Western University
Scholarship@Western

Chemistry Publications | Chemistry Department

8-4-2020

Formation of Gaseous Proteins via the Ion Evaporation Model (IEM) in Electrospray Mass Spectrometry.

Elnaz Aliyari

Lars Konermann

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

 Part of the Chemistry Commons

Citation of this paper:
Aliyari, Elnaz and Konermann, Lars, "Formation of Gaseous Proteins via the Ion Evaporation Model (IEM) in Electrospray Mass Spectrometry." (2020). *Chemistry Publications*. 242.
https://ir.lib.uwo.ca/chempub/242

Formation of Gaseous Proteins via the Ion Evaporation Model (IEM) in Electrospray Mass Spectrometry

Elnaz Aliyari and Lars Konermann*

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

* corresponding author: konerman@uwo.ca

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

ABSTRACT: The mechanisms whereby protein ions are released into the gas phase from charged droplets during electrospray ionization (ESI) continue to be controversial. Several pathways have been proposed. For native ESI the charged residue model (CRM) is favored; it entails the liberation of proteins via solvent evaporation to dryness. Unfolded proteins likely follow the chain ejection model (CEM), which involves the gradual expulsion of stretched-out chains from the droplet. According to the ion evaporation model (IEM) ions undergo electrostatically driven desorption from the droplet surface. The IEM is well supported for small precharged species such as Na^+. However, it is unclear whether proteins can show IEM behavior as well. We examined this question using molecular dynamics (MD) simulations, mass spectrometry (MS), and ion mobility spectrometry (IMS) in positive ion mode. Ubiquitin was chosen as model protein because of its structural stability which allows the protein charge in solution to be controlled via pH adjustment without changing the protein conformation. MD simulations on small ESI droplets (3 nm radius) showed CRM behavior regardless of the protein charge in solution. Surprisingly, many MD runs on larger droplets (5.5 nm radius) culminated in IEM ejection of ubiquitin, as long as the protein carried a sufficiently large positive solution charge. MD simulations predicted that nonspecific salt adducts are less prevalent for IEM-generated protein ions than for CRM products. This prediction was confirmed experimentally. Also, collision cross sections of MD structures were in good agreement with IMS data. Overall, this work reveals that the CRM, CEM, and IEM all represent viable pathways for generating gaseous protein ions during ESI. The IEM is favored for proteins that are tightly folded and highly charged in solution, and for droplets in a suitable size regime.

Almost four decades after the initial implementation of electrospray ionization (ESI) mass spectrometry (MS) by John Fenn,[1] the mechanism of the ESI process remains controversial.[2-9] Considering the central role of ESI-MS for countless applications, it is crucial to develop a comprehensive understanding of this area. This is especially true for proteins and protein complexes, for which ESI-MS represents one of the most important analytical techniques.[10-13] Only the early ESI steps are well understood; this includes the formation of micrometer-sized charged droplets at the outlet of the ESI or nanoESI emitter.[2] These droplets undergo solvent evaporation and jet fission, ultimately generating offspring droplets with radii of a few nanometers that are charged close to the Rayleigh limit.[2, 14, 15] Knowledge gaps persist regarding the mechanisms whereby gaseous analyte ions are released from these final ESI nanodroplets.

Early discussions of the ESI mechanism focused on two competing ideas, the ion evaporation model (IEM)[16] and the charged residue model (CRM).[17] The IEM was developed for small atomic ions such as Na^+ (Figure 1A).[16] Experiments and theoretical investigations suggest that such ions undergo desorption from the surface of highly charged droplets. The driving force for these IEM events is the electric field emanating from the droplet surface. IEM ejection requires the departing ion to overcome an activation barrier that arises from the interplay of several factors; these include surface deformation, rupture of solvent-solvent contacts, solvent polarization, and ion-droplet Coulombic repulsion.[2, 16, 18, 19]

It is an interesting question whether the IEM can also be operative for larger analytes such as globular proteins (Figure 1B). Most researchers argued that such a scenario is unlikely, and that large globular analytes instead follow the CRM (Figure 1C).[2, 7, 8, 20, 21] The defining characteristic of the CRM is that analyte ions are released via droplet evaporation to dryness. The CRM is supported by the observation that the charge states of globular proteins are close to the Rayleigh charge of

protein-sized water droplets.[6] Also, the well-known tendency of protein ions to undergo nonspecific adduction to salts and other solutes under native ESI conditions[10, 11] is consistent with solvent evaporation to dryness during the CRM.[2, 21-26]

The prevailing current view is that the IEM applies to small preformed ions, while the CRM describes the behavior of large globular species such as protein ions generated by native ESI.[2, 6, 7, 11, 21] However, the IEM vs. CRM debate is far from settled,[27-29] and there are unresolved issues that deserve closer attention. For example, the distinction between "small" and "large" analytes is contentious. Estimates of the IEM/CRM size cutoff vary widely between 100 Da,[8] 1 kDa,[30] 6 kDa,[6] and 5 MDa.[31] John Fenn himself was a proponent of the IEM; in 2007 he stated that "*for most, if not all, cases in which ESI is effective, gas-phase solute ions are formed from charged droplets according to the sequence of events described in the IEM*".[5] In a similar vein, Ogorzalek Loo et al. proposed the desorption of hydrated protein ions from deformed ESI droplets or from the tip of the Taylor cone.[4]

A third ESI scenario is the chain ejection model (CEM) which applies to unfolded proteins.[11, 21, 32-34] According to the CEM, exposed nonpolar residues drive protein chains to the droplet surface. The proteins are then gradually expelled from droplet via tadpole-shaped intermediates (Figure 1D). The high charge states of unfolded proteins[35-37] are attributed to the equilibration of mobile H^+ between the droplet and its protein tail, analogous to H^+ migration events that take place for collisionally heated protein complexes.[21, 38-40]

In recent years, molecular dynamics (MD) simulations of charged nanodroplets have emerged as an important tool for exploring ESI mechanisms.[21, 28, 32, 41] For example, it has become possible to obtain atomistic insights into the IEM ejection of small ions.[19, 21, 42, 43] Simulations on compact folded proteins,[21, 44] nucleic acid duplexes,[45] and peptides[46, 47] revealed CRM behavior,

while unfolded proteins were shown to follow the CEM.[21] Those MD studies provided strong support for the viability of the three ESI mechanisms depicted in Figure 1A, C, and D.

The current work takes a closer look at the feasibility of protein IEM events (Figure 1B), focusing on positive ion mode which is most widely used in experiments. We conducted MD simulations under conditions that have received little attention in previous studies, i.e., tightly folded proteins that are positively charged in solution and reside in relatively large (5.5 nm radius) ESI droplets. Surprisingly, we found that protein IEM under these conditions takes place readily. Our computational data are supported by experiments that explored the charge states, salt adduction behavior, and the conformations of electrosprayed protein ions. Thus, the current work demonstrates for the first time that all four mechanisms depicted in Figure 1 are viable. The prevalence of the different protein ESI scenarios (Figure 1B-D) depends on parameters such as analyte charge, conformation, and droplet size.

Materials and Methods

Bovine ubiquitin was from Sigma, St. Louis, MO. Native ESI samples contained 5 μM protein in 10 mM neutral aqueous ammonium acetate. Acidic or basic solutions were produced by addition of formic acid or ammonium hydroxide. Mass spectra and travelling wave ion mobility spectrometry (IMS) data were acquired on a Synapt G2 Si Q-TOF (Waters, Milford, MA). Solutions were infused at 5 μL min^{-1} using an ESI voltage of +2.8 kV. Data were acquired under gentle conditions using a cone voltage of 5 V, with source and desolvation temperatures of 25 °C and 40 °C, respectively. IMS arrival time distributions were converted to effective He Ω values.[48] Droplet MD simulations followed established protocols.[21] Briefly, we used Gromacs 2016[49] with the Charmm36 force field[50] and TIP4P/2005 water.[51] Different protonation patterns of side chains and termini were used (Table

