

9-21-2021

Atomistic Insights into the Formation of Nonspecific Protein Complexes during Electrospray Ionization.

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, "Atomistic Insights into the Formation of Nonspecific Protein Complexes during Electrospray Ionization." (2021). *Chemistry Publications*. 229.
<https://ir.lib.uwo.ca/chempub/229>

Atomistic Insights into the Formation of Nonspecific Protein Complexes During Electrospray Ionization

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: Native electrospray ionization (ESI) mass spectrometry (MS) is widely used for the detection and characterization of multi-protein complexes. A well-known problem with this approach is the possible occurrence of nonspecific protein clustering in the ESI plume. This effect can distort the results of binding affinity measurements, and it can even generate gas phase complexes from proteins that are strictly monomeric in bulk solution. By combining experiments and molecular dynamics (MD) simulations, the current work for the first time provides detailed insights into the ESI clustering of proteins. Using ubiquitin as a model system, we demonstrate how the entrapment of more than one protein molecule in an ESI droplet can generate nonspecific clusters (e.g., dimers or trimers) via solvent evaporation to dryness. These events are in line with earlier proposals, according to which protein clustering is associated with the charged residue model (CRM). MD simulations on cytochrome *c* (which carries a large intrinsic positive charge) confirmed the viability of this CRM avenue. In addition, the cytochrome *c* data uncovered an alternative mechanism where protein-protein contacts were formed early within ESI droplets, followed by cluster ejection from the droplet surface. This second pathway is consistent with the ion evaporation model (IEM). The observation of these IEM events for large protein clusters is unexpected, because the IEM has been thought to be associated primarily with low MW analytes. In all cases, our MD simulations produced protein clusters that were stabilized by intermolecular salt bridges. The MD-generated charge states agreed with experiments. Overall, this work reveals that ESI-induced protein clustering does not follow a tightly orchestrated pathway, but can proceed along different avenues.

Introduction

Multi-subunit proteins are involved in numerous biological tasks. These complexes range from dimers all the way to MDa assemblies that contain dozens of subunits.¹⁻⁴ Some protein complexes are stabilized by intermolecular disulfide bridges. More commonly, however, protein-protein binding involves only noncovalent contacts such as van der Waals interactions, H-bonds, and salt bridges. The hydrophobic effect can play a major role as well.⁵ The same types of noncovalent contacts can also mediate the formation of amyloid and other protein aggregates that are associated with various diseases.^{6, 7} The detection and characterization of all these complexes remains challenging.⁵ Available high-throughput methods include yeast-two-hybrid and affinity purification protocols. In addition, chromatographic, spectroscopic, and calorimetric techniques can be applied for targeted assays. X-ray diffraction provides high-resolution data, but often it is uncertain if protein contacts detected in this way are biologically relevant or if they are crystallization artifacts.⁵

Native electrospray ionization (ESI) mass spectrometry (MS) has emerged as another key tool for the characterization of protein complexes. This approach relies on the premise that noncovalent assemblies can be transformed into gaseous ions that retain many of their solution properties. Mass analysis of these ions reveals their subunit stoichiometry.⁸⁻¹² Native ESI-MS is attractive because of its conceptual simplicity, minimal sample preparation, high sensitivity, and short analysis time. Additional insights are obtainable by combining ESI-MS with ion mobility spectrometry (IMS),¹³⁻¹⁶ gas phase activation,¹⁷⁻¹⁹ and fragmentation techniques.²⁰⁻²⁴

ESI commences with a plume of charged droplets that emanates from the tip of a Taylor cone at the emitter outlet. These droplets undergo solvent evaporation and fission events, culminating in nanometer-sized progeny droplets. Droplets in the ESI plume are close to the Rayleigh limit $z_R = 8\pi/e \times (\epsilon_0 \gamma r^3)^{1/2}$ [r = radius, γ = surface tension, e = elementary charge].²⁵⁻²⁷

The release of analyte ions from ESI nanodroplets into the gas phase remains an active research area.²⁸⁻³³ It is believed that in native ESI most protein ions are released via droplet evaporation to dryness, in accordance with the charged residue model (CRM).³⁴⁻³⁷ The ion evaporation model (IEM) describes an alternative mechanism where ions are desorbed from the droplet surface. Although the IEM is invoked mostly for low MW ions,^{27, 38-42} it can also apply to larger species such as peptides⁴³ and some proteins.⁴⁴ Requirements for protein IEM include a relatively large droplet size, a compact conformation, and a sufficiently high intrinsic protein charge that can trigger electrostatic ejection from the droplet.⁴⁴ The chain ejection model (CEM) applies to unfolded proteins and therefore does not usually play a role in native ESI.³⁴

Despite the widespread use of native ESI-MS for studying protein complexes (and other noncovalently bound systems), this technique can be prone to artifacts.^{25, 45} False-positive results are obtained when mass spectra show complexes that did not exist in solution. Conversely, false-negative outcomes occur when complexes that exist in solution are unobservable by ESI-MS. More generally, there can be a range of undesirable scenarios where the free vs. bound ratio in solution differs from that in the gas phase.⁴⁵ Possible reasons for such discrepancies include concentration and pH changes during ESI,⁴⁶⁻⁴⁹ differences in the detection efficiencies of free vs. bound species,⁴⁵ and the dissociation of complexes on their way from solution into the gas phase.⁵⁰⁻⁵³

False-positive outcomes arise from ESI-induced nonspecific clustering. This phenomenon can manifest itself as complex formation from monomeric proteins,⁵⁴⁻⁵⁸ or the assembly of complexes into higher order oligomers.^{20, 59-61} There can also be a mix of specific solution binding and nonspecific clustering.^{45, 62} All of these clustering scenarios complicate the interpretation of native ESI-MS data. Various strategies have been proposed for mitigating this problem,^{24, 45, 61-71} but it is nonetheless challenging to distinguish specific from nonspecific complexes in a mass spectrum. Interestingly, there are also instances where nonspecific clustering is beneficial; for example, protein

clusters can serve as model systems for benchmarking mass analyzer performance at high m/z , and as a testbed for top-down dissociation experiments.^{20, 56, 57, 60}

Nonspecific clustering is usually attributed to the CRM, where a nanodroplet containing two or more analyte molecules causes these solutes to “stick” to one another as the droplet dries out. In addition to nonspecific protein-protein contacts, these conditions can cause adduction to other nonvolatile species.^{25, 45} Nonspecific clustering can be reduced by using narrow emitters that produce smaller initial droplets, thereby decreasing the number of analyte molecules in each droplet.^{24, 62, 66, 67, 70} Unfortunately, narrow emitters are prone to clogging.⁷² Similarly, one can lower the analyte concentration, such that the fraction of droplets containing more than one analyte molecule is lowered.^{61, 68} However, results obtained in this way can be ambiguous, because mass action dictates that lower concentrations also reduce specific binding in solution.⁷³

In summary, nonspecific protein-protein clustering represents an impediment for the interpretation of native ESI data. Part of the problem is that the mechanistic origins of cluster formation are poorly understood. Questions that have to be answered include the following: Is it true that clustering can always be attributed to ion formation via the CRM? How, when, and where are protein clusters formed during ESI? Will droplets that contain several proteins always generate a nonspecific cluster? What are the intermolecular contacts that mediate protein-protein clustering? Using a combination of ESI-MS experiments and molecular dynamics (MD) simulations, the current work addresses all of these questions. For the first time, we provide an atomistic view of the processes that culminate in the formation of ESI-generated protein clusters.

Materials and Methods

Horse heart cytochrome *c* (cyt *c*) and bovine ubiquitin were from Sigma (St. Louis, MO). Samples were prepared in LC grade water, adjusted to pH 7 using traces of ammonium hydroxide. Ammonium acetate or NaCl were added as noted below. Data were acquired on a Waters Synapt G2 Q-TOF using a standard ESI source at 5 $\mu\text{L min}^{-1}$ and +2.8 kV. The conditions were chosen to be as gentle (“native”) as possible, with a cone voltage of 5 V. Source and desolvation temperatures were 30 °C and 40 °C, respectively. IMS arrival time distributions were converted to He collision cross section (Ω) distributions using a calibration procedure that involved a number of reference ions (monomeric cyt *c*, ubiquitin, and myoglobin in various charge states).⁷⁴ Each reported value represents the maximum of the corresponding Ω distribution, averaged over three replicates.⁷⁵ The IMS parameters were DC entrance 6.7 V, He cell DC 10 V, He exit -5 V, bias 3 V, exit 0 V, wave height 8 V, wave velocity 450 m s^{-1} , N₂ IMS gas 3.61 mbar at 16 mL min^{-1} , He cell 7.41e2 mbar.

