes to ion formation pathways. The current work provides a new experimental tool that can help distinguish ESI mechanisms.

Our data demonstrate that progressive Lys acetylation does not affect the charge states of protein ions in native ESI, as long as the protein conformation in solution is unaffected by these chemical modifications. This behavior represents a hallmark of the CRM where charge states are governed by protein size rather than protein surface chemistry (eq. 1) [3,27,37,90]. In contrast, the ESI charge states of unfolded proteins decrease dramatically as more and more Lys are acetylated – an effect that is consistent with the CEM [6,26,27,46,47]. The CEM toy model of Figure 7 explains why Lys acetylation reduces the charge of electrosprayed protein ions; it is because lowering the number of basic sites compromises the ability of the departing chain to compete with the droplet for mobile H^+ . In summary, the response of electrosprayed proteins to the covalent blockage of basic sites is an interesting approach for ESI mechanistic studies, although the interpretation of such experiments can be complicated by protein conformational changes.

Acknowledgments. We thank Jonathan R. Adsetts from the laboratory of Prof. Zhifeng Ding, as well as Lee-Ann Briere for their help with the optical experiments of this work.

References

- [1] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* 246 (1989) 64-71.
- [2] P. Kebarle, L. Tang, From ions in solution to ions in the gas phase: The mechanism of electrospray mass spectrometry. *Anal. Chem.* 65 (1993) 972A-986A.
- [3] P. Kebarle, U.H. Verkerk, Electrospray: From Ions in Solutions to Ions in the Gas Phase, What We Know Now. *Mass Spectrom. Rev.* 28 (2009) 898-917.
- [4] M. Wilm, A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, M. Mann, Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 379 (1996) 466-469.
- [5] A.C. Leney, A.J.R. Heck, Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* 28 (2017) 5-13.
- [6] S. Mehmood, T.M. Allison, C.V. Robinson, Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annu. Rev. Phys. Chem.* 66 (2015) 453-474.
- [7] K.B. Shelimov, D.E. Clemmer, R.R. Hudgins, M.F. Jarrold, Protein Structure in Vacuo: The Gas-Phase Conformation of BPTI and Cytochrome c. *J. Am. Chem. Soc.* 119 (1997) 2240-2248.
- [8] T.M. Allison, P. Barran, S. Cianferani, M.T. Degiacomi, V. Gabelica, R. Grandori, E.G. Marklund, T. Menneteau, L.G. Migas, A. Politis, M. Sharon, F. Sobott, K. Thalassinou, J.L.P. Benesch, Computational Strategies and Challenges for Using Native Ion Mobility Mass Spectrometry in Biophysics and Structural Biology. *Anal. Chem.* 92 (2020) 10872-10880.
- [9] L.M. Young, J.C. Saunders, R.A. Mahood, C.H. Reville, R.J. Foster, L.-H. Tu, D.P. Raleigh, S.E. Radford, A.E. Ashcroft, Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry–mass spectrometry. *Nat. Chem.* 7 (2015) 73-81.
- [10] J.S. Brodbelt, Ion Activation Methods for Peptides and Proteins. *Anal. Chem.* 88 (2016) 30-51.
- [11] X. Han, M. Jin, K. Breuker, F.W. McLafferty, Extending Top-Down Mass Spectrometry to Proteins with Masses Greater Than 200 Kilodaltons. *Science* 314 (2006) 109-112.
- [12] J.G. Bonner, Y.A. Lyon, C. Nellesen, R.R. Julian, Photoelectron Transfer Dissociation Reveals Surprising Favorability of Zwitterionic States in Large Gaseous Peptides and Proteins. *J. Am. Chem. Soc.* 139 (2017) 10286-10293.
- [13] J. Seo, W. Hoffmann, S. Warnke, M.T. Bowers, K. Pagel, G. von Helden, Retention of Native Protein Structures in the Absence of Solvent: A Coupled Ion Mobility and Spectroscopic Study. *Angew. Chem.-Int. Edit.* 55 (2016) 14173-14176.
- [14] M.F. Czar, F. Zosel, I. Koenig, D. Nettels, B. Wunderlich, B. Schuler, A. Zarrine-Afsar, R.A. Jockusch, Gas-Phase FRET Efficiency Measurements To Probe the Conformation of Mass-Selected Proteins. *Anal. Chem.* 87 (2015) 7559-7565.
- [15] S. Daly, F. Rosu, V. Gabelica, Mass-resolved electronic circular dichroism ion spectroscopy. *Science* 368 (2020) 1465-+.

- [16] E.N. Kitova, A. El-Hawiet, P.D. Schnier, J.S. Klassen, Reliable Determinations of Protein–Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass Spectrom.* 23 (2012) 431-441.
- [17] G.T.H. Nguyen, T.N. Tran, M.N. Podgorski, S.G. Bell, C.T. Supuran, W.A. Donald, Nanoscale Ion Emitters in Native Mass Spectrometry for Measuring Ligand–Protein Binding Affinities. *ACS Centr. Sci.* 5 (2019) 308–318.
- [18] A.C. Susa, Z.J. Xia, E.R. Williams, Native Mass Spectrometry from Common Buffers with Salts That Mimic the Extracellular Environment. *Angew. Chem.-Int. Edit.* 56 (2017) 7912-7915.
- [19] J.E. Keener, D.E. Zambrano, G.Z. Zhang, C.K. Zak, D.J. Reid, B.S. Deodhar, J.E. Pemberton, J.S. Prell, M.T. Marty, Chemical Additives Enable Native Mass Spectrometry Measurement of Membrane Protein Oligomeric State within Intact Nanodiscs. *J. Am. Chem. Soc.* 141 (2019) 1054-1061.
- [20] J.V. Iribarne, B.A. Thomson, On the evaporation of small ions from charged droplets. *J. Chem. Phys.* 64 (1976) 2287-2294.
- [21] I.G. Loscertales, J.F. de la Mora, Experiments on the kinetics of field evaporation of small ions from droplets. *J. Chem. Phys.* 103 (1995) 5041-5060.
- [22] G. Wang, R.B. Cole, Charged residue versus ion evaporation for formation of alkali metal halide clusters ions in ESI. *Anal. Chim. Acta* 406 (2000) 53-65.
- [23] A.T. Iavarone, E.R. Williams, Mechanism of Charging and Supercharging Molecules in Electrospray Ionization. *J. Am. Chem. Soc.* 125 (2003) 2319-2327.
- [24] V. Znamenskiy, I. Marginean, A. Vertes, Solvated Ion Evaporation from Charged Water Droplets. *J. Phys. Chem. A* 107 (2003) 7406-7412.
- [25] C.J. Hogan, J.A. Carroll, H.W. Rohrs, P. Biswas, M.L. Gross, Combined Charged Residue-Field Emission Model of Macromolecular Electrospray Ionization. *Anal. Chem.* 81 (2009) 369-377.
- [26] M.T. Donor, S.A. Ewing, M.A. Zenaidee, W.A. Donald, J.S. Prell, Extended Protein Ions Are Formed by the Chain Ejection Model in Chemical Supercharging Electrospray Ionization. *Anal. Chem.* 89 (2017) 5107-5114.
- [27] L. Konermann, H. Metwally, Q. Duez, I. Peters, Charging and Supercharging of Proteins for Mass Spectrometry: Recent Insights into the Mechanisms of Electrospray Ionization. *Analyst* 144 (2019) 6157-6171.
- [28] E.I. Calixte, O.T. Liyanage, H.J. Kim, E.D. Ziperman, A.J. Pearson, E.S. Gallagher, Release of Carbohydrate-Metal Adducts from Electrospray Droplets: Insight into Glycan Ionization by Electrospray. *J. Phys. Chem. B* 124 (2020) 479-486.
- [29] V. Kwan, S. Consta, Molecular Characterization of the Surface Excess Charge Layer in Droplets. *J. Am. Soc. Mass Spectrom.* 32 (2021) 33-45.
- [30] D. Kim, N. Wagner, K. Wooding, D.E. Clemmer, D.H. Russell, Ions from Solution to the Gas Phase: A Molecular Dynamics Simulation of the Structural Evolution of Substance P during Desolvation of Charged Nanodroplets Generated by Electrospray Ionization. *J. Am. Chem. Soc.* 139 (2017) 2981-2988.
- [31] S.G. Kondalaji, M. Khakinejad, S.J. Valentine, Comprehensive Peptide Ion Structure Studies Using Ion Mobility Techniques: Part 3. Relating Solution-Phase to Gas-Phase Structures. *J. Am. Soc. Mass Spectrom.* 29 (2018) 1665-1677.

