7-15-2018

How to run molecular dynamics simulations on electrospray droplets and gas phase proteins: Basic guidelines and selected applications.

Lars Konermann

Haidy Metwally

Robert G McAllister

Vlad Popa

Invited Contribution

for

"MS-based Methods to Study Macromolecular Higher Order Structure and Interactions"

Edited by Igor A. Kaltashov

How to Run Molecular Dynamics Simulations On Electrospray Droplets and Gas Phase Proteins: Basic Guidelines and Selected Applications

Lars Konermann,* Haidy Metwally, Robert G. McAllister, Vlad Popa

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

* corresponding author, konerman@uwo.ca

Keywords: protein conformation; charged droplet; charged residue model; ion mobility spectrometry; Gromacs

Funding was provided by the Natural Sciences and Engineering Research Council of Canada (DG 217080-2013).

A B S T R A C T

The ability to transfer intact proteins and protein complexes into the gas phase by electrospray ionization (ESI) has opened up numerous mass spectrometry (MS)-based avenues for exploring biomolecular structure and function. However, many details regarding the ESI process and the properties of gaseous analyte ions are difficult to decipher when relying solely on experimental data. Molecular dynamics (MD) simulations can provide additional insights into the behavior of ESI droplets and protein ions. This review is geared primarily towards experimentalists who wish to adopt MD simulations as a complementary research tool. We touch on basic points such as force fields, the choice of a proper water model, GPU-acceleration, possible artifacts, as well as shortcomings of current MD models. Following this technical overview, we highlight selected applications. Simulations on aqueous droplets confirm that "native" ESI culminates in protein ion release via the charged residue model. MD-generated charge states and collision cross sections match experimental data. Gaseous protein ions produced by native ESI retain much of their solution structure. Moving beyond classical fixed-charge algorithms, we discuss a simple strategy that captures the mobile nature of H^+ within gaseous biomolecules. These mobile proton simulations confirm the high propensity of gaseous proteins to form salt bridges, as well as the occurrence of charge migration during collision-induced unfolding and dissociation. It is hoped that this review will promote the use of MD simulations in ESI-related research. We also hope to encourage the development of improved algorithms for charged droplets and gaseous biomolecular ions.

1. Introduction

Electrospray ionization (ESI) transforms solution phase proteins and other biomolecules into multiply charged gaseous ions for analysis by mass spectrometry (MS). “Native” ESI-MS [1-5] experiments aim to preserve solution structures and interactions throughout the transition from bulk solution into the vacuum of the mass spectrometer. This approach reports directly on protein-protein and protein-ligand binding stoichiometries [1-4,6,7]. Native ESI-MS requires non-denaturing aqueous solutions, along with ion sampling conditions that minimize collisional activation. This field continues to evolve rapidly, fueled by advances in ion sources, mass analyzers, experimental protocols, and computational/theoretical methods [8-12].

When operated in positive ion mode, an ESI source disperses protein solution into droplets that carry excess H^+, NH_4^+, or Na^+ [13,14]. Evaporation and fission events close to the Rayleigh limit produce successively smaller droplets [13,15]. The mechanisms of analyte ion release from these nanodroplets were shrouded in controversy for a long time [13,14,16-20]. In recent years molecular dynamics (MD) simulation techniques have contributed significantly to a better understanding of ESI droplets [21-27] and gas phase proteins [28-36]. The importance of MD techniques in this area stems from the fact that it is difficult to obtain high resolution structural information on ESI nanodroplets and large gaseous biomolecules. Ion mobility spectrometry (IMS) [37-39], optical techniques [40,41], and dissociation studies [42,43] provide partial insights, but they do not yield atomically resolved structural data. This is in contrast to the condensed phase, where detailed conformational information is available from X-ray crystallography, NMR spectroscopy, and cryo-electron microscopy [44,45].

This review is written for ESI-MS experimentalists who wish to adopt MD techniques for broadening the reach and scope of their activities. By sharing some of our own experiences, we hope to make MD simulations more accessible to this target group. Various MD simulation

packages are available, several of them as free open source downloads. Here we will focus on Gromacs [46,47], a program that has a large user community and offers impressive performance. We will briefly cover some technical aspects, before discussing MD data relevant to ESI droplets and the structure/dynamics of protein ions generated by native ESI.

2. Solution vs. Gas Phase Simulations

MD simulations [46] are a computational tool for modeling the temporal evolution of a system, e.g., the conformation of a protein, by iteratively integrating Newton's Second Law (*force* = *mass* × *acceleration*). The forces acting on every atom are calculated according to *force* = -grad $V(\mathbf{r})$, where $V(\mathbf{r})$ is the potential energy of an atom at position $\mathbf{r}$. $V(\mathbf{r})$ is determined by interactions with all other atoms. After calculating the acceleration for each atom, all atom coordinates are updated, and the process is repeated. Each of these iteration steps covers roughly 1 fs. The mathematical form of $V(\mathbf{r})$ is defined in the MD force field which includes atomic charges, Lennard-Jones parameters and torsional parameters (bond angles and lengths are often treated as constraints). Some widely used force fields are OPLS/AA [48], AMBER [49], and CHARMM [50]. Different versions of these force fields are available, and they remain under constant development. Compared to *ab initio* or DFT-based methods, the computational cost associated with MD simulations is low, allowing systems comprising many thousands of atoms to be studied on relatively long time scales. Nanoseconds to microseconds are easily accessible, even millisecond runs have been demonstrated [51,52].

Typical MD force fields such as those mentioned above are semi-empirical. Parameters are initially derived from *ab initio* or DFT calculations. Subsequent parameter fitting optimizes the match with condensed phase experimental data, including crystallographic conformational preferences and NMR dipolar couplings in solution [48,53]. Currently there is no comprehensive

force field that has been specifically designed for biomolecular systems in the gas phase. Instead, all of the commonly used force fields have been optimized for bulk solution simulations. It is common practice to apply these solution force fields for studies on ESI droplets [21-26] and electrosprayed proteins [28-36].

Is the use of solution force fields for gas phase simulations a valid approach? Under realistic conditions certain parameters such as charge distributions around surface-exposed atoms will differ in the two environments [54]. Polarizable models can capture some of these effects [55,56], but computational cost limits their usefulness. Luckily, even conventional (fixed charge) solution force fields can perform quite well in the gas phase, as exemplified by structural and energetic comparisons with *ab initio* vacuum data on small model compounds [48,53,57]. In addition, ESI droplets and native-like (compact) gaseous proteins have most of their atoms buried in the interior, such that they will not "feel" the gas phase environment to the same extent as would be the case for small model compounds *in vacuo*. Condensed phase force fields can be expected to provide a reasonable description of these buried moieties, whereas parameters derived for a genuine gas phase environment would likely be less appropriate.

In summary, the widely practiced use of solution force fields for gas phase simulations is borne out of necessity. Data generated in this way can nonetheless provide surprisingly good semi-quantitative descriptions of biomolecular gaseous systems. Ultimately, the results of *any* simulation have to be judged by the extent to which they offer new insights and make predictions that are verifiable in experiments. Future advances will hopefully produce computationally inexpensive models that are specifically designed for the gas phase. Some efforts in this direction are outlined below, such as the implementation of mobile proton algorithms [34,58].

3. Periodic Boundary Conditions in the Gas Phase?

MD simulations in solution usually employ periodic boundary conditions (PBC), i.e., the simulation box is surrounded by identical images of itself in all three dimensions [46]. The simulation box has to be neutralized by compensating the intrinsic protein charge via addition of electrolyte ions such as Na^+ and/or Cl^- to the solvent. PBC allow simulations on quasi-infinite bulk systems, without having to deal with surface effects (Fig. 1A). Electrostatic interactions within PBC simulations are commonly treated by splitting $V_{Coul}(\mathbf{r})$ into a short range term that is calculated explicitly, plus a long range term that is solved in reciprocal space using Particle Mesh Ewald (PME) summation [59,60]. Parameter files for such solution PBC/PME conditions are readily available [61].

For novices wanting to run gas phase MD simulations it is tempting to simply adapt the aforementioned solution protocol, turn the protein into a z+ ESI ion via a judicious choice of side chain charges, and use PBC/PME in the absence of solvent (Fig. 1B). Unfortunately, such PBC/PME strategies are not suitable for gas phase simulations. Subjecting a z+ simulation box to PBC implies a periodic system with infinite charge. The PME algorithm deals with this situation by adding a uniform neutralizing background charge that can induce simulation artifacts [62]. Also, the long-range nature of Coulomb interactions makes it unavoidable that PBC images interact with each other, unless special precautions are taken (see below). The proper way to run gas phase simulations is without PBC, without PME, and without cutoffs for Coulomb or Lennard-Jones interactions (Fig. 1C) [25].

The effects of an incorrectly implemented gas phase simulation can be demonstrated using a simple model system. We chose two gaseous Na^+ ions that were placed 1 nm apart at $t = 0$ in the vacuum. Once the MD simulation starts, the only physically reasonable outcome is that the two ions move away from each in monotonic fashion because of electrostatic repulsion. This behavior

is indeed observed when running the simulation under proper conditions, i.e., without PBC, without PME, and without cutoffs (Fig. 1D, blue line). In contrast, an unrealistic oscillatory motion is obtained when employing a typical solution-type MD protocol (without solvent, but with PBC, with PME, and with cutoffs), illustrated by the red line in Fig. 1D.