MD simulations of ESI droplets were conducted as described,⁴⁴ except that the initial droplets contained more than one protein molecule. Briefly, we used Gromacs 2018⁷⁶ with the Charmm36 force field.⁷⁷ The TIP4P/2005 model was used to adequately model the water surface tension.⁷⁸ The droplets had an initial radius of 5.5 nm (~23000 H₂O). Two or three protein molecules were inserted into these droplets, using the X-ray coordinates 1hrc⁷⁹ and 1ubq.⁸⁰ Titratable sites were set for pH 7 (N-terminus⁺, Arg⁺, Lys⁺, His⁰, Glu⁻, Asp⁻, C-terminus⁻), resulting in an “intrinsic” charge of 6+ for cyt *c* and zero for ubiquitin. Excess Na⁺ were added to bring the initial net droplet charge to $z_R = 46+$.²⁵ Four additional Na⁺ / Cl⁻ pairs were inserted to reflect the presence of some chloride counterions. For each run, Na⁺ and Cl⁻ were placed in random positions; proteins were inserted in random positions and orientations, and with random inter-protein distances. All runs used different initial atom positions and velocities. Simulations were performed at 370 K for 75 ns, then

the temperature was raised to 450 K for an additional 75 ns to promote the final steps of solvent evaporation. Charge states of MD-generated gaseous ions were determined by tallying the protein intrinsic charge and the charges of adducted Na⁺ and Cl⁻. Ω values of MD-generated protein structures were determined using Collidoscope⁸¹ for the final ($t = 150$ ns) desolvated species.

Results and Discussion

Protein Clustering Experiments. The first step in our efforts to understand ESI-induced protein clustering was the experimental characterization of simple test systems. We selected cyt *c* (12360 Da) and ubiquitin (8565 Da), both of which are common model proteins. Under physiological conditions both proteins adopt tightly folded monomeric structures; neither of them has a propensity to form noncovalent complexes in bulk solution.^{79, 80} While typical native ESI experiments try to suppress nonspecific clustering, we aimed to promote this phenomenon because our goal was to elucidate the clustering mechanism.

Native ESI of 5 μ M cyt *c* at pH 7 produced a narrow distribution of monomeric protein ions (Figure 1A). These data were acquired in the presence of ammonium acetate which is a standard background electrolyte for native ESI.²⁵ In contrast, aqueous solution without added background electrolyte resulted in nonspecific dimers and trimers (Figure 1B). The absence of these clusters in Figure 1A indicates that protein clustering may involve electrostatic contacts, keeping in mind that charge-charge interactions are weakened by dissolved electrolytes.⁸²

Next, we examined the concentration dependence of protein clustering in the absence of ammonium acetate. Nonspecific clustering was virtually absent for 2 μ M cyt *c* (Figure 1C), whereas an elevated protein concentration of 100 μ M generated abundant dimers, trimers, and tetramers

(Figure 1D). This trend is consistent with earlier observations.^{61, 68} A slight shift to lower charge states at elevated protein concentration (Figure 1C, D) may result from charge competition.⁸³

The solution conditions used for acquiring the data of Figure 1 resulted in $[M + zH]^{z+}$ ions that were almost free of adducts. Extensive protein cluster formation was also observed when electrospraying 100 μ M *cyt c* in the presence of 0.2 mM NaCl, with peak broadening due to salt adduction (Figure 2A). The clusters formed under these conditions had a $[M + (z-n+m)H + nNa + mCl]^{z+}$ composition, where a significant fraction of the overall charge resulted from sodiation (Figure S1).²⁵ Thus, protein clustering occurred regardless of the type of ESI charge carrier (H^+ or Na^+).

We also performed experiments on ubiquitin. For a protein concentration of 100 μ M in the presence of 0.2 mM NaCl, ubiquitin showed a behavior similar to that of *cyt c*. The ubiquitin spectra showed various clusters, ranging from dimers to pentamers (Figure 2B) with heterogeneous charging due to sodiation, protonation, and chloride binding (Figure S1). These data demonstrate that protein clustering takes place regardless of intrinsic protein charge, keeping in mind that at pH 7 *cyt c* carries a net positive charge ($pI \approx 10$),⁸⁴ while ubiquitin is neutral ($pI \approx 7$).⁸⁰

ESI Simulations – General Considerations. We and others^{34, 40, 41, 44, 51, 85-90} previously used MD simulations to gain insights into ESI mechanisms. The current work marks the first time that this approach was applied to nonspecific protein clustering, focusing on the two proteins introduced above. Realistic modeling of H^+ as ESI charge carrier is challenging. A workaround is to replace H^+ with low MW metal ions.³⁴ For the current simulations we therefore used droplets that were charged with Na^+ . Cl^- were added as well, to mimic the aqueous NaCl solutions used in the experiments of Figure 2 that produced $[M + (z-n+m)H + nNa + mCl]^{z+}$ gaseous ions.

As outlined in the Methods section, the charges on all titratable protein sites were set to their pH 7 values, in accordance with the composition of the bulk solution. ESI droplets can undergo acidification caused by electrochemical H⁺ production⁹¹ and evaporative shrinkage.⁴⁸ However, pH is a macroscopic property that is not necessarily meaningful in nanometer-sized droplets. Excess H⁺ are believed to stay preferentially on the droplet surface,⁹² such that ESI-induced acidification in the interior (where proteins reside throughout much of the process) is likely more moderate. Hence, the use of pH 7 charge patterns in our MD runs represents a reasonable approximation.

Our simulations relied on the premise that droplets containing more than one protein represent the prerequisite for nonspecific ESI clustering.^{24, 45, 61-70} Hence, the MD runs described below followed a strategy similar to earlier protein ESI simulations,^{34, 44} except that the initial droplets contained multiple protein molecules. The initial Rayleigh-charged droplets in our simulations had a 5.5 nm radius, consistent with the size in the ESI plume after several fission/evaporation cycles.²⁵

MD Simulations of Nonspecific Ubiquitin Clustering. A typical simulation run for an ESI droplet containing two ubiquitin molecules is depicted in Figure 3A. The droplets underwent gradual solvent evaporation, accompanied by occasional Na⁺ IEM events that kept the shrinking droplets close to the Rayleigh limit. Such IEM ejection of low MW ions is a common occurrence.^{34, 40, 41, 85, 87} One of these IEM events is highlighted in Figure 3A for $t = 0.5$ ns. The two ubiquitin molecules remained inside the droplet until all the water had evaporated. Initially ($t = 0.5$ ns), the proteins were well separated. At around $t \approx 33$ ns they established contact with one another. Ultimately, the proteins formed a gaseous dimer where the chains were noncovalently bound to one another. The dimer also contained a number of Na⁺ and Cl⁻, just like in the experimental spectra of Figures 2, S1. Tallying of all charged components revealed a cluster charge of 9+. Ten repeat runs yielded 9+ (6/10) and

10+ clusters (4/10), for an average charge of 9.4+. These MD-generated charge states fall within the range of the experimentally observed ubiquitin dimers (6+ to 11+, Figure 2B).

Similar events were observed in simulations on droplets containing three ubiquitin molecules, culminating in the formation of trimeric clusters with a number of Na⁺ and Cl⁻ attached (Figure 3B). Seven repeat runs produced trimers in charge states 10+ (1/7), 11+ (4/7), 12+ (1/7), and 13+ (1/7), for an average charge of 11.3+. These MD results coincide with the experimentally observed trimer charge states (Figure 2B).

Overall, the MD data of Figure 3 reveal that ESI-generated ubiquitin clusters form via the CRM. Two or more proteins that are entrapped in the same droplet associate with one another as the droplet dries out. The net charge of the resulting gaseous cluster is governed by the residual Na⁺ and Cl⁻ that bind to the cluster as the final solvent layers evaporate. This CRM cluster formation is consistent with mechanistic proposals that had been put forward in earlier studies.^{25, 45, 57, 61, 68, 70}

MD Simulations on Droplets Containing Two Cyt *c*. A key difference between ubiquitin and cyt *c* is that the former has an intrinsic charge of ~zero, while the cyt *c* intrinsic charge is 6+ (see Methods). Simulations analogous to those discussed in the preceding section were conducted for cyt *c* to examine whether intrinsic protein charge affects the clustering mechanism. Figure 4A displays MD snapshots for a cyt *c* dimer CRM trajectory similar to that seen for ubiquitin in Figure 3A. Both cyt *c* molecules stayed in the droplet, and they associated with one another as the solvent evaporated. The resulting ESI-generated dimer accommodated several Na⁺ and Cl⁻, for an overall charge state of 13+. Nine replicates were performed, and CRM behavior like that of Figure 4A was seen in 5/9 instances. The charge states formed in these runs were 12+ and 13+, for an average charge of 12.8+. These MD-generated dimer charge states are consistent with the experiments of Figure 2A.