- [32] P. Tiwari, M.F. Czar, R. Zenobi, Fluorescence-Based Detection of the Desolvation Process of Protein Ions Generated in an Aqueous Electrospray Plume. *Anal. Chem.* 93 (2021) 3635-3642.
- [33] R.R. Ogorzalek Loo, R. Lakshmanan, J.A. Loo, What Protein Charging (and Supercharging) Reveal about the Mechanism of Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* 25 (2014) 1675-1693.
- [34] J. Li, C. Santambrogio, S. Brocca, G. Rossetti, P. Carloni, R. Grandori, Conformational Effects in Protein Electrospray Ionization Mass Spectrometry. *Mass Spectrom. Rev.* 35 (2016) 111-122.
- [35] S. Nguyen, J.B. Fenn, Gas-phase ions of solute species from charged droplets of solutions. *Proc. Natl. Acad. Sci. U.S.A.* 104 (2007) 1111-1117.
- [36] R.L. Grimm, J.L. Beauchamp, Evaporation and Discharge Dynamics of Highly Charged Multicomponent Droplets Generated by Electrospray Ionization. *J. Phys. Chem. A* 114 (2010) 1411-1419.
- [37] J.F. de la Mora, Electrospray Ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta* 406 (2000) 93-104.
- [38] J.A. Loo, C.G. Edmonds, H.R. Udseh, R.D. Smith, Effect of Reducing Disulfide-Containing Proteins on Electrospray Ionisation Mass Spectra. *Anal. Chem.* 62 (1990) 693-698.
- [39] S.K. Chowdhury, V. Katta, B.T. Chait, Probing Conformational Changes in Proteins by Mass Spectrometry. *J. Am. Chem. Soc.* 112 (1990) 9012-9013.
- [40] A.J.H. Borysic, S.E. Radford, A.E. Ashcroft, Co-populated Conformational Ensembles of b2-Microglobulin Uncovered Quantitatively by Electrospray Ionization Mass Spectrometry. *J. Biol. Chem.* 279 (2004) 27069-27077.
- [41] R. Grandori, Detecting equilibrium cytochrome *c* folding intermediates by electrospray ionization mass spectrometry: Two partially folded forms populate the molten globule state. *Protein Sci.* 11 (2002) 453-458.
- [42] A. Dobo, I.A. Kaltashov, Detection of Multiple Protein Conformational Ensembles in Solution via Deconvolution of Charge-State Distributions in ESI MS. *Anal. Chem.* 73 (2001) 4763-4773.
- [43] D.W. Woodall, L.W. Henderson, S.A. Raab, K. Honma, D.E. Clemmer, Understanding the Thermal Denaturation of Myoglobin with IMS-MS: Evidence for Multiple Stable Structures and Trapped Pre-equilibrium States. *Journal of the American Society for Mass Spectrometry* 32 (2021) 64-72.
- [44] G. Wang, R.B. Cole, Disparity Between Solution-phase Equilibria and Charge State Distributions in Positive-ion Electrospray Mass Spectrometry. *Org. Mass Spectrom.* 29 (1994) 419-427.
- [45] I.A. Kaltashov, R.R. Abzalimov, Do Ionic Charges in ESI MS Provide Useful Information on Macromolecular Structure? *J. Am. Soc. Mass Spectrom.* 19 (2008) 1239-1246.
- [46] R. Beveridge, L.G. Migas, R.K. Das, R.V. Pappu, R.W. Kriwacki, P.E. Barran, Ion Mobility Mass Spectrometry Uncovers the Impact of the Patterning of Oppositely Charged Residues on the Conformational Distributions of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* 141 (2019) 4908-4918.
- [47] M.A. Zenaidee, M.G. Leeming, F.T. Zhang, T.T. Funston, W.A. Donald, Highly Charged Protein Ions: The Strongest Organic Acids to Date. *Angew. Chem.-Int. Edit.* 56 (2017) 8522-8526.