S1). Droplets with radius of 3 nm (~3500 H_2O) or 5.5 nm (~23000 H_2O) were built around the protein, using the X-ray coordinates 1UBQ[52] as starting conformation. Excess Na^+ were inserted in random positions to ensure an initial droplet charge at the Rayleigh limit. This initial charge included the protein and all ions (19+ for 3 nm, 46+ for 5.5 nm droplets). These two droplet sizes were chosen to allow comparisons with previous work.[21] Droplet simulations exhibit an unfavorable N^2 scaling of computer time with the number of atoms N,[53] such that our early studies were limited to droplet radii no larger than 3 nm. Subsequent advances in graphics processing unit (GPU) hardware and software[49] opened the door to larger systems, rendering MD studies on 5.5 nm droplets relatively straightforward.[21] Simulations were run at 370 K for 75 ns; then the temperature was raised to 450 K to promote evaporation of the final water molecules. Many runs were conducted in the presence of additional Na^+ and Cl^- ions. Ω calculations using Collidoscope[54] were performed for the final (t = 150 ns) time point by including the protein with its Na^+/Cl^- adducts. Root mean square deviation (RMSD) values were calculated for non-H atoms. The use of Na^+ as excess charge carrier in the MD runs sidesteps difficulties related to H^+ simulations.[21] Protein ions generated under these conditions are multiply charged due to sodiation instead of protonation. Experiments confirmed that the charge states and collision cross sections of $[M + zH]^{z+}$ and $[M + zNa]^{z+}$ ions are indistinguishable.[21]

Results and Discussion

Ubiquitin in Solution. Ubiquitin is a small (8565 Da, 76 residue) protein with a globular native structure and an isoelectric point close to seven.[52] It has served as model system for many earlier ESI-MS investigations.[55-59] Ubiquitin has a high resilience against pH-induced unfolding in aqueous solution; CD spectra show that its structure remains virtually unperturbed between pH 3 and pH 11 (Figure S1). This structural stability is due to the highly compact and tightly H-bonded fold of the

native state.[52] Only at pH 2 and in 50% methanol the native structure breaks down[60] (Figure S1). This remarkable stability offers the opportunity to alter the ubiquitin charge in solution by changing pH, while avoiding complications arising from structural perturbations. The ubiquitin solution charge is 6- at pH 11, zero at pH 7, 6+ at pH 4.4, and 12+ at pH 3 (Figure S2).

ESI Simulations: 3 nm Droplets. We previously conducted MD simulations to probe the behavior of ubiquitin in small aqueous ESI droplets (3 nm radius),[21] using a protein net charge of zero which mimics neutral pH. In that earlier work we found that all trajectories culminated in the formation of gaseous protein ions via the CRM (Figure 1C).[21] Here we repeated several runs under these conditions. We also conducted simulations on other solution charge states (6+ and 6-) to explore the effects of acidic and basic pH. The 3 nm droplet simulations of the current work all followed the CRM, regardless of the protein charge in solution (three runs for each solution charge state). Representative MD data are compiled in Figure S3.

All 3 nm droplet runs produced ubiquitin ions with a gas phase charge of 6+, regardless of the initial protein charge in solution. In other words, for solution charge states of 6-/0/6+ the CRM caused binding of 12/6/0 Na^+ to the protein. Any Na^+ that did not bind to the protein underwent IEM ejection, thereby ensuring that the shrinking droplets stayed close to the Rayleigh limit.[21] These MD data are consistent with the fact that 6+ is the dominant charge state in native ESI experiments.[21] Our data support the view that CRM charge states depend on the protein size, while being independent of the protein charge in solution.[2, 6, 20] Consistent with current mechanistic views of the ESI process,[2, 6, 21] the 6+ gas phase charge of the CRM-generated protein ions is close to the Rayleigh charge z_R of protein-sized water droplets (z_R = 6.6).[21]

ESI Simulations: 5.5 nm Droplets. The ESI plume in the ion source of a mass spectrometer comprises various droplet sizes. The 3 nm regime discussed above represents the lower end of this range.[2] To obtain a more comprehensive view of the ESI events we next examined larger droplets with 5.5 nm radius. All ubiquitin MD runs at basic and neutral pH (solution charge 6- and 0) exhibited CRM behavior (Figures 2A, 3A/B). As expected,[21] the shrinking droplets underwent IEM ejection of Na^+, exemplified in Figure 3B for t = 33.8 ns.

Surprisingly, simulations on 5.5 nm droplets with a ubiquitin solution charge of 6+ (pH 4.4) showed a different behavior. Most (6/7) MD runs under these conditions culminated in protein release via the IEM (Figure 3C). After migration of the protein to the surface (2.5 ns), the droplet was electrostatically distorted into a pear-like shape with the protein at its tip (5.5 ns). As the protein moved further outward, it briefly remained connected with the droplet through a water bridge (6.06 ns), before it finally broke free at 6.14 ns (Figure 3C, top to bottom). Immediately after ejection the protein retained ~200 water molecules, reflecting the presence of a tightly bound hydration shell.[61] Breaking these protein-H_2O contacts would be energetically costly (around 60 kJ mol^{-1} for each water molecule), while the rupture of water-water contacts farther away from the protein requires roughly half as much energy.[62, 63] Nonetheless, the heat supplied to the system subsequently caused the ejected proteins to dry out completely within tens of nanoseconds (Figures 2B).

Figure 3C marks the first observation of a *protein* IEM event in a MD simulation. This is in contrast to the well-established *small ion* IEM events[2, 16, 18, 19, 21, 42, 43] mentioned above. It is important to differentiate the protein IEM of Figure 3C from the CEM, as both mechanisms involve protein ejection from the droplet surface. The CEM applies to unfolded chains, and proceeds via intermediate stages where a stretched-out protein tail protrudes from the droplet (Figure 1D). In contrast, during the protein IEM of Figure 3C the polypeptide chain remains tightly folded while it leaves the droplet (matching the scenario of Figure 1B).

ESI Simulations: Salt-Containing 5.5 nm Droplets. The MD conditions of Figure 3 are somewhat unrealistic because they used droplets containing Na^+ without counterions. We proceeded to examine the behavior of 5.5 nm droplets in the presence of Cl^-, keeping in mind that NaCl is a ubiquitous contaminant in biological samples. Each Cl^- was balanced by inclusion of one additional Na^+ to ensure an initial droplet charge identical to the salt-free MD runs.

All ubiquitin simulations with a protein solution charge of zero in NaCl-containing 5.5 nm droplets followed the CRM, while ubiquitin^{6+} showed a mix of CRM and IEM behavior (Figure 2A). The CRM and IEM runs of Figure 4 (with 15 NaCl) both produced 6+ gaseous proteins. CRM runs caused the formation of massive NaCl adducts on the protein (15 NaCl bound in Figure 4A), while adduction was much less pronounced for the IEM trajectory (2 NaCl bound in Figure 4B). The kinetics of water evaporation and Na^+ IEM ejection were unaffected by the addition of salt (Figure 2B, C). However, the prevalence of protein IEM events was slightly lower in the presence of salt for ubiquitin^{6+}, from 86% without NaCl to 53% with 15 NaCl (Figure 2A). The latter trend may be linked to Cl^- mediated electrostatic screening.[64] Higher solution charge states (7+, 8+, 12+) overwhelmingly showed protein IEM (Figure 2A).

Comparison with Experimental ESI Charge States. Native ubiquitin ESI mass spectra were acquired in the presence of 1 mM NaCl, ensuring conditions comparable to our MD simulations, i.e., with ESI droplets that had Na^+ as the main charge carrier.[2] Mass spectra acquired between pH 11 and 3 showed very little change, with 6+ as the dominant charge state (Figure 5). This lack of major changes reflects the fact that ubiquitin retains its native structure in solution between pH 11 and 3 (Figure S1).[3, 35] The pH 3 spectrum (Figure 5D) exhibits a slight shift to higher charge states;

however, this alteration is minor compared to the spectrum of unfolded ubiquitin in acidified methanol/water, which has a maximum at 11+ and extends to 13+ (Figure S4).

All MD runs of this work on 5.5 nm droplets with ubiquitin^{6-} (pH 11) and ubiquitin^{0} (pH 7) followed the CRM, generating ESI charge states of (5.8 ± 0.5)+ and (6.1 ± 0.4)+, respectively. Simulations on ubiquitin^{6+} (pH 4.4) resulted in a mix of CRM (13/39) and IEM (26/39) runs; the former resulted in charge states of (6.5 ± 0.5)+ while the latter produced (6.3 ± 0.7)+. These MD charge states agree with the experimental 6+ maximum (Figure 5A-C).

MD runs on ubiquitin^{12+} (pH 3) in 5.5 nm droplets all followed the IEM. Gaseous proteins generated in these simulations from NaCl-containing droplets were bound to several Cl^-, generating ESI charge states of (8 ± 1)+ (Figure S5). This shift to slightly higher charge states matches the trend seen in the experimental data (Figure 5D). In summary, the MD-predicted ESI charge states agree well with experimental mass spectra over a wide range, from pH 11 down to 3.