Interestingly, the remaining 4/9 *cyt c* runs showed a different behavior. Instead of forming a dimer, they culminated in the IEM ejection of one protein, while the other protein remained in the droplet and followed the CRM (Figure 4B). Protein IEM has been examined in earlier MD work from our laboratory.⁴⁴ IEM ejection is driven by electrostatic repulsion between the analyte charge and the other charges within the droplet.⁴⁴ This explains why IEM ejection can take place for *cyt c* (intrinsic charge 6+), while it is not feasible for ubiquitin (intrinsic charge zero). IEM and CRM-generated monomeric *cyt c* ions in our MD runs had charge states of 7+ and 8+, in agreement with the experimental charge states (Figure 2A). Overall, the *cyt c* data of Figure 4 reveal that the entrapment of two proteins in the same ESI droplet does not necessarily have to culminate in a nonspecific cluster. In addition to CRM dimer formation (Figure 4A), we observed the formation of monomeric protein ions via IEM and CRM pathways (Figure 4B).

MD Simulations on Droplets Containing Three *Cyt c*. The capability of *cyt c* to undergo IEM ejection due to its high intrinsic charge gave rise to a variety of scenarios for droplets that initially contained three proteins. Nine runs were conducted for droplets of this type. The majority of these simulations (5/9) showed CRM behavior, where all three proteins stayed in the droplet and formed trimeric clusters (Figure 5A). The MD charge states of these trimers were between 15+ and 18+ for an average of 16.6+, consistent with the experimental trimer charge states (Figure 2A). Additionally, we observed instances where either a single protein, a protein dimer, or a protein trimer underwent IEM ejection (Figure 5B-D). Any proteins remaining in the droplet after these IEM events followed the CRM. Charge states of the protein clusters formed via these IEM/CRM avenues were indistinguishable from those discussed previously.

The ejection of *cyt c* dimers and trimers (Figure 5C, D) illustrates that the CRM is *not* the only mechanism that can produce nonspecific protein clusters. Instead, the IEM also presents a

viable pathway. Prerequisites for this IEM avenue are (i) protein cluster formation prior to (or during) ejection, and (ii) a sufficiently large intrinsic cluster charge that can trigger ejection via repulsion from the remaining droplet charge ($2 \times 6+$, or $3 \times 6+$ in Figure 5C, D).⁴⁴ Nonetheless, the canonical CRM scenario of Figure 5A remained the dominant clustering mechanism even for cyt *c*, evident from the fact that more than half of the trimer-containing droplets showed this behavior.

As noted, the IEM has traditionally been associated with low MW ions,^{27, 38-42} exemplified in Figure 3A (at $t = 0.5$ ns) for Na^+ . However, it has already been suggested that this mechanism can also apply to larger analytes such as peptides and proteins.⁴³ The largest species for which IEM behavior had previously been demonstrated is monomeric cyt *c*.⁴⁴ The data of Figure 5D now extend this IEM range to cyt *c* trimers with a mass of 37 kDa.

Cluster Formation Kinetics. After having established that most of the ubiquitin and cyt *c* MD runs followed the CRM, we examined at what point during droplet shrinkage the protein clusters were formed. To this end, we tracked protein-protein distances vs. time by focusing on specific marker atoms that were buried close to the protein center (the C_α atoms of V26 and L32 were chosen for ubiquitin and cyt *c*, respectively). The time point when the CRM clusters formed varied considerably between individual runs. For ubiquitin dimers, Figure 6B illustrates an instance where the two proteins assembled during the very final stages of solvent evaporation, around $t = 45$ ns. The trajectory in Figure 6C illustrates the opposite extreme, where the dimer formed already after ~ 15 ns when more than half of the solvent was still present.

Considerable temporal heterogeneity was also observed for ubiquitin trimers; in Figure 6E, formation of a dimeric complex was followed almost immediately by binding of the third protein. On the other hand, there were also examples of trajectories where a dimer formed instantaneously, followed by binding of the third chain at a much later stage (Figure 6F). A similar temporal

heterogeneity was also observed for cyt *c* dimers and trimers (Figure S2). We conclude that nonspecific CRM clustering does not follow a tightly scripted timeline. Instead, the exact cluster formation time point depends on the random rotational and translational diffusion of the proteins, as well as shape fluctuations of the droplet.

Protein-Protein Contacts. What are the types of interactions that link the components of ESI-generated protein clusters? In their native state, both ubiquitin and cyt *c* have globular structures with a hydrophobic core, while charged and other hydrophilic residues are found in the exterior. Inspection of the MD-generated clusters revealed that none of the protein chains underwent major structural changes during cluster formation, implying that nonpolar core residues remained inaccessible and were not available for intermolecular contacts. Instead, the cluster interfaces comprised a multitude of salt bridges among surface residues (Arg⁺, Lys⁺, Glu⁻, Asp⁻). These electrostatic networks also incorporated a number of Na⁺ and Cl⁻ ions (Figure S3).

The role of salt bridges as a dominant type of protein-protein contact in ESI-generated clusters is consistent with the experiments of Figure 1A, B, where 10 mM ammonium acetate suppressed cluster formation. We tentatively attribute this effect to salt-induced electrostatic screening,⁸² which weakens salt bridges and thereby interferes with the formation of protein-protein contacts in the droplets. In comparison, the presence of a low NaCl concentration (0.2 mM) was not sufficient for suppressing cluster formation (Figure 2). The existence of salt bridges in monomeric proteins has been noted earlier,^{93, 94} but the dominant role of these zwitterionic contacts for nonspecific protein clustering (as seen in Figure S3) had not been demonstrated previously.

Collision Cross Sections. The Ω values of MD-generated ubiquitin dimers and trimers were slightly (5 - 10%) lower than the corresponding experimental values (Figure S4A,B). Considering the

challenges associated with the accurate calculation of Ω values,¹³ this level of agreement is quite reasonable. For *cyt c* clusters, the discrepancy between MD-generated and experimental Ω values was somewhat larger, particularly for trimers (up to ~20%, Figure S4C,D). These deviations suggest that the experimentally generated clusters have less compact structures than those produced in our simulations. Almost all of the MD-generated *cyt c* trimers had a triangular arrangement (Figure 5A,D). Only one of the runs produced a *cyt c* trimer with a more elongated structure, where three globular chains were in a linear arrangement (Figure S4D). Interestingly, the calculated Ω of this linear trimer was in excellent agreement with the experimental value of ~38 nm². It is therefore possible that ESI clustering of *cyt c* under experimental conditions favors linear trimers, in contrast to the MD runs which mostly produced triangular complexes.

Conclusions

Nonspecific protein clustering is a well-known problem in native ESI. If not properly recognized, this process can mislead experimentalists into reporting erroneously high protein-protein binding affinities. Even worse, clustering can generate complexes that are completely artifactual, as demonstrated here for two proteins that are known to be monomeric in solution. The formation of such nonspecific clusters can be suppressed by ensuring that the initial ESI droplets contain no more than a single protein molecule, e.g., by using low analyte concentrations (Figure 1C/D)^{61, 68} and/or narrow emitter tips that produce smaller droplets.^{24, 62, 66, 67, 70} Also, the use of relatively high background electrolyte concentrations (e.g., ammonium acetate) can help suppress the formation of ESI-generated salt bridge contacts (Figures 1A/B, S3).

For the first time, this work provides an atomistic view of ESI-induced protein cluster formation. Our MD runs generated dimers and trimers in charge states that coincided with experimental values, attesting to the fidelity of the modeling strategy used. Overall, our simulations support the view^{24, 45, 61-70} that protein cluster formation is often linked to the CRM, where individual solute molecules are forced to “stick” to one another as solvent evaporates. However, we found that the CRM is not the only possible protein cluster formation pathway. For proteins that carry a sufficiently large intrinsic charge, clustering within the droplet can be followed by IEM cluster ejection. Similar IEM scenarios have previously been discussed for certain salt clusters,^{31, 95, 96} but the existence of this pathway for proteins is unexpected.

Regardless of the release mechanism (CRM or IEM), protein cluster assembly was shown to occur at various time points for different runs. Sometimes protein-protein contacts were formed very early when the droplet had lost very little solvent, while in other instances, clustering took place during the final stages of solvent evaporation. One might have expected that clustering is disfavored for proteins that are intrinsically charged, such that the chains repel each other. Interestingly, we found that intrinsic charge is irrelevant for protein clustering, since *cyt c* (intrinsic charge 6+) and ubiquitin (intrinsic charge zero) showed very similar behavior. We attribute this effect to the tendency of droplets to project their internal charge to the droplet surface via dipole ordering of the solvent, thereby creating a field-free region in the droplet interior.⁹⁷

All clusters in our MD runs were assembled from proteins that retained a native-like fold. Protein-protein contacts were mediated primarily by salt bridges involving surface residues. This retention of native-like structure might open up interesting avenues for emerging single-particle structure determination methods, where the deliberate creation of protein clusters could help boost signal intensities by increasing the number of X-ray scattering centers.⁹⁸ In addition, the combination of ESI clustering with soft-landing⁹⁹ could produce protein assemblies for future applications in

nanotechnology. Hence, while nonspecific clustering is a nuisance in native ESI, there could be scenarios where clustering is desirable. It is hoped that the mechanistic insights of this work will stimulate future studies on the behavior of proteins in charged droplets.