- [48] M. Chen, L.X. Zheng, B. Santra, H.Y. Ko, R.A. DiStasio, M.L. Klein, R. Car, X.F. Wu, Hydroxide diffuses slower than hydronium in water because its solvated structure inhibits correlated proton transfer. *Nat. Chem.* 10 (2018) 413-419.
- [49] E. Brini, C.J. Fennell, M. Fernandez-Serra, B. Hribar-Lee, M. Luksic, K.A. Dill, How Water's Properties Are Encoded in Its Molecular Structure and Energies. *Chem. Rev.* 117 (2017) 12385-12414.
- [50] Y.X. Peng, J.M.J. Swanson, S.G. Kang, R.H. Zhou, G.A. Voth, Hydrated Excess Protons Can Create Their Own Water Wires. *J. Phys. Chem. B* 119 (2015) 9212-9218.
- [51] S. Cukierman, Et tu, Grotthuss! and other unfinished stories. *Biochim. Biophys. Acta* 1757 (2006) 876-885.
- [52] R.K. Boyd, Á. Somogyi, The Mobile Proton Hypothesis in Fragmentation of Protonated Peptides: A Perspective. *J. Am. Soc. Mass Spectrom.* 21 (2010) 1275-1278.
- [53] A.R. Dongré, J.L. Jones, Á. Somogyi, V.H. Wysocki, Influence of Peptide Composition, Gas-Phase Basicity, and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model. *J. Am. Chem. Soc.* 118 (1996) 8365-8374.
- [54] T.J.D. Jørgensen, H. Gårdsvoll, M. Ploug, P. Roepstorff, Intramolecular Migration of Amide Hydrogens in Protonated Peptides upon Collisional Activation. *J. Am. Chem. Soc.* 127 (2005) 2785-2793.
- [55] J.Y. Li, W.P. Lyu, G. Rossetti, A. Konijnenberg, A. Natalello, E. Ippoliti, M. Orozco, F. Sobott, R. Grandori, P. Carloni, Proton Dynamics in Protein Mass Spectrometry. *J. Phys. Chem. Lett.* 8 (2017) 1105-1112.
- [56] N. Felitsyn, E.N. Kitova, J.S. Klassen, Thermal Decomposition of a Gaseous Multiprotein Complex Studied by Blackbody Infrared Radiative Dissociation. Investigating the Origin of the Asymmetric Dissociation Behavior. *Anal. Chem.* 73 (2001) 4647-4661.
- [57] J.C. Jurchen, E.R. Williams, Origin of Asymmetric Charge Partitioning in the Dissociation of Gas-Phase Protein Homodimers. *J. Am. Chem. Soc.* 125 (2003) 2817-2826.
- [58] J.L.P. Benesch, Collisional Activation of Protein Complexes: Picking Up the Pieces. *J. Am. Soc. Mass Spectrom.* 20 (2009) 341-348.
- [59] S.K. Fegan, M. Thachuk, A Charge Moving Algorithm for Molecular Dynamics Simulations of Gas-Phase Proteins. *J. Chem. Theory Comput.* 9 (2013) 2531-2539.
- [60] S.V. Sciuto, J. Liu, L. Konermann, An Electrostatic Charge Partitioning Model for the Dissociation of Protein Complexes in the Gas Phase. *J. Am. Soc. Mass Spectrom.* 22 (2011) 1679-1689.
- [61] E. Aliyari, L. Konermann, Formation of Gaseous Proteins via the Ion Evaporation Model (IEM) in Electrospray Mass Spectrometry. *Anal. Chem.* 92 (2020) 10807-10814.
- [62] P.D. Schnier, D.S. Gross, E.R. Williams, On the Maximum Charge State and Proton Transfer Reactivity of Peptide and Protein Ions Formed By Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* 6 (1995) 1086-1097.
- [63] V.L. Mendoza, R.W. Vachet, Probing Protein Structure by Amino Acid-specific Covalent Labeling and Mass Spectrometry. *Mass Spectrom. Rev.* 28 (2009) 785-815.
- [64] Y. Goto, N. Takahashi, A.L. Fink, Mechanism of Acid-Induced Folding of Proteins. *Biochemistry* 29 (1990) 3480-3488.

- [65] M. Graf, R.G. Garcia, H. Watzig, Protein adsorption in fused-silica and polyacrylamide-coated capillaries. *Electrophoresis* 26 (2005) 2409-2417.
- [66] K. Wada, K. Okunuki, Studies on Chemically Modified Cytochrome cI. The Acetylated Cytochrome c. *Journal of Biochemistry* 64 (1968) 667-&.
- [67] T. Mikami, T. Takao, K. Yanagi, H. Nakazawa, N (alpha) Selective Acetylation of Peptides. *Mass Spectrom. (Tokyo, Japan)* 1 (2012) A0010-A0010.
- [68] H.R. Bosshard, M. Zurrer, The conformation of cytochrome c in solution. Localization of a conformational difference between ferri-and ferrocytochrome c on the surface of the molecule. *J. Biol. Chem.* 255 (1980) 6694-6699.
- [69] G.W. Bushnell, G.V. Louie, G.D. Brayer, High-resolution Three-dimensional Structure of Horse Heart Cytochrome c. *J. Mol. Biol.* 214 (1990) 585-595.
- [70] B.T. Turner, T.M. Sabo, D. Wilding, M.C. Maurer, Mapping of factor XIII solvent accessibility as a function of activation state using chemical modification methods. *Biochemistry* 43 (2004) 9755-9765.
- [71] A.C. Susa, Z.J. Xia, E.R. Williams, Small Emitter Tips for Native Mass Spectrometry of Proteins and Protein Complexes from Nonvolatile Buffers That Mimic the Intracellular Environment. *Anal. Chem.* 89 (2017) 3116-3122.
- [72] J.C. May, E. Jurneczko, S.M. Stow, I. Kratochvil, S. Kalkhof, J.A. McLean, Conformational landscapes of ubiquitin, cytochrome c, and myoglobin: Uniform field ion mobility measurements in helium and nitrogen drift gas. *International Journal of Mass Spectrometry* 427 (2018) 79-90.
- [73] H. Zhang, W. Cui, M.L. Gross, Native electrospray ionization and electron-capture dissociation for comparison of protein structure in solution and the gas phase. *Int. J. Mass Spectrom.* 354-355 (2013) 288-291.
- [74] J.D. Carbeck, J.C. Severs, J.M. Gao, Q.Y. Wu, R.D. Smith, G.M. Whitesides, 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. *J. Phys. Chem. B* 102 (1998) 10596-10601.
- [75] B.F. Shaw, H. Arthanari, M. Narovlyansky, A. Durazo, D.P. Frueh, M.P. Pollastri, A. Lee, B. Bilgicer, S.P. Gygi, G. Wagner, G.M. Whitesides, Neutralizing Positive Charges at the Surface of a Protein Lowers Its Rate of Amide Hydrogen Exchange without Altering Its Structure or Increasing Its Thermostability. *J. Am. Chem. Soc.* 132 (2010) 17411-17425.
- [76] N.J. Greenfield, Methods to Estimate the Conformation of Proteins and Polypeptides from Circular Dichroism Data. *Anal. Biochem.* 235 (1996) 1 - 10.
- [77] H. Roder, K. Maki, H. Cheng, Early Events in Protein Folding Explored by Rapid Mixing Methods. *Chem. Rev.* 106 (2006) 1836-1861.
- [78] L. Konermann, D.J. Douglas, Acid-Induced Unfolding of Cytochrome c at Different Methanol Concentrations: Electrospray Ionization Mass Spectrometry Specifically Monitors Changes in the Tertiary Structure. *Biochemistry* 36 (1997) 12296-12302.
- [79] T.E. Creighton. *Proteins*; W. H. Freeman & Co: New York, 1993.