4. GPU-Accelerated Gas Phase Simulations in Gromacs: Pseudo-PBC

Gas phase MD simulations can be challenging because proper conditions without cutoffs (Fig. 1C) result in an awkward N^2 scaling [25], implying that runs become very slow when moving towards proteins and droplets with $N >> 10^3$ atoms. This problem can be mitigated by graphics processing unit (GPU) accelerated computing. In recent years molecular simulations have benefited tremendously from cheap high-performance GPUs that were originally developed for gaming (e.g., from NVIDIA). Gromacs and other MD packages can employ GPUs for non-bonded force calculations in a highly parallel fashion [47], boosting the performance of desktop workstations to levels that used to be restricted to high-end computer clusters [63].

Unfortunately, much of the GPU acceleration code in current MD packages is geared towards solution PBC/PME runs. For existing Gromacs versions (starting at 4.6) GPU acceleration is available only in conjunction with the Verlet neighbor search algorithm [64], and the latter has been implemented only for PBC. In other words, GPU acceleration in Gromacs requires the use of PBC. This is a problem for gas phase simulations, where PBC can cause artifacts as noted earlier (Fig. 1B). Luckily, there is a simple work-around for this issue.

GPU-accelerated gas phase simulations can be implemented in Gromacs by using a "pseudo-PBC" approach. For this purpose PME is switched off. The PBC box dimensions are increased to the maximum allowed value of $(999.9\ \text{nm})^3$. Coulomb and Lennard-Jones cutoffs are set to a value that is very large but does not exceed the box dimensions (we use 333.3 nm). Next,

the gas phase protein or droplet is placed in the center of the box and the MD run can commence. The basic idea behind this approach is that the simulation system will experience a proper vacuum environment as long as it remains within its 333.3 nm cutoff bubble (Fig. 1E). All Coulomb and Lennard-Jones interactions within the system are fully accounted for, but interactions between PBC images do not take place due to the combined effects of box size and cutoffs. It may be necessary to remove ejected solvent molecules or clusters from the simulation to ensure that PBC images remain isolated from one another. Gas phase data produced using the pseudo-PBC approach are indistinguishable from those generated in a genuine non-PBC vacuum environment (Fig. 1D), while shortening run times for large systems by one to two orders of magnitude.

5. Additional Considerations for ESI Droplets

Solvent evaporation from nanodroplets tends to lower the droplet temperature by evaporative cooling [65]. Under experimental conditions this effect is countered by heating in the ion source region. In simulations, evaporative cooling is problematic because it tends to freeze the system [25]. Traditional thermostating algorithms [66-68] only offer a partial solution to this problem [69,70]. To overcome this issue we developed a trajectory stitching approach [26,69], where long thermostated simulation runs are broken down into short (~250 ps) segments. Solvent molecules and charge carriers that have left the droplet are removed from the simulation after each segment. The remaining droplet is then re-centered in the pseudo-PBC box, and new atom velocities corresponding to the desired temperature are sampled from a Maxwell-Boltzmann distribution prior to beginning the next segment. In essence, this procedure represents a crude implementation of the Andersen thermostat [71]. In addition to stabilizing the temperature, trajectory stitching dramatically shortens the overall run time by successively reducing the number of molecules in the

simulation. For droplets containing sulfolane or other slowly evaporating species it may be advantageous to use trajectory stitching in conjunction with a forced evaporation protocol [72].

A number of droplet simulations in the literature have relied on three-site water models such as SPC, SPC/E, or TIP3P that are computationally cheap and perform well in bulk solution [52]. Unfortunately, these models yield a surface tension γ that is too low by roughly 30%. Considering the central role of γ for the ESI process [13,73,74], this can be a serious limitation. We advocate the use of TIP4P/2005 water which matches the experimental γ within ~1% over a wide temperature range [75]. TIP4P/2005 employs three fixed point charges and one Lennard-Jones center [76]. The presence of four sites renders TIP4P/2005 computationally more expensive than other models, but its superior performance ensures that simulation results can be directly compared to experimental data.

Accurately modeling the formation of $[M + zH]^{z+}$ analyte ions from aqueous droplets would require *ab initio* methods for dealing with Grotthus shuttling [77], i.e., water-water and water-protein proton transfer. Sadly, the time range and droplet size associated with ESI events are orders of magnitude beyond such high level approaches [78]. To sidestep this problem we resorted to simulate ESI droplets with Na^+ (instead of H^+) as excess charge carriers. Such simulations culminate in the formation of $[M + zNa]^{z+}$ protein ions [26]. Experiments have demonstrated that protonated and sodiated proteins generated by native ESI share very similar properties [26], implying that many aspects of their formation mechanisms are analogous to one another. Nonetheless, it is hoped that future developments will soon pave the way towards computational protocols for modeling proton transfer events for large solvent/protein systems on long time scales. First steps in this direction are already underway [32,34,58,79,80] (see also section 9).

6. Flying Ice Cubes

A well-known artifact that can arise during MD simulations is the so-called "flying ice cube effect" [81]. It describes a situation where the kinetic energy of the system gets converted largely into rotation around the center of mass and/or translation of the entire system. Under these conditions the total kinetic energy can be formally consistent with the user-selected temperature, while the kinetic energy stored in internal degrees of freedom is much lower than expected such that the system is internally cold. Also, centrifuge pseudoforces encountered during rapid rotation can trigger unrealistic events. Gas phase simulations are particularly vulnerable to such artifacts, because rotation/translation is not inhibited by friction against an external solvent. Different types of systems and different thermostats can be more or less prone to flying ice cube behavior [66-68,71]. It is therefore essential to periodically remove the net translation and rotation from the simulation, and to carefully inspect unprocessed trajectory data for possible artifacts.

7. Native ESI Simulations

The remainder of this review illustrates simulation results obtained using the MD strategies outlined above. We focus on native ESI conditions, where proteins in non-denaturing aqueous solution enter the ESI capillary in their biologically active, folded conformations. The current section and the two subsequent chapters will discuss three key issues (i) How are protein ions released from ESI nanodroplets during native ESI? (ii) Do electrosprayed proteins retain solution-like conformations under native ESI conditions? (iii) How do electrosprayed proteins respond to external perturbations such as collisional heating?

Fig. 2A shows MD data that describe the release of ubiquitin from an aqueous ESI nanodroplet that is charged with excess Na^+ [26]. Gradual droplet shrinkage as the result of water

evaporation is accompanied by the ejection of solvated Na^+ (Fig. 2B). These charge ejection events are well described by the ion evaporation model (IEM) [12,16]. As the final water molecules leave, all remaining Na^+ in the vanishing droplet associate with carboxylates on the protein surface, thereby producing a $[M + zNa]^{z+}$ gaseous ion. The charge states and collision cross sections predicted by MD simulations for ubiquitin and other globular proteins agree well with experimental data (Fig 2C). Also, the gaseous proteins retain conformations close to their initial solution structures [26]. Overall, these MD results confirm the view that native ESI produces solution-like gas phase proteins via the charged residue mechanism (CRM) [73,74]. IEM events play an ancillary role during this process by keeping the net charge of the shrinking droplet just below the Rayleigh limit [19,82]. These data [26] explain the experimental observation that protein ions produced by native ESI have charge states close to those of protein-sized aqueous droplets at the Rayleigh limit [73]. This conclusion holds regardless of the initial protein charge state in solution [26].

8. Mobile Proton MD Simulations on Gaseous Proteins

Typical MD force fields employ *static* charges on all titratable side chains and termini. For solution simulations at neutral pH this is quite adequate because pK_a values dictate that Glu, Asp, and C-termini (CT) will be negatively charged, whereas each Arg, Lys, and N-terminus (NT) carries a positive charge. Histidine with its pK_a in the near-neutral range is usually treated as a neutral, although it can also be protonated [83].

The situation is more challenging when modeling the behavior of electrosprayed proteins. In traditional gas phase MD runs the user has to manually select the protonation state of each titratable site such that the overall charge state is consistent with that of the experimentally formed $[M + zH]^{z+}$ ion. Some studies make the simplifying assumption that all carboxylates are neutral

(protonated), and that excess protons can reside only on Arg, Lys, His, and NT [28,33,84]. However, even under such simplified conditions there are countless possibilities to distribute z excess protons on the protein. More importantly, it is incorrect to assume that acidic sites in $[M + zH]^{z+}$ ions are generally neutral. Experiments have shown that gaseous biomolecular systems can contain a large number of $R\text{-}COO^-$ sites that are involved in salt bridges [42,80,85-89]. Another complication is the fact that protons in electrosprayed biomolecular ions are highly mobile, and that the sites of preferred proton occupancy may change as the protein structure evolves [32,80,90-92]. Practitioners wanting to simulate the behavior of electrosprayed proteins thus face two dilemmas: (i) It is unclear how to choose the most appropriate spatial protonation pattern, considering the astronomical number of possible H^+ locations. (ii) Conventional force fields will retain the initially chosen protonation pattern throughout the entire simulation.