Supporting Information. Figure S1: Close-up view of selected protein signals in the presence of NaCl. Figure S2: Protein-protein distances during cluster formation. Figure S3: Protein-protein contacts. Figure S4: Collision cross sections of experimental and MD-generated clusters.

Acknowledgment. We thank Prof. Vicki Wysocki and her students for stimulating discussions.

References

- (1) Mei, G.; Di Venere, A.; Rosato, N.; Finazzi-Agro, A. The importance of being dimeric. *FEBS J.* **2005**, *272*, 16-27.
- (2) Saibil, H. Chaperone machines for protein folding, unfolding and disaggregation. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 630-642.
- (3) Zhou, A. N.; Rohou, A.; Schep, D. G.; Bason, J. V.; Montgomery, M. G.; Walker, J. E.; Grigorieff, N.; Rubinstein, J. L. Structure and conformational states of the bovine mitochondrial ATP synthase by cryo-EM. *Elife* **2015**, *4*.
- (4) Rossmann, M. G. Structure of viruses : a short history. *Q. Rev. Biophys.* **2013**, *46*, 133-180.
- (5) Keskin, O.; Tuncbag, N.; GURSOY, A. Predicting Protein-Protein Interactions from the Molecular to the Proteome Level. *Chem. Rev.* **2016**, *116*, 4884-4909.
- (6) Balchin, D.; Hayer-Hartl, M.; Hartl, F. U. In vivo aspects of protein folding and quality control. *Science* **2016**, *353*, 42.
- (7) Chiti, F.; Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu. Rev. Biochem.* **2017**, *86*, 27-68.
- (8) 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.
- (9) Leney, A. C.; Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 5-13.
- (10) Cubrilovic, D.; Barylyuk, K.; Hofmann, D.; Walczak, M. J.; Graber, M.; Berg, T.; Wider, G.; Zenobi, R. Direct monitoring of protein-protein inhibition using nano electrospray ionization mass spectrometry. *Chem. Sci.* **2014**, *5*, 2794-2803.
- (11) Sharon, M. How Far Can We Go with Structural Mass Spectrometry of Protein Complexes? *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 487-500.
- (12) Lutomski, C. A.; Lykтей, N. A.; Zhao, Z. C.; Pierson, E. E.; Zlotnick, A.; Jarrold, M. F. Hepatitis B Virus Capsid Completion Occurs through Error Correction. *J. Am. Chem. Soc.* **2017**, *139*, 16932-16938.
- (13) Allison, T. M.; Barran, P.; Cianferani, S.; Degiacomi, M. T.; Gabelica, V.; Grandori, R.; Marklund, E. G.; Menneteau, T.; Migas, L. G.; Politis, A.; Sharon, M.; Sobott, F.; Thalassinou, K., et al. Computational Strategies and Challenges for Using Native Ion Mobility Mass Spectrometry in Biophysics and Structural Biology. *Anal. Chem.* **2020**, *92*, 10872-10880.
- (14) Ashcroft, A. E. Mass Spectrometry and the Amyloid Problem - How Far Can We Go in the Gas Phase? *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1087-1096.
- (15) Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T. Ion mobility-mass spectrometry reveals a conformational conversion from random assembly to β -sheet in amyloid fibril formation. *Nat. Chem.* **2011**, *3*, 172-177.
- (16) Jurneczko, E.; Barran, P. E. How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. *Analyst* **2011**, *136*, 20-28.
- (17) Zhong, Y.; Han, L.; Ruotolo, B. T. Collisional and Coulombic unfolding of gas-phase proteins: high correlation to their domain structures in solution. *Angew. Chem.* **2014**, *53*, 9209-9212.
- (18) Yang, Y.; Niu, C. D.; Bobst, C. E.; Kaltashov, I. A. Charge Manipulation Using Solution and Gas-Phase Chemistry to Facilitate Analysis of Highly Heterogeneous Protein Complexes in Native Mass Spectrometry. *Anal. Chem.* **2021**, *93*, 3337-3342.

- (19) Felitsyn, N.; Kitova, E. N.; Klassen, J. S. Thermal Decomposition of a Gaseous Multiprotein Complex Studied by Blackbody Infrared Radiative Dissociation. Investigating the Origin of the Asymmetric Dissociation Behavior. *Anal. Chem.* **2001**, 73, 4647-4661.
- (20) Li, H. L.; Nguyen, H. H.; Ogorzalek Loo, R. R.; Campuzano, I. D. G.; Loo, J. A. An integrated native mass spectrometry and top-down proteomics method that connects sequence to structure and function of macromolecular complexes. *Nat. Chem.* **2018**, 10, 139-148.
- (21) Zhang, H.; Wen, J.; Blankenship, R. E.; Gross, M. L. Native Electrospray and Electron-Capture Dissociation FTICR Mass Spectrometry for Top-Down Studies of Protein Assemblies. *Anal. Chem.* **2011**, 83, 5598-5606.
- (22) Morrison, L. J.; Brodbelt, J. S. 193 nm Ultraviolet Photodissociation Mass Spectrometry of Tetrameric Protein Complexes Provides Insight into Quaternary and Secondary Protein Topology. *J. Am. Chem. Soc.* **2016**, 138, 10849-10859.
- (23) Lermyte, F.; Konijnenberg, A.; Williams, J. P.; Brown, J. M.; Valkenburg, D.; Sobott, F. ETD Allows for Native Surface Mapping of a 150 kDa Noncovalent Complex on a Commercial Q-TWIMS-TOF Instrument. *J. Am. Soc. Mass Spectrom.* **2014**, 25, 343-350.
- (24) Panczyk, E. M.; Gilbert, J. D.; Jagdale, G. S.; Stiving, A. Q.; Baker, L. A.; Wysocki, V. H. Ion Mobility and Surface Collisions: Submicrometer Capillaries Can Produce Native-like Protein Complexes. *Anal. Chem.* **2020**, 92, 2460-2467.
- (25) 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.
- (26) 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.
- (27) Harper, C. C.; Brauer, D. D.; Francis, M. B.; Williams, E. R. Direct observation of ion emission from charged aqueous nanodrops: effects on gaseous macromolecular charging. *Chem. Sci.* **2021**, 12, 5185-5195.
- (28) 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.
- (29) 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.
- (30) 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.
- (31) 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.
- (32) Hamdy, O. M.; Julian, R. R. Reflections on Charge State Distributions, Protein Structure, and the Mystical Mechanism of Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* **2012**, 23, 1-6.
- (33) Tiwari, P.; Czar, M. F.; Zenobi, R. Fluorescence-Based Detection of the Desolvation Process of Protein Ions Generated in an Aqueous Electrospray Plume. *Anal. Chem.* **2021**, 93, 3635-3642.
- (34) 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.