CRM vs. IEM Salt Adduction. During the CRM the concentration of nonvolatile solutes increases as droplets shrink to dryness,[65] ultimately forming nonspecific adducts that can be detected by MS.[22] Salt adduction in native ESI experiments on NaCl-containing solution is a well-known example of this process.[2] Our MD data confirmed massive NaCl adduction in CRM simulations on NaCl-containing droplets. Rather than being spread over the protein surface, most of the adducted Na^+ and Cl^- formed small clusters (Figures 4A, S6). CRM runs on ubiquitin^{6+} in droplets containing 15 Na^+/Cl^- resulted in adduction of all these salt cations/anions to the protein. IEM runs showed much less adduction (1 ± 1 NaCl, Figure 4B), because most of the salt stayed behind in the droplet during protein ejection. Our simulations therefore predict that salt adduction provides information on the ESI mechanism, with CRM ions carrying much more salt than IEM-generated ions. This prediction

comes with the caveat that Cl^- can leave via neutral loss as HCl prior to MS detection,[2] such that Na^+ adducts will be more prevalent than Cl^- adducts.

The aforementioned predictions agree with our experimental data. Panels on the right in Figure 5 show close-ups of the 7+ and 6+ signals. Extensive Na^+ adduction at pH 11 and pH 7 confirms that these ions are CRM products, in agreement with our MD data (Figures 2A, and 3A,B). Spectra at pH 4.4 show less adduction, particularly for 7+ ions (Figure 5C). This trend is in line with the MD prediction that pH 4.4 causes a mix of IEM and CRM events (Figures 2A, 4). Finally, pH 3 causes significantly reduced adduction for 6+ ions while 7+ ions are almost adduct-free (Figure 5D), supporting the MD prediction that protein ions at pH 3 are IEM products (Figure 2A). In summary, experimental salt adduction patterns support the view that the IEM plays a significant role for the formation of gaseous ubiquitin at low pH, while neutral and basic conditions cause CRM behavior. Putting forward this proposal on the basis of salt adduction experiments *alone* might be precarious. However, together with our MD data the experiments of Figure 5 yield a consistent picture, providing strong support for the proposed IEM vs. CRM distinction.

Ubiquitin Gas Phase Conformations. IMS was used to probe the gas phase conformation of ubiquitin. The dominant 6+ charge state had a monomodal Ω distribution peaking at 1050 $Å^2$, while 7+ ions displayed a bimodal pattern with maxima at 1100 $Å^2$ and 1300 $Å^2$ (Figure 6). Evidently, the additional charge caused partial unfolding in some of the 7+ ions, in agreement with previous reports.[55] Salt adducts had only minor effects on the IMS behavior (Figure S7).

According to the data presented in the preceding sections, the ions in Figure 6 were generated by different mechanisms, from pure CRM at pH 7 to pure IEM at pH 3 (Figure 2A). The fact that different pH values produced virtually identical experimental Ω distributions is consistent with our MD data which predict that CRM and IEM are similarly "gentle" in terms of preserving protein

conformations during ESI. Inspection of the simulated gaseous proteins revealed that most of them retained a native-like fold, with its characteristic β-sheet and α-helix (e.g., Figure 3A-C, bottom row). RMSDs relative to the native ubiquitin X-ray structure were (1.32 ± 0.2) nm and (1.33 ± 0.1) nm for CRM runs and IEM runs, respectively (for a ubiquitin^{6+} solution charge).

19/39 runs with a ubiquitin^{6+} solution charge generated 6+ gaseous ions; only one of these showed appreciable unfolding. The average Ω for the simulated 6+ ions was (1040 ± 30) Å^2, in good agreement with experiments (Figure 6A). 17/39 runs generated 7+ ions. Due to their higher charge, a larger fraction (4/17) of these ions underwent partial unfolding. When tallying the Ω values for folded and partially unfolded 7+ conformers, averages of (1063 ± 30) Å^2 and (1250 ± 100) Å^2 were obtained which mirror the bimodal experimental distribution of Figure 6B.

For those MD runs that involved partial unfolding, it is interesting to examine when exactly structural transitions took place. Figure 6C shows snapshots taken from an IEM trajectory that generated a 7+ ion, starting from ubiquitin^{6+} in solution. The protein remained folded during ejection. Unfolding only took place while the final solvent molecules evaporated (6.0 ns in Figure 6C). Runs conducted under more acidic conditions followed the same pattern (Figure S5).

Protein IEM Requires Folded Chains with High Solution Charge. The driving force for IEM events is the electrostatic repulsion between a pre-charged analyte and the other charges in the droplet.[2, 9, 16, 18, 19] Therefore, IEM ejection only occurs for proteins that carry a large, positive solution charge (e.g., 6+, 7+, 8+, 12+ for our simulations). Ubiquitin solution charges of 6- and zero caused CRM behavior due to a lack of repulsion between protein and droplet (Figure 2A).

Some other proteins (those with pI > 7) are positively charged even in neutral solution. An example of such a protein is cyt *c*, which carries a 6+ solution charge at pH 7. Our simulations

confirmed that cyt *c* can undergo IEM ejection at neutral pH. Cyt *c* simulations in 5.5 nm droplets using a solution charge state of 6+ under various salt conditions produced 5/14 IEM runs and 9/14 CRM runs. These and other cyt *c* results are summarized in Figure S8. All of these cyt *c* data support the conclusions obtained for ubiquitin.

The prevalence of IEM events can be enhanced by raising the protein charge in solution via acidification, as demonstrated here for ubiquitin. However, this strategy is only feasible for proteins that retain a compact structure at low pH. For ubiquitin this is the case because the protein is exceptionally stable (Figure S1).[52] Most other proteins are prone to acid-unfolding,[66] and those unfolded chains will leave the droplet via the CEM.[21] In summary, IEM behavior will only take place for proteins that carry a high solution charge while maintaining a tightly folded structure.

Protein Surface Affinity. Prerequisite for IEM ejection is migration of the protein to the droplet surface. This is exemplified in Figure 3C at $t = 2.5$ ns, where ubiquitin^{6+} resides at the surface before undergoing IEM ejection. However, protein migration to the surface is *not* a unique IEM feature. For example, in the CRM trajectory of Figure 3A ubiquitin^{6-} maintains positions close to the surface while the droplet shrinks to dryness. Many of the CRM runs on smaller (3 nm) droplets showed a similar behavior (Figure S3A, B). Overall, our MD data showed protein migration to the surface in most of the runs, regardless of solution charge and ESI mechanism. Pushing the protein to the surface allows the droplet to maximize enthalpically favorable water-water hydrogen bonds.[67] In the preferred surface positioning the ubiquitin α-helix points toward the interior, while a patch of exposed hydrophobic residues points to the outside (Figure S9). This orientation reflects the known affinity of hydrophobic moieties for the droplet surface.[67]

Droplet Size Effects. Most earlier IEM-related studies focused on the ejection of small singly charged ions.[2, 9, 16, 18] The repulsive electric field experienced by a $+e$ test charge at the surface of a

Rayleigh-charged droplet increases with decreasing droplet radius (Figure S10). It is therefore often implied that IEM events are favored by small droplets.[2, 9, 16, 18] This expectation appears to be at odds with the data presented in this work, where ubiquitin^{6+} did *not* undergo IEM ejection in 3 nm droplets, while a radius of 5.5 nm readily produced protein IEM (Figure 2A).

The conundrum can be resolved by noting that IEM ejection is driven by the repulsion between protein^{z+} and the non-protein charges in the droplet. For small droplets, protein^{z+} represents a major fraction of the droplet charge, such that there will not be enough other charges to push the protein^{z+} into the gas phase. For example, an r = 1.8 nm ESI droplet carries a Rayleigh charge of 8+ (Figure S10A). If this droplet contains ubiquitin^{6+} the non-protein charge is (8+) – (6+) = (2+). Clearly, these two charges will be incapable of pushing the 6+ protein out of the droplet, rendering an IEM scenario impossible. Instead, ESI will proceed via the CRM.[68] Electrostatic modeling predicts that IEM ejection of ubiquitin^{6+} is most favored for droplet radii around ~5.5 nm. Smaller droplets experience less internal repulsion, such that protein IEM becomes less likely (Figure S10C-E). These considerations explain why our simulations produced IEM events only for relatively large (5.5 nm) droplets, while smaller droplets followed the CRM.

Conclusions

What is the mechanism of protein ESI? The current work adds another layer of complexity to this long-standing question[2, 4-6, 8, 9, 21] (Figure 1). Over the past 15+ years the view has emerged that native ESI is dominated by the CRM, while the CEM is operative under denaturing conditions.[11, 21, 33, 34] The IEM has found near-universal acceptance for small precharged ions,[2, 16, 18] but for proteins the IEM has only had relatively few proponents.[4, 5]

Here we do *not* challenge the CRM as the dominant native ESI mechanism, and we do *not* dispute the validity of the CEM for unfolded proteins. However, the current work marks the first time that atomistic MD simulations confirm protein IEM as an additional viable process. In other words, all three scenarios depicted in Figure 1B-D can take place during protein ESI, depending on the experimental conditions and the physicochemical properties of the protein.

