

Electronic Thesis and Dissertation Repository

---

9-18-2018 10:00 AM

# Mechanism of Protein Charging and Supercharging in Electrospray Ionization: Molecular Dynamics Simulations and Experimental Investigations

Haidy S. Metwally, *The University of Western Ontario*

Supervisor: Dr. Lars Konermann, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Chemistry

© Haidy S. Metwally 2018

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

 Part of the [Analytical Chemistry Commons](#), and the [Physical Chemistry Commons](#)

---

## Recommended Citation

Metwally, Haidy S., "Mechanism of Protein Charging and Supercharging in Electrospray Ionization: Molecular Dynamics Simulations and Experimental Investigations" (2018). *Electronic Thesis and Dissertation Repository*. 5725.  
<https://ir.lib.uwo.ca/etd/5725>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

## Abstract

Electrospray ionization mass spectrometry (ESI-MS) is a powerful technique for investigating protein structures, conformations, and interactions. Despite its widespread use, many fundamental aspects of ESI remain poorly understood. In this thesis, we use a combination of molecular dynamics (MD) simulations and experiments to gain insights into the hidden complexities of ESI-MS.

Chapter 2 discusses the topic of salt-induced protein signal degradation. Salts such as NaCl, CsCl, and tetrabutylammonium chloride (NBu<sub>4</sub>Cl) interfere with MS data acquisition, leading to adduct formation and signal suppression. MD simulations provide an explanation for these salt interferences. Signal suppression can be broken down into two effects, i.e., i) peak splitting due to adduction, ii) “genuine” signal suppression. The results obtained may be helpful to anticipate solution conditions for improved protein analyses by ESI-MS.

The two subsequent Chapters examine the mechanism of native protein supercharging, which represents a highly contentious topic. Chapter 3 uses MD simulations along with ion mobility mass spectrometry (IMS/MS). Holo-myoglobin (hMb) serves as a model protein, along with the two most common supercharging agents (SCAs), sulfolane and *m*-nitrobenzyl alcohol (*m*-NBA). Our data show that supercharging is caused by ‘charge trapping’ that arises from solvent segregation in the droplets, resulting in the formation of SCA-enriched surface layer and an aqueous core. The key factor to charge trapping is the differential solubility of charge carriers (such as Na<sup>+</sup> or NH<sub>4</sub><sup>+</sup>) in water compared to the exterior SCA layer. After complete water evaporation, residual SCA molecules impede charge carrier release from the droplet, and any remaining charge carriers will bind to the protein. Slow SCA evaporation eventually releases a highly charged protein into the gas phase that may undergo Coloumbic unfolding. These findings represent the first atomistic view of protein supercharging.

In Chapter 4, we explore the mechanism of native protein supercharging from a different perspective using a crown ether (18C6). 18C6 selectively binds  $\text{Na}^+/\text{NH}_4^+$  and enhances their solubility in the SCA layer. This facilitates the release of 18C6-bound charge carriers from the droplet. As a result, 18C6 suppressed supercharging effect, as confirmed both in MD simulations and experimentally. These data support the proposed charge trapping mechanism for both proteins and dendrimers.

A chain ejection model (CEM) has been proposed to account for the protein ESI behavior under such non-native conditions. The CEM envisions that unfolded proteins are driven to the droplet surface by hydrophobic and electrostatic factors, followed by gradual ejection via intermediates where droplets carry extended protein tails. Thus far it has not been possible to support the CEM through MD simulations. In Chapter 5 we overcome these difficulties and use MD simulations along with ion mobility experiments to confirm CEM as an ejection mechanism for unfolded proteins. Overall, the modeling and experimental work in this thesis provide unprecedented insights into the mechanism of protein charging and supercharging during ESI.

**Keywords:** electrospray ionization mass spectrometry | molecular dynamics (MD) simulations | supercharging | charging | protein ion.

## Co-Authorship Statement

The work in Chapters 2, 3, 4 and 5 was published in the following articles:

Metwally, H., McAllister, R.G., & Konermann, L. (2015) Exploring the Mechanism of Salt-Induced Signal Suppression in Protein Electrospray Mass Spectrometry Using Experiments and Molecular Dynamics Simulations. *Anal. Chem.* 87:2434-2442. Reproduced with permission. ©American Chemical Society.

Metwally, H., McAllister, R.G., Popa, V., & Konermann, L. (2016) Mechanism of Protein Supercharging by Sulfolane and *m*-NBA: Molecular Dynamics Simulations of the Electrospray Process. *Anal. Chem.* 88:5345-5354. Reproduced with permission. ©American Chemical Society.

Metwally, H., & Konermann, L. (2018) Crown Ethers Modulate the Location of Charge Carriers in Electrospray Droplets: Implications for the Mechanism of Protein Charging and Supercharging *Anal. Chem.* 90:4126-4134. Reproduced with permission. ©American Chemical Society.

Metwally, H., Duez, Q., & Konermann, L. (2018) Chain Ejection Model for Electrospray Ionization of Unfolded Proteins: Evidence from Atomistic Simulations and Ion Mobility Spectrometry. *Anal. Chem.* 90:10069-10077. Reproduced with permission. ©American Chemical Society.

The original draft for each of the above articles was prepared by the author (H.M). Subsequent revisions were done by the author and Dr. Lars Konermann. MD simulations in Chapter 2 were performed and analyzed with the help of Mr. Robert McAllister. MD simulations in Chapter 3 were done by the author with the help of Mr. Robert McAllister. Mr. Vlad Popa participated with the author in developing a script used in the analysis of the radius of gyration in Chapter 3. The setup for Collidoscope in Chapter 5 was done by the author with the help of Quentin Duez. All experimental work was performed and analyzed by the author under the supervision of Dr. Lars Konermann. Fortran codes for trajectory stitching, forced evaporation and building up the droplets used in Chapters 3 and 4 was written by Dr. Konermann.

## Acknowledgments

First of all, I'm grateful to Allah for the successful completion of this degree.

The work in this thesis would not have been possible without the continuous encouragement of many wonderful and influential people in my life. Starting with my advisor Prof. Lars Konermann who is an amazing mentor. Dr. Konermann has taught me so much not only about mass spectrometry but also how to think critically about my work and how valuable it is to communicate my results effectively to my peers. I will always be grateful to Dr. Konermann for accepting me in the lab and dedicating a lot of his time and ideas to make my work successful and meaningful. Being a mom is not an easy task along with doing graduate studies, but Lars has been always supportive and helpful. I really enjoyed my time in the Konermann lab and I'm so sad that my journey came to an end.

I would like to thank my committee members Dr. Keith Griffiths and Dr. Elizabeth Gillies for devoting hours of their time to help me and providing me with insightful comments. I would like to thank my thesis examiners Dr. Gary Shaw and Dr. Philip Britz-McKibbin.

Many thanks to all of the past and present group members (Antony, Dr. Yue, Lauren, Claire, Courtney, Katja, Tilo, Samuel, Victor, Maryam, Angela, Quentin, and Vincent). I really appreciate your valuable opinions in my work. Special thanks to Dr. Siavash Vahidi for teaching me how to maintain and troubleshoot lab equipment. Both Siavash and Antony helped me when I first started and introduced me to the world of proteins. Thanks to Dr. Modupeola Sowole for her friendship and teaching me the art of performing HDX experiments. Robert McAllister, we worked together on so many projects and to a large extent, you are the motivator of all the MD work in my thesis. Thanks, Ming for your help in troubleshooting the mass spectrometer, HDX manager. I'm grateful to Ming for sharing his HDX experience and useful insights with me.

I would like to acknowledge with gratitude the support and love of my wonderful family especially my mom, dad, and my sister. A very special thanks to my

husband (Dr. Husam Osman), you have been a huge support through these past years. My sincere thanks to Brittany Vaughan for taking care of my lovely daughter while I'm working late. I would like to thank my pretty girl Malak Osman; without your smile, I would not be able to go through my day.

# Table of Contents

<b>Abstract</b> .....	i
<b>Co-Authorship Statement</b> .....	iii
<b>Acknowledgments</b> .....	iv
<b>List of Figures</b> .....	x
<b>List of Symbols and Abbreviations</b> .....	xiv
<b>Chapter 1. Introduction</b> .....	1
1.1. Protein Structure-Function Relationship.....	1
1.2. Factors that Contribute to the Stability of Native Protein Structures in Solution .....	2
1.3. Methods for Studying of Protein Structure and Dynamics .....	4
1.3.1. X-ray Crystallography .....	4
1.3.2. Nuclear Magnetic Resonance (NMR) Spectroscopy .....	5
1.3.3. Cryogenic Electron Microscopy (Cryo-EM).....	5
1.3.4. Optical Methods .....	6
1.4. Mass Spectrometry.....	7
1.4.1. Ion Source .....	7
1.4.2. Mass Analyzers .....	8
1.4.3. Quadrupole .....	9
1.5. MS Techniques for Studying of Protein Structure and Dynamics.....	10
1.5.1. Collision-Induced Dissociation and Tandem MS .....	10
1.5.2. HDX-MS, Covalent Labelling, and Cross-Linking .....	10
1.5.3. Ion Mobility Spectrometry (IMS) .....	12
1.6. Electrospray Ionization (ESI).....	15
1.6.1. Mechanism of Gas Phase Ion Release During ESI .....	16
1.6.2. Nanoelectrospray Ionization (NanoESI) .....	19
1.7. Protein ESI under Native and Denaturing Conditions .....	19
1.8. Proteins in the Gas Phase vs. in Solution.....	21
1.9. Computer Simulations .....	23
1.10. Molecular Dynamics (MD) Simulations.....	23
1.10.1. Initial Coordinates and Velocities .....	26
1.10.2. Force Fields .....	26

1.10.3. Integration Algorithms .....	29
1.10.4. Energy Minimization.....	31
1.10.5. Water Models .....	31
1.10.6. Thermostats .....	32
1.10.7. Periodic Boundary Conditions (PBC) .....	33
1.10.8. Treatment of Non-Bonded Interactions.....	34
1.10.9. Neighbor Lists .....	35
1.11. Scope of the Thesis .....	36
1.12 References .....	37
<b>Chapter 2. Exploring the Mechanism of Salt-Induced Signal Suppression in Protein Electrospray Mass Spectrometry Using Experiments and Molecular Dynamics Simulations.....</b>	<b>47</b>
2.1. Introduction.....	47
2.2. Materials and Methods.....	50
2.2.1. Protein Solutions .....	50
2.2.2. Mass Spectrometry (ESI-MS) .....	50
2.2.3. Experimental Design and Data Analysis.....	52
2.2.4. Molecular Dynamics Simulations .....	53
2.3. Results and Discussion .....	54
2.3.1. NaCl Effects on Protein Mass Spectra .....	54
2.3.2. Comparison with Other Salts .....	59
2.3.3. MD Simulations of Salt-Containing Droplets.....	65
2.3.4. Mechanism of Protein Ion Suppression .....	68
2.4. Conclusions .....	71
2.5. References .....	73
<b>Chapter 3. Mechanism of Protein Supercharging by Sulfolane and m-NBA: Molecular Dynamics Simulations of the Electrospray Process.....</b>	<b>77</b>
3.1. Introduction.....	77
3.2. Materials and Methods.....	79
3.2.1. Electrospray Ionization Mass Spectrometry and Ion Mobility Spectrometry ...	79
3.2.2. MD Simulations .....	80

3.3. Results and Discussion .....	82
3.3.1. Supercharging of Holo-Myoglobin .....	82
3.3.2. Choice of Charge Carrier for Simulations.....	83
3.3.3. Charge Carrier Behavior in Different Solvents.....	84
3.3.4. ESI Simulations.....	87
3.3.5. Anatomy of the ESI Process.....	91
3.3.6. Electrospray Ionization Supercharging via Charge Trapping .....	94
3.3.7. Surface Charge of ESI Droplets .....	95
3.3.8. Supercharging and Protein Unfolding.....	97
3.4. Conclusions.....	100
3.5. References .....	103
<b>Chapter 4. Crown Ethers Modulate the Location of Charge Carriers in Electrospray Droplets: Implications for the Mechanism of Protein Charging and Supercharging</b> .....	<b>108</b>
4.1. Introduction.....	108
4.2. Materials and Methods.....	110
4.2.1. Proteins and Reagents .....	110
4.2.2. Mass Spectrometry .....	111
4.2.3. MD Simulations .....	111
4.2.4. Accelerated MD with Forced Evaporation.....	112
4.3. Results and Discussion .....	114
4.3.1. Effects of 18C6 on ESI Charge States .....	114
4.3.2. 18C6 Suppresses Supercharging .....	114
4.3.3. 18C6 Effects in NaCl-Containing Solutions .....	116
4.3.4. Comparing MD and Experimental Results .....	119
4.3.5. Common Features of MD Trajectories.....	120
4.3.6. Principles that Govern the ESI Charge States of Proteins .....	123
4.3.7. MD Trajectories Reveal the Basis of Charging and Supercharging .....	123
4.3.8. Relationship between Supercharging and Unfolding.....	127
4.3.9. Dendrimer Supercharging .....	128
4.4. Conclusions.....	131

4.5. References .....	134
<b>Chapter 5. Chain Ejection Model for Electrospray Ionization of Unfolded Proteins: Evidence from Atomistic Simulations and Ion Mobility Spectrometry</b> .....	139
5.1. Introduction .....	139
5.2. Materials and Methods .....	143
5.2.1. Protein, Reagents and Mass Spectrometry .....	143
5.2.2. MD Simulations .....	143
5.3. Results and Discussion .....	144
5.3.1. ESI-MS and IMS Experiments .....	144
5.3.2. ESI Modeling Strategy .....	147
5.3.3. MD Simulations Confirm CEM Behavior .....	151
5.3.4. Charge Loss at the Rayleigh Limit .....	156
5.3.5. Protein Conformations after CEM Ejection .....	160
5.4. Conclusions .....	162
5.5. References .....	164
<b>Chapter 6: Conclusions and Future Work</b> .....	168
6.1 Conclusions .....	168
6.2 Future Directions .....	170
6.2.1 Studying the Effects of Anions on Protein Ion Signals in Negative Ion Polarity .....	170
6.2.2 Investigating the Effects of Salts on Protein Supercharging .....	170
6.2.3 Additional Supercharging Agents .....	170
6.2.4 Examining the Effects of Supercharging Agents in Negative Ion Mode .....	171
6.2.5 Uncovering the Mechanism of Supercharging Under Denaturing Conditions .....	171
6.2.6 Can the Chain Ejection Model (CEM) be applied to Intrinsically Disordered Proteins (IDPs)? .....	171
6.3 References .....	172
<b>Appendix I-Permissions</b> .....	174
<b>Curriculum Vitae</b> .....	179

# List of Figures

## Chapter 1.

- Figure 1.1.** A polypeptide showing a peptide bond in red. ....2
- Figure 1.2.** Cartoon representation showing quadrupole operation.....10
- Figure 1.3.** Schematic layout of a typical “Synapt” Q-TOF mass spectrometer.....14
- Figure 1.4.** Schematic representation of ESI source operated in positive ion mode.....16
- Figure 1.5.** Cartoon representation of proposed ESI models.....18
- Figure 1.6.** Bonded interactions in a typical forcefield between atoms.....28
- Figure 1.7.** Periodic boundary conditions in two dimensions.....34

## Chapter 2.

- Figure 2.1.** Schematic protein ESI mass spectra.....48
- Figure 2.2.** Cone voltage effects on ESI mass spectra of *cyt c*.....51
- Figure 2.3.** ESI mass spectra of *cyt c* acquired at pH 7 in the presence of various NaCl concentrations.....55
- Figure 2.4.** Salt dependence of the integrated ion intensity  $R(C)$  relative to NaCl-free samples.....56
- Figure 2.5.** ESI mass spectra of Ubq and Lyz acquired at pH 7 .....57
- Figure 2.6.** ESI mass spectra of Cyt *c*, Ubq and rLyz acquired under denaturing solvent conditions.....58
- Figure 2.7.** ESI mass spectra of Cyt *c* acquired at pH 7 in the presence of different CsCl and NBu<sub>4</sub>Cl concentrations .....60
- Figure 2.8.** Close-up views of the Cyt *c* .....61
- Figure 2.9.** ESI mass spectrum of Cyt *c* acquired at pH 7 in the presence of 10 mM NBu<sub>4</sub>Cl.....62
- Figure 2.10.** ESI mass spectra of Cyt *c* acquired at pH 2 in the presence of different concentrations of CsCl and NBu<sub>4</sub>Cl.....63