- (35) 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.
- (36) Iavarone, A. T.; Williams, E. R. Mechanism of Charging and Supercharging Molecules in Electrospray Ionization. *J. Am. Chem. Soc.* **2003**, 125, 2319-2327.
- (37) Kaltashov, I. A.; Mohimen, A. Estimates of Protein Surface Area in Solution by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2005**, 77, 5370-5379.
- (38) Iribarne, J. V.; Thomson, B. A. On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **1976**, 64, 2287-2294.
- (39) 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.
- (40) Consta, S.; Mainer, K. R.; Novak, W. Fragmentation mechanisms of aqueous clusters charged with ions. *J. Chem. Phys.* **2003**, 119, 10125-10132.
- (41) Znamenskiy, V.; Marginean, I.; Vertes, A. Solvated Ion Evaporation from Charged Water Droplets. *J. Phys. Chem. A* **2003**, 107, 7406-7412.
- (42) McQuinn, K.; Hof, F.; McIndoe, J. S. Direct observation of ion evaporation from a triply charged nanodroplet. *Chem. Commun.* **2007**, 2007, 4099-4101.
- (43) 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.
- (44) Aliyari, E.; Konermann, L. Formation of Gaseous Proteins via the Ion Evaporation Model (IEM) in Electrospray Mass Spectrometry. *Anal. Chem.* **2020**, 92, 10807-10814.
- (45) 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.
- (46) 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.
- (47) McDonald, L. W.; Campbell, J. A.; Clark, S. B. Failure of ESI Spectra to Represent Metal-Complex Solution Composition: A Study of Lanthanide-Carboxylate Complexes. *Anal. Chem.* **2014**, 86, 1023-1029.
- (48) Zhou, S.; Prebyl, B. S.; Cook, K. D. Profiling pH Changes in the Electrospray Plume. *Anal. Chem.* **2002**, 74, 4885-4888.
- (49) Van Berkel, G. J.; Asano, K. G.; Schnier, P. D. Electrochemical Processes in a Wire-in-a-Capillary Bulk-Loaded, Nano-Electrospray Emitter. *J. Am. Soc. Mass Spectrom.* **2001**, 12, 853-862.
- (50) Bich, C.; Baer, S.; Jecklin, M. C.; Zenobi, R. Probing the Hydrophobic Effect of Noncovalent Complexes by Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2010**, 21, 286-289.
- (51) Oh, M. I.; Consta, S. Stability of a Transient Protein Complex in a Charged Aqueous Droplet with Variable pH. *J. Phys. Chem. Lett.* **2017**, 8, 80-85.
- (52) Xiao, H.; Kaltashov, I. A.; Eyles, S. J. Indirect Assessment of Small Hydrophobic Ligand Binding to a Model Protein Using a Combination of ESI MS and HDX/ESI MS. *J. Am. Soc. Mass Spectrom.* **2003**, 14, 506-515.
- (53) Wigger, M.; Eyler, J. R.; Benner, S. A.; Li, W.; Marshall, A. G. Fourier transform-ion cyclotron resonance mass spectrometric resolution, identification, and screening of non-covalent complexes of Hck Src homology 2 domain receptor and ligands from a 324-member peptide combinatorial library. *J. Am. Soc. Mass Spectrom.* **2002**, 13, 1162-1169.

- (54) Winger, B. E.; Light-Wahl, K. J.; Ogorzalek Loo, R. R.; Udseth, H. R.; Smith, R. D. Observation and implications of high mass-to-charge ratio ions from electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 536-545.
- (55) Servage, K. A.; Silveira, J. A.; Fort, K. L.; Clemmer, D. E.; Russell, D. H. Water-Mediated Dimerization of Ubiquitin Ions Captured by Cryogenic Ion Mobility-Mass Spectrometry. *J. Phys. Chem. Lett.* **2015**, *6*, 4947-4951.
- (56) Mistarz, U. H.; Chandler, S. A.; Brown, J. M.; Benesch, J. L. P.; Rand, K. D. Probing the Dissociation of Protein Complexes by Means of Gas-Phase H/D Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 45-57.
- (57) Jurchen, J. C.; Williams, E. R. Origin of Asymmetric Charge Partitioning in the Dissociation of Gas-Phase Protein Homodimers. *J. Am. Chem. Soc.* **2003**, *125*, 2817-2826.
- (58) Abzalimov, R. R.; Frimpong, A. K.; Kaltashov, I. A. Gas-phase processes and measurements of macromolecular properties in solution: On the possibility of false positive and false negative signals of protein unfolding. *Int. J. Mass Spectrom.* **2006**, *253*, 207-216.
- (59) van Berkel, W. J. H.; van den Heuvel, R. H. H.; Versluis, C.; Heck, A. J. R. Detection of intact megaDalton protein assemblies of vanillyl-alcohol oxidase by mass spectrometry. *Protein Sci.* **2000**, *9*, 435-439.
- (60) Sever, A. I. M.; Yin, V.; Konermann, L. Interrogating the Quaternary Structure of Noncanonical Hemoglobin Complexes by Electrospray Mass Spectrometry and Collision-Induced Dissociation. *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 270-280.
- (61) Lane, L. A.; Ruotolo, B. T.; Robinson, C. V.; Favrin, G.; Benesch, J. L. P. A Monte Carlo approach for assessing the specificity of protein oligomers observed in nano-electrospray mass spectra. *Int. J. Mass Spectrom.* **2009**, *283*, 169-177.
- (62) Xia, Z. J.; Williams, E. R. Effect of droplet lifetime on where ions are formed in electrospray ionization. *Analyst* **2019**, *144*, 237-248.
- (63) Qian, C.; Fu, H. Q.; Kovalchik, K. A.; Li, H. H.; Chen, D. D. Y. Specific Binding Constant and Stoichiometry Determination in Free Solution by Mass Spectrometry and Capillary Electrophoresis Frontal Analysis. *Anal. Chem.* **2017**, *89*, 9483-9490.
- (64) Zhang, H.; Lu, H. Y.; Chingin, K.; Chen, H. W. Stabilization of Proteins and Noncovalent Protein Complexes during Electrospray Ionization by Amino Acid Additives. *Anal. Chem.* **2015**, *87*, 7433-7438.
- (65) Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Gas phase stabilization of noncovalent protein complexes formed by electrospray ionization. *Anal. Chem.* **2009**, *81*, 7801-7806.
- (66) Hu, J.; Guan, Q. Y.; Wang, J.; Jiang, X. X.; Wu, Z. Q.; Xia, X. H.; Xu, J. J.; Chen, H. Y. Effect of Nanoemitters on Suppressing the Formation of Metal Adduct Ions in Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2017**, *89*, 1838-1845.
- (67) 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.
- (68) Reardon, P. N.; Jara, K. A.; Rolland, A. D.; Smith, D. A.; Hoang, H. T. M.; Prell, J. S.; Barbar, E. J. The dynein light chain 8 (LC8) binds “in-register” to multivalent intrinsically disordered partners. *J. Biol. Chem.* **2020**, *295*, 4912-4922.
- (69) Daubenfeld, T.; Bouin, A.-P.; van der Rest, G. A Deconvolution Method for the Separation of Specific Versus Nonspecific Interactions in Noncovalent Protein-Ligand Complexes Analyzed by ESI-FT-ICR Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1239-1248.

- (70) Davidson, K. L.; Oberreit, D. R.; Hogan, C. J.; Bush, M. F. Nonspecific aggregation in native electrokinetic nanoelectrospray ionization. *Int. J. Mass Spectrom.* **2017**, 420, 35-42.
- (71) Muneeruddin, K.; Thomas, J. J.; Salinas, P. A.; Kaltashov, I. A. Characterization of Small Protein Aggregates and Oligomers Using Size Exclusion Chromatography with Online Detection by Native Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2014**, 86, 10692-10699.
- (72) Gibson, G. T. T.; Mugo, S. M.; Oleschuk, R. D. Nanoelectrospray Emitters: Trends and Perspective. *Mass Spectrom. Rev.* **2009**, 28, 918-936.
- (73) Kenakin, T. The mass action equation in pharmacology. *Br. J. Clin. Pharmacol.* **2016**, 81, 41-51.
- (74) 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.
- (75) Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H., et al. Recommendations for reporting ion mobility Mass Spectrometry measurements. *Mass Spectrom. Rev.* **2019**, 38, 291-320.
- (76) 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.
- (77) 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.
- (78) Abascal, J. L. F.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, 123, 234505.
- (79) Bushnell, G. W.; Louie, G. V.; Brayer, G. D. High-resolution Three-dimensional Structure of Horse Heart Cytochrome c. *J. Mol. Biol.* **1990**, 214, 585-595.
- (80) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. Structure of Ubiquitin Refined at 1.8 Å Resolution. *J. Mol. Biol.* **1987**, 194, 531-544.
- (81) 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.
- (82) Perez-Jimenez, R.; Godoy-Ruiz, R.; Ibarra-Molero, B.; Sanchez-Ruiz, J. M. The efficiency of different salts to screen charge interactions in proteins: A Hofmeister effect? *Biophys. J.* **2004**, 86, 2414-2429.
- (83) Pan, P.; McLuckey, S. A. Electrospray Ionization of Protein Mixtures at Low pH. *Anal. Chem.* **2003**, 75, 1491-1499.
- (84) Graf, M.; Garcia, R. G.; Watzig, H. Protein adsorption in fused-silica and polyacrylamide-coated capillaries. *Electrophoresis* **2005**, 26, 2409-2417.
- (85) Caleman, C.; van der Spoel, D. Evaporation from water clusters containing singly charged ions. *Phys. Chem. Chem. Phys.* **2007**, 9, 5105-5111.
- (86) 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.
- (87) 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.