We addressed both of these issues by developing a mobile proton MD protocol for gas phase simulations [34,58]. Initial efforts in this direction were described by Thachuk et al. [32], but their approach used a simplified coarse-grained force field and did not allow for negative charges and/or salt bridges. The approach outlined here [34,58] is compatible with any atomistic force field (we chose OPLS/AA), and it allows for protonation changes of all titratable sites ($NT^{0/+}$, $Lys^{0/+}$, $Arg^{0/+}$, $His^{0/+}$, $Asp^{0/-}$, $Glu^{0/-}$, and $CT^{0/-}$. Key element of this mobile proton method is a steepest-descent search protocol that quickly scans thousands of protonation patterns within a $[M + zH]^{z+}$ ion. The algorithm works by minimizing the energy E_{tot} that has an electrostatic term V_{Coul} and a proton affinity contribution E_{PAint} according to [32,34,58,84]

$$E_{tot} = V_{Coul} + E_{PAint} \qquad (1)$$

where

$$E_{PA\,\mathrm{int}} = -\sum_{x=1}^{N} PA_{\mathrm{int}}(x) \times \delta_x \qquad (2)$$

and where $PA_{int}(x)$ represents the intrinsic proton affinities ($PA_{int}(x) > 0$) of the $x = 1 \ldots N$ titratable sites, with $\delta_x = 1$ for protonated sites, and $\delta_x = 0$ for deprotonated sites. Two algorithms for V_{Coul} were implemented; a simple version that only considers interactions among sites with net charge [34], and a more sophisticated version that includes intramolecular solvation by considering electrostatic contributions from all atoms [58]. Similar to the trajectory stitching approach discussed in section 5, long MD runs were broken down into brief segments during which the protein structure was allowed to evolve with fixed charges. After each segment the protons were redistributed using equations 1 and 2, and subsequently the simulation was restarted at the desired temperature.

The results of a mobile proton MD simulation are illustrated in Fig. 3, using 16+ ions of the homo-tetrameric protein avidin at 300 K, i.e., under gentle thermal conditions [58]. The run was started with the X-ray coordinates of the complex. All 16 charges were placed on the same subunit, while all other sites were initially neutral. Evidently, this particular starting pattern is not realistic; it was chosen only to illustrate the performance of the algorithm and the absence of memory effects. Immediately after starting the simulation the algorithm produced protonation patterns where the 16 excess protons were spread roughly evenly across all four subunits. After 200 ns of mobile proton MD numerous additional positive charges had formed, compensated by the presence of deprotonated carboxylates. This proton migration caused an energetically favorable change in V_{Coul}, reflecting newly formed positive/negative charge-charge interactions. At the same time, there were unfavorable changes in E_{PAint} caused by proton transfer from -COOH

groups to sites with lower PA_{int}. These two opposing trends were of similar magnitude (~10^4 kJ mol^{-1}), resulting in E_{tot} values that remained roughly constant throughout the simulations.

In the final avidin MD structures after 200 ns the positively and negatively charged side chains had reoriented such that an intricate salt bridge network was generated at the protein surface (Fig. 3B), in line with recent experimental data on smaller systems [42,80,85-89]. Comparison of the MD structure after 200 ns and the initial X-ray coordinates reveals that gas phase avidin retained most of its structural features, despite some loop rearrangements and alterations in the orientation of surface residues (Fig. 3C). These mobile proton MD data support the basic premise of native ESI-MS experiments, i.e., the view that electrosprayed proteins can retain much of their solution structure [1-4]. We hypothesize that the formation of a salt bridge network at the protein surface (Fig. 3B) assists in the kinetic trapping of these solution-like conformers [58].

9. Collision-Induced Unfolding and Dissociation

In addition to studying "unperturbed" gaseous proteins, it is interesting to probe how electrosprayed ions respond to destabilizing external factors. Collisions with background gas can be implemented on most MS instruments. The resulting analyte heating causes collision-induced unfolding (CIU) and/or collision-induced dissociation (CID) [43,93]. Experiments demonstrated that multi-protein complexes typically undergo CID via ejection of a single subunit with a disproportionally large amount of charge [93-96]. For example, $[M_4 + 14H]^{14+}$ transthyretin tetramers produce M^{7+} or M^{8+} monomers along with the complementary trimers [97]. In this case the monomeric CID products carry roughly ~50% of the total charge but only 25% of the total mass.

A tentative model has emerged to explain the asymmetric charge partitioning of collisionally activated complexes [32,93-96,98]. The model envisions that the complex initially retains a compact structure with excess protons that are spread more or less evenly across the surface. Collisional activation induces unfolding of a single subunit. Subsequently, this subunit is ejected from the complex. It is assumed that unfolding is accompanied by proton migration towards the unravelling subunit. This charge transfer takes place up to the point where the departing subunit separates from the complex [32,93-95,98,99]. The proposed unfolding/charge migration model relies on the high mobility of protons (see section 8), and it is supported by the observation of semi-unfolded intermediates in IMS experiments [98,100]. However, the structures of these intermediates remain nebulous, and there are also alternative proposals that envision completely different mechanistic avenues [101]. MD simulations offer an obvious approach to explore the CID mechanism, but until recently [32] the fixed charge nature of traditional MD force fields has precluded meaningful computational studies in this area.

Mobile proton MD runs of the CID process for transthyretin $[M_4 + 14H]^{14+}$ are illustrated in Fig. 4 [34]. The algorithm used for these runs was similar to that discussed in section 8. However, for the scenario considered here the effects of collisional heating were mimicked by gradually raising the temperature from 350 K to 600 K. Simulations with an explicit collision gas would also have been possible [102]. For the data considered here, the MD runs showed that transthyretin initially retained a compact conformation with roughly the same net charge on each subunit. Subsequently, one of the subunits started to unfold. This process occurred in conjunction with H^+ migration onto the unraveling chain, driven by electrostatic repulsion within the complex. Ultimately, the unfolded chain was ejected as an elongated 8+ ion (Fig. 4A). The fluctuating number of protons on each subunit, and the close correlation between unfolding and charge enrichment of subunit C is evident from Fig. 4B (top two panels). Similar to avidin (section 8),

native-like transthyretin initially formed numerous salt bridges, evident from the large number of negative charges ($R\text{-}COO^-$ moieties) at the onset of the run. Most of these carboxylates gradually disappeared as salt bridges were disrupted by heat-induced structural changes (Fig. 4B, bottom panel). The simulations [34] also revealed unexpected phenomena, such as instances where two chains started to unfold simultaneously and competed for excess charge until one of them underwent ejection (with subsequent collapse of the charge-depleted trimer into a compact structure). The average charge states and overall conformations of the CID products from these simulations closely matched experimental observations, implying that the MD data capture the atomistic details of the CID process quite well. In conclusion, the MD results of Fig. 4 support the unfolding/charge migration model of the CID process [32,93-96,98], In addition, the simulations provide mechanistic insights that are not directly accessible by MS and IMS [34].

9. Conclusions

MD simulations are increasingly being used for investigating charged droplets, electrosprayed proteins, and other gaseous biomolecular species. Limitations associated with current gas phase MD simulations include the following: (i) Commonly used force fields are designed for condensed phase applications; the parameters are not optimized for a gas phase environment [48-50]. (ii) Most MD runs are much shorter than the micro- to millisecond residence times of proteins within ESI droplets and inside typical MS/IMS instruments [13]. (iii) Mobile proton effects are often ignored, or they are treated via exceedingly simple algorithms instead of using proper *ab initio* methods [58]. (iv) While classical MD force fields describe orientational polarization quite well, electronic polarization effects are not considered in detail [12]. (v) The droplet simulations

discussed above employed Na^+ as charge carriers, producing $[M + zNa]^{z+}$ ions rather than protonated species that are of greater analytical relevance [26].

The issues summarized above provide a to-do list for computational chemists. It is hoped that many of these issues will be addressed in the near future through the development of better algorithms and faster computer hardware. How much trust can researchers put into gas phase data generated with *currently existing* MD methods? From the examples discussed above it is evident that the current methods already perform quite well. MD data confirmed that proteins can survive the ESI process without major structural perturbations [1-5]. Charge states and collision cross sections of MD-generated ions were in excellent agreement with experimental data [26]. Simple mobile proton algorithms [58] support experimental evidence that gaseous biomolecules tend to be zwitterionic with a large number of salt bridges [42,80,85-89]. CID simulations on multiprotein complexes correctly reproduce the charge states and conformational features of the dissociation products [34].

The value of any simulation strategy has to be judged by how well its predictions match experimental results, and by the extent to which it provides mechanistic insights that are otherwise inaccessible. Currently available MD techniques already have considerable predictive power for gas phase simulations. Considering the ongoing advances in this area, it seems certain that the role of MD simulations and related computational approaches will continue to expand. It is hoped that this review will encourage a greater number of experimentalists to include MD simulations in their tool box, for providing MS and IMS data with an atomistic structural underpinning.