<b>Figure 2.11.</b> Salt dependence of the integrated ion intensity $R(C)$ for Cyt <i>c</i> .....	64
<b>Figure 2.12.</b> MD simulation results for the evaporation of ESI droplets.....	67
<b>Figure 2.13.</b> Schematic of droplet evolution within the ESI plume.....	69
<b>Chapter 3.</b>	
<b>Figure 3.1.</b> ESI mass spectra of hMb in aqueous solution at pH 7.....	83
<b>Figure 3.2.</b> ESI mass spectrum obtained after infusion of holo-myoglobin in an aqueous solution containing 0.1 mM sodium acetate and 1% <i>m</i> -NBA.....	85
<b>Figure 3.3.</b> Quasi-equilibrium MD data for ESI droplets containing hMb.....	86
<b>Figure 3.4.</b> MD snapshots illustrating the temporal evolution of hMb.....	89
<b>Figure 3.5.</b> Details of hMb interactions with selected Na <sup>+</sup> ions.....	90
<b>Figure 3.6.</b> ESI charge states of hMb .....	90
<b>Figure 3.7.</b> MD data, illustrating the evaporation of hMb-containing ESI droplets.....	93
<b>Figure 3.8.</b> Electrostatic potential $V(r)$ of nanodroplets containing hMb.....	96
<b>Figure 3.9.</b> The radius of gyration ( $R_g$ ) as a function of time for ESI droplets .....	99
<b>Figure 3.10.</b> ESI in water without and with supercharging.....	101
<b>Chapter 4.</b>	
<b>Figure 4.1.</b> Illustrative results of COM (center of mass) pulling simulations for hMb-containing ESI droplets.....	113
<b>Figure 4.2.</b> Mass spectra acquired after electrospraying holo-myoglobin without and with 1 mM 18C6.....	115
<b>Figure 4.3.</b> Mass spectra acquired after electrospraying holo-myoglobin without and with 18C6 in neutral aqueous solution containing 1% sulfolane .....	117
<b>Figure 4.4.</b> Mass spectra acquired after electrospraying ubiquitin without and with 18C6.....	118
<b>Figure 4.5.</b> Mass spectra acquired after electrospraying holo-myoglobin in a neutral aqueous solution containing 1% sulfolane, 1 mM NaCl.....	119

<b>Figure 4.6.</b> ESI charge states of hMb .....	120
<b>Figure 4.7.</b> Snapshots are taken from MD trajectories that culminate in the production of desolvated hMb ions from ESI nanodroplets.....	121
<b>Figure 4.8.</b> Mass spectrum obtained upon electrospraying 1 mM aqueous ammonium acetate containing 1 mM 18C6.....	122
<b>Figure 4.9.</b> MD data for four types of hMb-containing ESI droplets.....	124
<b>Figure 4.10.</b> Protein interactions with Na <sup>+</sup> , 18C6, and [18C6 + Na] <sup>+</sup> .....	125
<b>Figure 4.11.</b> Protein compactness during ESI, reported as the radius of gyration.....	128
<b>Figure 4.12.</b> ESI mass spectra of G5 PAMAM dendrimer in aqueous solution containing ammonium acetate.....	130
<b>Figure 4.13</b> ESI mass spectra of G5 PAMAM dendrimer in aqueous solution containing ammonium acetate and 1 mM NaCl.....	130
<b>Figure 4.14.</b> Cartoon summary of MD and experimental results.....	133
<b>Chapter 5.</b>	
<b>Figure 5.1.</b> Cartoon depiction of two analogous gas phase processes. (a) CID of a noncovalent protein complex.....	142
<b>Figure 5.2.</b> Experimental ESI mass spectra for aMb recorded in aqueous solution at pH 4 and pH 2.....	145
<b>Figure 5.3.</b> $\Omega$ distributions measured by IMS.....	146
<b>Figure 5.4.</b> Calculated aMb net charge in solution vs. pH.....	149
<b>Figure 5.5.</b> aMb charge patterns .....	150
<b>Figure 5.6.</b> Typical snapshots from CEM simulation runs on Rayleigh-charged aqueous droplets containing unfolded aMb.....	152
<b>Figure 5.7.</b> Snapshots from CEM simulation run on aMb.....	153
<b>Figure 5.8.</b> CEM ejection aMb from aqueous ESI droplets.....	154
<b>Figure 5.9.</b> CEM ejection of aMb in charge states (a) 27 <sup>+</sup> and (b) 33 <sup>+</sup> from the ESI droplet.....	155
<b>Figure 5.10.</b> Snapshots from simulation runs on three protonation patterns.....	156

**Figure 5.11.** MD simulation data for typical CEM runs on aMb.....158  
**Figure 5.12.** Formation of small progeny droplets via jet fission .....159  
**Figure 5.13.** Experimental IMS data measured .....161

## List of Symbols and Abbreviations

18C6 – 18-crown-6

aMb – apo-myoglobin

AMBER – assisted model building with energy refinement

ADC – analog-to-digital converter count

CCS – collision cross section

CD – circular dichroism

CEM – chain ejection model

CHARMM – chemistry at Harvard molecular mechanics

CID – collisional induced dissociation

CRM – charged residue model

CsCl – cesium chloride

Cyt *c* – cytochrome *c*

Da – dalton

DC – direct current

DFT – density functional theory

DTT – 1,4 dithiothreitol

*e* – the elementary charge ( $1.602 \times 10^{-19}$  coulombs)

ESI – electrospray ionization

FT-ICR – Fourier transform- ion cyclotron resonance

HDX – hydrogen/deuterium exchange

hMb – holo-myoglobin

IEM – ion ejection model

IMS – ion mobility spectrometry

LC – liquid chromatography

Lyz – lysozyme

MALDI – matrix-assisted laser desorption/ionization

*m*-NBA – *m*-nitrobenzyl alcohol

MS – mass spectrometry

MS/MS – tandem mass spectrometry

*m/z* – mass-to-charge ratio

N – native state

NMR – nuclear magnetic resonance

•OH – hydroxyl radical

OPLS/AA – optimized potential for liquid Simulation – all-atom

PBC – periodic boundary conditions

PME – particle mesh Ewald

PDB – protein data bank

RF – radio frequency

SCA– supercharging agent

TOF – time-of-flight

TROSY – transverse relaxation optimized spectroscopy

TWIM – traveling-wave ion mobility

Ubq – ubiquitin

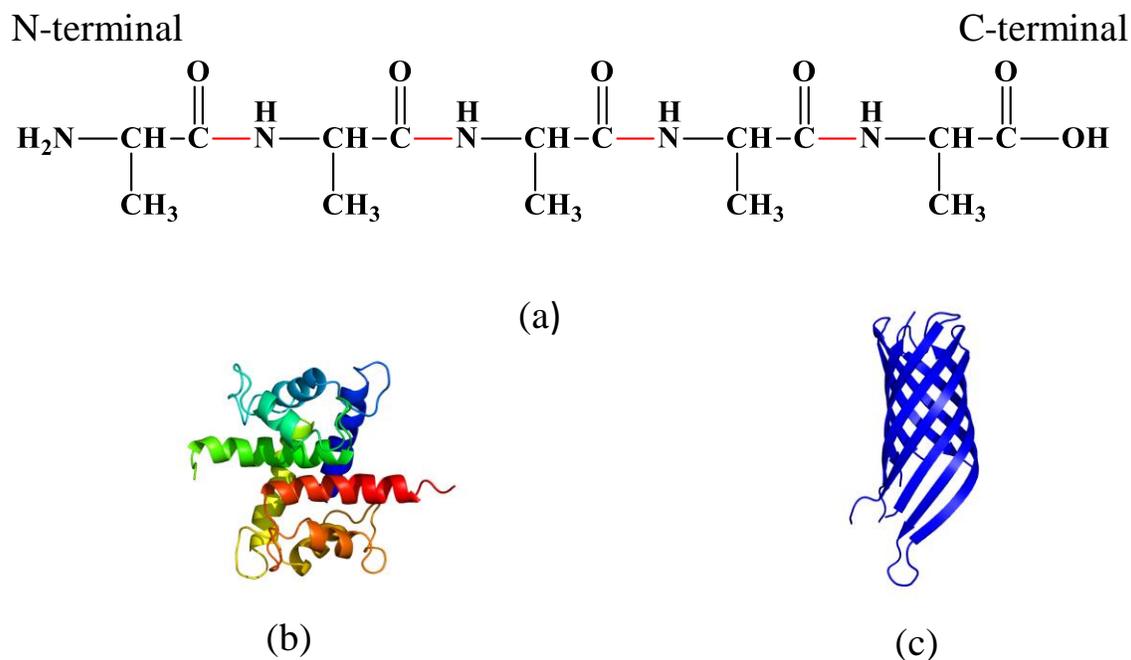
UV-Vis – ultraviolet-visible

## Chapter 1. Introduction

### 1.1. Protein Structure-Function Relationship

Proteins are biological polymers that are involved in all physiological processes. There are structural proteins (viral coat proteins, epidermal keratin); catalytic proteins (enzymes); transport and storage proteins (hemoglobin, myoglobin, ferritin); regulatory proteins (hormones, transcription regulators), and proteins of the immune system. Their size ranges from a few thousand Daltons (Da) all the way to the megadalton (MDa) range.<sup>1</sup>

Proteins adopt their higher order structure through folding (Figure 1.1). Twenty naturally occurring amino acids represent the basic building blocks during protein biosynthesis. These amino acids form polypeptide chains via peptide bonds; this is known as the primary structure (Figure 1.1a). Secondary structure forms via intramolecular hydrogen bonds that organize the backbone into  $\alpha$ -helices and  $\beta$ -sheets. The arrangement of secondary motifs results in the formation of tertiary structure that is stabilized by non-covalent interactions (hydrogen bonds, van der Waals forces, ionic interactions and hydrophobic packing) (Figure 1.1b, c). Quaternary structure refers to an assembly of two or more protein chains.<sup>2</sup>



**Figure 1.1.** a) A polypeptide showing a peptide bond in red. b) NMR structure of S100B (PDB: 1UWO), showing  $\alpha$ -helices c) Tertiary structure of outer membrane protein A (PDB: 1QJP), illustrating  $\beta$ -sheets.

## 1.2. Factors that Contribute to the Stability of Native Protein Structures in Solution

In their biologically active (“native”) state, most proteins are folded into highly ordered compact structures. There is a delicate balance between the folded native conformation (N) and the unfolded state (U). Folding transitions are often discussed in terms of a simple two-state equilibrium  $N \rightleftharpoons U$ , and the corresponding equilibrium constant  $K = [U]/[N]$  can be affected by changes in temperature or solvent properties.

The Gibb’s free energy change  $\Delta G^\circ = -RT \ln K$  provides a framework for discussing protein stability; it includes enthalpic ( $\Delta H$ ) and entropic ( $\Delta S$ ) contributions arising from the protein and the surrounding solvent according to

$$\Delta G^\circ_{N \rightarrow U} = \Delta H^\circ_{N \rightarrow U} - T \Delta S^\circ_{N \rightarrow U} \quad (1.1)$$

For N to be stable,  $\Delta G^\circ$  has to be positive. To decipher the basics of protein stability, it is essential to identify the types of interactions that affect  $\Delta H$  and  $\Delta S$ .<sup>3</sup>

Hydrogen bonds (H-bonds) are one of the primary stabilizing factors for secondary structure ( $\alpha$ -helices and  $\beta$ -sheets). H-bonds are non-covalent interactions that occur between a hydrogen donor and acceptor e.g., in  $\alpha$ -helices the amide hydrogen donor (N-H) is hydrogen bonded to the oxygen of the carbonyl group. In addition to these backbone contacts, H-bonds can form among side chains, or between side chains and the backbone.<sup>4-5</sup>

Van der Waals interactions take place between atoms that are in close contact with one another. The protein has favorable internal van der Waals interactions, but there are also favorable van der Waals contacts with the solvent.<sup>6</sup>

Salt bridges form due to electrostatic interactions between the positively charged side chain of lysine (Lys) and arginine (Arg) and the carboxylate group of glutamic acid (Glu) and aspartic acid (Asp). Salt bridges are rarely located within the protein as it is energetically unfavorable to desolvate two charged groups. Mostly, salt bridges are located on the surface where they are solvated by water. Therefore, the contribution of salt bridges to protein stability is limited.<sup>7</sup>

Disulfide (S-S) bridges are covalent contacts formed between pairs of cysteine (Cys) side chains. These bonds are found in many proteins such as lysozyme and insulin. They contribute to the stabilization of native state by decreasing the conformational entropy of the unfolded state.<sup>8</sup>

The hydrophobic effect is the most dominant factor in protein stability. Nonpolar amino acids tend to sequester themselves in the protein interior to avoid solvation by water. This phenomenon can be explained thermodynamically. The surface of the protein is decorated with hydrophilic residues that are solvated by water which is enthalpically favorable. The hydrophobic residues that cannot form H-bonds are buried inside which will decrease the entropic penalty as described by the "iceberg model". In liquid water, the hydrogen bond network is very dynamic due to Brownian motions. Placing a non-polar molecule in bulk water will cause the formation of immobilized "water cages" around them. This iceberg water has lower entropy than bulk water.

To avoid this undesirable effect, water molecules decrease their contact with non-polar molecules. This is achieved by the burial of hydrophobic molecules within the protein core.<sup>9-10</sup>

The entropy of the polypeptide chain decreases dramatically upon folding. This will reduce the entropy of the protein ( $\Delta S < 0$ ) which is thermodynamically unfavorable. At the same time, hydrophobic interactions will compensate for this effect by increasing the entropy of the surrounding water. When the hydrophobic chains are clustered together, the water molecules surrounding these non-polar patches will be released in bulk water and participate into the H-bond network of the solvent. Overall, the magnitude and the sign of  $\Delta G^\circ$  in equation 1.1 represents a delicate balance between numerous competing factors and interactions.

## 1.3. Methods for Studying of Protein Structure and Dynamics

### 1.3.1. X-ray Crystallography

X-ray crystallography represents the most important tool for determining high-resolution 3D protein structures. Thousands of protein structures in the Protein Data Bank (PDB) have been obtained via X-ray crystallography. Under optimal conditions, this technique can yield structural data with a resolution of about 1 Å which is on the order of atomic bond lengths.<sup>11</sup>

The birth of X-ray crystallography started with the work of Nobel prize winner Max von Laue, who discovered the interaction of X-ray photons with crystals. In 1915, Nobel Prize winners Lawrence Bragg and his father William Bragg developed Bragg's law that is used to analyze the diffraction patterns produced by the crystal lattice upon exposure to X-ray. In X-ray crystallography, protein crystals are exposed to an X-ray beam. The X-ray photons are diffracted by the electrons in the crystal lattice which constructively or destructively interfere with each other. From the resulting diffraction pattern, it is possible to produce a three-dimensional electron density map, from which atomic positions and bonds in the crystal can be obtained.<sup>12</sup>

Proteins undergo incessant thermal motions and fluctuate between different conformations.<sup>13</sup> The crystallographic B-factors report on the thermal displacement of atoms from their average positions, but they also have contributions from crystal imperfections.<sup>14-15</sup>

Overall, dynamic information obtained from B-factors is limited, calling for the use of complementary techniques such as nuclear magnetic resonance spectroscopy.<sup>16-17</sup>

### 1.3.2. Nuclear Magnetic Resonance (NMR) Spectroscopy

The development of NMR spectroscopy started in the 1940s by the Rabi, Purcell, and Bloch groups.<sup>18-19</sup> All of them received the Nobel Prize for their achievements. Since then, NMR has evolved into a valuable tool that is used for a wide range of applications, from the characterization of small molecules to the conformational determination of biomolecules. NMR probes nuclei that have a spin such as  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{17}\text{O}$ ,  $^{33}\text{S}$ . Unlike X-ray crystallography, NMR is conducted in aqueous solution and it provides a near physiological environment for protein analyses.