- [80] P.D. Schnier, D.S. Gross, E.R. Williams, Electrostatic Forces and Dielectric Polarizability of Multiply Protonated Gas-Phase Cytochrome c Ions Probed by Ion/Molecule Chemistry. *J. Am. Chem. Soc.* 117 (1995) 6747-6757.
- [81] D.R. Lide. *CRC Handbook of Chemistry and Physics* 82nd ed.; CRC Press: Boca Raton, London, New York, Washington, 2001.
- [82] M. Samalikova, R. Grandori, Role of opposite charge in protein electrospray ionization mass spectrometry. *J. Mass Spectrom.* 38 (2003) 941-947.
- [83] M.C. Kuprowski, L. Konermann, Signal Response of Co-Existing Protein Conformers in Electrospray Mass Spectrometry. *Anal. Chem.* 79 (2007) 2499-2506.
- [84] A.P. Null, A.I. Nepomuceno, D.C. Muddiman, Implications of Hydrophobicity and Free Energy of Solvation for Characterization of Nucleic Acids by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* 75 (2003) 1331-1339.
- [85] I. Peters, H. Metwally, L. Konermann, Mechanism of Electrospray Supercharging for Unfolded Proteins: Solvent-Mediated Stabilization of Protonated Sites During Chain Ejection. *Anal. Chem.* 91 (2019) 6943-6952.
- [86] L. Konermann, A.D. Rodriguez, J. Liu, On the Formation of Highly Charged Gaseous Ions from Unfolded Proteins by Electrospray Ionization. *Anal. Chem.* 84 (2012) 6798-6804.
- [87] S. Consta, J.K. Chung, Charge-Induced Conformational Changes of PEG-(Na⁺)_n in Vacuum and Aqueous Nanodroplets. *J. Phys. Chem. B* 115 (2011) 10447-10455.
- [88] A. Patriksson, E. Marklund, D. van der Spoel, Protein Structures under Electrospray Conditions. *Biochemistry* 46 (2007) 933-945.
- [89] M. Porrini, F. Rosu, C. Rabin, L. Darre, H. Gomez, M. Orozco, V. Gabelica, Compaction of Duplex Nucleic Acids upon Native Electrospray Mass Spectrometry. *ACS Central Sci.* 3 (2017) 454-461.
- [90] I.A. Kaltashov, A. Mohimen, Estimates of Protein Surface Area in Solution by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* 77 (2005) 5370-5379.
- [91] S.A. Ryce, R.R. Wyman, Two sphere model for the asymmetric division of electrically charged liquid drops. *Can. J. Phys.* 48 (1970) 2571-2576.
- [92] L. Konermann, A Simple Model for the Disintegration of Highly Charged Solvent Droplets during Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* 20 (2009) 496-506.
- [93] K.J. Mark, D.J. Douglas, Coulomb effects in binding of heme in gas-phase ions of myoglobin. *Rapid Commun. Mass Spectrom.* 20 (2006) 111-117.
- [94] A. Moser, K. Range, D.M. York, Accurate Proton Affinity and Gas-Phase Basicity Values for Molecules Important in Biocatalysis. *J. Phys. Chem. B* 114 (2010) 13911-13921.
- [95] Z. Markovic, J. Tosovic, D. Milenkovic, S. Markovic, Revisiting the solvation enthalpies and free energies of the proton and electron in various solvents. *Comput. Theor. Chem.* 1077 (2016) 11-17.

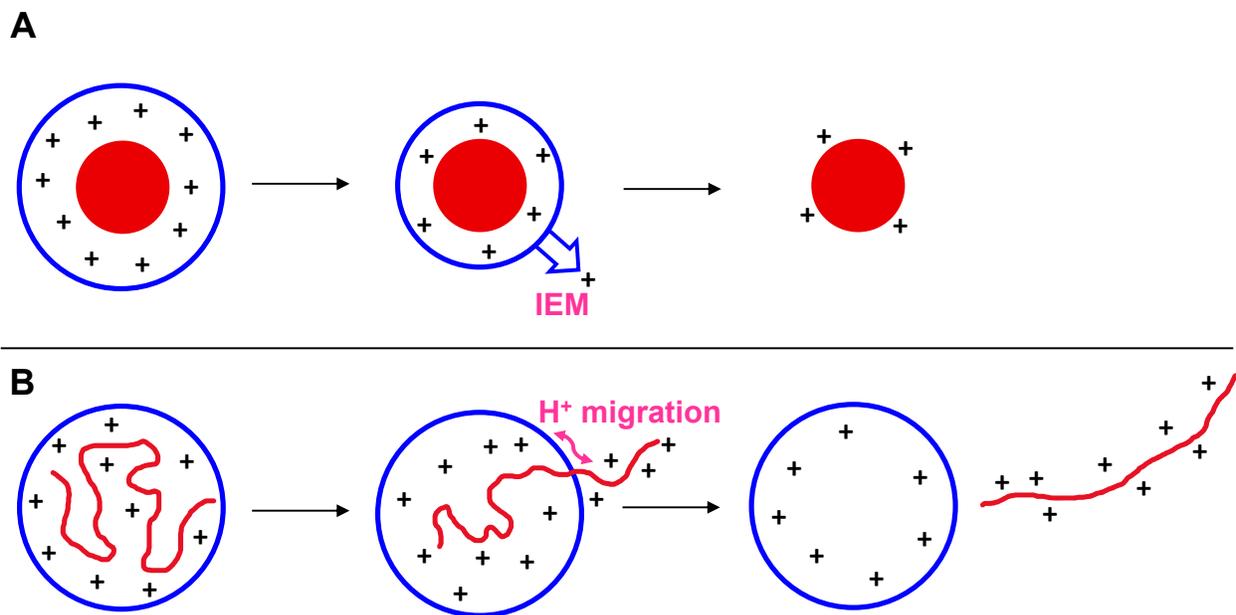


Figure 1. Cartoon depiction of two protein ESI models. ESI nanodroplets are shown in blue. The protein is depicted in red, excess H^+ are indicated as “+”. (A) CRM, where compact proteins are released via droplet evaporation to dryness. IEM events of H^+ and other small charge carries keep the shrinking droplet close to the Rayleigh limit. (B) CEM where stretched-out chains are ejected from the droplet as highly charged ions. The high charge states are caused by electrostatically-driven migration of mobile H^+ .