- (88) 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.
- (89) 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.
- (90) 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.
- (91) Van Berkel, G. J.; Kertesz, V. Using the Electrochemistry of the Electrospray Ion Source. *Anal. Chem.* **2007**, 79, 5511-5520.
- (92) 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.
- (93) 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.
- (94) Marchese, R.; Grandori, R.; Carloni, P.; Raugei, S. On the Zwitterionic Nature of Gas-Phase Peptides and Protein Ions. *PLoS Comput. Biol.* **2010**, 6, e1000775.
- (95) 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.
- (96) Gamero-Castano, M.; de la Mora, J. F. Modulations in the Abundance of Salt Clusters in Electrosprays. *Anal. Chem.* **2000**, 72, 1426-1429.
- (97) Ahadi, E.; Konermann, L. Surface charge of electrosprayed water nanodroplets: A molecular dynamics study. *J. Am. Chem. Soc.* **2010**, 132, 11270-11277.
- (98) Bielecki, J.; Hantke, M. F.; Daurer, B. J.; Reddy, H. K. N.; Hasse, D.; Larsson, D. S. D.; Gunn, L. H.; Svenda, M.; Munke, A.; Sellberg, J. A.; Flueckiger, L.; Pietrini, A.; Nettelblad, C., et al. Electrospray sample injection for single-particle imaging with x-ray lasers. *Sci. Adv.* **2019**, 5, eaav8801.
- (99) Mikhailov, V. A.; Mize, T. H.; Benesch, J. L. P.; Robinson, C. V. Mass-Selective Soft-Landing of Protein Assemblies with Controlled Landing Energies. *Anal. Chem.* **2014**, 86, 8321-8328.

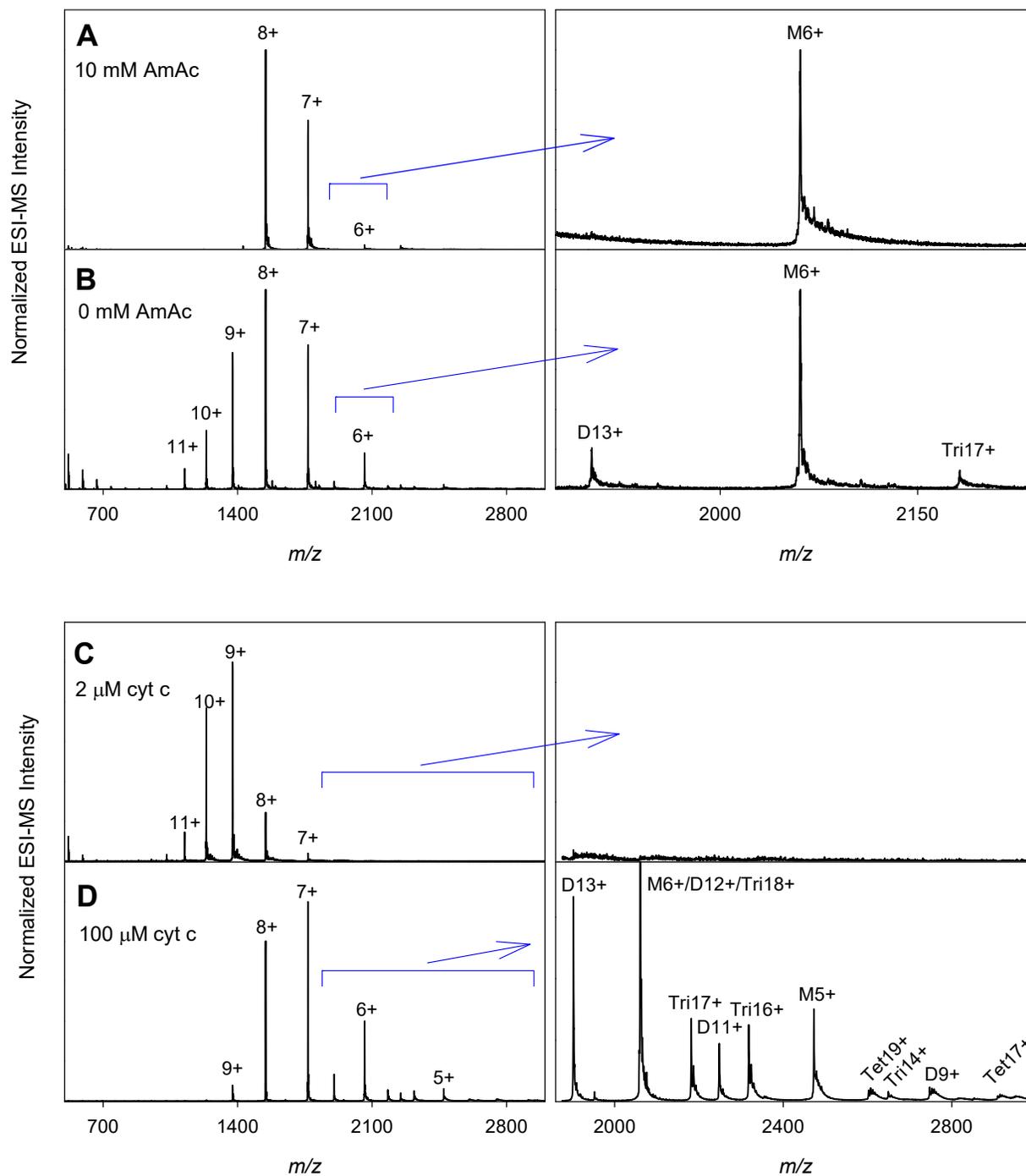


Figure 1. ESI mass spectra of *cyt c* acquired at pH 7 under different solvent conditions. Panels on the right zoom into regions of interest, with signals annotated as M (monomer), D (dimer), Tri (trimer), and Tet (tetramer) along with the corresponding charge states. (A) 5 μ M protein in water containing 10 mM ammonium acetate. (B) 5 μ M protein in water. (C) 2 μ M protein in water. (D) 100 μ M protein in water.

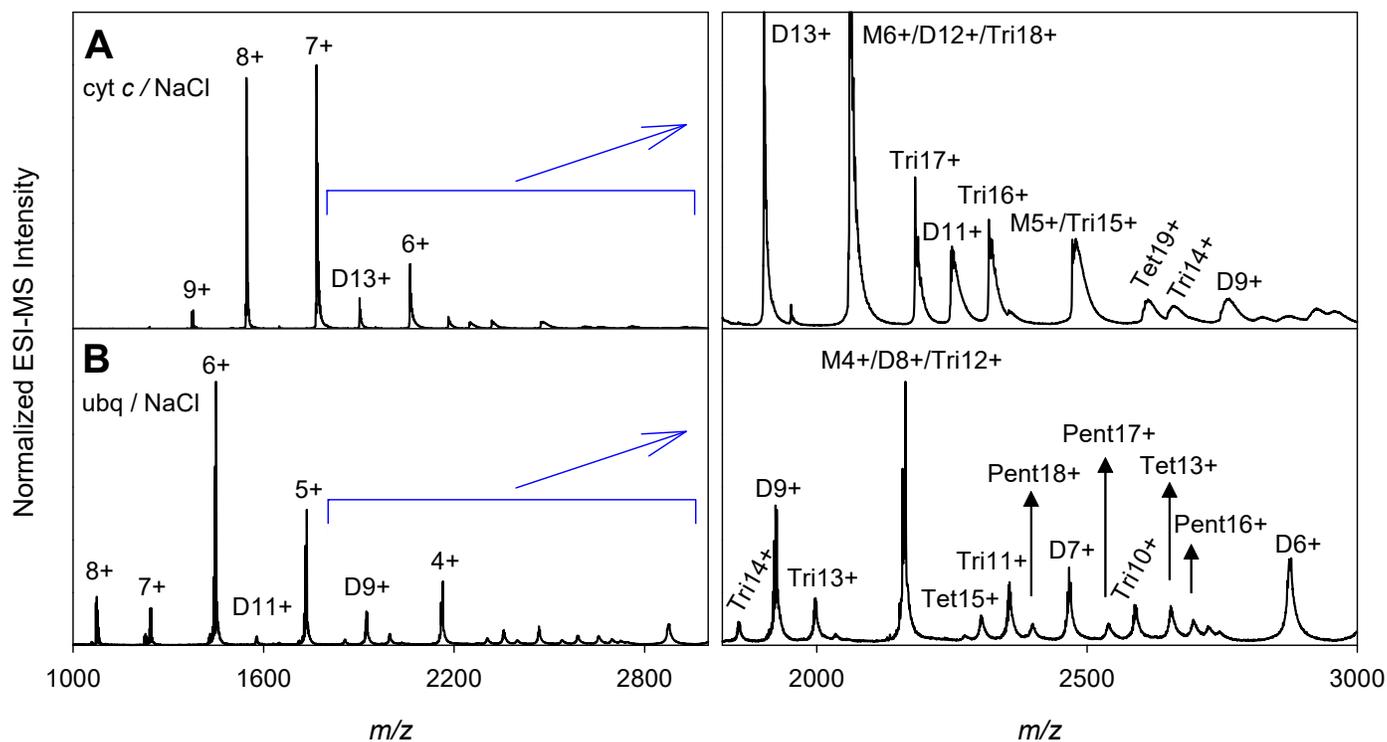


Figure 2. ESI mass spectra acquired in aqueous solution containing 100 μ M protein and 0.2 mM NaCl. (A) *cyt c*, (B) ubiquitin. Panels on the right zoom into regions of interest, with signals annotated as M (monomer), D (dimer), Tri (trimer), Tet (tetramer), and Pent (pentamer) along with the corresponding charge states. Peak broadening is due to salt adduction caused by the addition of NaCl (see Figure S1 for details).