References

[1] A.C. Leney, A.J.R. Heck, Native Mass Spectrometry: What is in the Name?, J. Am. Soc. Mass Spectrom. 28 (2017) 5-13, https://doi.org/10.1007/s13361-016-1545-3

[2] M.T. Marty, H. Zhang, W.D. Cui, R.E. Blankenship, M.L. Gross, S.G. Sligar, Native Mass Spectrometry Characterization of Intact Nanodisc Lipoprotein Complexes, Anal. Chem. 84 (2012) 8957-8960, https://doi.org/10.1021/ac302663f

[3] J. Marcoux, C.V. Robinson, Twenty Years of Gas Phase Structural Biology, Structure 21 (2013) 1541-1550, https://doi.org/10.1016/j.str.2013.08.002

[4] C.S. Kaddis, J.A. Loo, Native Protein MS and Ion Mobility: Large Flying Proteins with ESI, Anal. Chem. 79 (2007) 1779-1784, https://doi.org/10.1021/ac071878c

[5] G. Bonvin, C.E. Bobst, I.A. Kaltashov, Interaction of transferrin with non-cognate metals studied by native electrospray ionization mass spectrometry, Int. J. Mass Spectrom. 420 (2017) 74-82, https://doi.org/10.1016/j.ijms.2017.01.014

[6] J.M. Daniel, S.D. Friess, S. Rajagopalan, S. Wendt, R. Zenobi, Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry, Int. J. Mass Spectrom. 216 (2002) 1-27, https://doi.org/10.1016/S1387-3806(02)00585-7

[7] A. Dyachenko, R. Gruber, L. Shimon, A. Horovitz, M. Sharon, Allosteric mechanisms can be distinguished using structural mass spectrometry, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 7235-7239, https://doi.org/10.1073/pnas.1302395110

[8] M.F. Bush, Z. Hall, K. Giles, J. Hoyes, C.V. Robinson, B.T. Ruotolo, Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology, Anal. Chem. 82 (2010) 9667-9565, https://doi.org/10.1021/ac1022953

[9] E.N. Kitova, A. El-Hawiet, P.D. Schnier, J.S. Klassen, Reliable Determinations of Protein–Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet?, J. Am. Soc. Mass Spectrom. 23 (2012) 431-441, https://doi.org/10.1007/s13361-011-0311-9

[10] R.J. Rose, E. Damoc, E. Denisov, A. Makarov, A.J.R. Heck, High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies, Nature Methods 9 (2012) 1084-+, https://doi.org/10.1038/nmeth.2208

[11] A.C. Susa, Z.J. Xia, E.R. Williams, Native Mass Spectrometry from Common Buffers with Salts That Mimic the Extracellular Environment, Angew. Chem.-Int. Edit. 56 (2017) 7912-7915, https://doi.org/10.1002/anie.201702330

[12] L. Konermann, E. Ahadi, A.D. Rodriguez, S. Vahidi, Unraveling the Mechanism of Electrospray Ionization, Anal. Chem. 85 (2013) 2-9, https://doi.org/10.1021/ac302789c

[13] P. Kebarle, U.H. Verkerk, Electrospray: From Ions in Solutions to Ions in the Gas Phase, What We Know Now, Mass Spectrom. Rev. 28 (2009) 898-917, https://doi.org/10.1002/mas.20247

[14] N.B. Cech, C.G. Enke, Practical Implication of Some Recent Studies in Electrospray Ionization Fundamentals, Mass Spectrom. Rev. 20 (2001) 362-387, https://doi.org/10.1002/mas.10008

[15] R.L. Grimm, J.L. Beauchamp, Evaporation and Discharge Dynamics of Highly Charged Multicomponent Droplets Generated by Electrospray Ionization, J. Phys. Chem. A 114 (2010) 1411-1419, https://doi.org/10.1021/jp907162w

[16] J.V. Iribarne, B.A. Thomson, On the evaporation of small ions from charged droplets, J. Chem. Phys. 64 (1976) 2287-2294, https://doi.org/10.1063/1.432536

[17] R.B. Cole, Some tenets pertaining to electrospray ionization mass spectrometry, J. Mass. Spectrom. 35 (2000) 763-772, https://doi.org/10.1002/1096-9888(200007)35:7<763::AID-JMS16>3.0.CO;2-%23

[18] I.G. Loscertales, J.F. de la Mora, Experiments on the kinetics of field evaporation of small ions from droplets, J. Chem. Phys. 103 (1995) 5041-5060, https://doi.org/10.1063/1.470591

[19] C.J. Hogan, J.A. Carroll, H.W. Rohrs, P. Biswas, M.L. Gross, Combined Charged Residue-Field Emission Model of Macromolecular Electrospray Ionization, Anal. Chem. 81 (2009) 369-377, https://doi.org/10.1021/ac8016532

[20] S. Nguyen, J.B. Fenn, Gas-phase ions of solute species from charged droplets of solutions, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 1111-1117, https://doi.org/10.1073/pnas.0609969104

[21] D. Kim, N. Wagner, K. Wooding, D.E. Clemmer, D.H. Russell, Ions from Solution to the Gas Phase: A Molecular Dynamics Simulation of the Structural Evolution of Substance P during Desolvation of Charged Nanodroplets Generated by Electrospray Ionization, J. Am. Chem. Soc. 139 (2017) 2981-2988, https://doi.org/10.1021/jacs.6b10731

[22] M.I. Oh, S. Consta, What factors determine the stability of a weak protein-protein interaction in a charged aqueous droplet?, Phys. Chem. Chem. Phys. 19 (2017) 31965-31981, https://doi.org/10.1039/C7CP05043G

[23] M. Porrini, F. Rosu, C. Rabin, L. Darre, H. Gomez, M. Orozco, V. Gabelica, Compaction of Duplex Nucleic Acids upon Native Electrospray Mass Spectrometry, ACS Central Sci. 3 (2017) 454–461, https://doi.org/10.1021/acscentsci.7b00084

[24] V. Znamenskiy, I. Marginean, A. Vertes, Solvated Ion Evaporation from Charged Water Droplets, J. Phys. Chem. A 107 (2003) 7406-7412, https://doi.org/10.1021/jp034561z

[25] C. Caleman, D. van der Spoel, Temperature and structural changes of water clusters in vacuum due to evaporation, J. Chem. Phys. 125 (2006) 1545081-1545089, https://doi.org/10.1063/1.2357591

[26] R.G. McAllister, H. Metwally, Y. Sun, L. Konermann, Release of Native-Like Gaseous Proteins from Electrospray Droplets via The Charged Residue Mechanism: Insights from Molecular Dynamics Simulations, J. Am. Chem. Soc. 137 (2015) 12667-12676, https://doi.org/10.1021/jacs.5b07913

[27] H. Higashi, T. Tokumi, C.J. Hogan, H. Suda, T. Seto, Y. Otani, Simultaneous ion and neutral evaporation in aqueous nanodrops: experiment, theory, and molecular dynamics simulations, Phys. Chem. Chem. Phys. 17 (2015) 15746-15755, https://doi.org/10.1039/C5CP01730K

[28] Y. Mao, M.A. Ratner, M.F. Jarrold, Molecular Dynamics Simulations of the Charge-Induced Unfolding and Refolding of Unsolvated Cytochrome c, J. Phys. Chem. B 103 (1999) 10017-10021, https://doi.org/10.1021/jp991093d

[29] A. Baumketner, S.L. Bernstein, T. Wyttenbach, G. Bitan, D.B. Teplow, M.T. Bowers, J.E. Shea, Amyloid β-protein monomer structure: a computational and experimental study, Protein Sci. 15 (2006) 420-428, https://doi.org/10.1110/ps.051762406

[30] A. Patriksson, C.M. Adams, F. Kjeldsen, R.A. Zubarev, D. van der Spoel, A direct comparison of protein structure in the gas and solution phase: The TRP-cage, J. Phys. Chem. B 111 (2007) 13147-13150, https://doi.org/10.1021/jp709901t

[31] M.Z. Steinberg, R. Elber, F.W. McLafferty, R.B. Gerber, K. Breuker, Early Structural Evolution of Native Cytochrome c After Solvent Removal, ChemBioChem 9 (2008) 2417-2423, https://doi.org/10.1002%2Fcbic.200800167

[32] S.K. Fegan, M. Thachuk, A Charge Moving Algorithm for Molecular Dynamics Simulations of Gas-Phase Proteins, J. Chem. Theory Comput. 9 (2013) 2531-2539, https://doi.org/10.1021/ct300906a

[33] Z. Hall, A. Politis, M.F. Bush, L.J. Smith, C.V. Robinson, Charge-State Dependent Compaction and Dissociation of Protein Complexes: Insights from Ion Mobility and Molecular Dynamics, J. Am. Chem. Soc. 134 (2012) 3429-3438, https://doi.org/10.1021/ja2096859

[34] V. Popa, D.A. Trecroce, R.G. McAllister, L. Konermann, Collision-Induced Dissociation of Electrosprayed Protein Complexes: An All-Atom Molecular Dynamics Model with Mobile Protons, J. Phys. Chem. B 120 (2016) 5114-5124, https://doi.org/10.1021/acs.jpcb.6b03035

[35] S.G. Kondalaji, M. Khakinejad, A. Tafreshian, S.J. Valentine, Comprehensive Peptide Ion Structure Studies Using Ion Mobility Techniques: Part 1. An Advanced Protocol for

Molecular Dynamics Simulations and Collision Cross-Section Calculation, J. Am. Soc. Mass Spectrom. 28 (2017) 947-959, https://doi.org/10.1007/s13361-017-1599-x

[36] T. Ly, R.R. Julian, Elucidating the Tertiary Structure of Protein Ions in Vacuo with Site Specific Photoinitiated Radical Reactions, J. Am. Chem. Soc. 132 (2010) 8602-8609, https://doi.org/10.1021/ja910665d