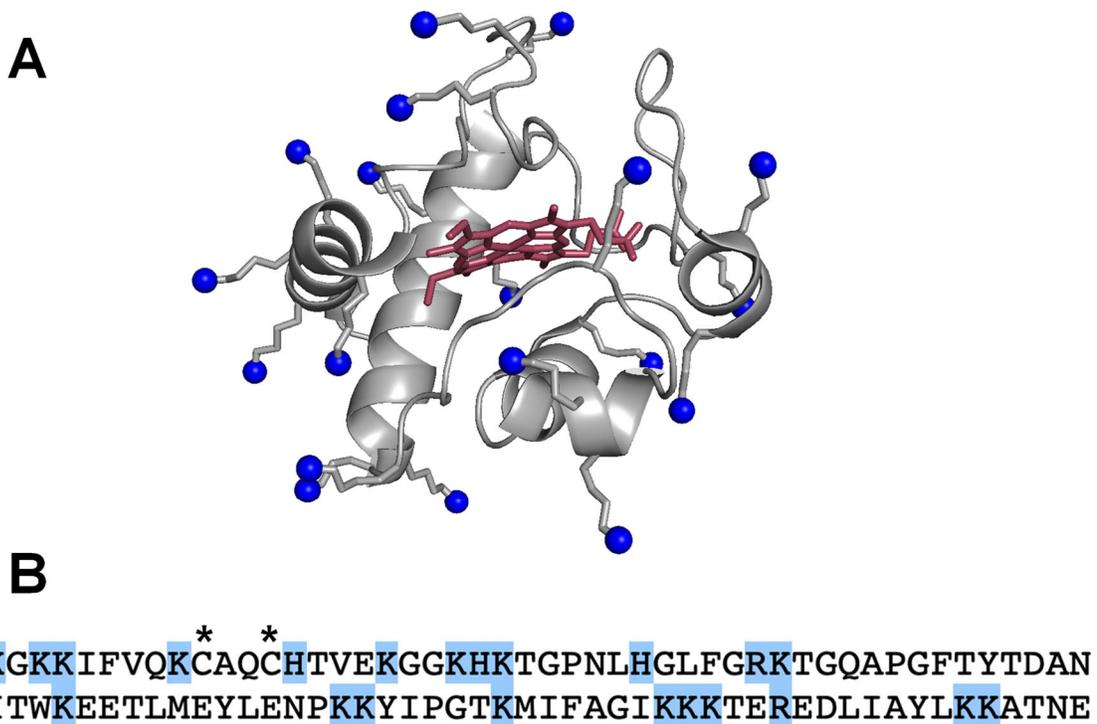


Figure 2. (A) Crystal structure of cyt *c* (1HRC). The 19 Lys side chains are shown as sticks, with ϵ -NH₂ groups (the sites of acetylation after acetic anhydride exposure) highlighted as blue spheres. (B) Cyt *c* sequence, with basic residues highlighted in blue. Asterisks indicate the sites of covalent heme attachment (Cys14 and Cys17).

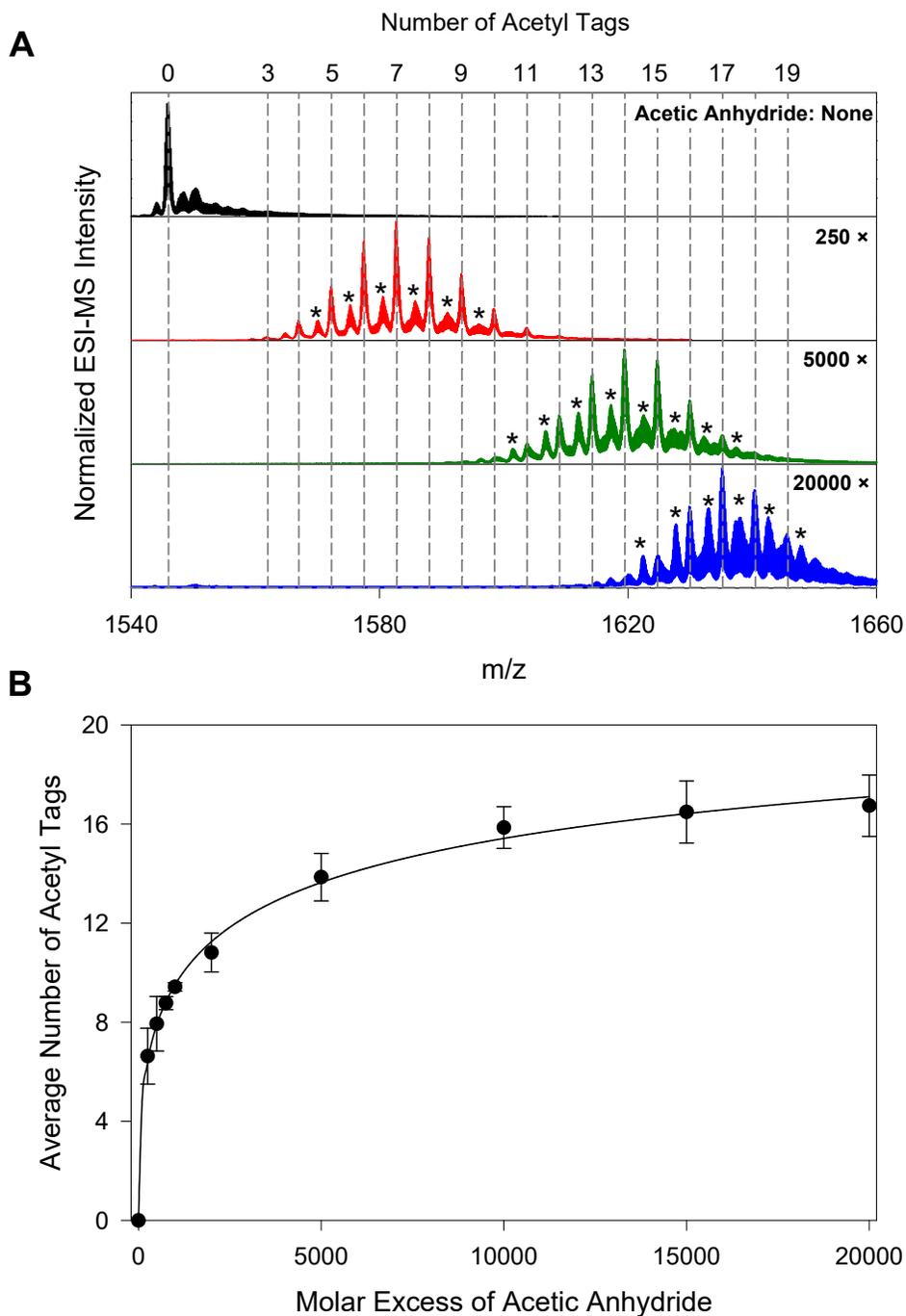


Figure 3. (A) Partial ESI mass spectra, showing the number of +42 Da acetyl tags for *cyt c 8+* at pH 7. Data are displayed for an unlabeled control, and after exposure to 250, 5000, and 20000-fold molar excess of acetic anhydride (top to bottom). Asterisks indicate Na adducts. (B) Average number of acetyl tags (eq. 2) vs. acetic anhydride molar excess. Experimental data (circles) are the average of three independent replicates, error bars represent standard deviations. An interpolation curve is included to guide the eye.