Various NMR techniques are available for probing of protein structures<sup>20</sup> and dynamics at atomic resolution.<sup>21</sup> NMR spectra often suffer from spectral congestion and decreased intensity, especially for larger proteins. Replacement of hydrogen with deuterium in conjunction with transverse relaxation- optimized spectroscopy (TROSY) experiments is one way to address this difficulty, thereby extending the size range accessible by NMR methods.<sup>22</sup>

### 1.3.3. Cryogenic Electron Microscopy (Cryo-EM)

In recent years, cryo-EM has become a powerful tool in structural biology. In 2017 the Nobel Prize in chemistry was granted to Dubochet, Frank, and Henderson for the development of cryo-EM which helps in obtaining structural information for non-crystallized proteins.<sup>23</sup> This technique has proved to be an excellent tool in case of multimeric proteins and membrane proteins. Cryo-EM uses frozen protein samples.<sup>24</sup> The continuous developments of cryo-EM made it possible to resolve challenging structures that have been unamenable to X-ray crystallography, e.g., the yeast exocyst that is responsible for vesicle transport<sup>25</sup> and the structure of actin bound to myosin which is important for muscle contraction.<sup>26</sup>

### 1.3.4. Optical Methods

Optical methods are easy to use, quick and sensitive, but they do not provide high-resolution structural information. Examples include UV-Visible (UV-Vis) absorption, circular dichroism (CD) and fluorescence spectroscopy.

UV-Vis absorption spectroscopy can be used to probe protein conformations as a peptide bond absorbs light in the far UV range around 195 nm. A second absorption maximum around 280 nm arises from aromatic amino acids such as tyrosine, tryptophan, and phenylalanine. Sometimes prosthetic groups have a strong absorption band, e.g., the spectroscopic properties of heme groups depend on the oxidation state, the presence of ligands, and the surrounding environment.<sup>27</sup> In addition, protein concentrations can be calculated using the Beer-Lambert law

$$A = \epsilon C d \quad (1.2)$$

where  $A$  is the absorbance,  $\epsilon$  is the molar absorption coefficient as a function of wavelength,  $C$  is the concentration, and  $d$  is the path length of the cuvette.<sup>28</sup>

Circular Dichroism spectroscopy is useful for interrogating chromophores that contain chiral centers (such as the alpha carbons along the protein backbone). A CD spectrum represents the absorption difference between left and right circularly polarized light. This technique is commonly used to probe changes in protein secondary structure. For example, a dominant CD band at 222 nm is a hallmark of  $\alpha$ -helical structure, while a band around 215 nm corresponds to  $\beta$ -sheets. The random coil signal appears at  $\sim 200$  nm.<sup>29-30</sup>

Fluorescence spectroscopy is routinely used for protein structural analyses. An excited fluorophore emits photons due to relaxation to the ground state. Tryptophan (Trp) is the most intense fluorophore in proteins.<sup>31</sup> The chromophore environment has a significant effect on fluorescence. Förster Resonance Energy Transfer (FRET) is a valuable method for characterizing protein conformations. FRET depends on the distance between the excited donor and acceptor. One of the consequences of FRET is quenching. Quenching takes place if a fluorophore is close to a non-fluorescent acceptor, such as the heme group in cytochrome *c*, where the energy transfers from Trp59 to heme takes place in native conformation, so Trp59 becomes non-

fluorescent. In contrast, for unfolded cytochrome *c* Trp59 is highly fluorescent, thereby providing a tool to monitor conformational changes.<sup>32</sup>

## 1.4. Mass Spectrometry

Mass spectrometry (MS) has become an indispensable tool in the field of structural biology and proteomics. Its applications include measuring mass to charge ratio ( $m/z$ ), as well as providing information on structures, protein-ligand interactions and post-translational modifications. MS can be used either as an isolated technique or coupled to liquid chromatography (LC), gas chromatography (GC), capillary electrophoresis (CE) or ion mobility spectrometry (IMS).

### 1.4.1. Ion Source

Ionization of analytes takes place in the ion source. There are different methods of ionization in MS, such as electron ionization (EI), fast atom bombardment (FAB) and chemical ionization (CI). These “traditional” techniques tend to cause fragmentation, particularly for larger analytes.<sup>33</sup> The analysis of biomolecules such as proteins requires gentle ionization techniques such as electrospray ionization (ESI) and matrix-assisted laser absorption/desorption ionization (MALDI). In 2002, Fenn<sup>34</sup> and Tanaka won the Chemistry Nobel Prize for developing ESI and MALDI, respectively. In the positive ion mode, both ESI and MALDI predominantly form protonated analyte ions with  $m/z$  that is given by

$$\frac{m}{z} = \frac{[M + z \times 1.008]}{z} \quad (1.3)$$

where  $M$  is the mass of the neutral analyte and 1.008 represents the proton mass. The mechanism of ESI will be discussed in more detail below (section 1.6).

MALDI is a gentle surface desorption ionization technique that was introduced by Hillenkamp and Karas in 1985.<sup>35-36</sup> The protein sample is mixed with a matrix, for example, 2,4-dihydroxybenzoic acid. The matrix absorbs UV laser energy and then desorbs from the surface together with the protein. The charge is transferred from the matrix to the gaseous analyte within the MALDI plume.<sup>37-38</sup> One of the drawbacks of MALDI is that the extent of charging is limited, such that analyte ions tend to appear at very high  $m/z$ , beyond the range accessible to some mass

analyzers. For this reason, MALDI is often coupled with time of flight mass analyzers (TOF) which have a very wide  $m/z$  range. Also, MALDI imaging-MS has become a widely used technique for direct, label-free detection of proteins, lipids, and metabolites in tissue sections.<sup>39-40</sup>

### 1.4.2. Mass Analyzers

A mass analyzer is the part of the mass spectrometer that separates gas phase ions according to their  $m/z$ . There are several mass analyzers such as quadrupoles, ion traps, TOFs and Fourier-transform ion cyclotron resonance (FT-ICR) instruments. Only the quadrupole and TOF operation will be briefly discussed due to their use in the experiments of the current thesis.

TOF analyzers have an excellent ion transmission, resolution and an extended mass range. The TOF operation is straightforward; ions with different  $m/z$  are accelerated by the same voltage supplied by an ion pusher. The voltage pulse DELTA  $U$  supplies all ions with potential energy which is then converted into kinetic energy.

$$ze\Delta U = \frac{1}{2}mv^2 \quad (1.4)$$

$m, v, ze$  represent the mass, velocity and ion charge. rearrangement of equation 1.4 yields

$$v = \sqrt{\frac{2ze\Delta U}{m}} \quad (1.5)$$

The time for an ion to reach the detector  $t$  is given by

$$t = \frac{l}{v} = \frac{l\sqrt{m}}{\sqrt{2ze\Delta U}} = \frac{l}{\sqrt{2e\Delta U}} \sqrt{\frac{m}{z}} \quad (1.6)$$

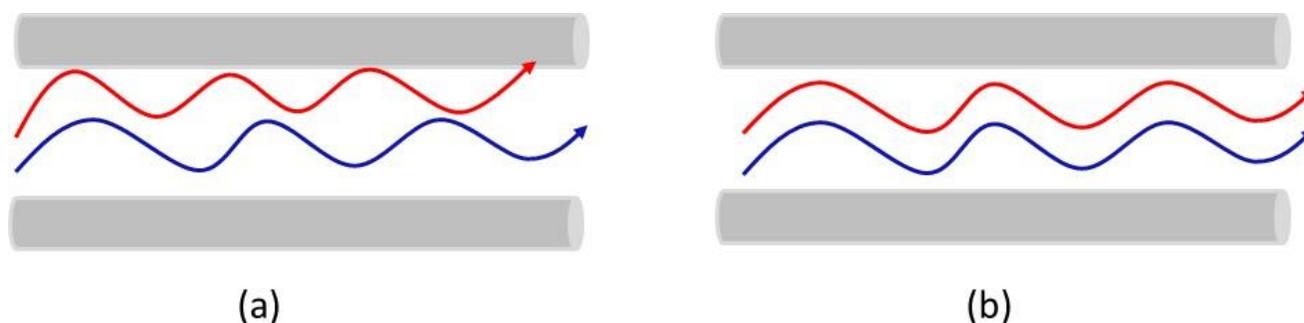
Equation 1.6 shows that the flight time  $t$  for an ion depends on  $m/z$ . Therefore, a mass spectrum can be recorded as a function of flight time. Ions with different  $m/z$  will have variable

velocities and different flight times. However, under realistic conditions some ions with the same  $m/z$  may have slightly different kinetic energy because they do not all experience exactly the same acceleration potential  $\Delta U$ . These conditions cause peak broadening and decreased resolution.<sup>41</sup> Reflectrons are used to correct this issue. The reflectron provides an electrostatic field to reverse the direction of the ion trajectory. If two ions have the same  $m/z$  but different velocities, the faster one penetrates deeper into the reflectron before being directed back toward the detector. A reflectron increase the TOF resolution up to 40,000, but at the same time sensitivity (ion transmission) is decreased.<sup>42-43</sup> TOF mass spectrometers can easily be combined with a pulsed ionization technique such as MALDI. However, it is also possible to use continuous ionization techniques such as ESI (see below).

### 1.4.3. Quadrupole

One of the most commonly used mass analyzers is the quadrupole. It is composed of four parallel cylindrical rods. Both radiofrequency (RF) and direct current (DC) electric potentials are applied to the rods if the quadrupole is used as a ‘mass filter’. At a constant DC/RF ratio, ions of a certain  $m/z$  travel through the quadrupole and reach the detector. Any ions with other  $m/z$  values will have unstable trajectories and hit one of the rods, i.e. they will not be transmitted (Figure 1.2a). In the absence of DC, all ions will pass through the quadrupole regardless of their  $m/z$ . In this ‘RF-only’ mode, the quadrupole acts as an ion guide (Figure 1.2b). RF-only quadrupole can also be used as collision cell in the presence of a suitable background gas and an accelerating bias.<sup>44</sup>

Quadrupole mass analyzers have a relatively low resolution (~4000) but offer excellent sensitivity. Usually, they are combined with a TOF (Q-TOF) or in a series of three quadrupoles known as “triple quads” for tandem MS applications, where the second quadrupole acts as a collision cell.<sup>43</sup>



**Figure 1.2.** Cartoon representation showing quadrupole operation. a) Application of DC voltage locks the quadrupole so only specific  $m/z$  can pass (blue line). b) RF only where all ions can pass.

## 1.5. MS Techniques for Studying of Protein Structure and Dynamics

### 1.5.1. Collision-Induced Dissociation and Tandem MS

Fragmentation experiments are essential for deciphering the chemical composition of gaseous analytes such as peptides and proteins. The bottom-up methodologies are very common, they involve enzymatic digestion of the protein prior to gas phase fragmentation and MS analysis.<sup>45</sup> On the other hand, in the top-down approach, intact proteins undergo fragmentation inside the mass spectrometer without prior enzymatic digestion.<sup>46</sup>

In collision-induced dissociation (CID), protein or peptide ions are accelerated by an electric voltage in a region filled with inert gas. Collisions with this background gas will gradually increase the internal energy, ultimately resulting in the rupture of noncovalent contacts and covalent bonds. CID can be conducted directly in the ion sampling interface (in-source CID), or in an RF-only ion guide after quadrupole-based precursor selection. The latter technique is known as tandem MS (MS/MS) which is essential for analyte mixtures.<sup>47</sup>

### 1.5.2. HDX-MS, Covalent Labelling, and Cross-Linking

Hydrogen/deuterium exchange (HDX) MS is commonly employed to detect protein-ligand interactions and to explore protein conformational dynamics. HDX uses deuterium ( $D_2O$ ) as a labeling reagent.<sup>48</sup> HDX exploits the exchange of labile protein hydrogen atoms with

deuterium from the solvent. Backbone NH sites and those of side chains tend to exchange rapidly. However, hydrogens that are H-bonded (in  $\alpha$ -helices and  $\beta$ -sheets) and/or those that are buried will exchange much more slowly. Deuteration at these protected sites is mediated by thermal opening/closing transitions, i.e. by the conformational dynamics of the protein.<sup>49</sup> HDX experiments were initially conducted with NMR spectroscopic detection,<sup>50</sup> later the use of MS for this purpose has become the more widely used approach.<sup>48, 51</sup>

The HDX workflow starts with the mixing of the protein solution with D<sub>2</sub>O based buffer for various time intervals (continuous labeling), or just for a fixed labeling period (pulsed HDX). This labeling step is followed by quenching at pH ~2.5 and at low temperature, to avoid deuterium back exchange. Information about the exchange can be obtained either from intact protein mass analyses (global HDX) or after proteolytic digestion and LC-MS of the resulting peptides.<sup>49, 52-53</sup>

Covalent labeling is a powerful technique that can probe the surface accessibility of amino acids in solution via irreversible chemical modifications.<sup>54-55</sup> The principle of covalent labeling is related to HDX, the main difference is that HDX reports on backbone H-bonds while covalent labeling targets protein side chains. Hydroxyl radical ( $\cdot$ OH) is a commonly employed highly reactive covalent label that modifies side chains via oxidation.  $\cdot$ OH radical can be generated by X-ray radiolysis of water or by photolysis of hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) and Fenton chemistry. In the last method, a metal is used such as Fe (II) to reduce H<sub>2</sub>O<sub>2</sub>.<sup>56</sup> UV photolysis of H<sub>2</sub>O<sub>2</sub> is a method developed by Hambly and Gross, known as fast photochemical oxidation of proteins (FPOP).<sup>57-58</sup> To determine the location and extent of oxidation, protein digestion is used along with MS/MS. Caution must be taken to avoid extensive oxidation, low concentrations of H<sub>2</sub>O<sub>2</sub> must be employed and control experiments should be done with no laser treatment to evaluate the extent of oxidation. Radical scavengers such as glutamine are used to quench the reaction to decrease the  $\cdot$ OH lifetime. A recent study by our group has shown that FPOP can be complicated by the presence of secondary radicals that can form during FPOP which extend the time of the reaction and give rise to artifacts.<sup>59</sup> Despite that, FPOP holds great promise as a technique to probe protein-protein or protein-ligand interactions. In studies done by Stocks and Konermann, FPOP was used to probe short-lived intermediates in protein folding and unfolding.<sup>60-62</sup>

Chemical cross-linking is another widely used approach. The choice of a cross-linking reagent relies on the distance between two peptide residues. The crosslinker is a molecule with two reactive groups separated by a spacer of a certain length. They can be of different types such as homobifunctional molecules that target the same groups on the protein, or heterobifunctional molecules that bind different amino acids.<sup>63</sup> The main target residues for crosslinking are the primary amines and NS carboxyls.<sup>64</sup> Cross-linked proteins are analyzed by enzymatic digestion followed by LC-MS/MS. The resulting data are used to identify the modified peptide and the sites of linkage. Cross-linking provides a distance constraint that limits the number of possible structures and helps to draw the topology of subunits in complexes which promotes its use in computer modeling. Interpretation of data is quite challenging as there may be too many cross-links.<sup>65-67</sup> This can be resolved by comparing the possible cross-linked peptides against database searches.<sup>68</sup>

### 1.5.3. Ion Mobility Spectrometry (IMS)

IMS is a technique for separating analytes that have different gas-phase structures.<sup>15</sup> The history of IMS can be traced back to the early 20<sup>th</sup> century.<sup>69</sup> In recent years, the coupling of IMS with soft ionization techniques such as ESI and MALDI has opened up new avenues for the analysis of biomolecules.<sup>70-73</sup> In IMS, protein ion separation depends on size, charge and conformation.