[37] T.J. El-Baba, D.W. Woodall, S.A. Raab, D.R. Fuller, A. Laganowsky, D.H. Russell, D.E. Clemmer, Melting Proteins: Evidence for Multiple Stable Structures upon Thermal Denaturation of Native Ubiquitin from Ion Mobility Spectrometry-Mass Spectrometry Measurements, J. Am. Chem. Soc. 139 (2017) 6306-6309, https://doi.org/10.1021/jacs.7b02774

[38] T. Wyttenbach, C. Bleiholder, M.T. Bowers, Factors contributing to the collision cross section of polyatomic ions in the kilodalton to gigadalton range: application to ion mobility measurements, Anal. Chem. 85 (2013) 2191-2199, https://doi.org/10.1021/ac3029008

[39] E.G. Marklund, M.T. Degiacomi, C.V. Robinson, A.J. Baldwin, J.L.P. Benesch, Collision Cross Sections for Structural Proteomics, Structure 23 (2015) 791-799, https://doi.org/10.1016/j.str.2015.02.010

[40] F.O. Talbot, A. Rullo, H. Yao, R.A. Jockusch, Fluorescence Resonance Energy Transfer in Gaseous, Mass-Selected Polyproline Peptides, J. Am. Chem. Soc. 132 (2010) 16156-16164, https://doi.org/10.1021/ja1067405

[41] A.I.G. Florez, E. Mucha, D.S. Ahn, S. Gewinner, W. Schollkopf, K. Pagel, G. von Helden, Charge-Induced Unzipping of Isolated Proteins to a Defined Secondary Structure, Angew. Chem.-Int. Edit. 55 (2016) 3295-3299, https://doi.org/10.1002/anie.201510983

[42] J.G. Bonner, Y.A. Lyon, C. Nellessen, R.R. Julian, Photoelectron Transfer Dissociation Reveals Surprising Favorability of Zwitterionic States in Large Gaseous Peptides and Proteins, J. Am. Chem. Soc. 139 (2017) 10286-10293, https://doi.org/10.1021/jacs.7b02428

[43] Y. Zhong, L. Han, B.T. Ruotolo, Collisional and coulombic unfolding of gas-phase proteins: high correlation to their domain structures in solution, Angew. Chem. 53 (2014) 9209-12, https://doi.org/10.1002/anie.201403784

[44] H. van den Bedem, J.S. Fraser, Integrative, dynamic structural biology at atomic resolution-it's about time, Nature Methods 12 (2015) 307-318, https://doi.org/10.1038/nmeth.3324

[45] X.-C. Bai, G. McMullan, S.H.W. Scheres, How cryo-EM is revolutionizing structural biology, Trends Biochem. Sci. 40 (2015) 49-57, https://doi.org/10.1016/j.tibs.2014.10.005

[46] E. Lindahl. Molecular Dynamics Simulations. In *Molecular Modeling of Proteins. Methods in Molecular Biology (Methods and Protocols);* (ed), K. A. Ed.; Humana Press: New York, NY, 2015; Vol. 1215.

[47] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19-25, https://doi.org/10.1016/j.softx.2015.06.001

[48] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, J. Phys. Chem. B 105 (2001) 6474-6487, https://doi.org/10.1021/jp003919d

[49] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, D.E. Shaw, Improved side-chain torsion potentials for the Amber ff99SB protein force field, Proteins 78 (2010) 1950-1958, https://doi.org/10.1002/prot.22711

[50] J. Huang, A.D. MacKerell, CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data, J. Comput. Chem. 34 (2013) 2135-2145, https://doi.org/10.1002/jcc.23354

[51] M.M. Sultan, R.A. Denny, R. Unwalla, F. Lovering, V.S. Pande, Millisecond dynamics of BTK reveal kinome-wide conformational plasticity within the apo kinase domain, Scientific Reports 7 (2017) https://doi.org/10.1038/s41598-017-10697-0

[52] S. Piana, K. Lindorff-Larsen, D.E. Shaw, Atomic-level description of ubiquitin folding, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 5915-5920, https://doi.org/10.1073/pnas.1218321110

[53] W.L. Jorgensen, D.S. Maxwell, J. TiradoRives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, J. Am. Chem. Soc. 118 (1996) 11225-11236, https://doi.org/10.1021/ja9621760

[54] J.W. Ponder, D.A. Case. Force Fields for Protein Simulations. In *Advances in Protein Chemistry*; Elsevier, 2003; Vol. 66; pp. 27-84.

[55] G.A. Kaminski, H.A. Stern, B.J. Berne, R.A. Friesner, Y.X.X. Cao, R.B. Murphy, R.H. Zhou, T.A. Halgren, Development of a polarizable force field for proteins via ab initio quantum chemistry: First generation model and gas phase tests, J. Comput. Chem. 23 (2002) 1515-1531, https://doi.org/10.1002/jcc.10125

[56] J.A. Lemkul, J. Huang, B. Roux, A.D. MacKerell, An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications, Chem. Rev. 116 (2016) 4983-5013, https://doi.org/10.1021/acs.chemrev.5b00505

[57] M.D. Beachy, D. Chasman, R.B. Murphy, T.A. Halgren, R.A. Friesner, Accurate ab initio quantum chemical determination of the relative energetics of peptide conformations and assessment of empirical force fields, J. Am. Chem. Soc. 119 (1997) 5908-5920, https://doi.org/10.1021/ja962310g

[58] L. Konermann, Molecular Dynamics Simulations on Gas-Phase Proteins with Mobile Protons: Inclusion of All-Atom Charge Solvation, J. Phys. Chem. B 121 (2017) 8102-8112, https://doi.org/10.1021/acs.jpcb.7b05703

[59] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A Smooth Particle Mesh Ewald Method, J. Chem. Phys. 103 (1995) 8577-8593, https://doi.org/10.1063/1.470117

[60] G.A. Cisneros, M. Karttunen, P.Y. Ren, C. Sagui, Classical Electrostatics for Biomolecular Simulations, Chem. Rev. 114 (2014) 779-814, https://doi.org/10.1021/cr300461d

[61] A. Kukol. *Molecular Modeling of Proteins*; Humana Press, 2015.

[62] J.S. Hub, B.L. de Groot, H. Grubmuller, G. Groenhof, Quantifying Artifacts in Ewald Simulations of Inhomogeneous Systems with a Net Charge, J. Chem. Theory Comput. 10 (2014) 381-390, https://doi.org/10.1021/ct400626b

[63] C. Kutzner, S. Pall, M. Fechner, A. Esztermann, B.L. de Groot, H. Grubmuller, Best bang for your buck: GPU nodes for GROMACS biomolecular simulations, J. Comput. Chem. 36 (2015) 1990-2008, https://doi.org/10.1002/jcc.24030

[64] S. Pall, B. Hess, A flexible algorithm for calculating pair interactions on SIMD architectures, Comput. Phys. Commun. 184 (2013) 2641-2650, https://doi.org/10.1016/j.cpc.2013.06.003

[65] S.C. Gibson, C.S. Feigerle, K.D. Cook, Fluorometric Measurement and Modeling of Droplet Temperature Changes in an Electrospray Plume, Anal. Chem. 86 (2014) 464-472, https://doi.org/10.1021/ac402364g

[66] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007) 0141011-0141017, https://doi.org/10.1063/1.2408420

[67] H.J.C. Berendsen, J.P.M. Postma, W.F. Vangunsteren, A. Dinola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684-3690, https://doi.org/10.1063/1.448118

[68] W.G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Phys. Rev. A 31 (1985) 1695-1697, https://doi.org/10.1103/PhysRevA.31.1695

[69] L. Konermann, R.G. McAllister, H. Metwally, Molecular Dynamics Simulations of the Electrospray Process: Formation of NaCl Clusters via the Charged Residue Mechanism, J. Phys. Chem. B 118 (2014) 12025-12033, https://doi.org/10.1021/jp507635y

[70] A. Mor, G. Ziv, Y. Levy, Simulations of proteins with inhomogeneous degrees of freedom: The effect of thermostats, J. Comput. Chem. 29 (2008) 1992-1998, https://doi.org/10.1002/jcc.20951

[71] H.C. Andersen, Molecular dynamics simulations at constant pressure and/or temperature, J. Chem. Phys. 72 (1980) 2384-2393, https://doi.org/10.1063/1.439486

[72] H. Metwally, R.G. McAllister, V. Popa, L. Konermann, Mechanism of Protein Supercharging by Sulfolane and m-Nitrobenzyl Alcohol: Molecular Dynamics Simulations of the Electrospray Process, Anal. Chem. 88 (2016) 5345-5354, https://doi.org/10.1021/acs.analchem.6b00650

[73] F.J. de la Mora, Electrospray Ionization of large multiply charged species proceeds via Dole's charged residue mechanism, Anal. Chim. Acta 406 (2000) 93-104, https://doi.org/10.1016/S0003-2670(99)00601-7

[74] A.T. Iavarone, E.R. Williams, Mechanism of Charging and Supercharging Molecules in Electrospray Ionization, J. Am. Chem. Soc. 125 (2003) 2319-2327, https://doi.org/10.1021/ja021202t

[75] C. Vega, E. de Miguel, Surface tension of the most popular models of water by using the test-area simulation method, J. Chem. Phys. 126 (2007) 154707, https://doi.org/10.1063/1.2715577