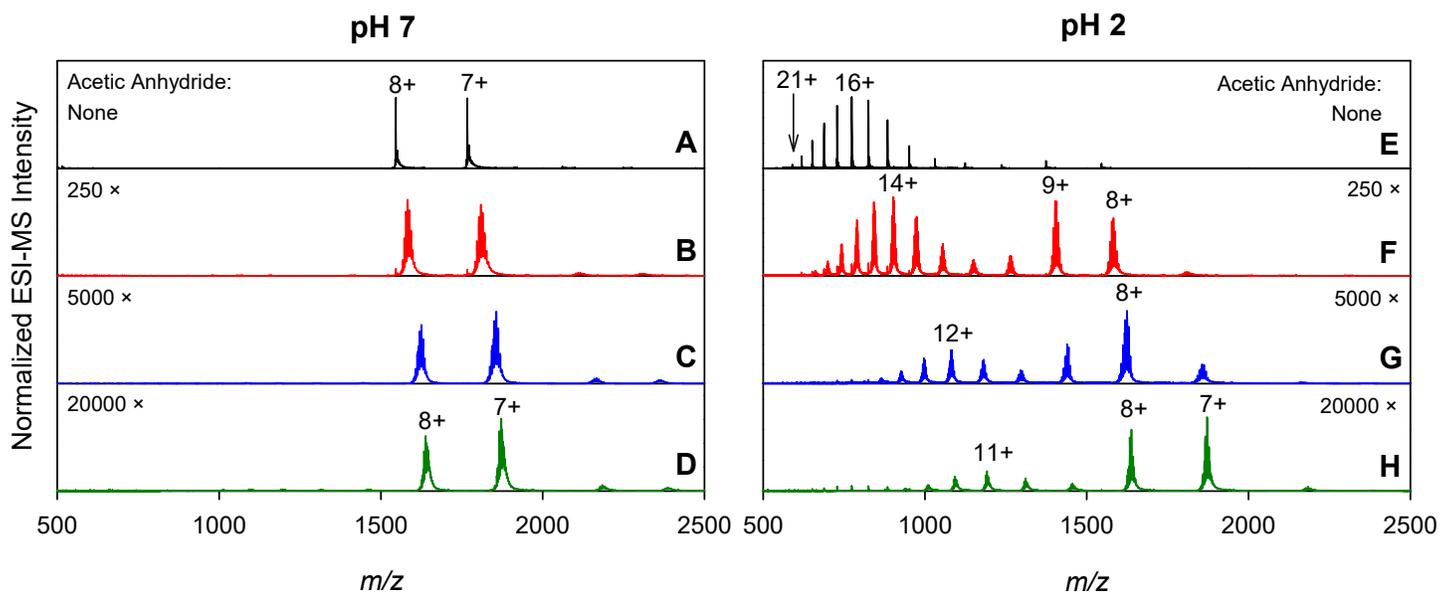


Figure 4. (A-D) Native ESI mass spectra of cyt *c* (aqueous solution at pH 7) after incubation with increasing amounts of acetic anhydride. (E-H) Mass spectra of the same samples electrosprayed under denaturing conditions at pH 2. The molar excess of acetic anhydride (relative to the protein) is indicated in each panel. Also indicated are the charge states of selected protein ions.

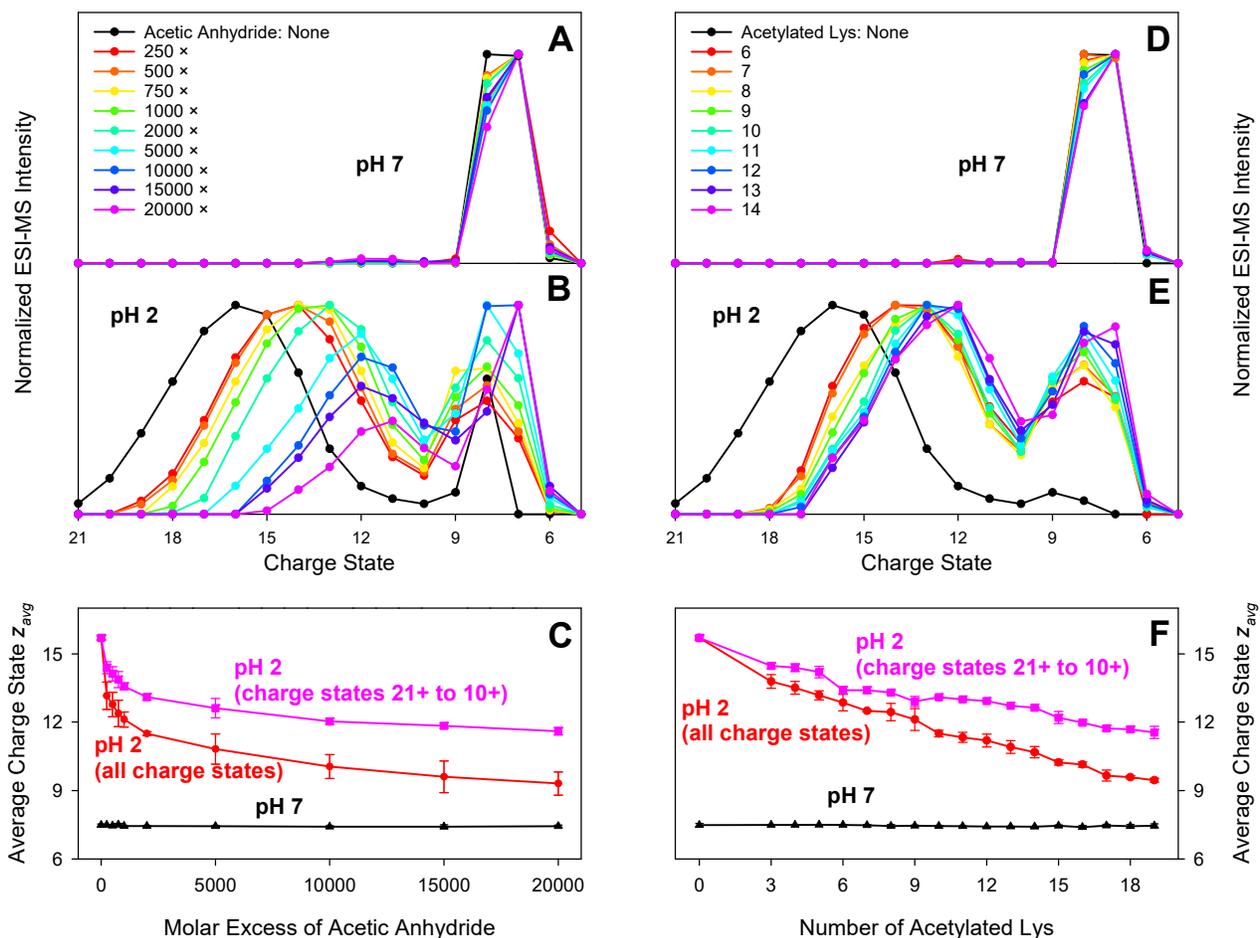


Figure 5. Panels on the left show *cyt c* charge state distributions (after adding the signal intensities of all acetylation levels for any given value of z) at (A) pH 7 and (B) pH 2 vs. acetic anhydride concentration. (C) Dependence of average charge state (eq. 3) on acetic anhydride concentration. Panels on the right show charge state distributions at (D) pH 7 and (E) pH 2 for protein ions carrying different numbers of acetylated Lys, extracted from experiments that used 2000 \times acetic anhydride. (F) Average charge state (eq. 3) vs. number of acetylated Lys, extracted from experiments that used between zero and 10000 \times acetic anhydride. Analysis of the pH 2 data in panels C and F was performed twice: for the entire charge state distributions (red) and only for the high charge states (21+ to 10+, magenta).