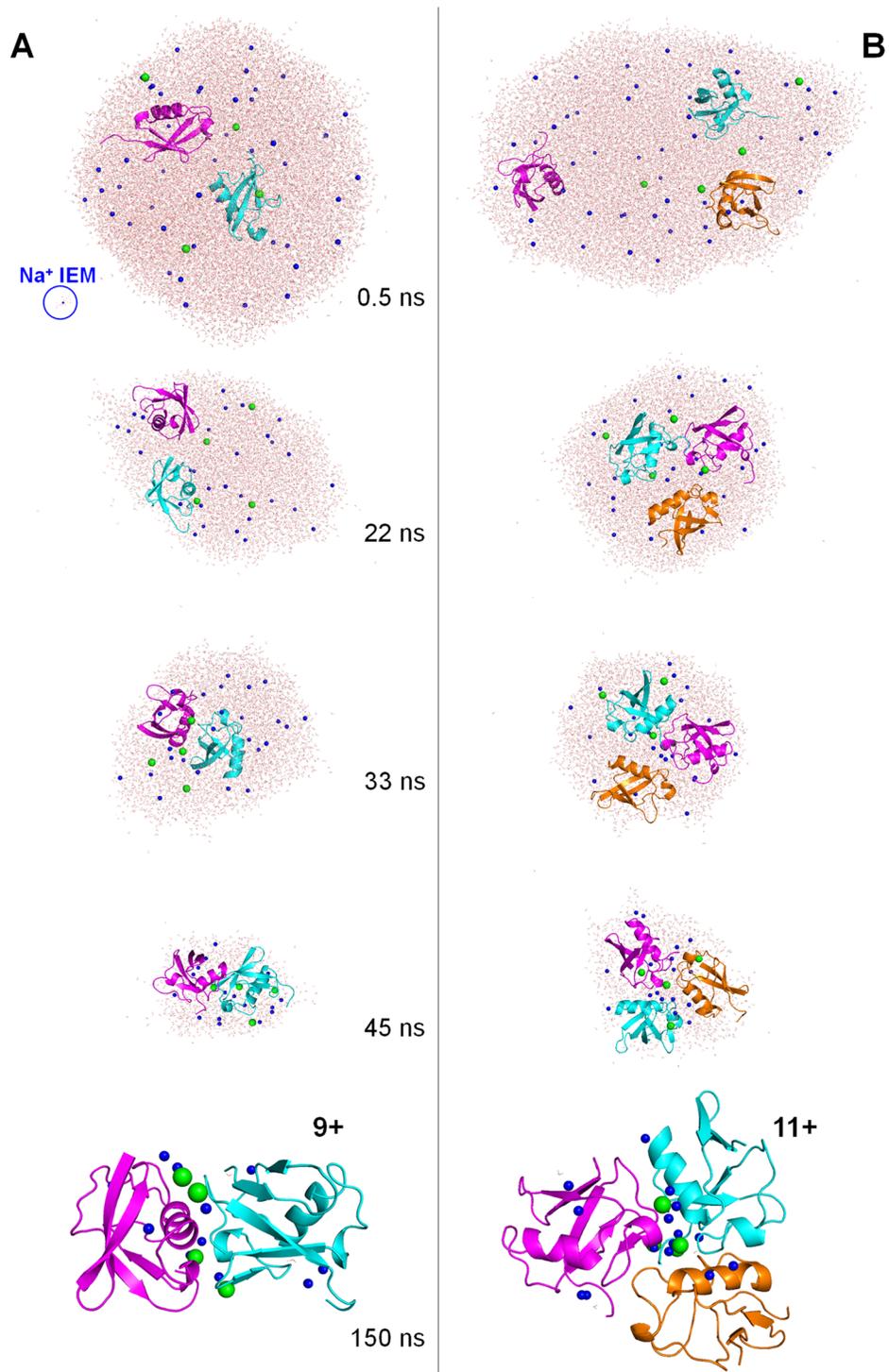


Figure 3. ESI simulation snapshots for two MD runs that show CRM behavior. (A) Droplet containing two ubiquitin molecules, forming a protein dimer. (B) Droplet containing three ubiquitin molecules, forming a protein trimer. The charge states of the nonspecific clusters are indicated in the final frame. Protein chains are depicted in different colors, water oxygen is red, Na^+ (blue) and Cl^- (green) are shown as spheres. Time points are identical for panels on the left and right. The blue circle in the top left panel highlights the IEM ejection of a solvated Na^+ .

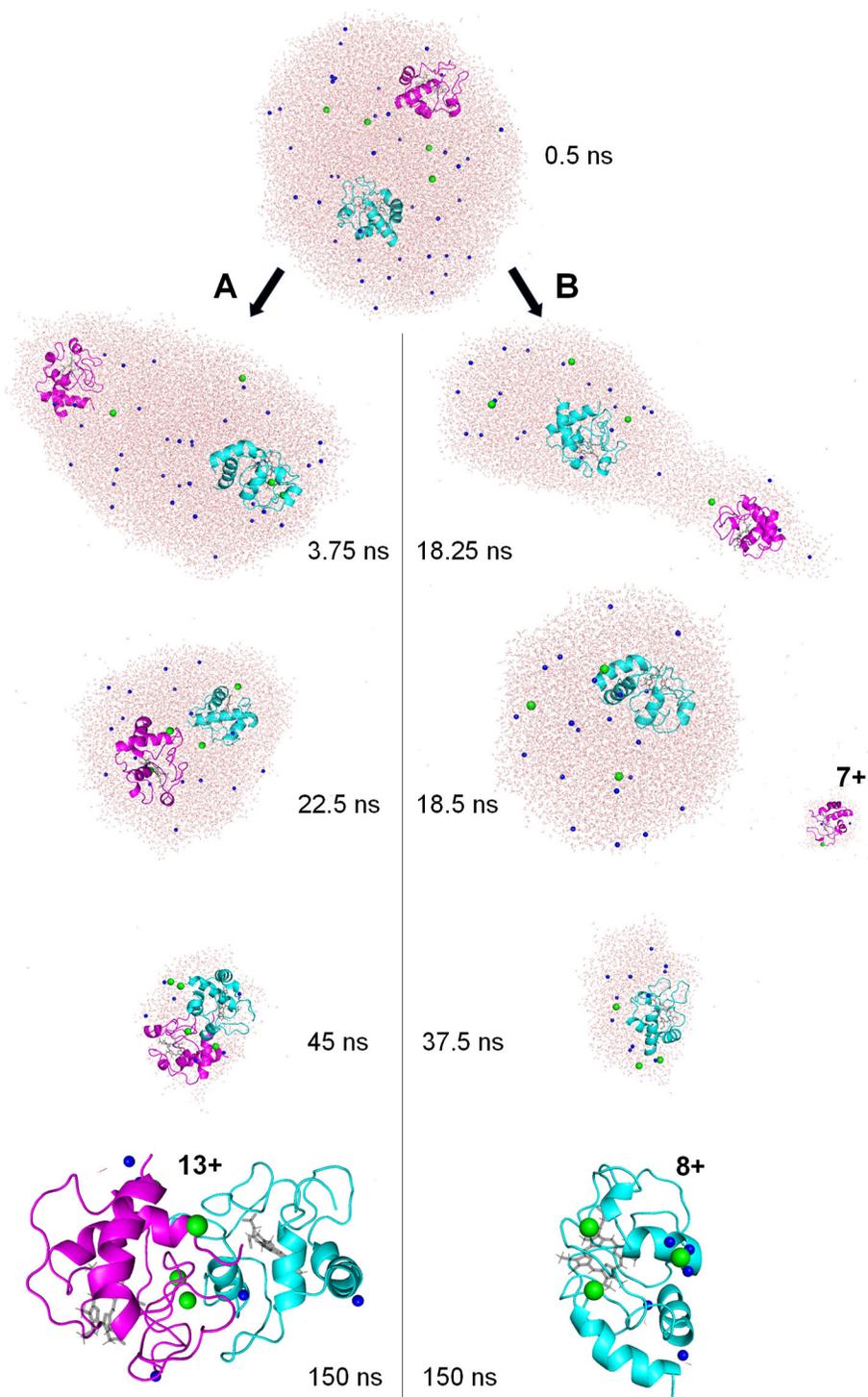


Figure 4. ESI simulation snapshots for two MD runs, both of which started with a droplet that initially contained two *cyt c* molecules (top). (A) The two proteins undergo nonspecific clustering, and a 13+ dimer is formed via the CRM. (B) One protein (magenta) undergoes IEM ejection at ~18.25 ns, liberating a 7+ monomer. The other protein (cyan) forms an 8+ monomer via the CRM. Element coloring is as in the preceding figure.