IEM events can take place for proteins that carry a large positive charge in solution, while maintaining a compact folded conformation. A large protein charge in solution can be implemented by acidification of the solvent, with the caveat that many proteins unfolded under such conditions.[66] Acid-induced unfolding would cause the protein to follow the CEM.[21] Only proteins that are resilient to acid unfolding can be coaxed into IEM ejection by this approach, demonstrated here for ubiquitin. Alternatively, IEM events can take place for basic proteins (pI >> 7) which already carry a large charge in pH-neutral solution, demonstrated here for cyt *c*.

Several aspects of the protein IEM remain to be explored in future studies. (i) It is possible that protein IEM events are accompanied by charge equilibration between the droplet and the departing protein, analogous to H^+ transfer events associated with the CEM (Figure 1D). (ii) The upper size limit of IEM events is unclear; the current work demonstrated the viability of IEM ejection only for proteins up to 12.4 kDa. (iii) The relationship between protein charge in solution and pH can be complicated by pH changes during droplet shrinkage.[69] The definition $pH = -\log[H^+]$ becomes ambiguous in nanometer-sized droplets, e.g., because hydrated H^+ likely undergo enrichment at the surface instead of being distributed throughout the droplet.[70, 71] The effective pH experienced by a protein may therefore depend on the positioning of titratable side chains in the droplet. (iv) The viability of IEM events depends on droplet size. For droplets that are too small there are not enough non-protein charges to expel the protein, while for droplets that are too large the surface electric field will be insufficient to trigger IEM events (Figure S10). Considering that

the ESI plume comprises a wide range of droplet sizes (μm to nm),[2] it is not easy to predict which of the competing ESI mechanisms in Figure 1 is prevalent under a given set of conditions.

It is customary to consider the IEM and CRM as mutually exclusive mechanisms.[2] To some extent, the current work reconciles the two models. IEM events release proteins with a solvent shell of 10^2-10^3 water molecules (Figures 3C, 4B, 6C, S8B). One could interpret these nascent IEM products as protein-containing offspring droplets, and their evaporation to dryness would then be a CRM process. This view is consistent with the proposal that in addition to the "pure" ESI mechanisms of Figure 1 there can be various hybrid scenarios[4, 33] such that, in the words of Richard Cole, the distinction between mechanisms "*can become foggy*".[7]

Supporting Information. Supporting Table S1: Protonation patterns for ubiquitin simulations. Supporting Figures: Ubiquitin CD spectra, ubiquitin solution charge vs. pH, MD data for ubiquitin in 3 nm droplets, ubiquitin mass spectra, MD on ubiquitin^{12+}, ubiquitin CRM salt adducts, salt-dependent IMS data, MD data for cyt *c* in 5.5 nm droplets, orientation of ubiquitin at droplet surface, electrostatic model of IEM forces.

References

(1) Yamashita, M.; Fenn, J. B. Electrospray Ion Source. Another variation on the Free-Jet Theme. *J. Phys. Chem.* **1984**, 88, 4451-4459.

(2) Kebarle, P.; Verkerk, U. H. Electrospray: From Ions in Solutions to Ions in the Gas Phase, What We Know Now. *Mass Spectrom. Rev.* **2009**, 28, 898-917.

(3) Li, J.; Santambrogio, C.; Brocca, S.; Rossetti, G.; Carloni, P.; Grandori, R. Conformational Effects in Protein Electrospray Ionization Mass Spectrometry. *Mass Spectrom. Rev.* **2016**, 35, 111-122.

(4) Ogorzalek Loo, R. R.; Lakshmanan, R.; Loo, J. A. What Protein Charging (and Supercharging) Reveal about the Mechanism of Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* **2014**, 25, 1675-1693.

(5) Nguyen, S.; Fenn, J. B. Gas-phase ions of solute species from charged droplets of solutions. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, 104, 1111-1117.

(6) de la Mora, J. F. Electrospray Ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta* **2000**, 406, 93-104.

(7) Cole, R. B. Some tenets pertaining to electrospray ionization mass spectrometry. *J. Mass. Spectrom.* **2000**, 35, 763-772.

(8) Iavarone, A. T.; Williams, E. R. Mechanism of Charging and Supercharging Molecules in Electrospray Ionization. *J. Am. Chem. Soc.* **2003**, 125, 2319-2327.

(9) Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. Combined Charged Residue-Field Emission Model of Macromolecular Electrospray Ionization. *Anal. Chem.* **2009**, 81, 369-377.

(10) Leney, A. C.; Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* **2017**, 28, 5-13.

(11) Mehmood, S.; Allison, T. M.; Robinson, C. V. Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annu. Rev. Phys. Chem.* **2015**, 66, 453-474.

(12) Nilsson, T.; Mann, M.; Aebersold, R.; Yates, J. R.; Bairoch, A.; Bergeron, J. J. M. Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat. Methods* **2010**, 7, 681-685.

(13) Kaur, U.; Johnson, D. T.; Chea, E. E.; Deredge, D. J.; Espino, J. A.; Jones, L. M. Evolution of Structural Biology through the Lens of Mass Spectrometry. *Anal. Chem.* **2019**, 91, 142-155.

(14) Gomez, A.; Tang, K. Charge and fission of droplets in electrostatic sprays. *Phys. Fluids* **1994**, 6, 404-414.

(15) Grimm, R. L.; Beauchamp, J. L. Evaporation and Discharge Dynamics of Highly Charged Multicomponent Droplets Generated by Electrospray Ionization. *J. Phys. Chem. A* **2010**, 114, 1411-1419.

(16) Iribarne, J. V.; Thomson, B. A. On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **1976**, 64, 2287-2294.

(17) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. Molecular beams of macroions. *J. Chem. Phys.* **1968**, 49, 2240-2249.

(18) Loscertales, I. G.; de la Mora, J. F. Experiments on the kinetics of field evaporation of small ions from droplets. *J. Chem. Phys.* **1995**, 103, 5041-5060.

(19) Consta, S. Fragmentation reactions of charged aqueous clusters. *J. Mol. Struct. (Theochem)* **2002**, 591, 131-140.

(20) Kaltashov, I. A.; Mohimen, A. Estimates of Protein Surface Area in Solution by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2005**, 77, 5370-5379.
(21) Konermann, L.; Metwally, H.; Duez, Q.; Peters, I. Charging and Supercharging of Proteins for Mass Spectrometry: Recent Insights into the Mechanisms of Electrospray Ionization. *Analyst* **2019**, 144, 6157-6171.
(22) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Reliable Determinations of Protein–Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass Spectrom.* **2012**, 23, 431-441.
(23) DeMuth, J. C.; McLuckey, S. A. Electrospray Droplet Exposure to Organic Vapors: Metal Ion Removal from Proteins and Protein Complexes. *Anal. Chem.* **2015**, 87, 1210-1218.
(24) Hopper, J. T. S.; Oldham, N. J. Alkali Metal Cation-Induced Destabilization of Gas-Phase Protein-Ligand Complexes: Consequences and Prevention. *Anal. Chem.* **2011**, 83, 7472-7479.
(25) Nguyen, G. T. H.; Tran, T. N.; Podgorski, M. N.; Bell, S. G.; Supuran, C. T.; Donald, W. A. Nanoscale Ion Emitters in Native Mass Spectrometry for Measuring Ligand–Protein Binding Affinities. *ACS Centr. Sci.* **2019**, 5, 308–318.
(26) Cleary, S. P.; Prell, J. S. Liberating Native Mass Spectrometry from Dependence on Volatile Salt Buffers by Use of Gabor Transform. *Chemphyschem* **2019**, 20, 519-523.
(27) Spencer, E. A. C.; Ly, T.; Julian, R. K. Formation of the serine octamer: Ion evaporation or charge residue? *Int. J. Mass Spectrom.* **2008**, 270, 166-172.
(28) Calixte, E. I.; Liyanage, O. T.; Kim, H. J.; Ziperman, E. D.; Pearson, A. J.; Gallagher, E. S. Release of Carbohydrate-Metal Adducts from Electrospray Droplets: Insight into Glycan Ionization by Electrospray. *J. Phys. Chem. B* **2020**, 124, 479-486.
(29) Wang, G.; Cole, R. B. Charged residue versus ion evaporation for formation of alkali metal halide clusters ions in ESI. *Anal. Chim. Acta* **2000**, 406, 53-65.
(30) Wilm, M. Principles of Electrospray Ionization. *Mol. Cell. Proteomics* **2011**, 10, 0094071-0094078.
(31) Fenn, J. B. Ion Formation from Charged Droplets: Roles of Geometry, Energy, and Time. *J. Am. Soc. Mass Spectrom.* **1993**, 4, 524-535.
(32) Consta, S.; Chung, J. K. Charge-Induced Conformational Changes of PEG-(Na+)n in Vacuum and Aqueous Nanodroplets. *J. Phys. Chem. B* **2011**, 115, 10447-10455.
(33) Beveridge, R.; Phillips, A. S.; Denbigh, L.; Saleem, H. M.; MacPhee, C. E.; Barran, P. E. Relating gas phase to solution conformations: Lessons from disordered proteins. *Proteomics* **2015**, 15, 2872-2883.
(34) Donor, M. T.; Ewing, S. A.; Zenaidee, M. A.; Donald, W. A.; Prell, J. S. Extended Protein Ions Are Formed by the Chain Ejection Model in Chemical Supercharging Electrospray Ionization. *Anal. Chem.* **2017**, 89, 5107-5114.
(35) Dobo, A.; Kaltashov, I. A. Detection of Multiple Protein Conformational Ensembles in Solution via Deconvolution of Charge-State Distributions in ESI MS. *Anal. Chem.* **2001**, 73, 4763-4773.
(36) Chowdhury, S. K.; Katta, V.; Chait, B. T. Probing Conformational Changes in Proteins by Mass Spectrometry. *J. Am. Chem. Soc.* **1990**, 112, 9012-9013.
(37) Kharlamova, A.; Prentice, B. M.; Huang, T.-Y.; McLuckey, S. A. Electrospray Droplet Exposure to Gaseous Acids for the Manipulation of Protein Charge State Distributions. *Anal. Chem.* **2010**, 82, 7422-7429.
(38) Fegan, S. K.; Thachuk, M. A Charge Moving Algorithm for Molecular Dynamics Simulations of Gas-Phase Proteins. *J. Chem. Theory Comput.* **2013**, 9, 2531-2539.