[76] J.L.F. Abascal, C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005, J. Chem. Phys. 123 (2005) 234505, https://doi.org/10.1063/1.2121687

[77] S. Cukierman, Et tu, Grotthuss! and other unfinished stories, Biochim. Biophys. Acta 1757 (2006) 876-885, https://doi.org/10.1016/j.bbabio.2005.12.001

[78] M. Chen, H.Y. Ko, R.C. Remsing, M.F.C. Andrade, B. Santra, Z.R. Sun, A. Selloni, R. Car, M.L. Klein, J.P. Perdew, X.F. Wu, Ab initio theory and modeling of water, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) 10846-10851, https://doi.org/10.1073/pnas.1712499114

[79] H. Nakai, A.W. Sakti, Y. Nishimura, Divide-and-Conquer-Type Density-Functional Tight-Binding Molecular Dynamics Simulations of Proton Diffusion in a Bulk Water System, J. Phys. Chem. B 120 (2016) 217-221, https://doi.org/10.1021/acs.jpcb.5b12439

[80] J.Y. Li, W.P. Lyu, G. Rossetti, A. Konijnenberg, A. Natalello, E. Ippoliti, M. Orozco, F. Sobott, R. Grandori, P. Carloni, Proton Dynamics in Protein Mass Spectrometry, J. Phys. Chem. Lett. 8 (2017) 1105-1112, https://doi.org/10.1021/acs.jpclett.7b00127

[81] S.C. Harvey, R.K.Z. Tan, T.E. Cheatham, The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition, J. Comput. Chem. 19 (1998) 726-740, https://doi.org/10.1002/(sici)1096-987x(199805)19:7<726::aid-jcc4>3.0.co;2-s

[82] S.J. Allen, A.M. Schwartz, M.F. Bush, Effects of Polarity on the Structures and Charge States of Native-Like Proteins and Protein Complexes in the Gas Phase, Anal. Chem. 85 (2013) 12055−12061, https://doi.org/10.1021/ac403139d

[83] T.E. Creighton. *Proteins*; W. H. Freeman & Co: New York, 1993.

[84] P.D. Schnier, D.S. Gross, E.R. Williams, Electrostatic Forces and Dielectric Polarizability of Multiply Protonated Gas-Phase Cytochrome c Ions Probed by Ion/Molecule Chemistry, J. Am. Chem. Soc. 117 (1995) 6747-6757, https://doi.org/10.1021/ja00130a015

[85] M.W. Forbes, M.F. Bush, N.C. Polfer, J. Oomens, R.C. Dunbar, E.R. Williams, R.A. Jockusch, Infrared spectroscopy of arginine cation complexes: Direct observation of gas-phase zwitterions, J. Phys. Chem. A 111 (2007) 11759-11770, https://doi.org/10.1021/jp074859f

[86] Z. Zhang, S.J. Browne, R.W. Vachet, Exploring Salt Bridge Structures of Gas-Phase Protein Ions using Multiple Stages of Electron Transfer and Collision Induced Dissociation, J. Am. Soc. Mass Spectrom. 25 (2014) 604-613, https://doi.org/10.1007/s13361-013-0821-8

[87] K. Breuker, S. Brüschweiler, M. Tollinger, Electrostatic Stabilization of a Native Protein Structure in the Gas Phase, Angew. Chem. Int. Ed. 50 (2011) 873-877, https://doi.org/10.1002/anie.201005112

[88] J. Seo, S. Warnke, K. Pagel, M.T. Bowers, G. von Helden, Infrared spectrum and structure of the homochiral serine octamer-dichloride complex, Nat. Chem. 9 (2017) 1263-1268, https://doi.org/10.1038/nchem.2821

[89] H.J. Yoo, N. Wang, S.Y. Zhuang, H.T. Song, K. Hakansson, Negative-Ion Electron Capture Dissociation: Radical-Driven Fragmentation of Charge-Increased Gaseous Peptide Anions, J. Am. Chem. Soc. 133 (2011) 16790-16793, https://doi.org/10.1021/ja207736y

[90] A.R. Dongré, J.L. Jones, Á. Somogyi, V.H. Wysocki, Influence of Peptide Composition, Gas-Phase Basicity, and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model, J. Am. Chem. Soc. 118 (1996) 8365-8374, https://doi.org/10.1021/ja9542193

[91] R.K. Boyd, Á. Somogyi, The Mobile Proton Hypothesis in Fragmentation of Protonated Peptides: A Perspective, J. Am. Soc. Mass Spectrom. 21 (2010) 1275-1278, https://doi.org/10.1016/j.jasms.2010.04.017

[92] J. Cautereels, F. Blockhuys, Quantum Chemical Mass Spectrometry: Verification and Extension of the Mobile Proton Model for Histidine, J. Am. Soc. Mass Spectrom. 28 (2017) 1227-1235, https://doi.org/10.1007/s13361-017-1636-9

[93] J.L.P. Benesch, B.T. Ruotolo, D.A. Simmons, C.V. Robinson, Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics, Chem. Rev. 107 (2007) 3544-3567, https://doi.org/10.1021/cr068289b

[94] K.J. Light-Wahl, B.L. Schwartz, R.D. Smith, Observation of the Noncovalent Quaternary Association of Proteins by Electrospray Ionization Mass Spectrometry, J. Am. Chem. Soc. 116 (1994) 5271-5278, https://doi.org/10.1021/ja00091a035

[95] J.C. Jurchen, E.R. Williams, Origin of Asymmetric Charge Partitioning in the Dissociation of Gas-Phase Protein Homodimers, J. Am. Chem. Soc. 125 (2003) 2817-2826, https://doi.org/10.1021/ja0211508

[96] N. Felitsyn, E.N. Kitova, J.S. Klassen, Thermal Decomposition of a Gaseous Multiprotein Complex Studied by Blackbody Infrared Radiative Dissociation. Investigating the Origin of the Asymmetric Dissociation Behavior, Anal. Chem. 73 (2001) 4647-4661, https://doi.org/10.1021/ac0103975

[97] L. Han, S.-J. Hyung, B.T. Ruotolo, Bound Cations Significantly Stabilize the Structure of Multiprotein Complexes in the Gas Phase, Angew. Chem. Int. Ed. 51 (2012) 5692-5695, https://doi.org/10.1002/anie.201109127

[98] R.S. Quintyn, M. Zhou, J. Yan, V.H. Wysocki, Surface-Induced Dissociation Mass Spectra as a Tool for Distinguishing Different Structural Forms of Gas-Phase Multimeric Protein Complexes, Anal. Chem. 87 (2015) 11879-11886, https://doi.org/10.1021/acs.analchem.5b03441

[99] S.V. Sciuto, J. Liu, L. Konermann, An Electrostatic Charge Partitioning Model for the Dissociation of Protein Complexes in the Gas Phase, J. Am. Soc. Mass Spectrom. 22 (2011) 1679-1689, https://doi.org/10.1007/s13361-011-0205-x

[100] B.T. Ruotolo, S.-J. Hyung, P.M. Robinson, K. Giles, R.H. Bateman, C.V. Robinson, Ion Mobility–Mass Spectrometry Reveals Long-Lived, Unfolded Intermediates in the Dissociation of Protein Complexes, Angew. Chem. Int. Ed. 46 (2007) 8001-8004, https://doi.org/10.1002/anie.200702161

[101] R.R. Ogorzalek Loo, J.A. Loo, Salt Bridge Rearrangement (SaBRe) Explains the Dissociation Behavior of Noncovalent Complexes, J. Am. Soc. Mass Spectrom. 27 (2016) 975-990, https://doi.org/10.1007/s13361-016-1375-3

[102] C.D. Daub, N.M. Cann, How Are Completely Desolvated Ions Produced in Electrospray Ionization: Insights from Molecular Dynamics Simulations, Anal. Chem. 83 (2011) 8372-8376, https://doi.org/10.1021/ac202103p

Fig. Captions

Fig. 1. Proper and improper use of PBC and PME for MD simulations. (A) Typical setup for solution runs, with a protein **P** and solvent molecules **s**. The simulation box is surrounded by identical images of itself. (B) Improper use of PBC/PME for gas phase simulations on a protein ion $\mathbf{P}^{z+}$. (C) Proper setup for gas phase simulations, where $\mathbf{P}^{z+}$ is treated without PBC, without PME, and without cutoffs. (D) MD data for two Na^+ ions that were initially placed 1 nm apart in vacuum. Under proper conditions (blue line, as in panel C) the Na^+ move away from each other. Under improper conditions (red line, as in panel B, PBC box width 13 nm, PME cutoff 1 nm) repulsion among PBC images causes the Na^+ to oscillate. (E) Proper "pseudo-PBC" setup for gas phase simulations using Gromacs with GPU acceleration without PME. Coulomb and Lennard-Jones cutoffs (dotted circles) are chosen such that all atoms within the protein interact with each other, while excluding image interactions. MD data produced under these conditions are included in panel D as open circles.