Drift tubes represent the conceptually most straightforward type of IMS device. Ions move into a background gas (typically helium) under the influence of a weak constant electric field  $E$  in a radial direction towards the mass analyzer. The ions are separated based on their mobility  $K$

$$K = \frac{v}{E} = \frac{L}{t_D E} \quad (1.7)$$

In this equation  $L$  refers to the length of the drift tube,  $t_D$  is the drift time and  $v$  is the velocity. The reduced ion mobility  $K_0$  can be calculated according to standard conditions of temperature and pressure of the buffer gas to compensate for differences between laboratories and instruments.

$$K_0 = K \frac{T_0 P}{T P_0} \quad (1.8)$$

where  $T_0=273.2$  K, and  $P_0 = 760$  Torr. Drift tubes have high-resolving power. Early versions suffered from poor ion transmission, but continuous improvements and the use of ion guides before the drift cell helps in ion storage and focusing.<sup>15</sup>

The collision cross section (CCS) represents a type of projection area that arises from the average of all possible orientations involved in the gas phase collisions. The CCS depends on the shape of ion and its interaction with the background gas. The dependence between the CCS ( $\Omega$ ) and the measured drift time  $t_D$  is given by equation 1.9<sup>74</sup>

$$\Omega = t_{DZ} \frac{eE}{16NL} \sqrt{\frac{18\pi}{\mu k_B T} \frac{P_0}{P} \frac{T}{T_0}} \quad (1.9)$$

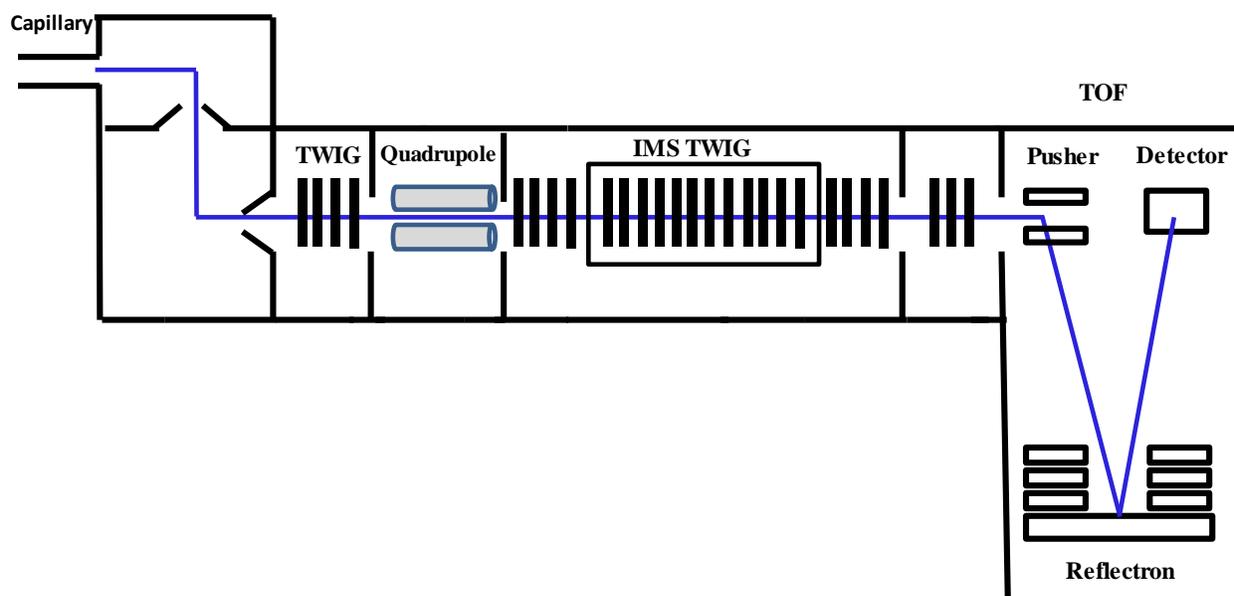
where  $N$  is the number density of the gas,  $\mu$  is the reduced mass of ion and gas.  $k_B$  is the Boltzmann constant.<sup>74-75</sup> Large ions have a longer drift time due to their more frequent collisions with the background gas, while smaller ions will experience a less frictional force. Ions with more charges will reach the detector faster, i.e., they will have higher mobilities.

The introduction of the Synapt platform by Waters in the early 2000s for the first time allowed the use of a different type of IM separation. This approach involves the use of traveling wave ion guides (TWIGs), and the technique is known as traveling wave ion mobility spectrometry (TWIMS). The main difference between drift tube IMS and TWIMS is that drift tubes have a static electric field, while TWIMS technology uses “peristaltic” traveling waves. More specifically, ions travel through a stack of ring electrodes at which opposite phases of radiofrequency (RF) are applied to adjacent electrodes. This creates a radial confining potential well. Ions are trapped inside the well. When a transient direct current (DC) voltage that jumps from ring to ring is superimposed on the RF voltage ions will travel through the ion guide by “surfing” on these waves. Ions with high mobility will tend to be swept along with the wave,

while ions of low mobility will occasionally roll over and slip behind the crest of the wave.<sup>76-77</sup> TWIMS generally has a lower resolution than drift tube IMS, but its performance can be enhanced by using a He cell prior to the TWIMS device (triwave cell), or by manipulating the wave amplitude and velocity.<sup>77</sup>

The nature of the TWIMS separation process is not fully understood, so calibration methods have been developed to extract numerically accurate CCSs from TWIMS experiments.<sup>78</sup>

CCS values can also be predicted on the basis of 3D protein conformations. Different algorithms have developed for this purpose. These include the projection approximation (PA) which represents the simplest and least accurate of all available methods.<sup>79</sup> The exact hard sphere (EHSS) is computationally fast and considers the scattering effects of colliding ions, but does not include the long-range interactions.<sup>80</sup> The later effects are being included in the trajectory method (TM) which by far is the most accurate and expensive method.<sup>81</sup>



**Figure 1.3.** Schematic layout of a typical “Synapt” Q-TOF mass spectrometer with a quadrupole, TOF and ion mobility (TWIMS) separator. The blue line represents the ion trajectory.

## 1.6. Electrospray Ionization (ESI)

ESI is the most common atmospheric pressure ionization technique. The idea behind ESI started with Malcolm Dole who was the first to exploit the idea of electrostatic spray painting for electro spraying kDa polymers. Dole's results were the starting point for the subsequent "ESI revolution".<sup>82</sup> In the 1980s, ESI was further developed by John Fenn. Since then, this ionization technique has been proven to be effective for a wide range of analytes including proteins, amino acids, DNA, lipids, and sugars.<sup>34</sup>

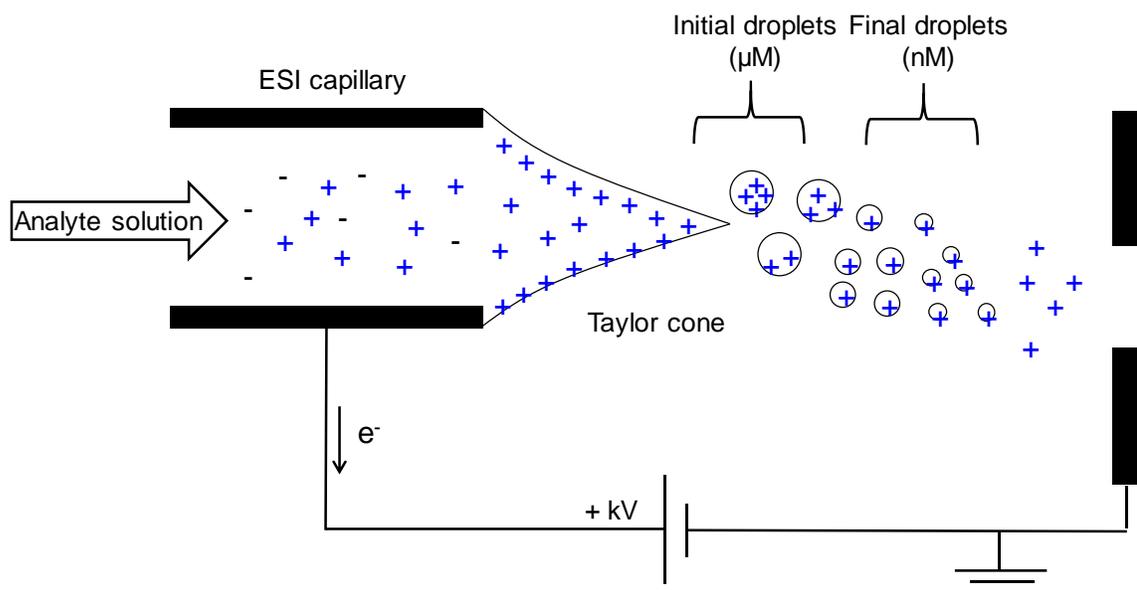
ESI has many advantages such as the possibility for coupling with LC, and the production of multiply charged ions which allows the detection of high molecular weight biomolecules on analyzers with limited  $m/z$  range. ESI cannot tolerate high salt concentrations except when the samples are desalted (e.g., by LC, see Chapter 2).<sup>83</sup>

ESI starts with analyte solution that is infused through a metal capillary (typically ~ 0.1 mm I.D.) held at an electric potential of several kV. The following discussion will be limited to positive ion mode as this is most commonly used for a wide range of applications. Charge separation takes place by removal of electrons from solution through oxidation (e.g.  $2\text{H}_2\text{O} \rightarrow \text{O}_2 + 4\text{H}^+ + 4\text{e}^-$ ). This will induce charge accumulation at the tip of the capillary, forming a Taylor cone that emits a mist of  $\mu\text{m}$ -sized droplets. The droplets will be charged with excess positive ions (such as  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{H}^+$ ,  $\text{NH}_4^+$ ). Acceleration potentials between the capillary exit and orifice will drive the droplets toward the ion sampling interface. During their journey, the droplets undergo evaporation assisted by the heating elements in most commercial ion sources. Evaporation enhances the charge density on the shrinking droplets until Coulombic repulsive forces exceed surface tension; this is known as the Rayleigh limit.<sup>84-85</sup> At this point, the number  $z_R$  of elementary charges  $e$  on the droplet is given by

$$z_R = \frac{8\pi\sqrt{\varepsilon_0\gamma R^3}}{e} \quad (1.10)$$

where  $R$  is the droplet radius,  $\varepsilon_0$  is the vacuum permittivity, and  $\gamma$  is the surface tension.

At the Rayleigh limit, the droplet becomes unstable and undergoes jet fission, resulting in the formation of smaller droplets. Repeated evaporation/fission cycles generate nanodroplets from which gas phase analyte ions are released. These ions will travel to the vacuum interface where any remaining solvent clusters will be removed by collisional activation.<sup>83, 86-88</sup>



**Figure 1.4.** Schematic representation of ESI source operated in positive ion mode.

### 1.6.1. Mechanism of Gas Phase Ion Release During ESI

The ESI mechanism remains controversial; three main models have been suggested to describe how protein ion is ejected from the droplet, known as charged residue model (CRM), ion evaporation model (IEM) and chain ejection model (CEM).

In the IEM, successive evaporation and fission events will increase the repulsion of charges on the surface of the droplet; this will allow the ejection of small analytes (Figure 1.5a) when the electrostatic repulsion allows the ions to overcome the free energy barrier for ejection.<sup>34, 89-91</sup> The IEM was initially suggested for small inorganic ions such as  $\text{Na}^+$  and  $\text{NH}_4^+$ . However, some authors suggest that the IEM also applies to macromolecules.<sup>92</sup>



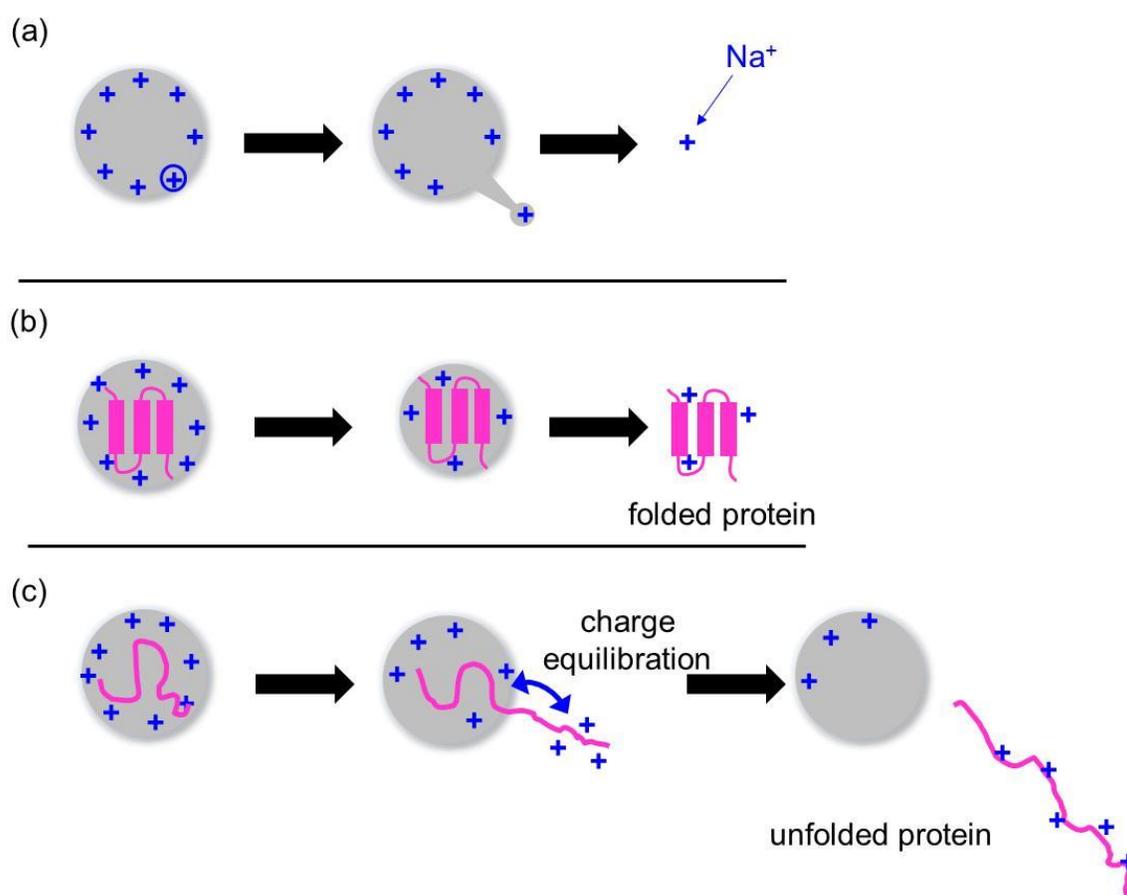
Ion release in IEM depends mainly on the ability of a solvated ion to cross the free energy barrier, so if the ion resides close to the liquid-vacuum surface, then this will enhance its ejection probability. De la Mora's work provided experimental support for the IEM<sup>93-94</sup> while MD simulations have shed light on how small solvated ions get released from the surface of the droplet.<sup>83,95</sup>

Large globular analytes such as natively folded proteins are believed to follow the CRM, where evaporation and Coulombic fission occur until a single analyte molecule remains inside the droplet. This molecule gets released into the gas phase upon droplet evaporation to dryness. Any remaining charges inside the vanishing droplet transfer to the analyte (Figure 1.5b).<sup>96</sup> MD simulations and experimental work support the CRM. The CRM/Rayleigh charge model shows good agreement between the expected final charge state of protein ions using equation 1.10 and the charge produced under experimental conditions.<sup>97-98</sup> This charge is exactly what would be expected for droplets that dry out until they have reached the same size as the protein. Very likely, salt clusters like  $[Na_nCl_{n-1}]$  are also formed by the CRM.<sup>99</sup>

A related framework is known as the combined charged residue-field emission model (CCRFEM). It envisions that the number of charges the protein holds is related to the release of low molecular weight (MW) charge carriers from the droplet, i.e., any ion that fails to leave the droplet will end up on the macromolecule. It can be argued that CCRFEM is very similar to the CRM but one difference between both models is that the final analyte charge state does not depend on the Rayleigh limit but on the rate of low MW ion departure from the droplet.<sup>100-101</sup>

One key factor that determines the mechanism whereby macromolecules are released from ESI nanodroplets is their conformation. In native proteins, hydrophobic amino acids are buried inside the core while hydrophilic groups are on the surface. These conditions favor the CRM. However, under denaturing conditions (i.e. high organic, low pH) the hydrophobic core is exposed to solvent. This is energetically unfavorable as described in section 1.2. To understand how disordered protein chains are released from the droplet, it is essential to examine the behavior of macromolecules such as polyethylene glycol (PEG) and polypropylene glycol (PPG). They are conceptually similar to unfolded proteins in being randomly coiled. Simulations using PEG as a model by Consta suggest that PEG travels to the surface of the droplet. One chain

terminus then erupts from the droplet and can remain attached until complete dryness of the droplet or it can separate from it.<sup>102-103</sup> Our laboratory proposed that a similar ESI mechanism applies to unfolded proteins named chain ejection model (CEM).<sup>83</sup> In the CEM, the protein travels to the droplet surface and then undergoes stepwise ejection via hydrated ‘tadpole-like’ conformers where the droplet carries an extended protein tail. This chain ejection is enhanced by electrostatic repulsion of the charge on the protruding tail and the remaining charges on the droplet along with hydrophobicity. CEM is mechanistically similar to CID processes of multi-subunit systems.<sup>104-105</sup> So far it was not feasible to simulate the CEM for unfolded proteins due to difficulties associated with the droplet size and charge migration. Details about how we were able to deal with those difficulties are discussed in Chapter 5.