(39) Ruotolo, B. T.; Hyung, S.-J.; Robinson, P. M.; Giles, K.; Bateman, R. H.; Robinson, C. V. Ion Mobility–Mass Spectrometry Reveals Long-Lived, Unfolded Intermediates in the Dissociation of Protein Complexes. *Angew. Chem. Int. Ed.* **2007**, 46, 8001-8004.
(40) Quintyn, R. S.; Zhou, M.; Yan, J.; Wysocki, V. H. Surface-Induced Dissociation Mass Spectra as a Tool for Distinguishing Different Structural Forms of Gas-Phase Multimeric Protein Complexes. *Anal. Chem.* **2015**, 87, 11879-11886.
(41) Patriksson, A.; Marklund, E.; van der Spoel, D. Protein Structures under Electrospray Conditions. *Biochemistry* **2007**, 46, 933-945.
(42) Znamenskiy, V.; Marginean, I.; Vertes, A. Solvated Ion Evaporation from Charged Water Droplets. *J. Phys. Chem. A* **2003**, 107, 7406-7412.
(43) Higashi, H.; Tokumi, T.; Hogan, C. J.; Suda, H.; Seto, T.; Otani, Y. Simultaneous ion and neutral evaporation in aqueous nanodrops: experiment, theory, and molecular dynamics simulations. *Phys. Chem. Chem. Phys.* **2015**, 17, 15746-15755.
(44) Beveridge, R.; Migas, L. G.; Das, R. K.; Pappu, R. V.; Kriwacki, R. W.; Barran, P. E. 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.* **2019**, 141, 4908-4918.
(45) Porrini, M.; Rosu, F.; Rabin, C.; Darre, L.; Gomez, H.; Orozco, M.; Gabelica, V. Compaction of Duplex Nucleic Acids upon Native Electrospray Mass Spectrometry. *ACS Central Sci.* **2017**, 3, 454–461.
(46) Kim, D.; Wagner, N.; Wooding, K.; Clemmer, D. E.; Russell, D. H. 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.* **2017**, 139, 2981-2988.
(47) Kondalaji, S. G.; Khakinejad, M.; Valentine, S. J. Comprehensive Peptide Ion Structure Studies Using Ion Mobility Techniques: Part 3. Relating Solution-Phase to Gas-Phase Structures. *J. Am. Soc. Mass Spectrom.* **2018**, 29, 1665-1677.
(48) Sun, Y.; Vahidi, S.; Sowole, M. A.; Konermann, L. Protein Structural Studies by Traveling Wave Ion Mobility Spectrometry: A Critical Look at Electrospray Sources and Calibration Issues. *J. Am. Soc. Mass Spectrom.* **2016**, 27, 31-40.
(49) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, 1–2, 19-25.
(50) Huang, J.; MacKerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, 34, 2135-2145.
(51) Abascal, J. L. F.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, 123, 234505.
(52) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. Structure of Ubiquitin Refined at 1.8 A Resolution. *J. Mol. Biol.* **1987**, 194, 531-544.
(53) Caleman, C.; van der Spoel, D. Temperature and structural changes of water clusters in vacuum due to evaporation. *J. Chem. Phys.* **2006**, 125, 1545081-1545089.
(54) Ewing, S. A.; Donor, M. T.; Wilson, J. W.; Prell, J. S. Collidoscope: An Improved Tool for Computing Collisional Cross-Sections with the Trajectory Method. *J. Am. Soc. Mass Spectrom.* **2017**, 28, 587-596.
(55) Wyttenbach, T.; Bowers, M. T. Structural Stability from Solution to the Gas Phase: Native Solution Structure of Ubiquitin Survives Analysis in a Solvent-Free Ion Mobility–Mass Spectrometry Environment. *J. Phys. Chem. B* **2011**, 115, 12266-12275.

(56) Skinner, O. S.; McLafferty, F. W.; Breuker, K. How Ubiquitin Unfolds after Transfer into the Gas Phase. *J. Am. Soc. Mass Spectrom.* **2012**, 23, 1011-1014.
(57) Laszlo, K. J.; Munger, E. B.; Bush, M. F. Folding of Protein Ions in the Gas Phase after Cation-to-Anion Proton-Transfer Reactions. *J. Am. Chem. Soc.* **2016**, 138, 9581-9588.
(58) El-Baba, T. J.; Woodall, D. W.; Raab, S. A.; Fuller, D. R.; Laganowsky, A.; Russell, D. H.; Clemmer, D. E. Melting Proteins: Evidence for Multiple Stable Structures upon Thermal Denaturation of Native Ubiquitin from Ion Mobility Spectrometry-Mass Spectrometry Measurements. *J. Am. Chem. Soc.* **2017**, 139, 6306-6309.
(59) Bonner, J. G.; Lyon, Y. A.; Nellessen, C.; Julian, R. R. Photoelectron Transfer Dissociation Reveals Surprising Favorability of Zwitterionic States in Large Gaseous Peptides and Proteins. *J. Am. Chem. Soc.* **2017**, 139, 10286-10293.
(60) Brutscher, B.; Brüwchweiler, R.; Ernst, R. R. Backbone Dynamics and Structural Characterization of the Partially Folded A State of Ubiquitin by 1H, 13C and 15N Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **1997**, 36, 13043-13053.
(61) Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D. A unified model of protein dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, 106, 5129-5134.
(62) Liu, D. F.; Wyttenbach, T.; Barran, P. E.; Bowers, M. T. Sequential hydration of small protonated peptides. *J. Am. Chem. Soc.* **2003**, 125, 8458-8464.
(63) Leforestier, C.; Szalewicz, K.; van der Avoird, A. Spectra of water dimer from a new ab initio potential with flexible monomers. *J. Chem. Phys.* **2012**, 137, 17.
(64) Norouzy, A.; Assaf, K. I.; Zhang, S.; Jacob, M. H.; Nau, W. M. Coulomb Repulsion in Short Polypeptides. *J. Phys. Chem. B* **2015**, 119, 33-43.
(65) Wortmann, A.; Kistler-Momotova, A.; Zenobi, R.; Heine, M. C.; Wilhelm, O.; Pratsinis, S. E. Shrinking Droplets in Electrospray Ionization and Their Influence on Chemical Equilibria. *J. Am. Soc. Mass Spectrom.* **2007**, 18, 385-393.
(66) Creighton, T. E. *Proteins*; W. H. Freeman & Co: New York, 1993.
(67) Null, A. P.; Nepomuceno, A. I.; Muddiman, D. C. Implications of Hydrophobicity and Free Energy of Solvation for Characterization of Nucleic Acids by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2003**, 75, 1331-1339.
(68) Daub, C. D.; Cann, N. M. How Are Completely Desolvated Ions Produced in Electrospray Ionization: Insights from Molecular Dynamics Simulations. *Anal. Chem.* **2011**, 83, 8372-8376.
(69) Zhou, S.; Prebyl, B. S.; Cook, K. D. Profiling pH Changes in the Electrospray Plume. *Anal. Chem.* **2002**, 74, 4885-4888.
(70) Iyengar, S. S.; Day, T. J. F.; Voth, G. A. On the amphiphilic behavior of the hydrated proton: an ab initio molecular dynamics study. *Int. J. Mass Spectrom.* **2005**, 241, 197-204.
(71) Hub, J. S.; Wolf, M. G.; Caleman, C.; Maaren, P. J.; Groenhof, G.; van der Spoel, D. Thermodynamics of hydronium and hydroxide surface solvation. *Chem. Sci.* **2014**, 5, 1745-1749.