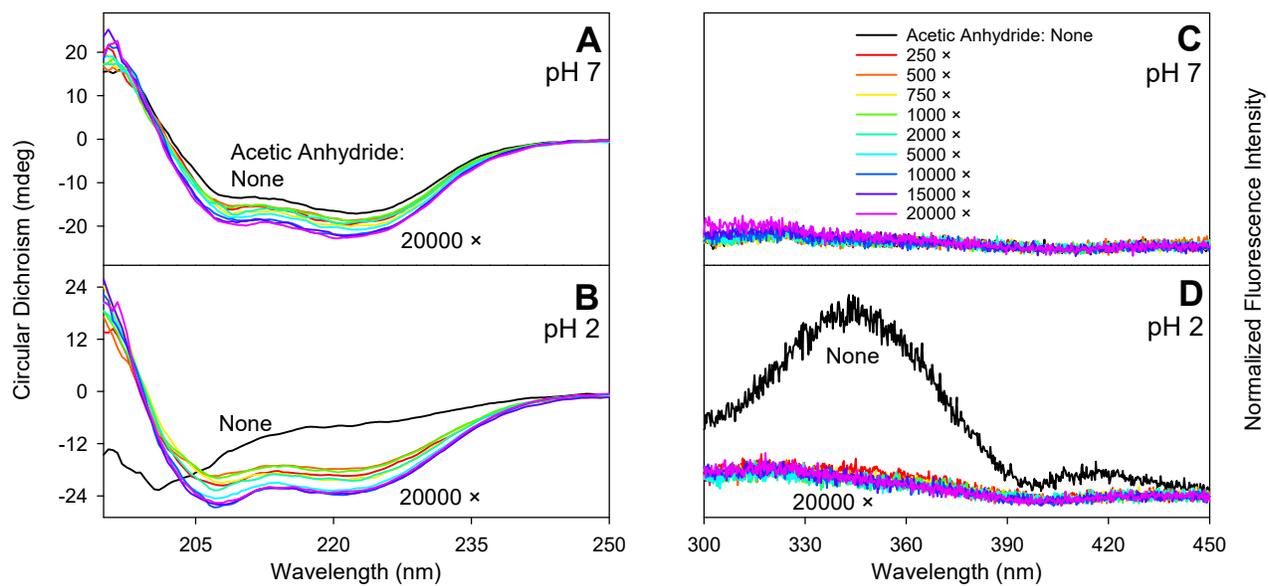


Figure 6. Acetylation-dependent changes of *cyt c* structure in bulk solution probed by optical spectroscopy. (A) CD spectra acquired at pH 7 and (B) at pH 2. (C) Trp fluorescence data recorded at pH 7 and (D) at pH 2.

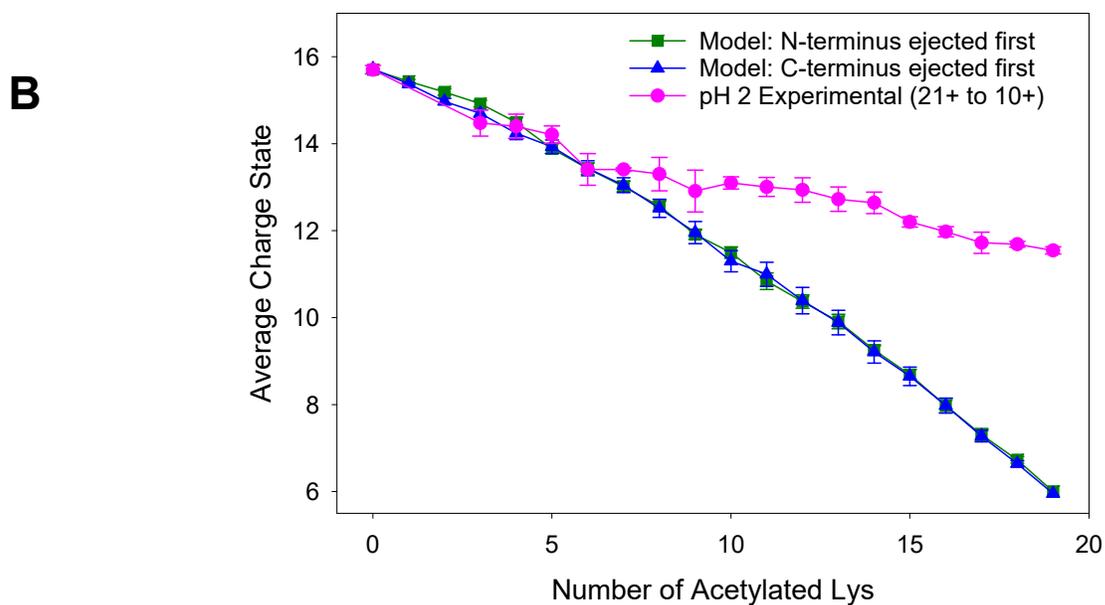
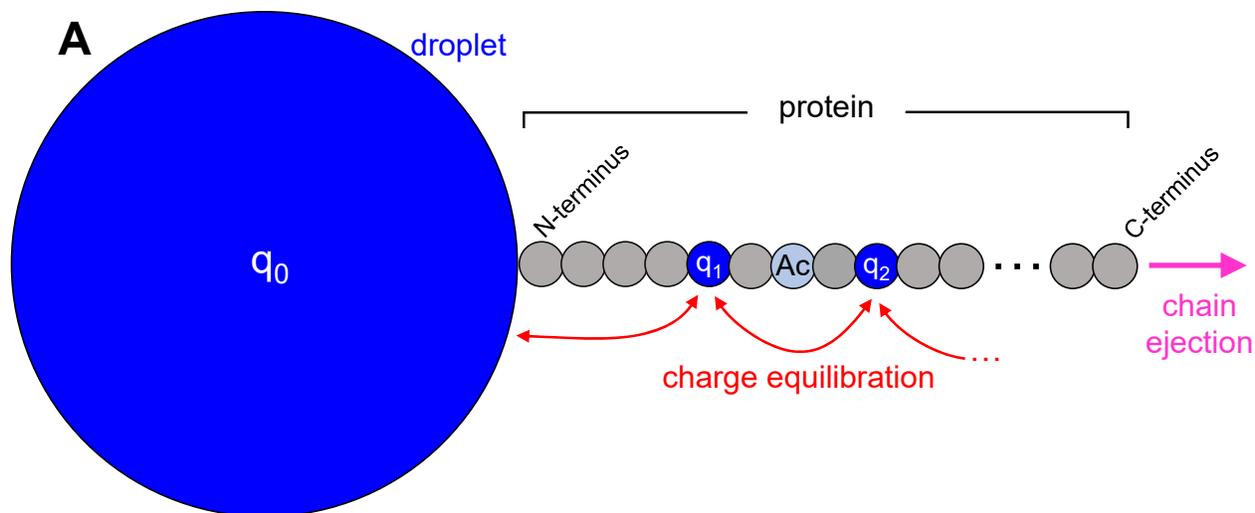
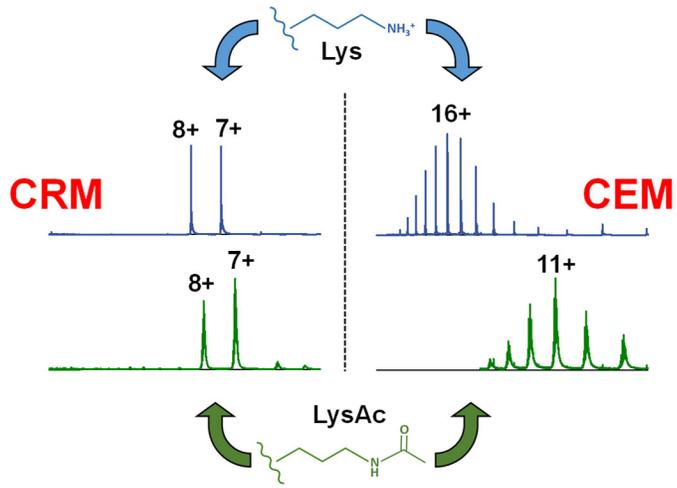