**Figure 1.5.** Cartoon representation of proposed ESI models. a) The ion evaporation model is proposed for small ions, b) The charged residue model is suggested for globular folded proteins, c) The chain ejection model is envisioned for polymers and unfolded proteins.

### 1.6.2. Nanoelectrospray Ionization (NanoESI)

NanoESI is a variation of ESI that was developed in the mid-1990s by Wilm and Mann.<sup>106</sup> Both nanoESI and conventional ESI follow the same fundamental steps. NanoESI uses a lower flow rate, typically less than  $100 \text{ nL min}^{-1}$ , while conventional ESI operates in the  $\mu\text{L min}^{-1}$  range. The emitter tip diameter in nanoESI is typically on the order of  $1 \mu\text{m}$ . The analyte can be in contact with high voltage at the tip of the emitter, or with a metal (Pt) wire that is placed inside the needle. The need to initiate and maintain a constant flow rate can be achieved by using gentle  $\text{N}_2$  back-pressure. The low flow rate is responsible for the high ionization efficiency. One of the advantages of nanoESI is its higher salt tolerance. Because the initial droplets formed in nanoESI are smaller, the salt concentration after evaporative droplet shrinkage tends to be lower. NanoESI is also beneficial because of its much lower sample consumption.<sup>107-108</sup> It has been proposed that nanoESI is even more gentle than regular ESI for the preservation of non-covalent complexes<sup>109</sup> but these claims are not universally accepted.<sup>110-112</sup> In recent years, nanoESI with very small emitter diameters have been used to increase charge states,<sup>113</sup> to observe membrane proteins,<sup>114</sup> and to study proteins in physiological buffers by direct infusion nanoESI-MS.<sup>115</sup>

## 1.7. Protein ESI under Native and Denaturing Conditions

Electrosprayed proteins are always multiply charged. The number of charges can be described by charge state ( $z$ ), and the protein charge state distribution (CSD) depends critically on the protein conformation in solution.

Globular proteins are characterized by a hydrophobic core and a hydrophilic surface. Their tightly folded structure gives rise to CSDs centered at low charged states. Direct analysis of these globular folded proteins by MS was developed in the 1990s by Katta<sup>116</sup> and Ganem,<sup>117</sup> this was later known as 'native ESI-MS'. The name originates from the ability of the technique to preserve protein-ligand and protein-protein interactions in the gas phase.<sup>118-120</sup> Both Robinson<sup>121</sup> and Heck<sup>122</sup> have contributed to the development of native ESI-MS in investigating mega Dalton proteins, viruses or their capsid shells.<sup>123-128</sup> This field has benefited greatly from the development of nanoESI, the inception of novel mass analyzers, and the development of commercially available IMS devices.<sup>129-131</sup>

Controlling the instrumental parameters along with using suitable buffers helps in the analysis of protein and protein complexes under native ESI-MS. Typically, ammonium acetate is used as a buffer in native ESI-MS to mimic the cellular environment and decrease the probability of structural changes. However, ammonium acetate is not an actual ‘buffer’ at pH 7 because it is not composed of a weak acid and its conjugate base. Ammonium acetate is nonetheless used due to its volatile character which facilitates the formation of adduct-free protein ions. Conventional pH 7 buffers (e.g. phosphate buffer) cannot be used for direct infusion experiments because they would trigger extensive adduct formation.<sup>132</sup> Ammonium carbonate is a volatile pH 7 buffer, but it can destabilize proteins under ESI conditions.<sup>133-134</sup> Substantial evidence has been accumulated that globular proteins generated by native ESI are formed via the CRM. As outlined above, the low charge states of these ions are close to the Rayleigh charge of a protein-sized water droplet.

The ESI behavior of unfolded proteins is completely different, these species show wide CSDs that are shifted to much higher  $z$  values than in native ESI. Any non-covalent interactions are disrupted. This effect can be triggered by solution unfolding after exposure to acid,<sup>135-136</sup> base,<sup>137</sup> heat,<sup>138</sup> or reduction of disulfide bonds.<sup>139</sup> Instrumental parameters such as increasing the temperature inside the ion source can also induce unfolding.<sup>140</sup> As noted above, unfolded proteins are likely transferred into the gas phase via the CEM, but this concept remains controversial (see Chapter 5).

High protein charge states can also be achieved by supercharging agents (SCA) in both native and denatured protein solutions.<sup>141-143</sup> SCAs are added to the sample at low concentration that do not affect protein structure.<sup>144</sup> Nonetheless, SCAs increase protein charging during ESI. SCAs have a high dipole moment and low volatility.<sup>145</sup> They concentrate inside the ESI droplet, therefore the final nanodroplets contain significantly enhanced SCA concentrations compared to the initial bulk solution. The mechanism of supercharging is highly controversial.<sup>146</sup> (see Chapters 3 and 4). High protein charge states formed after unfolding and/or in the presence of SCAs enhance fragmentation in top-down experiments.<sup>147</sup> Also, increased charges are beneficial for mass analyzers that have limited  $m/z$ , or that have  $m/z$ -dependent performance characteristics (resolution and sensitivity) such as FT-ICR or orbitrap mass analyzers.

## 1.8. Proteins in the Gas Phase vs. in Solution

It is common to think of proteins as biomolecules that are surrounded by a buffered aqueous environment. Water is essential for proteins to function and to maintain the native conformation. The presence of salts along with water dipoles screens electrostatic interactions and decreases the magnitude of Coulombic interactions. In addition, water plays a role in protein folding by enhancing hydrophobic collapse. Ideally, experiments should be done under native conditions to avoid any structural changes. However, MS experiments involve desolvated protein ions in a vacuum, usually in charge states that are very different from those in solution. The absence of water will decrease hydrophobic interactions and electrostatic interactions (both attractive and repulsive) will be enhanced.<sup>148-149</sup>

Electrostatic interactions between two charges  $q_i$  and  $q_j$  that are separated by distance  $r$  having a potential  $V_{ij}$  are described by Coulomb's law:

$$V_{ij} = \frac{q_i q_j}{4\pi\epsilon_0 k_e r} \quad (1.12)$$

$\epsilon_0$  is the vacuum permittivity and  $k_e$  is the dielectric constant of the medium. Water has  $k_e \approx 80$ . Water dipoles arrange themselves around charges inside the medium. Therefore, interactions between  $q_1$  and  $q_2$  tend to be quite weak. In contrast, vacuum has  $k_e = 1$  so interactions between charges are much more pronounced.

In solution, pK<sub>a</sub> values of the main amino acids such as Lys (K), Arg (R), His (H), Glu (E), Asp (D), added to N-terminus and C-terminus govern the titration behavior and the net charge of the protein.<sup>150</sup> Generally, pK<sub>a</sub> characterizes the protonation status in solution which is affected by pH changes and the surrounding environment. At pH 7 Asp (pK<sub>a</sub>~ 4), Glu (pK<sub>a</sub>~4.4) and C-terminal carboxylic acids (pK<sub>a</sub> ~3.9) are expected to be negatively charged while Lys, Arg, and N-terminus are protonated.<sup>151</sup> In the gas phase, proton affinity or gas phase basicity controls the protonation of amino acids. Gas phase basicity indicates the transfer of H<sup>+</sup> from gaseous biomolecular ions to another gas phase species and vice versa. For electrosprayed ions of

the same charge states, unfolded proteins tend to have a higher gas phase basicity than folded ones due to electrostatic interactions in globular proteins.<sup>152</sup>

Some studies have attempted to explain the factors that affect the charging of protein in solution<sup>150</sup> or the gas phase. Suggestions include increase in the accessibility of possible charge sites,<sup>116, 153</sup> decreased of Coulombic repulsion between charges, disruption of interactions between COO<sup>-</sup> and positive sites<sup>154</sup> and formation of salt bridges.<sup>155</sup> Besides, ESI process plays a role in determining the charge state of native protein according to CRM, where the droplet transfers its remaining charges to contained macromolecules just at complete evaporation.

Despite the vastly different environment experienced by proteins in solution and in the gas phase, it is well documented that protein ions produced by “native ESI” generally retain solution-like properties. In other words, gas phase conformers appear to be kinetically trapped in conformations that resemble those in bulk solution. Evidence to support this view includes the observation of CCSs that are consistent with crystal structures<sup>15, 110, 127, 156-157</sup>, as well as the survival of intricate multi-component protein complexes in the gas phase.<sup>126, 158-160</sup>

## 1.9. Computer Simulations

Computational simulations enable us to study system properties and model phenomena that are difficult to observe experimentally. Simulations can act as a link between theory and experiment. To perform simulations, we need to use algorithms that provide numerical solutions for some of the fundamentals of physical laws. Recent advances related to software development and the ever-increasing speed of computers continue to push this entire field forward.

Exact solutions of the Schrödinger equation exist only for a small number of systems such as the particle in a box, harmonic oscillator, and hydrogen atom. Because of that, approximative *ab initio* methods have been developed to give the closest solution to the Schrödinger equation. These methods do not rely on experimental parameters, but they are extremely computationally expensive and can only be applied to small systems.

Molecular mechanics (MM) is a modeling tool that is used for simulating large molecules. MM is not based on quantum mechanics calculations. Instead, it uses classical methods to calculate the potential energy of the system. These methods are known as force field methods which contain a set of parameters that are derived from experimental data or *ab initio* methods.<sup>161</sup>

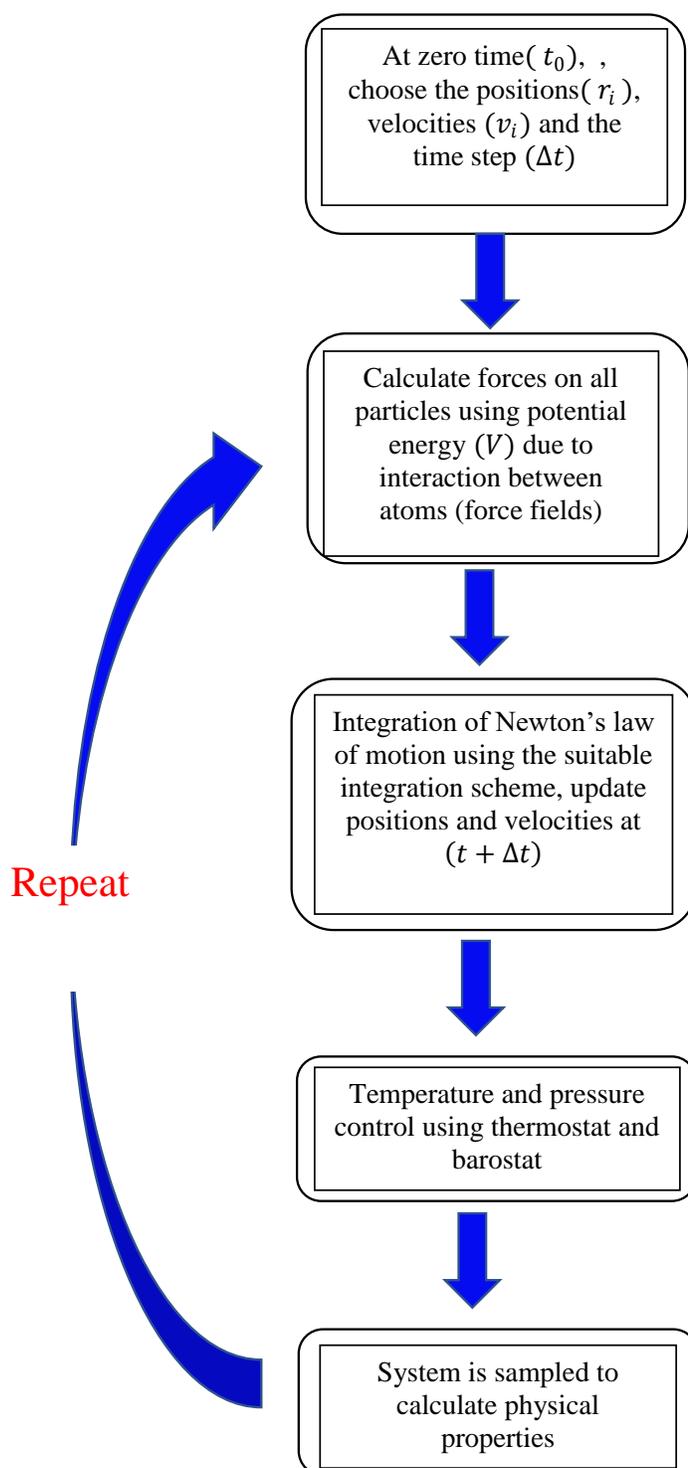
Monte Carlo methods generate different configurations for the system randomly, with relative occupancies that are governed by free energies. Monte Carlo methods can provide a good conformational sampling of proteins and other large systems, but they cannot study the system in a time-dependent fashion.<sup>161</sup>

## 1.10. Molecular Dynamics (MD) Simulations

The development of MD simulations started in the 50s with work by Alder, Wainwright, and Rahman.<sup>162-163</sup> Ever since the field has advanced rapidly. Today, MD simulations are being applied in many fields such as modeling of biomolecules, drug discovery and materials science.<sup>164</sup> MD simulations model the motion of atoms by solving the classical (Newtonian) laws of motion. Many MD methods employ the fact that the statistical ensemble average is equal to the time average of a single system according to the ergodic principle. The ensemble average

corresponds to a series of microstates of the system under study. The macroscopic properties of the system under equilibrium can be predicted from these microstates. There are different ensembles that can be modeled based on the experimental conditions. They are named according to what is kept constant for example, the *microcanonical ensemble* (NVE) can be used for isolated system simulations in which the number of particles, volume, and energy is kept constant; in *isobaric-isothermal ensemble* (NPT), the temperature and pressure are constant; the *canonical ensemble* (NVT) with constant number of particles, volume and temperature.<sup>165</sup>

A typical MD simulation has the following scheme:



### 1.10.1. Initial Coordinates and Velocities

As outlined in the above scheme, the first step in MD simulations is to find the starting positions and velocities. Three-dimensional X-ray structures are used to obtain the primary coordinates. Initial velocities can be assigned randomly from *Maxwell-Boltzmann* distribution at a specific temperature.<sup>166</sup>

$$p(v) = \sqrt{\frac{m_i}{2\pi k_B T}} \exp\left[-\frac{m_i v^2}{2k_B T}\right] \quad (1.13)$$

Where  $p(v)$  is the probability distribution of atoms having velocities between  $v$  and  $v + \Delta v$ ,  $v$  represents velocity in three directions,  $k_B$  is the Boltzmann constant,  $T$  is the temperature and  $m$  is the mass. The temperature of the system during simulation will not be constant due to randomization of velocity but this can be adjusted using thermostats (section 1.10.6)

### 1.10.2. Force Fields

The idea behind the use of forcefields for macromolecules is based on the work of Warshel, Levitt, and Karplus who shared the 2013 Chemistry Nobel Prize.<sup>167</sup> Force fields consist of parameters that are obtained from *ab-initio* methods or DFT. These calculations are quite expensive for large molecules such as proteins. Therefore, the interactions between atoms are modeled semi-empirically using MM force field. Force fields include terms that are derived from interactions between atoms in the microscopic system to form molecules. These interactions contribute to the total potential of the system. The total potential energy is derived from  $N$  interacting atoms as a function of their position  $V(\vec{r}_1 \dots \vec{r}_N)$ . From this  $V$  term the force acting on each atom  $i$  can be calculated as the gradient with respect to the atom's position

$$\vec{F}_i = -\nabla V(\vec{r}_1 \dots \vec{r}_N) \quad (1.14)$$

$V(\vec{r}_1 \dots \vec{r}_N)$  is the sum of bonded and non-bonded interactions at any point in time.