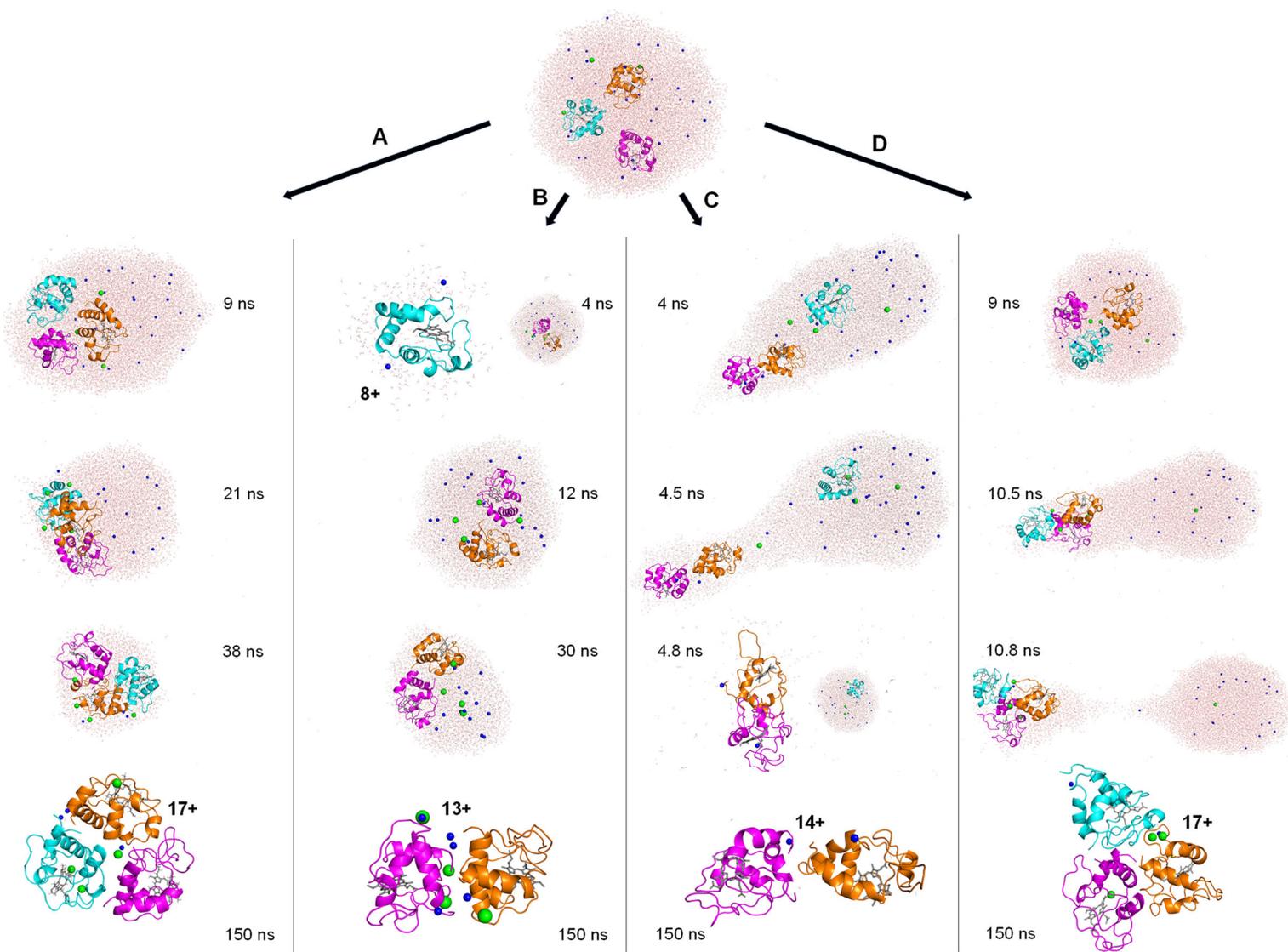


Figure 5. ESI simulation snapshots for four MD runs, all of which started with a droplet that initially contained three cyt *c* molecules (top). (A) All three proteins undergo clustering, and a 17+ trimer is formed via the CRM. (B) One protein (cyan) undergoes IEM ejection at 4 ns, liberating a 8+ monomer. The remaining two proteins form a 13+ dimer via the CRM. (C) A 14+ dimer undergoes IEM ejection at 4.8 ns. (D) A 17+ trimer is IEM ejected at 10.5 ns. Element coloring is as in the preceding figures.

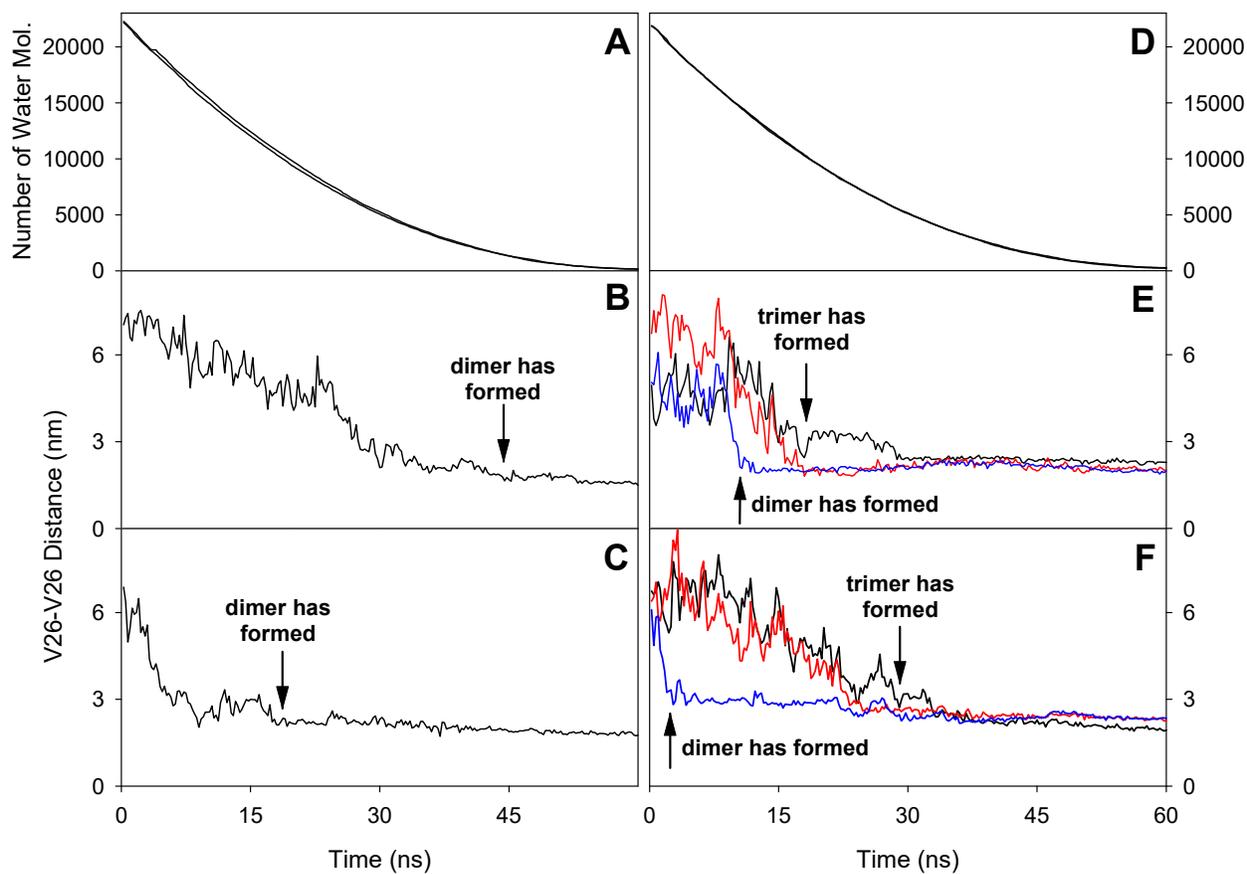


Figure 6. Examples of ubiquitin MD time profiles, illustrating cluster formation for droplets that initially contained two (A-C) or three (D-F) proteins. (A, D) Number of water molecules during droplet shrinkage for the data shown in the subsequent panels. (B, C) Distance between V26 C_{α} of the two proteins, representing a marker atom that is buried in the ubiquitin core. (E, F) Distances between V26 C_{α} atoms of the three proteins; the three traces represent distances between proteins 1-2, 1-3, and 2-3.

For Table of Contents Only