Figure 7. CEM model for exploring the effects of Lys acetylation on the ESI charge state of *cyt c*. (A) Model layout. A stretched-out protein (chain of small beads) is about to separate from the ESI droplet (large blue sphere). The droplet and basic residues (small blue beads: Lys, Arg, His) can carry charge. Acetylated Lys (pale blue, marked “Ac”) cannot carry charge. Also, all other residues (gray) cannot carry charge. The net charge of the system equilibrates (red arrows) between droplet and unmodified basic residues, such that Coulomb energy is minimized. (B) Average ESI charge of the protein vs. number of acetylated Lys. Experimental data are in magenta (pH 2, high charge state signals from Figure 5F). Green and blue symbols are model predictions for randomly acetylated Lys, for proteins being ejected with the N-terminus or C-terminus first.

Table of Contents Graphic



SUPPORTING INFORMATION

**Using Covalent Modifications to Distinguish Protein
Electrospray Mechanisms: Charged Residue Model (CRM)
vs. Chain Ejection Model (CEM)**

Douglas J. D. Pimlott and Lars Konermann*

*Department of Chemistry, The University of Western Ontario, London, Ontario,
N6A 5B7, Canada.*

This file contains:

CEM Electrostatic Toy Model: Implementation Details

CEM Electrostatic Toy Model: Implementation Details

The toy model of Figure 7A was used to qualitatively illustrate the impact of Lys acetylation on the charge states of protein ions that leave the ESI droplet via the CEM. The model represents extension of work by Ryce and Wyman [91], and it has previously been used in a CEM context [85,86], as well as for droplet fission [92], and for the collision-induced dissociation of protein complexes [60]. The model is structured as follows:

(i) The protein chain is ejected from the ESI droplet in a linear (fully stretched) conformation.

This assumption is consistent with Monte Carlo simulations [86] and MD results [27].

(ii) Charge can partition freely between the ESI droplet and the protruding chain; this charge partitioning takes place up to the point when droplet and protein separate from one another. This aspect of the model reflects the highly mobile nature of H^+ in water [48-51] and in gaseous proteins [52-55].

(iii) The droplet and all protein residues are treated as spherical beads with a density of 1 g cm^{-3} and radii that are denoted as r_i [37]; $i = 0$ represents the droplet, while $i = 1-105$ (104 residues plus the N-terminal acetyl group) represent the protein. Beads 1-105 all have the same size, while bead 0 is much larger (see details below). We assume that protein charge can only reside on the basic amino acids Lys, Arg, and His [59,62]. The positions of these basic beads are defined by the cyt *c* sequence [69]; giving rise to two scenarios where the chain leaves the droplet with the N-terminus or the C-terminus first. The midpoint distance between the protein beads is 4 \AA , which coincides with the C_α distance in a polypeptide chain [86]. The protein mass of 12360 Da corresponds to bead radii r_1 to r_{105} of 0.361 nm. The electrostatic framework used here assumes that all charge-carrying beads are physically separated [91], however, the bond distance and r_i values outlined above would result in partial overlap of some adjacent basic residues. To solve this problem

we slightly modified the placement of a few basic residue in the cyt *c* sequence (highlighted in color, below) by one position:

Original sequence:

AC-GDVEKGKKIF VQKCA $\color{red}{\text{Q}}$ CHTV EKGKHKHTGP NLHGLFGRKT GQAPGFTYTD
 ANKNKGITWK EETLMEYLEN PKKYIPGTM IFAGIKKKTE REDLIAYLKK ATNE

Modified sequence:

AC-GDVEKGKIKF VQKCA $\color{red}{\text{Q}}$ CHTV EKGKGHTKGP NLHGLFRGKT GQAPGFTYTD
 ANKNKGITWK EETLMEYLEN PKYKIPGTM IFAGKIKIKE REDLIAYLKA KTNE

In addition, we assigned a permanent 1+ charge to Q16, to account for the presence of one positive charge in this CA $\color{red}{\text{Q}}$ C region to account for the ferri-heme⁺ in cyt *c* [93].

(iv) The ESI droplet was assumed to be at the Rayleigh limit prior to ejection of the chain (eq. 1), with $\gamma = 0.05891 \text{ N m}^{-1}$ [81]. To match the experimentally observed charge state of 15.7+ for the non-acetylated protein, the droplet radius was set to $r_0 = 5.3 \text{ nm}$. This represents the only adjustable parameter of the model. This value of r_0 corresponds to a total charge of $q_{total} = 44+$ (eq. 1), which subsequently partitions over the droplet and all of the basic residues.

(v) The droplet and all basic residues were considered to be electric conductors. Each basic residue can carry a fractional charge $q_i \leq 1 e$. Charge partitioning is governed by the tendency of the system to minimize its overall electrostatic energy V

$$V = \frac{1}{4\pi\epsilon_0} \left(\frac{1}{2} \sum_{i=0}^N \frac{q_i^2}{r_i} + \sum_{i=0}^{N-1} \sum_{j=i+1}^N \frac{q_i q_j}{r_{ij}} \right) \quad (4)$$

where the index i refers to chargeable beads (droplet and basic residues) only. Image charges were not considered because they have been shown to only have very minor effects [91]. Similarly, proton

hydration energies and proton affinities were not included in the model because these factors are of very similar magnitude (around 1000 kJ mol^{-1}) such that they cancel out when comparing the energies of different charge partitioning patterns [94,95]. Lys acetylation was modeled by holding the charge of the corresponding beads at zero. Minimization of the electrostatic energy in eq. 4 was performed by adjusting the charges q_i on each chargeable bead using Microsoft Excel Solver. Predicted protein charge states $q_{protein}$ correspond to the sum of q_i for all basic residues, and $q_{droplet} = q_{protein} - q_{total}$.