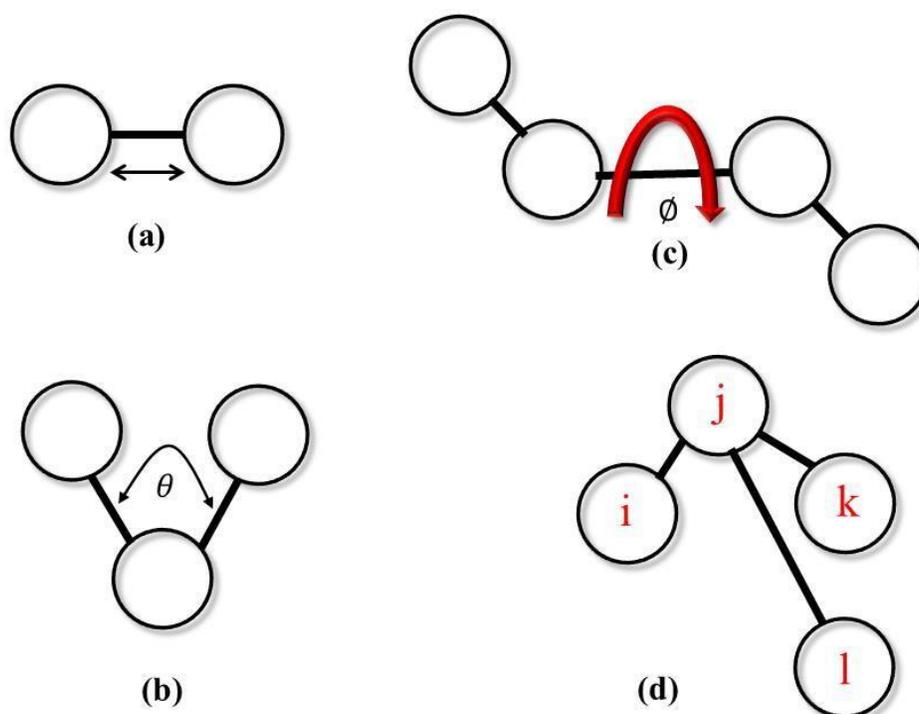
$$V_{total} = V_{bonded} + V_{non-bonded} \quad (1.15)$$

Bonded interactions ( $V_{bonded}$ ) arise from covalent bonds. They consist of a sum of terms related to bond stretching, change in angles, and torsion terms.

$$V_{bonded} = \sum_{bonds} \frac{1}{2} k_b (b - b_0)^2 + \sum_{angles} \frac{1}{2} k_\theta (\theta - \theta_0)^2 \quad (1.16)$$

$$+ \sum_{dihedrals} \frac{1}{2} k_n [1 + \cos(n\theta - \delta)] + \sum_{impropers} k_\psi (\psi - \psi_0)^2$$

In equation 1.16 the first term involves changes in the bond energy that is taken to have harmonic form,  $k_b$  is the force constant of the bond,  $b$  is current bond length,  $b_0$  is the equilibrium bond length. The second term represents energies associated with changes in bond angles. It is also using harmonic potentials, where  $k_\theta$  is the force constant of angle,  $\theta$  is the distorted angle,  $\theta_0$  is the equilibrium value.



**Figure 1.6.** Interactions in a typical forcefield between atoms a) Bond stretch, b) Angle rotation, c) Dihedral angle rotation, or torsion angle applies where two pairs of covalently bonded atoms are joined by another covalent bond. It is the angle between the planes formed by two atom pairs. d) Improper dihedral angle is the angle between the planes formed by atoms i, j, l and j, k, l.

Non-bonded interactions ( $V_{non-bonded}$ ) refer to contacts between atoms or molecules that are not directly bonded. These interactions are very essential for the stability of biological macromolecules. From a computational point of view, the most important ones are the short-range interactions which are described by *Lennard Jones* (LJ) potentials ( $V_{LJ}$ ) and electrostatic interactions represented by Coulomb potential.

*Lennard Jones* (LJ) potentials are weak and proportional to a distance of  $\frac{1}{d^6}$ , where  $A$  is the repulsion component which falls off with  $d^{-12}$  and  $B$  is the attractive component due to dipole-interactions that also degenerates after  $d^{-6}$ .

$$V_{LJ} = 4\epsilon \left( \frac{A}{d^{12}} - \frac{B}{d^6} \right) \quad (1.17)$$

In Coulomb potentials, the electrostatic energy is a function of charges on non-bonded atoms along with the distance separating them. Dielectric constant accounts for increasing or decreasing the strength of interactions (section 1.8).

Some of the most commonly used force fields are Optimized Potential for Liquid Simulations-All Atoms (OPLS/AA)<sup>168</sup>, Assisted Model Building with Energy Refinement (AMBER), and The Chemistry at Harvard Molecular Mechanics (CHARMM).<sup>169</sup>

### 1.10.3. Integration Algorithms

MD simulations study the time-dependent behavior of a microscopic system using classical laws of motion. This is achieved by integration of Newton's second law of motion.

$$\vec{F}_i = m_i a_i = m_i \frac{d^2 \vec{r}_i}{dt^2} = - \frac{\partial V(\vec{r}_i \dots \vec{r}_N)}{\partial r_i} \quad (1.18)$$

Where,  $\vec{F}_i$  is the net force acting on atom  $i$ ,  $a_i$  is the corresponding acceleration,  $m_i$  is the mass,  $\vec{r}_i$  is the position, and  $V$  is the potential energy.

$V$  is a function of the  $3N$  atomic positions ( $x, y, z$ ) of all  $N$  atoms in the simulation. Because of that there is no explicit solution to the equation of motion, and numerical algorithms must be used. These algorithms should allow for a long integration time step ( $\Delta t$ ), and to achieve that bond constraints such as SHAKE, LINCS and SETTLE are commonly used to eliminate fast vibrational bond motions.<sup>161</sup> The positions, velocities and accelerations of atoms at points in  $(t + \Delta t)$  or  $(t - \Delta t)$  can be approximated by a Taylor series expansion. After calculating the acceleration for each atom, coordinates are updated, and the process is repeated.

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \Delta v_i(t) \cdot \Delta t + \frac{\Delta t^2}{2} a_i(t) + O(\Delta t^3) \quad (1.19)$$

$$\vec{r}_i(t - \Delta t) = \vec{r}_i(t) - \Delta v_i(t) \cdot \Delta t + \frac{\Delta t^2}{2} a_i(t) + O(\Delta t^3) \quad (1.20)$$

The  $O$  indicates higher order terms that decays to zero and can be neglected

Adding and rearranging equations 1.19 and 1.20 gives the Verlet algorithm,

$$\vec{r}_i(t + \Delta t) \simeq 2\vec{r}_i(t) - \vec{r}_i(t - \Delta t) + \Delta t^2 a_i(t) \quad (1.21)$$

The Verlet algorithm is easy to apply and efficient, but an explicit velocity term is missing. A modification of the Verlet algorithm can solve the velocity problem; this is known as velocity Verlet algorithm.

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \Delta t v_i(t) + \frac{\Delta t^2}{2} a_i(t) \quad (1.22)$$

$$v_i(t + \Delta t) = v_i(t) + \frac{\Delta t}{2} (a_i(t) + a_i(t + \Delta t)) \quad (1.23)$$

The velocity Verlet algorithm calculates the position and velocity at the same time. Another important tool is the leapfrog algorithm that is commonly used in MD simulations due to simplicity. Velocity calculations are done at halftime step interval while the positions are calculated at full-time step. Within the leapfrog scheme, both positions and velocities leap over each other.

$$v_i\left(t + \frac{\Delta t}{2}\right) = v_i\left(t - \frac{\Delta t}{2}\right) + \Delta t a_i(t) \quad (1.24)$$

Velocities at time  $t + \frac{\Delta t}{2}$  are calculated from velocities at time  $\left(t - \frac{\Delta t}{2}\right)$  and accelerations at the time  $(t)$ .

The position is calculated in full-time step

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \Delta t v_i\left(t + \frac{\Delta t}{2}\right) \quad (1.25)$$

#### 1.10.4. Energy Minimization

At the beginning of each simulation, the internal configuration (usually obtained from X-ray or NMR structures) may involve unfavorable local contacts between atoms of the system. The purpose of energy minimization is to perform a simple MM relaxation procedure that will eliminate direct clashes and moves the system into a local energy minimum. The most common method for this purpose is the steepest descent algorithm, which is an iterative method to obtain an approximate solution.

$$\vec{r}_{n+1} = \vec{r}_n - K_n \nabla V(r(n)) \quad (1.26)$$

$\vec{r}_n$  is a vector representing the position of all particles in the system,  $K_n$  is the step size towards the lowest energy minima and  $\nabla V$  is the gradient of potential energy.<sup>161, 165</sup>

#### 1.10.5. Water Models

Water is very important as most chemical reactions take place in aqueous media. The quality of macromolecular simulations depends on the choice of water model. Although water molecules are very small there are different models to describe it.

Water models can be classified into rigid models where only non-bonded interactions considered and the bonded interactions are restrained. Polarizable models include explicit parameters for polarization. Rigid models are usually preferred for MD simulations of large systems because of their low computational cost. Rigid models differ in their interaction sites, e.g. the SPC (simple point charge model) and TIP3P (transferable intermolecular 3-point potential) where point charges are placed on each of the three H<sub>2</sub>O atoms. Both are computationally cheap and perform well in bulk solution. However, their surface tension properties are poor compared to experimentally measured values.<sup>170</sup> The TIP4P and TIP4P/2005 models have four interaction sites; the negative charge of the oxygen is moved slightly and is placed on a virtual point on the bisector of the HOH angle which improves electrostatic properties.<sup>170</sup> TIP4P/2005 is mostly used in the work described in the thesis as it gives the best representation of surface tension compared to all other commonly used models.<sup>171-172</sup>

### 1.10.6. Thermostats

Thermostat algorithms are indispensable in MD simulations to keep the temperature of the system close to the desired one for sampling in  $NVT$  or  $NPT$  ensemble

Temperature is related to the average kinetic energy or velocity (equation 1.27).<sup>161, 173</sup>

$$K_{avg} = \sum_{i=1}^N \frac{m_i v^2}{2} = \frac{1}{2} k_B T N_f \quad (1.27)$$

Where  $K_{avg}$  is the average kinetic energy,  $N_f$  is the number of degrees of freedom in 3-dimensional space,  $k_B$  is the Boltzmann constant,  $T$  is the temperature.

One of the ways to adjust temperature is to rescale velocity known as the velocity-rescaling temperature coupling. In this algorithm, the velocity is rescaled by multiplying the velocity of each particle with rescaling factor  $\lambda$ .

$$\lambda = \sqrt{\frac{T_{target}}{T_{current}}} \quad (1.28)$$

This is an easy and straightforward method but temperature fluctuations are expected. Another way to control the temperature is to couple the system to heat bath at a fixed temperature where the velocity is rescaled at each step by supplying or removing heat from the system like in the Berendsen thermostat.<sup>174</sup> The rate of temperature change is kept proportional to the temperature difference between the bath and system. However, the Berendsen thermostat represents a weak coupling of the system with a heat bath so the system will spend a long time to reach the desired temperature.

The Nosé-Hoover thermostat is an extended method algorithm which provides a rigorous way of a system and heat bath coupling. Both the system and the heat bath exchange kinetic energy. The average kinetic energy of the particles in the system does not change at each

time step while the temperature is kept constant by scaling of velocity. This can be represented by introducing an additional parameter ( $\xi$ ) to the equation of motion. This introduced parameter ( $\xi$ ) reflects the frictional or drag force which can slow down or accelerates the particles in the system until the desired temperature is reached. The Nosé-Hoover equations of motion showing an additional frictional term ( $\xi$ ) :

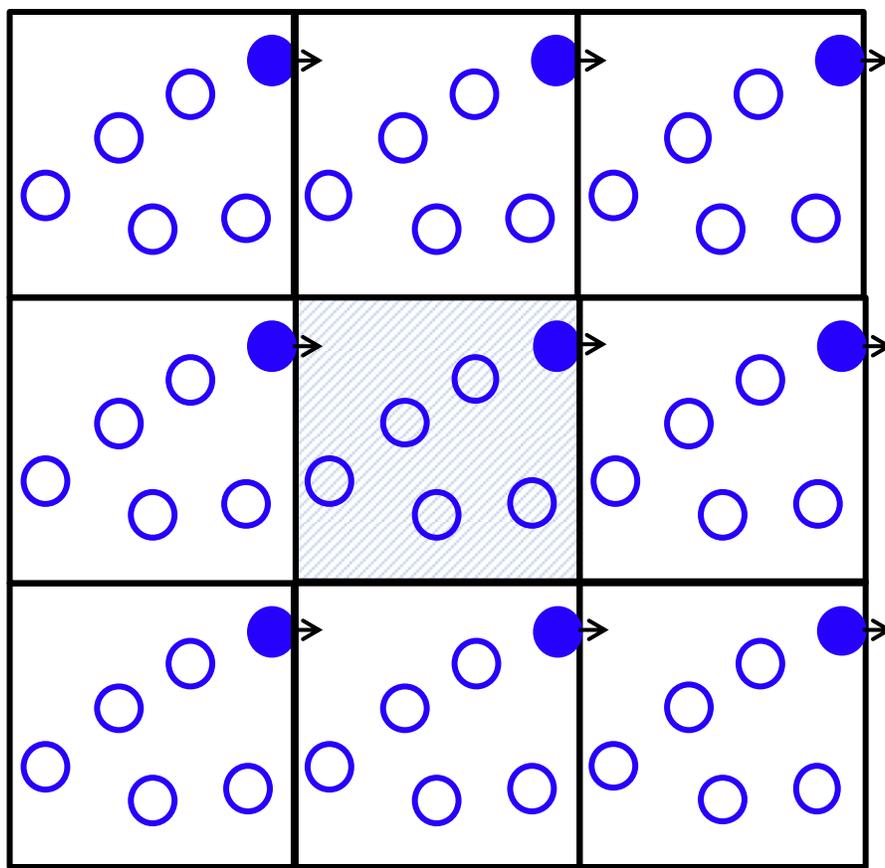
$$\xi = \frac{\sum_i P_i/m_i - 3Nk_B T}{Q} \quad (1.29)$$

Equation 1.29 shows  $Q$  which determines the strength of coupling between the heat bath and the system,  $T$  is the desired temperature and  $P_i$  is the momentum. This thermostat reproduces the  $NVT$  ensemble, time-reversible but computationally expensive.<sup>161</sup>

### 1.10.7. Periodic Boundary Conditions (PBC)

Modeling the bulk properties and the behavior of massive systems such as proteins or DNA is done in an explicit solvent which increases their modeling complexity. Periodic boundary conditions can be used to avoid the finite-size problem and to eliminate surface artifacts.<sup>175</sup>

PBC use a simulation box that contains the system under study, and this box is repeated on all sides such that the simulated system is surrounded by identical copies of itself. Any particle that leaves the box from one side, will re-enter the box from the opposite side. In this way, the number of particles in the simulation box remains the same, and there are no physical boundaries or surface molecules. In protein simulations, ions such as  $\text{Na}^+$  or  $\text{Cl}^-$  are used to compensate the intrinsic charge of protein and neutralize the box. Different box shapes can be used, such as cubic, hexagonal, etc.<sup>176</sup> It is necessary to choose a suitable box shape that is compatible with the geometry of the system and also minimize the number of solvent molecules to save the computer time.