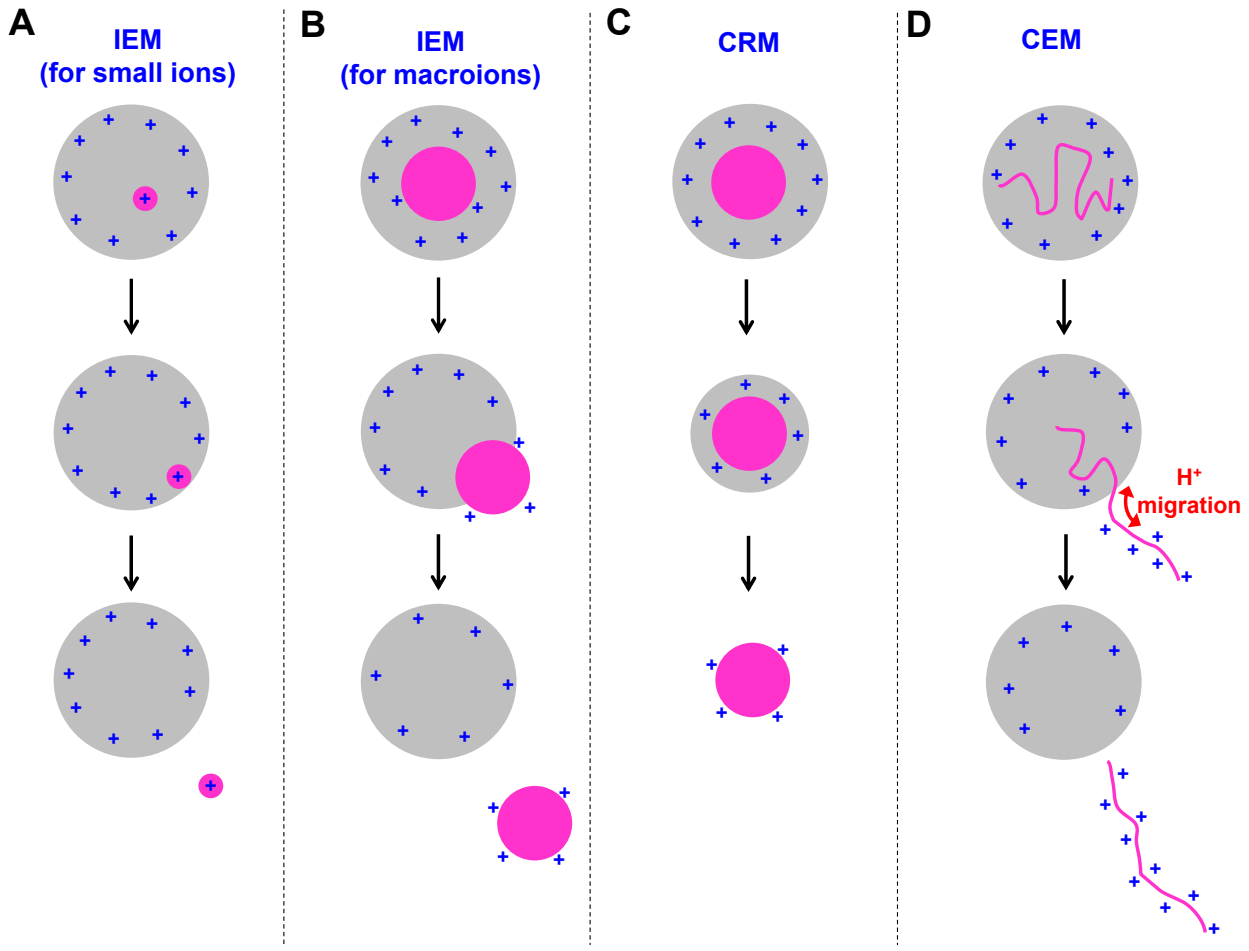

FIGURE 1. Cartoon depiction of different ESI mechanisms. IEM, ion evaporation model; CRM, charged residue model; CEM, chain ejection model. All scenarios start with a solvent droplet (gray) that contains an analyte (pink) and excess charge carriers (such as Na^+, NH_4^+, H_3O^+, shown as "+"). All mechanisms culminate in the formation of gaseous analyte ions.

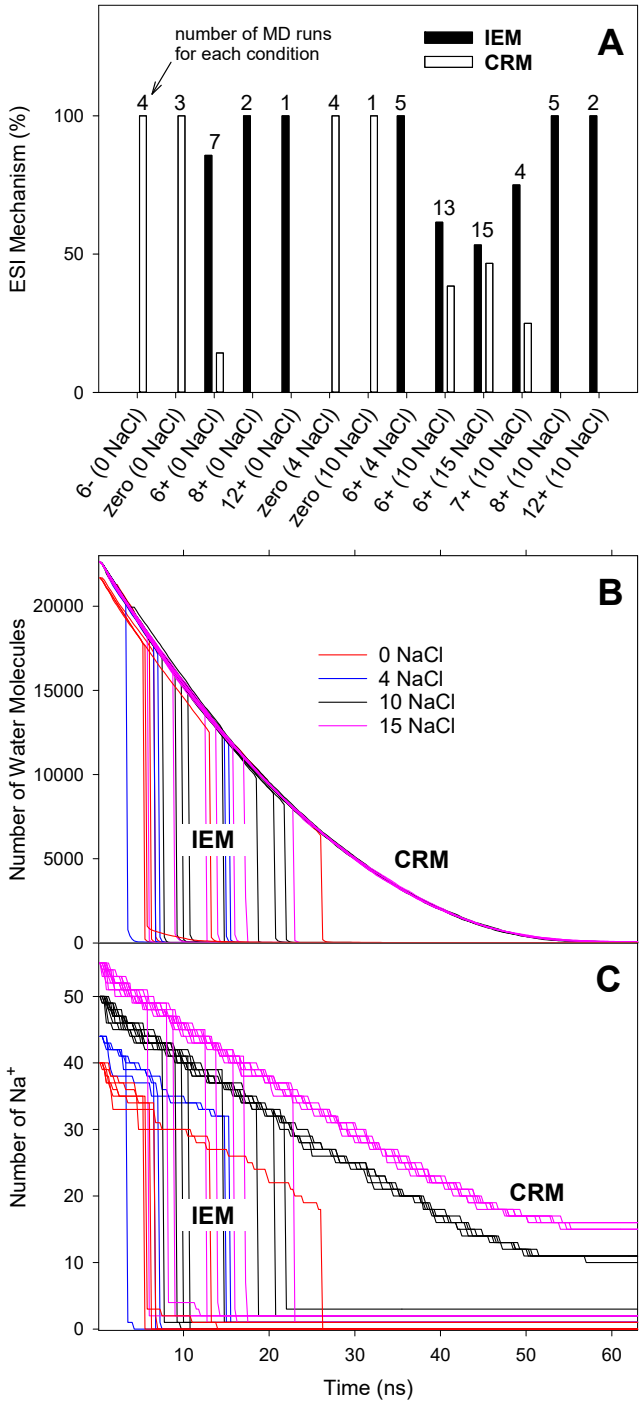

FIGURE 2. MD data for ESI droplets with 5.5 nm radius. (A) Tally of IEM vs. CRM events for ubiquitin in solution charge states 6-, zero, 6+, 8+, and 12+ (as shown along the *x*-axis). Also indicated is the number of additional Na^+ Cl^- ion pairs in the droplets. (B) Number of water molecules in the protein-containing droplet for a ubiquitin solution charge state of 6+. CRM trajectories show evaporation to dryness. Sudden downward transitions indicate protein ejection from the droplet (protein IEM). Panel C: Same as in B, but for Na^+ ions in the protein-containing droplet.