Fig. 2. MD Simulation data, describing the formation of gaseous ubiquitin ions from an aqueous nanodroplet under native ESI conditions. (A) Protein release via droplet evaporation to dryness. Blue spheres represent Na^+. "IEM" at $t = 0.73$ exemplifies a charge carrier ejection event. (B) Details of charge carrier ejection, i.e., number of Na^+ in the droplet and relative droplet charge vs. time. (C) Comparison of experimental vs. MD-generated charge states and collision cross sections (Ω) for ubiquitin (ubq) cytochrome *c* (cyt *c*), and holo-myoglobin (hMb). Reproduced with permission from ref. [26]. Copyright 2015, American Chemical Society.

Fig. 3. Mobile proton results for avidin $[M_4 + 16H]^{16+}$ in the gas phase. The four subunits are

depicted in different colors. (A) Starting conformation with an unrealistic initial charge pattern, where all 16 excess H^+ were placed on side chains (blue) within the same subunit. After 200 ns of mobile proton MD, charge was distributed more or less evenly over all subunits, and abundant salt bridges had formed. Positively and negatively charged residues are highlighted in blue and red, respectively (B) Close-up of a salt bridge cluster in one of the subunits at t = 200 ns. (C) Comparison of X-ray coordinates and MD structure at t = 200 ns, revealing that the gas phase protein retained a native-like structure. Reproduced with permission from ref. [58]. Copyright 2017, American Chemical Society.

Fig. 4. Collision-induced dissociation of a gaseous transthyretin $[M_4 + 14H]^{14+}$ tetramer studied by mobile proton MD simulations. Both panels show data for the same simulation run. (A) Structural snapshots, illustrating different time points during the dissociation process. Also indicated is the net charge on each subunit for the various time points. (B) Temporal profiles for various parameters, from top to bottom: net charge on each subunit, radius of gyration (R_g), and total number of positive and negative charges in the complex. The dotted vertical line at t = 15.13 ns indicates the point where the departing chain separates from the rest of the complex. Reproduced with permission from ref. [34]. Copyright 2016, American Chemical Society.

Fig. 1

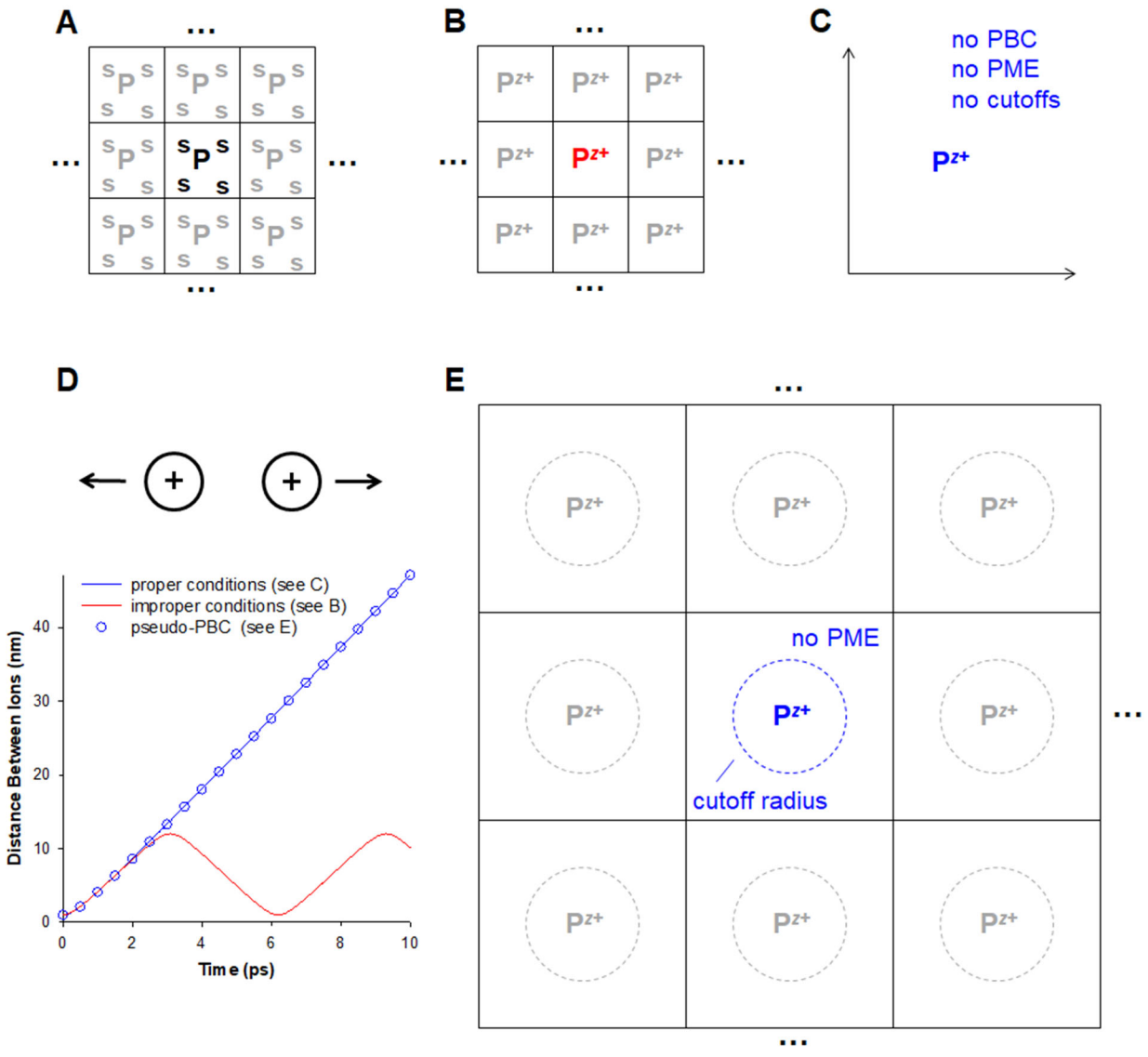

A
B
C
no PBC
no PME
no cutoffs
D
proper conditions (see C)
improper conditions (see B)
pseudo-PBC (see E)
Distance Between Ions (nm)
Time (ps)
E
no PME
cutoff radius

Fig. 2

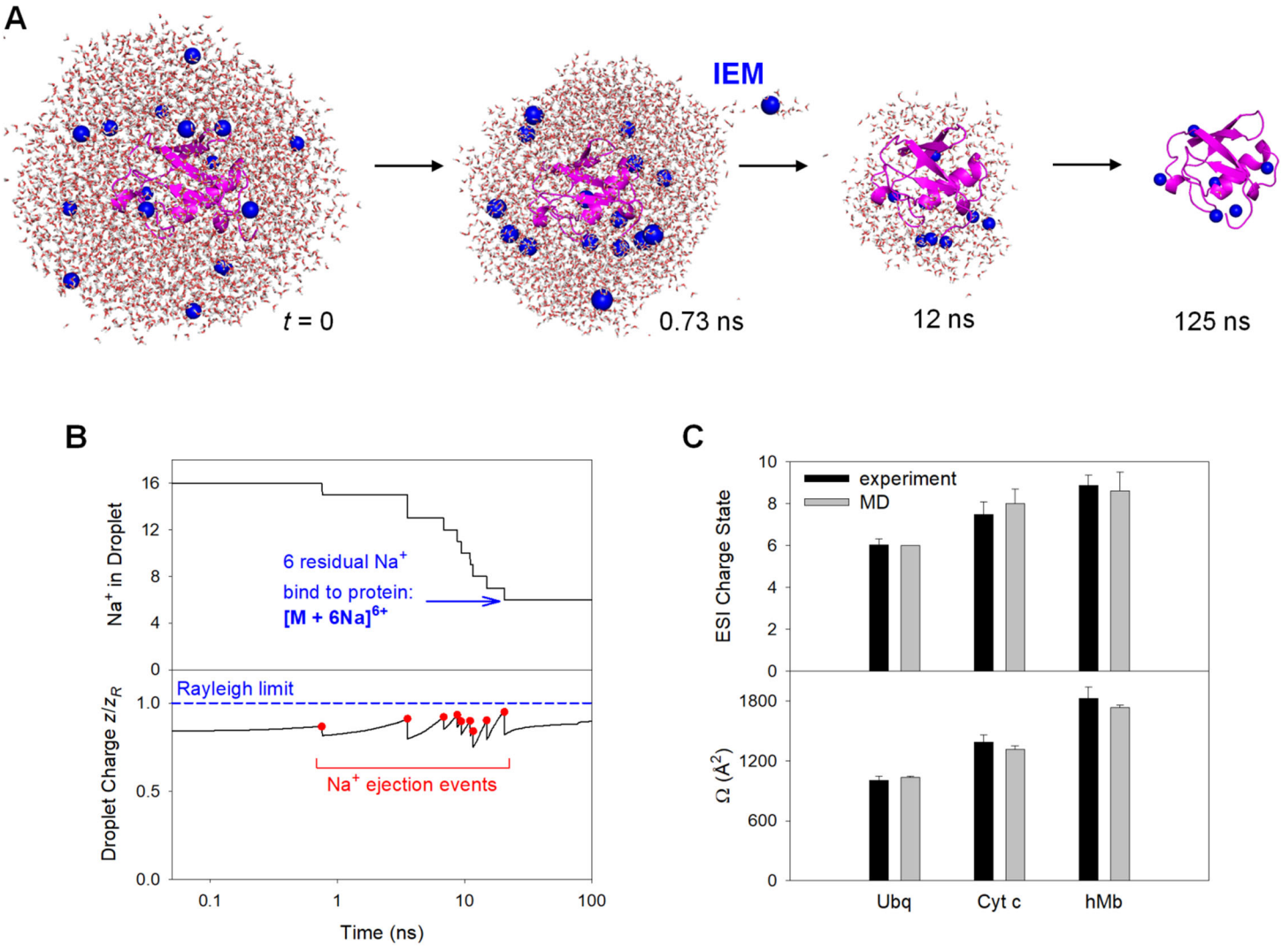

A
IEM
t = 0
0.73 ns
12 ns
125 ns
B
Na+ in Droplet
6 residual Na+
bind to protein:
[M + 6Na]6+
Rayleigh limit
Droplet Charge z/zR
Na+ ejection events
Time (ns)
C
ESI Charge State
experiment
MD
Ω (Å2)
Ubq
Cyt c
hMb