**Figure 1.7.** Periodic boundary conditions in two dimensions. The shaded middle cell represents the simulation cell, where the filled blue atom moves out of the box and all images move in the same manner.

### 1.10.8. Treatment of Non-Bonded Interactions

Non-bonded interactions are the most time-consuming part of MD simulations. To reduce computational cost, cut-offs are used. This means that the interaction between two atoms separated by more than the cut-off distance is ignored. For short range (LJ) interactions, the cut-off is simply introduced to potential ( $V_{LJ}$ ) and any interactions beyond this cut-off are set to zero without causing noticeable perturbations. A different approach is needed for the Coulomb potential which represents a long-range interaction. In PBC simulations this problem is solved by using Particle Mesh Ewald (PME) summation. This method is based on fast Fourier transform algorithms that calculate the potential for charges beyond a certain cutoff and maps them on a mesh.<sup>177</sup> A disadvantages of using PBCs with PME is that the approach is only suitable for neutral systems as applying it to the charged system causes artifacts.<sup>178-180</sup>

Using PBC/PME with cutoffs is not suitable for gas phase simulations due to the strength of electrostatic interactions in the vacuum, plus the fact that the vacuum ions (or droplets) considered below carry a net charge. Also, periodic images would feel each other due to Coulomb interactions. Nowadays, the speed of the modeling is controlled by graphics processing units (GPU). Unfortunately, widely used MD packages such as Gromacs that run with GPU acceleration were designed for use with PBC/PME and cutoffs. This problem can be circumvented by using a ‘pseudo-PBC’ method that allows gas phase simulations using GPU-acceleration without PBC/PME artifacts. Generally, the simulation system under study is placed in the center of the PBC box with large dimensions ( $999.9\text{nm}^3$ ), PME is switched off, cutoffs are set less than the dimensions of the box. The atoms contained inside the box will then interact among themselves, while interactions with periodic images are absent.<sup>181</sup> Effectively, this method provides an environment that is indistinguishable from a genuine vacuum environment without cutoffs.

### 1.10.9. Neighbor Lists

The use of cutoff may not reduce the time spent to calculate non-bonded interactions because the distance between every pair of atoms must be computed to decide whether they are close enough to calculate their interaction energy. This pair distance calculations can be time-consuming, but the problem can be addressed by using a Verlet neighbor list. This list sorts all atoms that are potential partners and lie within the cutoff distance or slightly further from the cutoff distance. Initially, the neighbor list is constructed for each pair of particles. Only the pairs in the list are checked for force calculations for some integration steps, if the particle moves by more than the cutoff distance then an algorithm automatically update the neighbor list in a way that it is not too fast or too slow.<sup>161, 175</sup>

## 1.11. Scope of the Thesis

In this dissertation, we use ESI-MS experiments and MD simulations to answer several controversial questions related to protein ESI.

Sample preparation is one of the most challenging aspects of protein ESI-MS. The presence of salts like NaCl can severely degrade the quality of the data obtained. Even after sample cleanup using different desalting techniques, the presence of residual salt can still be problematic. The mechanism of salt interferences in protein ESI-MS is not fully understood. Previous work relied on equilibrium-partitioning and charge competition to explain the behavior of mono-cations.<sup>86, 182</sup> In Chapter 2, we investigate the concept of charge competition between proteins and cations like tetrabutylammonium chloride and CsCl and we clarify that the concept of salt-induced signal suppression has different facets that have been under-appreciated in the previous literature.

Supercharging agents offer an attractive approach for increasing the protein charge under “native” ESI conditions, where SCAs are added to the bulk solution in low concentration where they do not affect protein structure or stability. Several proposals exist in the literature that aim to explain how SCAs work. These include surface tension effects and protein unfolding within droplets, but none of these explanation attempts offers a consistent picture. Chapters 3 and 4 scrutinize the native ESI supercharging mechanism from a new perspective. In Chapter 3, we uncover the first atomistic view of the supercharging mechanism. Our findings indicate that protein supercharging takes place due to “charge trapping” afforded by a SCA peripheral layer. This mechanism is further tested and confirmed in Chapter 4, where we use crown ether (18C6) to test the charge trapping hypothesis on both proteins and dendrimers.

Chapter 5 provides a detailed study about the behavior of unfolded proteins in ESI nanodroplets. For the first time, we apply a combination of MD simulations and ESI-MS/IMS to address this topic. We show unequivocally that the CEM is the actual ESI mechanism for acid-unfolded proteins in high charge states.

Finally, Chapter 6 highlights possible future research directions, outlining how the combination of MD simulations and ESI-IM/MS will continue to advance the field.

## 1.12 References

1. Brooks, C. L., *Proteins : a theoretical perspective of dynamics, structure, and thermodynamics*. John Wiley: New York, **1990**.
2. Voet, D., Pratt, C.W, *Fundamentals of biochemistry*. Wiley: Chichester, 2016.
3. Dill, K. A.; MacCallum, J. L., *Science* **2012**, 338, 1042-1046.
4. Nick Pace, C.; Scholtz, J. M.; Grimsley, G. R., *FEBS Letters* **2014**, 588, 2177-2184.
5. Sticke, D. F.; Presta, L. G.; Dill, K. A.; Rose, G. D., *Journal of Molecular Biology* **1992**, 226, 1143-1159.
6. Chen, J.; Stites, W. E., *Biochemistry* **2001**, 40, 15280-15289.
7. Hendsch, Z. S.; Tidor, B., *Protein Science : A Publication of the Protein Society* **1994**, 3, 211-226.
8. Wedemeyer, W. J.; Welker, E.; Narayan, M.; Scheraga, H. A., *Biochemistry* **2000**, 39, 4207-4216.
9. Custer, G. S.; Das, P.; Matysiak, S., *The Journal of Physical Chemistry B* **2017**, 121, 2731-2738.
10. Tanford, C., *Journal of the American Chemical Society* **1962**, 84, 4240-4247.
11. Shi, Y., *Cell* **159**, 995-1014.
12. Woolfson, M. M., *Physica Scripta* **2018**, 93, 032501.
13. Frauenfelder, H.; Sligar, S.; Wolynes, P., *Science* **1991**, 254, 1598-1603.
14. Carugo, O., *BMC Bioinformatics* **2018**, 19, 61.
15. Wyttenbach, T.; Pierson, N. A.; Clemmer, D. E.; Bowers, M. T., *Annual Review of Physical Chemistry* **2014**, 65, 175-196.
16. Bill, R. M.; Henderson, P. J. F.; Iwata, S.; Kunji, E. R. S.; Michel, H.; Neutze, R.; Newstead, S.; Poolman, B.; Tate, C. G.; Vogel, H., *Nature Biotechnology* **2011**, 29, 335.
17. Fenwick, R. B.; van den Bedem, H.; Fraser, J. S.; Wright, P. E., *Proceedings of the National Academy of Sciences* **2014**, 111, E445-E454.
18. Becker, E. D., *Analytical Chemistry* **1993**, 65, 295A-302A.
19. Purcell, E. M.; Torrey, H. C.; Pound, R. V., *Physical Review* **1946**, 69, 37-38.
20. Beasley, S. A.; Hristova, V. A.; Shaw, G. S., *Proceedings of the National Academy of Sciences* **2007**, 104, 3095-3100.

21. Sprangers, R.; Kay, L. E., *Nature* **2007**, *445*, 618.
22. Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K., *Proceedings of the National Academy of Sciences* **1997**, *94*, 12366-12371.
23. Raunser, S., *Angew. Chem. Int. Ed.* **2017**, *56*, 16450-16452.
24. Cheng, Y. F.; Glaeser, R. M.; Nogales, E., *Cell* **2017**, *171*, 1229-1231.
25. Mei, K.; Li, Y.; Wang, S.; Shao, G.; Wang, J.; Ding, Y.; Luo, G.; Yue, P.; Liu, J.-J.; Wang, X.; Dong, M.-Q.; Wang, H.-W.; Guo, W., *Nature Structural & Molecular Biology* **2018**, *25*, 139-146.
26. Mentes, A.; Huehn, A.; Liu, X.; Zwolak, A.; Dominguez, R.; Shuman, H.; Ostap, E. M.; Sindelar, C. V., *Proceedings of the National Academy of Sciences* **2018**, *115*, 1292-1297.
27. Lykkegaard, M. K.; Ehlerding, A.; Hvelplund, P.; Kadhane, U.; Kirketerp, M.-B. S.; Nielsen, S. B.; Panja, S.; Wyer, J. A.; Zettergren, H., *Journal of the American Chemical Society* **2008**, *130*, 11856-11857.
28. Schmid, F.-X., Biological Macromolecules: UV - visible Spectrophotometry. In *eLS*.
29. Sharon, M. K.; Nicholas, C. P., *Current Protein & Peptide Science* **2000**, *1*, 349-384.
30. Padmanabhan, S.; Marqusee, S.; Ridgeway, T.; Laue, T. M.; Baldwin, R. L., *Nature* **1990**, *344*, 268-270.
31. Chen, Y.; Barkley, M. D., *Biochemistry* **1998**, *37*, 9976-9982.
32. Shastry, M. C. R.; Roder, H., *Nature Structural Biology* **1998**, *5*, 385.
33. de Hoffmann, E.; Stroobant, V., *Mass Spectrometry Principles and Applications*. John Wiley: New York, **2013**.
34. Fenn, J. B., *Angew. Chem. Int. Ed.* **2003**, *42*, 3871-3894.
35. Karas, M.; Bachmann, D.; Hillenkamp, F., *Analytical Chemistry* **1985**, *57*, 2935-2939.
36. Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F., *International Journal of Mass Spectrometry and Ion Processes* **1987**, *78*, 53-68.
37. Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T., *Analytical Chemistry* **1991**, *63*, 1193A-1203A.
38. Renato, Z.; Richard, K., *Mass Spectrometry Reviews* **1998**, *17*, 337-366.
39. Caughlin, S.; Maheshwari, S.; Agca, Y.; Agca, C.; Harris, A. J.; Jurcic, K.; Yeung, K. K. C.; Cechetto, D. F.; Whitehead, S. N., *Biochimica et Biophysica Acta (BBA) - General Subjects* **2018**, *1862*, 1327-1338.
40. Cornett, D. S.; Reyzer, M. L.; Chaurand, P.; Caprioli, R. M., *Nature Methods* **2007**, *4*, 828.

41. Igor. A. Kaltashov, S., J. Eyles, **2005**.
42. Verentchikov, A. N.; Ens, W.; Standing, K. G., *Analytical Chemistry* **1994**, *66*, 126-133.
43. Xian, F.; Hendrickson, C. L.; Marshall, A. G., *Analytical Chemistry* **2012**, *84*, 708-719.
44. V., C. I.; V., L. A.; A., T. B., *Journal of Mass Spectrometry* **2001**, *36*, 849-865.
45. Chait, B. T., *Science* **2006**, *314*, 65-66.
46. Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W., *Journal of the American Chemical Society* **1999**, *121*, 806-812.
47. McLafferty, F., *Science* **1981**, *214*, 280-287.
48. Viswanatham, K.; T., C. B.; Steven, C., *Rapid Communications in Mass Spectrometry* **1991**, *5*, 214-217.
49. E., W. T.; R., E. J., *Mass Spectrometry Reviews* **2006**, *25*, 158-170.
50. Roder, H.; Elöve, G. A.; Englander, S. W., *Nature* **1988**, *335*, 700-704.
51. Zhongqi, Z.; L., S. D., *Protein Science* **1993**, *2*, 522-531.
52. Engen, J. R., *Analytical Chemistry* **2009**, *81*, 7870-7875.
53. Konermann, L.; Pan, J.; Liu, Y.-H., *Chemical Society Reviews* **2011**, *40*, 1224-1234.
54. Lars, K.; B., S. B.; Yan, P.; Xin, T., *Mass Spectrometry Reviews* **2010**, *29*, 651-667.
55. Wang, L.; Chance, M. R., *Analytical Chemistry* **2011**, *83*, 7234-7241.
56. Kiselar, J. G.; Chance, M. R., *Journal of mass spectrometry : JMS* **2010**, *45*, 1373-1382.
57. Hambly, D. M.; Gross, M. L., *Journal of the American Society for Mass Spectrometry* **2005**, *16*, 2057-2063.
58. Gau, B. C.; Sharp, J. S.; Rempel, D. L.; Gross, M. L., *Analytical Chemistry* **2009**, *81*, 6563-6571.
59. Vahidi, S.; Konermann, L., *Journal of The American Society for Mass Spectrometry* **2016**, *27*, 1-9.
60. Stocks, B. B.; Konermann, L., *Analytical Chemistry* **2009**, *81*, 20-27.
61. Stocks, B. B.; Konermann, L., *Journal of Molecular Biology* **2010**, *398*, 362-373.
62. Stocks, B. B.; Sarkar, A.; Wintrode, P. L.; Konermann, L., *Journal of Molecular Biology* **2012**, *423*, 789-799.
63. Brodie, N. I.; Makepeace, K. A. T.; Petrotchenko, E. V.; Borchers, C. H., *Journal of Proteomics* **2015**, *118*, 12-20.
64. Rappsilber, J., *Journal of Structural Biology* **2011**, *173*, 530-540.

65. Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R., *Molecular & Cellular Proteomics : MCP* **2010**, *9*, 1634-1649.
66. Brodie, N. I.; Popov, K. I.; Petrotchenko, E. V.; Dokholyan, N. V.; Borchers, C. H., *Science Advances* **2017**, *3*, 1-8.
67. Rozbeský, D.; Rosůlek, M.; Kukačka, Z.; Chmelík, J.; Man, P.; Novák, P., *Analytical Chemistry* **2018**, *90*, 1104-1113.
68. Yang, B.; Wu, Y.-J.; Zhu, M.; Fan, S.-B.; Lin, J.; Zhang, K.; Li, S.; Chi, H.; Li, Y.-X.; Chen, H.-F.; Luo, S.-K.; Ding, Y.-H.; Wang, L.-H.; Hao, Z.; Xiu, L.-Y.; Chen, S.; Ye, K.; He, S.-M.; Dong, M.-Q., *Nature Methods* **2012**, *9*, 904-910.
69. May, J. C.; McLean, J. A., *Analytical Chemistry* **2015**, *87*, 1422-1436.
70. Gillig, K. J.; Ruotolo, B.; Stone, E. G.; Russell, D. H.; Fuhrer, K.; Gonin, M.; Schultz, A. J., *Analytical Chemistry* **2000**, *72*, 3965-3971.
71. Myung, S.; Wiseman, J. M.; Valentine, S. J.; Takáts, Z.; Cooks, R. G.; Clemmer, D. E., *The Journal of Physical Chemistry B* **2006**, *110*, 5045-5051.
72. von Helden, G.; Wyttenbach, T.; Bowers, M. T., *Science* **1995**, *267*, 1483-1485.
73. Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F., *Journal of the American Chemical Society* **1995**, *117*, 10141-10142.
74. Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill, H. H., *Journal of Mass Spectrometry* **2008**, *43*, 1-22.
75. Cumeras, R.; Figueras, E.; Davis, C. E.; Baumbach, J. I.; Gràcia, I., *The Analyst* **2015**, *140*, 1376-1390.
76. Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H., *Rapid Communications in Mass Spectrometry* **2004**, *18*, 2401-2414.
77. Giles, K.; Williams, J. P.; Campuzano, I., *Rapid Communications in Mass Spectrometry* **2011**, *25*, 1559-1566.
78. Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V., *Nat. Protocols* **2008**, *3*, 1139-1152.
79. Wyttenbach, T.; von Helden, G.; Batka, J. J.; Carlat, D.; Bowers, M. T., *Journal of the American Society for Mass Spectrometry* **1997**, *8*, 275-282.
80. Shvartsburg, A. A.; Jarrold, M. F., *Chemical Physics Letters* **1996**, *261*, 86-91.