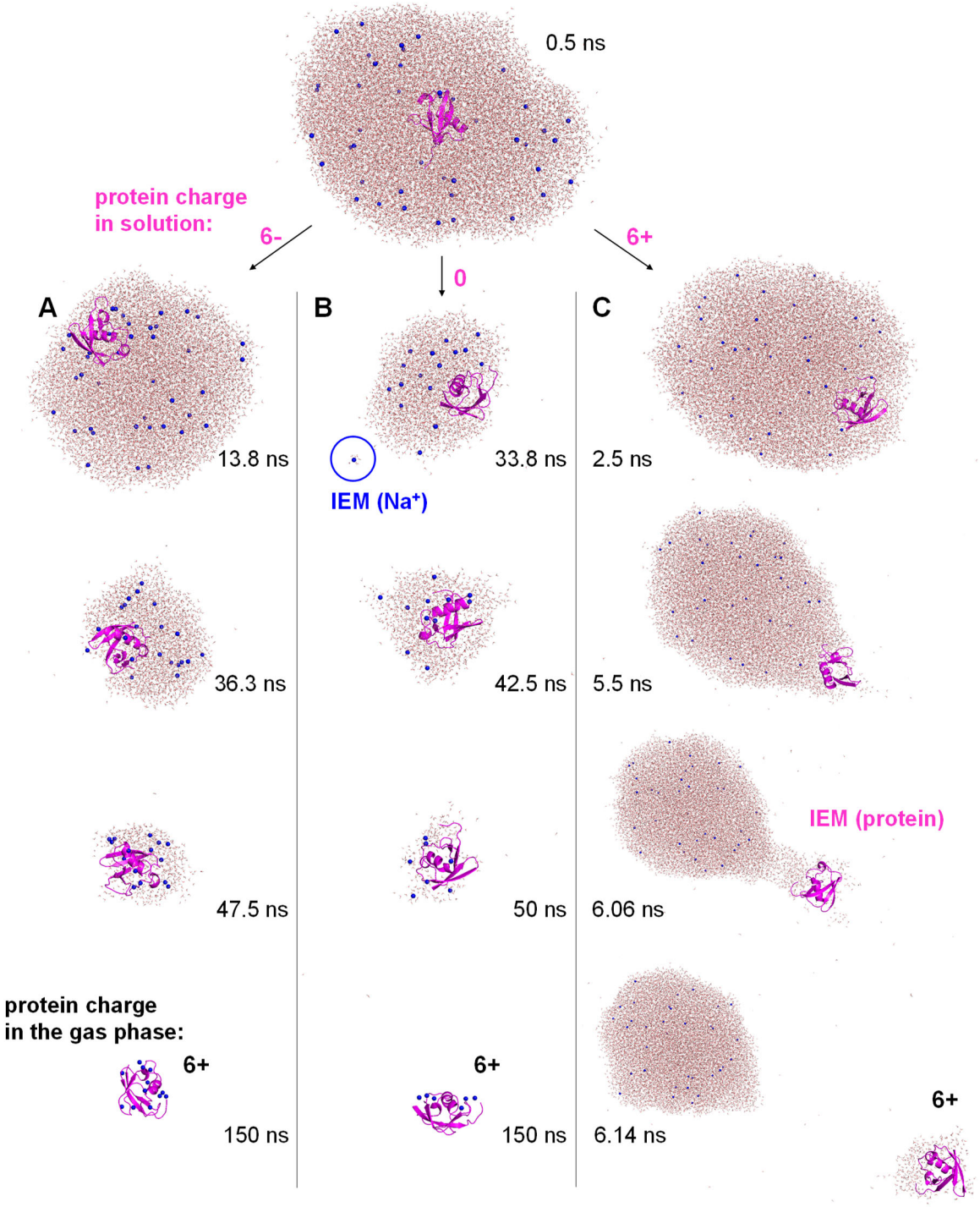

FIGURE 3. MD simulation snapshots for aqueous ESI droplets containing ubiquitin. Initial droplet radius = 5.5 nm. Initial droplet charge = 46+. (A) Protein charge in solution = 6-. (B) Protein charge in solution = 0. (C) Protein charge in solution = 6+. In panels A, B gaseous protein ions are produced via the CRM; in panel C the gaseous protein ion is produced via protein IEM. The final protein charge state is 6+ for all three runs. A Na^+ IEM event is highlighted in panel B (blue circle at 33.8 ns). Coloring: protein, pink; water oxygen, red; Na^+, blue.

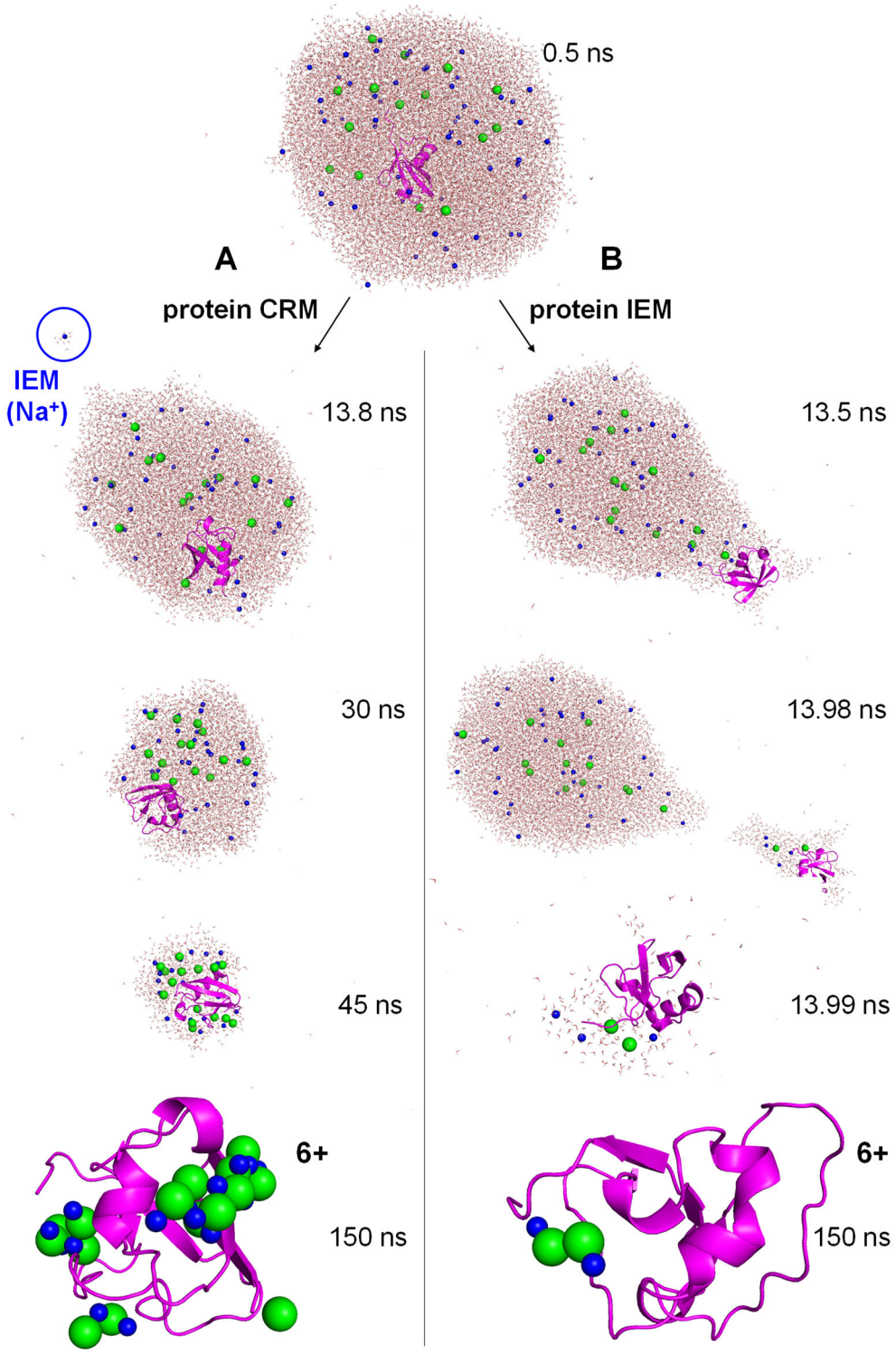

FIGURE 4. MD simulation snapshots for salt-containing ESI droplets. The initial droplets (5.5 nm radius, top) contained 15 Na^+ / Cl^- ion pairs. Cl^- is depicted in green. Initial protein charge in solution = 6+. All other conditions are as in Figure 2. For some runs, gaseous protein ions were formed via the CRM (A), other runs showed IEM behavior (B). Note the different level of salt adduction at 150 ns, i.e., 15 NaCl for the CRM run in panel A, 2 NaCl for the IEM run in panel B.