Fig. 3

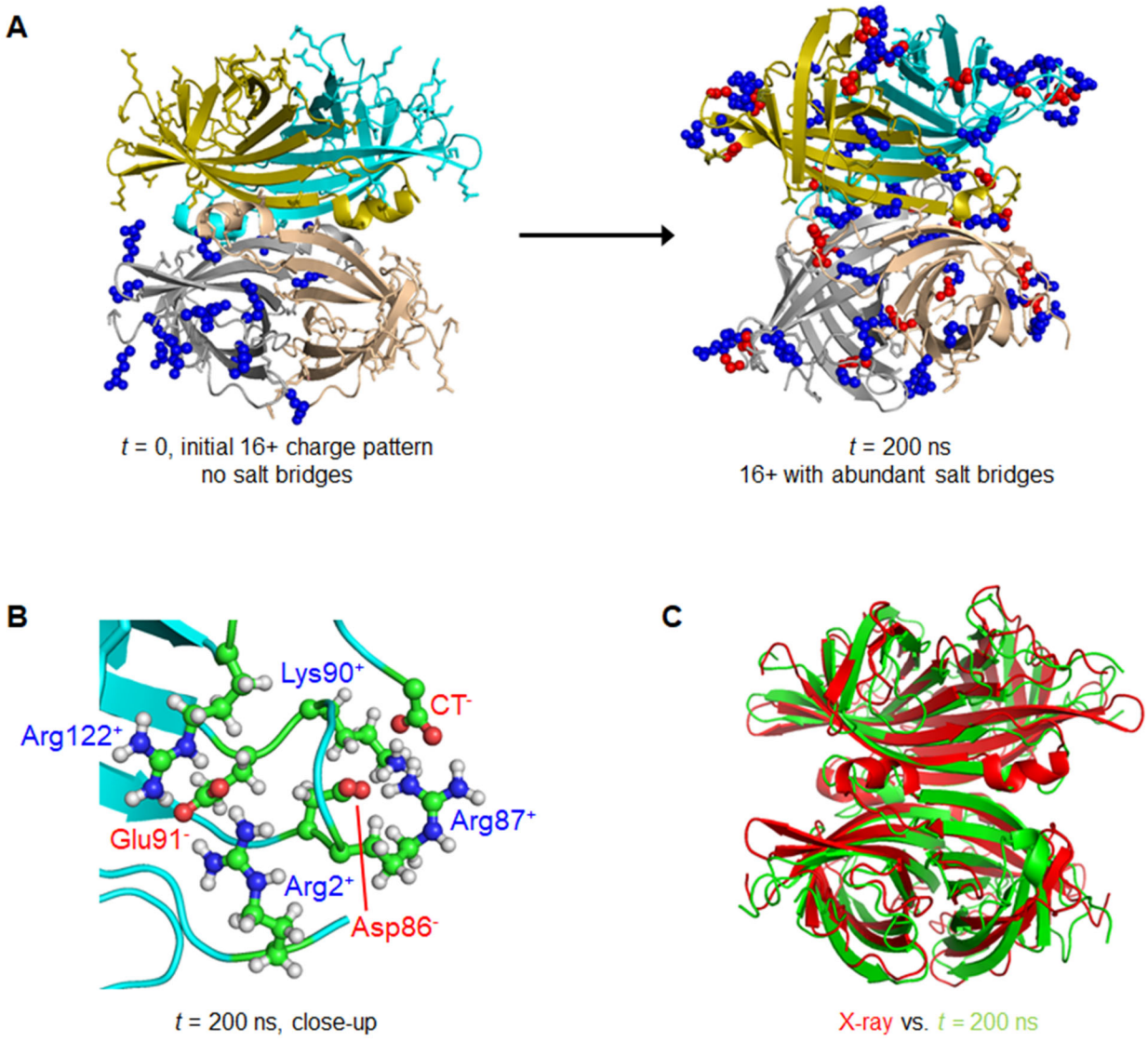
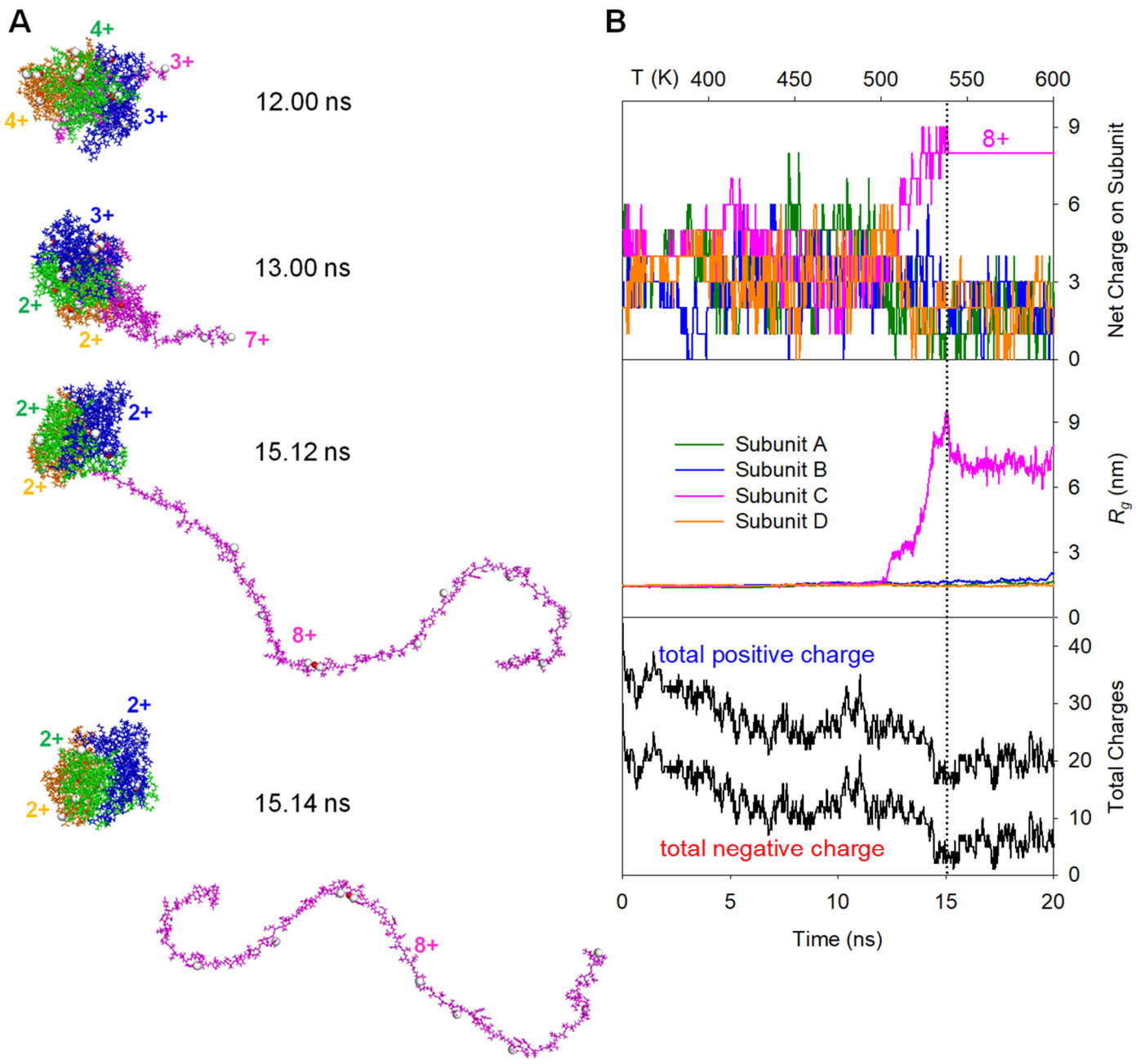

A
t = 0, initial 16+ charge pattern
no salt bridges
t = 200 ns
16+ with abundant salt bridges
B
Lys90+
CT-
Arg122+
Arg87+
Glu91-
Arg2+
Asp86-
t = 200 ns, close-up
C
X-ray vs. t = 200 ns

Fig. 4


A
4+
3+
4+
3+
12.00 ns
3+
2+
2+
7+
13.00 ns
2+
2+
2+
15.12 ns
8+
2+
2+
2+
15.14 ns
8+
B
T (K) 400 450 500 550 600
8+
Net Charge on Subunit
Subunit A
Subunit B
Subunit C
Subunit D
R_g (nm)
total positive charge
total negative charge
Total Charges
0 5 10 15 20
Time (ns)