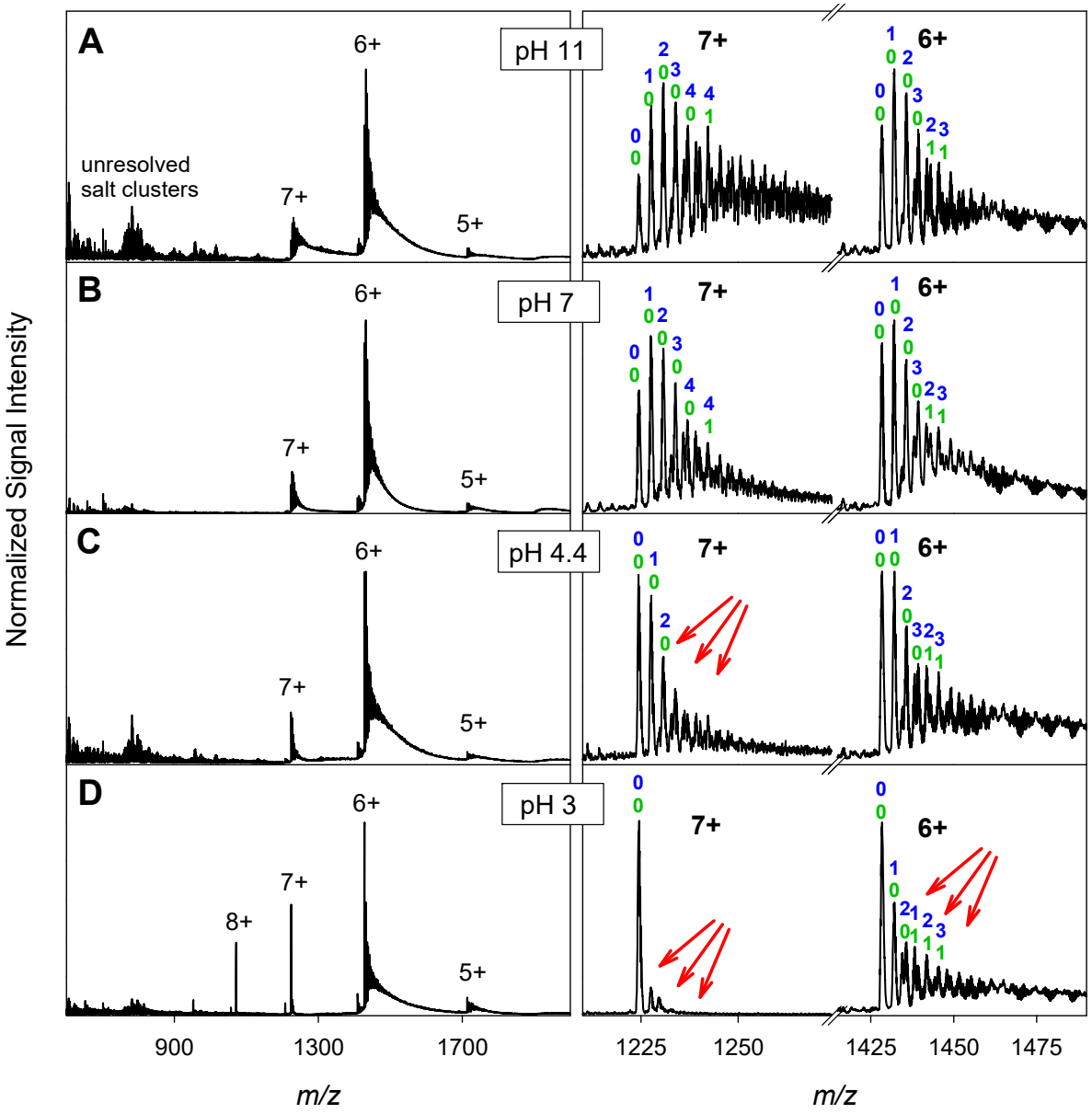

FIGURE 5. ESI mass spectra of ubiquitin in aqueous solution at different pH. The low signal-to-noise ratio is caused by the addition of 1 mM NaCl which results in salt adducts. Panels on the right zoom in on the 7+ and 6+ signals. Number pairs identify some salt adducts, indicating how many Na^+ (blue, top) and Cl^- (green, bottom) are bound to the protein ions. Red arrows indicate the significantly reduced salt adduction under acidic (IEM) conditions.

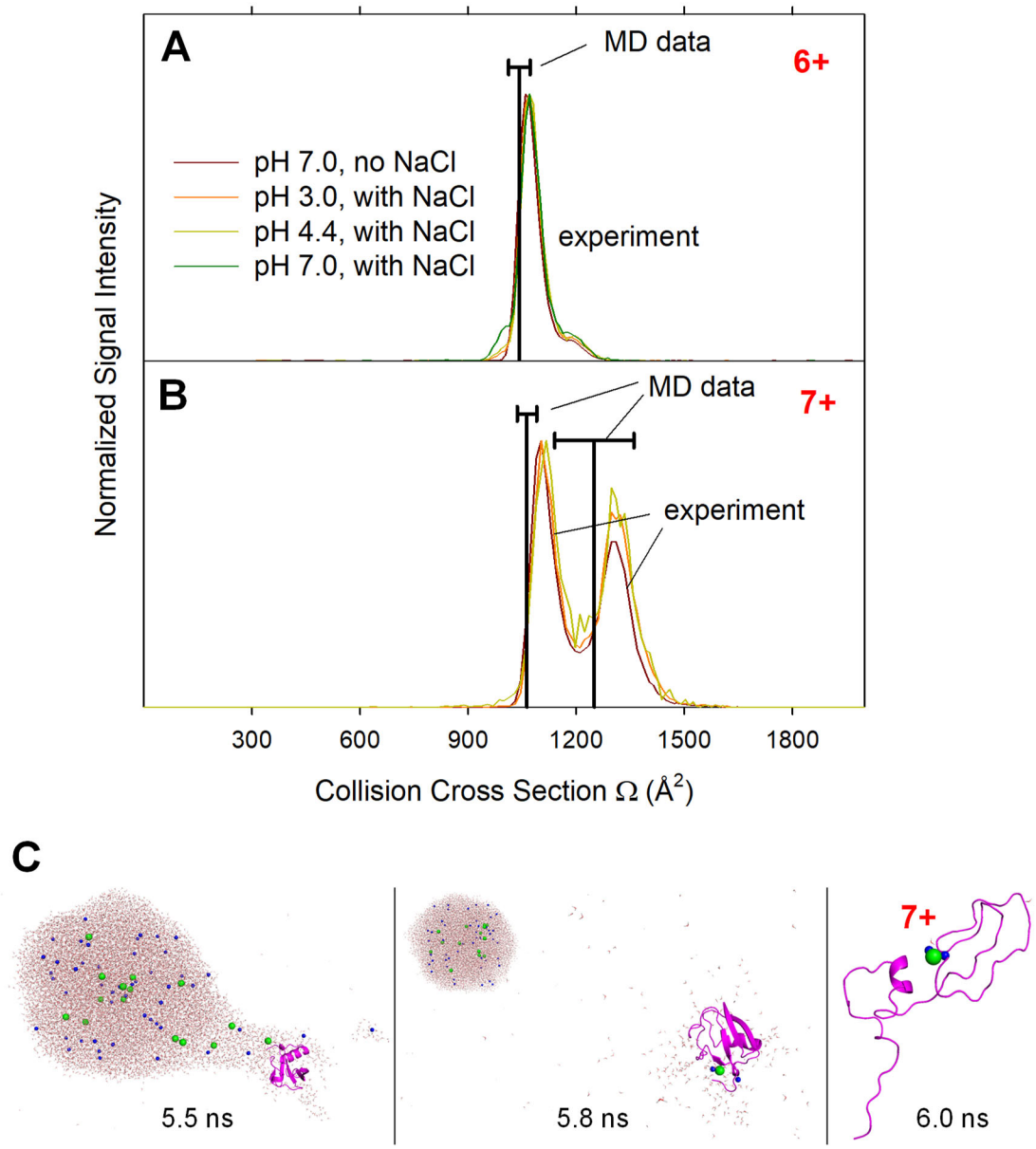

FIGURE 6. Comparison of experimental IMS data and Ω values of MD-generated gas phase ions, for (A) 6+ and (B) 7+ ubiquitin ions. Vertical lines are averages from all runs with and without NaCl. Horizontal error bars represent standard deviations. (C) MD snapshots of a trajectory generated under the conditions of Figure 4B, except that the $[\text{ubiquitin}^{6+} + 2\text{Na}^{+} + \text{Cl}^{-}]^{7+}$ ion undergoes partial unfolding after IEM ejection.

For Table of Contents Only