81. Ewing, S. A.; Donor, M. T.; Wilson, J. W.; Prell, J. S., *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 587-596.
82. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B., *The Journal of Chemical Physics* **1968**, *49*, 2240-2249.
83. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., *Analytical Chemistry* **2013**, *85*, 2-9.
84. Rayleigh, L., *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1882**, *14*, 184-186.
85. Duft, D.; Achtzehn, T.; Müller, R.; Huber, B. A.; Leisner, T., *Nature* **2003**, *421*, 128.
86. Kebarle, P.; Tang, L., *Analytical Chemistry* **1993**, *65*, 972A-986A.
87. Van Berkel, G. J.; Kertesz, V., *Analytical Chemistry* **2007**, *79*, 5510-5520.
88. Kebarle, P.; Verkerk, U. H., **2012**, 1-48.
89. Kebarle, P.; Verkerk, U. H., *Mass Spectrometry Reviews* **2009**, *28*, 898-917.
90. Iribarne, J. V.; Thomson, B. A., *The Journal of Chemical Physics* **1976**, *64*, 2287-2294.
91. Labowsky, M., *Rapid Communications in Mass Spectrometry* **2010**, *24*, 3079-3091.
92. Nguyen, S.; Fenn, J. B., *Proceedings of the National Academy of Sciences* **2007**, *104*, 1111-1117.
93. Gamero - Castaño, M.; Mora, J. F. d. l., *Journal of Mass Spectrometry* **2000**, *35*, 790-803.
94. Gamero-Castaño, M.; Mora, J. F. d. l., *The Journal of Chemical Physics* **2000**, *113*, 815-832.
95. Ahadi, E.; Konermann, L., *Journal of the American Chemical Society* **2011**, *133*, 9354-9363.
96. McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L., *Journal of the American Chemical Society* **2015**, *137*, 12667-12676.
97. Fernandez de la Mora, J., *Analytica Chimica Acta* **2000**, *406*, 93-104.
98. Larriba, C.; de la Mora, J. F.; Clemmer, D. E., *Journal of The American Society for Mass Spectrometry* **2014**, *25*, 1332-1345.
99. Konermann, L.; McAllister, R. G.; Metwally, H., *The Journal of Physical Chemistry B* **2014**, *118*, 12025-12033.

100. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L., *Analytical Chemistry* **2009**, *81*, 369-377.
101. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L., *Journal of the American Chemical Society* **2008**, *130*, 6926-6927.
102. Chung, J. K.; Consta, S., *The Journal of Physical Chemistry B* **2012**, *116*, 5777-5785.
103. Consta, S.; Oh, M. I.; Sharawy, M.; Malevanets, A., *The Journal of Physical Chemistry A* **2018**, *122*, 5239–5250 .
104. Pagel, K.; Hyung, S.-J.; Ruotolo, B. T.; Robinson, C. V., *Analytical Chemistry* **2010**, *82*, 5363-5372.
105. Popa, V.; Trecroce, D. A.; McAllister, R. G.; Konermann, L., *The Journal of Physical Chemistry B* **2016**, *120*, 5114-5124.
106. Wilm, M.; Mann, M., *Analytical Chemistry* **1996**, *68*, 1-8.
107. Rosati, S.; Yang, Y.; Barendregt, A.; Heck, A. J. R., *Nat. Protocols* **2014**, *9*, 967-976.
108. Rose, R. J.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A. J. R., *Nat Meth* **2012**, *9*, 1084-1086.
109. Benesch, J. L. P.; Ruotolo, B. T.; Simmons, D. A.; Robinson, C. V., *Chem.Rev* **2007**, *107*,3544-3567.
110. Sun, Y.; Vahidi, S.; Sowole, M. A.; Konermann, L., *Journal of The American Society for Mass Spectrometry* **2015**, *27*, 31-40.
111. Loo, J. A., *International Journal of Mass Spectrometry* **2000**, *200*, 175-186.
112. Liu, J.; Konermann, L., *Journal of The American Society for Mass Spectrometry* **2011**, *22*, 408-417.
113. Fisher, C. M.; Kharlamova, A.; McLuckey, S. A., *Analytical Chemistry* **2014**, *86*, 4581-4588.
114. Susa, A. C.; Lippens, J. L.; Xia, Z.; Loo, J. A.; Campuzano, I. D. G.; Williams, E. R., *Journal of The American Society for Mass Spectrometry* **2018**, *29*, 203-206.
115. Susa, A. C.; Xia, Z.; Williams, E. R., *Analytical Chemistry* **2017**, *89*, 3116-3122..
116. Katta, V.; Chait, B. T., *Journal of the American Chemical Society* **1991**, *113*, 8534-8535.
117. Ganem, B.; Li, Y. T.; Henion, J. D., *Journal of the American Chemical Society* **1991**, *113*, 6294-6296.

118. Leney, A. C.; Heck, A. J. R., *Journal of The American Society for Mass Spectrometry* **2017**, *28*, 5-13.
119. Campuzano, I. D. G.; Netirojjanakul, C.; Nshanian, M.; Lippens, J. L.; Kilgour, D. P. A.; Van Orden, S.; Loo, J. A., *Analytical Chemistry* **2018**, *90*, 745-751.
120. Chandler, S. A.; Benesch, J. L. P., *Current Opinion in Chemical Biology* **2018**, *42*, 130-137.
121. Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M., *Journal of the American Chemical Society* **1996**, *118*, 8646-8653.
122. Snijder, J.; Rose, R. J.; Veessler, D.; Johnson, J. E.; Heck, A. J. R., *Angew. Chem. Int. Ed.* **2013**, *52*, 4020-4023.
123. Gault, J.; Donlan, J. A. C.; Liko, I.; Hopper, J. T. S.; Gupta, K.; Housden, N. G.; Struwe, W. B.; Marty, M. T.; Mize, T.; Bechara, C.; Zhu, Y.; Wu, B.; Kleanthous, C.; Belov, M.; Damoc, E.; Makarov, A.; Robinson, C. V., *Nature methods* **2016**, *13*, 333-336.
124. Snijder, J.; van de Waterbeemd, M.; Damoc, E.; Denisov, E.; Grinfeld, D.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A.; Heck, A. J. R., *Journal of the American Chemical Society* **2014**, *136*, 7295-7299.
125. Snijder, J.; Kononova, O.; Barbu, I. M.; Utrecht, C.; Rurup, W. F.; Burnley, R. J.; Koay, M. S. T.; Cornelissen, J. J. L. M.; Roos, W. H.; Barsegov, V.; Wuite, G. J. L.; Heck, A. J. R., *Biomacromolecules* **2016**, *17*, 2522-2529.
126. Marcoux, J.; Robinson, Carol V., *Structure* **2013**, *21*, 1541-1550.
127. Seo, J.; Hoffmann, W.; Warnke, S.; Bowers, M. T.; Pagel, K.; von Helden, G., *Angew. Chem. Int. Ed.* **2016**, *55*, 14173-14176.
128. Heck, A. J. R., *Nature Methods* **2008**, *5*, 927-933.
129. Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F., *Journal of the American Chemical Society* **1997**, *119*, 2240-2248.
130. Vahidi, S.; Stocks, B. B.; Konermann, L., *Analytical Chemistry* **2013**, *85*, 10471-10478.
131. Chen, S.-H.; Russell, D. H., *Journal of The American Society for Mass Spectrometry* **2015**, *26*, 1433-1443.
132. Konermann, L., *Journal of The American Society for Mass Spectrometry* **2017**, *28*, 1827-1835.

133. Hedges, J. B.; Vahidi, S.; Yue, X.; Konermann, L., *Analytical Chemistry* **2013**, *85*, 6469-6476.
134. Xiao, Y.; Konermann, L., *Protein Science* **2015**, *24*, 1247-1256.
135. Kharlamova, A.; Prentice, B. M.; Huang, T.-Y.; McLuckey, S. A., *Analytical Chemistry* **2010**, *82*, 7422-7429.
136. Banerjee, S., *Journal of Mass Spectrometry* **2013**, *48*, 193-204.
137. Kharlamova, A.; McLuckey, S. A., *Analytical chemistry* **2011**, *83*, 431-437.
138. Wang, G.; Abzalimov, R. R.; Kaltashov, I. A., *Analytical Chemistry* **2011**, *83*, 2870-2876.
139. Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D., *Analytical Chemistry* **1990**, *62*, 693-698.
140. El-Baba, T. J.; Woodall, D. W.; Raab, S. A.; Fuller, D. R.; Laganowsky, A.; Russell, D. H.; Clemmer, D. E., *Journal of the American Chemical Society* **2017**, *139*, 6306-6309
141. Iavarone, A. T.; Williams, E. R., *Journal of the American Chemical Society* **2003**, *125*, 2319-2327.
142. Lomeli, S. H.; Peng, I. X.; Yin, S.; Ogorzalek Loo, R. R.; Loo, J. A., *Journal of the American Society for Mass Spectrometry* **2011**, *21*, 127-131.
143. Lomeli, S. H.; Yin, S.; Ogorzalek Loo, R. R.; Loo, J. A., *Journal of the American Society for Mass Spectrometry* **2009**, *20*, 593-596.
144. Hall, Z.; Robinson, C. V., *Journal of The American Society for Mass Spectrometry* **2012**, *23*, 1161-1168.
145. Zenaidee, M. A.; Donald, W. A., *Analyst* **2015**, *140*, 1894-1905.
146. Ogorzalek Loo, R. R.; Lakshmanan, R.; Loo, J. A., *Journal of The American Society for Mass Spectrometry* **2014**, *25*, 1675-1693.
147. Sterling, H. J.; Williams, E. R., *Analytical Chemistry* **2010**, *82*, 9050-9057.
148. Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R., *International Journal of Mass Spectrometry* **2002**, *216*, 1-27.
149. Meyer, T.; Gabelica, V.; Grubmüller, H.; Orozco, M., *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2013**, *3*, 408-425.
150. Guevremont, R.; Siu, K. W. M.; Le Blanc, J. C. Y.; Berman, S. S., *Journal of the American Society for Mass Spectrometry* **1992**, *3*, 216-224.

151. Pace, C. N.; Grimsley, G. R.; Scholtz, J. M., *The Journal of Biological Chemistry* **2009**, *284*, 13285-13289.
152. Schnier, P. D.; Gross, D. S.; Williams, E. R., *Journal of the American Chemical Society* **1995**, *117*, 6747-6757.
153. Chowdhury, S. K.; Katta, V.; Chait, B. T., *Journal of the American Chemical Society* **1990**, *112*, 9012-9013.
154. Rita, G., *Journal of Mass Spectrometry* **2003**, *38*, 11-15.
155. Loo, R. R. O.; Loo, J. A., *Journal of The American Society for Mass Spectrometry* **2016**, *27*, 975-990.
156. Shelimov, K. B.; Jarrold, M. F., *Journal of the American Chemical Society* **1997**, *119*, 2987-2994.
157. Scarff, C. A.; Thalassinou, K.; Hilton, G. R.; Scrivens, J. H., *Rapid Communications in Mass Spectrometry* **2008**, *22*, 3297-3304.
158. Heuvel, R. H. H. v. d.; Heck, A. J. R., *Current Opinion in Chemical Biology* **2004**, *8*, 519-526.
159. Kuhlen, L.; Abrusci, P.; Johnson, S.; Gault, J.; Deme, J.; Caesar, J.; Dietsche, T.; Mebrhatu, M. T.; Ganief, T.; Macek, B.; Wagner, S.; Robinson, C. V.; Lea, S. M., *Nature Structural & Molecular Biology* **2018**, *25*, 583-590.
160. van de Waterbeemd, M.; Tamara, S.; Fort, K. L.; Damoc, E.; Franc, V.; Bieri, P.; Itten, M.; Makarov, A.; Ban, N.; Heck, A. J. R., *Nature Communications* **2018**, *9*, 2493.
161. Leach, A. R., *Molecular modelling : principles and applications*. Pearson/Prentice Hall: Harlow, **2009**.
162. Alder, B. J.; Wainwright, T. E., *The Journal of Chemical Physics* **1957**, *27*, 1208-1209.
163. Rahman, A., *Physical Review* **1964**, *136*, A405-A411.
164. Karplus, M.; McCammon, J. A., *Nature Structural Biology* **2002**, *9*, 646.
165. Frenkel, D.; Smit, B., *Understanding molecular simulation from algorithms to applications*. Academic Press : Elsevier e-books: San Diego, **2014**.
166. Field, M. J., *A practical introduction to the simulation of molecular systems*. Cambridge Univ. Press, **2009**.
167. Hodak, H., *Journal of Molecular Biology* **2014**, *426*, 1-3.

168. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., *The Journal of Physical Chemistry B* **2001**, *105*, 6474-6487.
169. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M., *The Journal of Physical Chemistry B* **1998**, *102*, 3586-3616.
170. Abascal, J. L. F.; Vega, C., *The Journal of Chemical Physics* **2005**, *123*, 234505.
171. Caleman, C.; Spoel, D. v. d., *The Journal of Chemical Physics* **2006**, *125*, 154508.
172. Vega, C.; Abascal, J. L.; Conde, M. M.; Aragones, J. L., *Faraday Discuss* **2009**, *141*, 251-76.
173. Eastwood, M. P.; Stafford, K. A.; Lippert, R. A.; Jensen, M. Ø.; Maragakis, P.; Predescu, C.; Dror, R. O.; Shaw, D. E., *Journal of Chemical Theory and Computation* **2010**, *6*, 2045-2058.
174. Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. v.; DiNola, A.; Haak, J. R., *The Journal of Chemical Physics* **1984**, *81*, 3684-3690.
175. Kukul, A., *Molecular Modeling of Proteins*. Springer New York.: 2015.
176. A., W. T.; E., M. A., *Journal of Computational Chemistry* **2006**, *27*, 316-325.
177. Darden, T.; York, D.; Pedersen, L., *The Journal of Chemical Physics* **1993**, *98*, 10089-10092.
178. Hub, J. S.; de Groot, B. L.; Grubmüller, H.; Groenhof, G., *Journal of Chemical Theory and Computation* **2014**, *10*, 381-390.
179. Reis, P. B. P. S.; Vila-Viçosa, D.; Campos, S. R. R.; Baptista, A. M.; Machuqueiro, M., *ACS Omega* **2018**, *3*, 2001-2009.
180. Vila-Viçosa, D.; Teixeira, V. H.; Santos, H. A. F.; Baptista, A. M.; Machuqueiro, M., *Journal of Chemical Theory and Computation* **2014**, *10*, 5483-5492.
181. Konermann, L.; Metwally, H.; McAllister, R. G.; Popa, V., *Methods* **2018**, *144*, 104-112.
182. Cech, N. B.; Enke, C. G., *Mass Spectrometry Reviews* **2001**, *20*, 362-387.

## Chapter 2. Exploring the Mechanism of Salt-Induced Signal Suppression in Protein Electrospray Mass Spectrometry Using Experiments and Molecular Dynamics Simulations

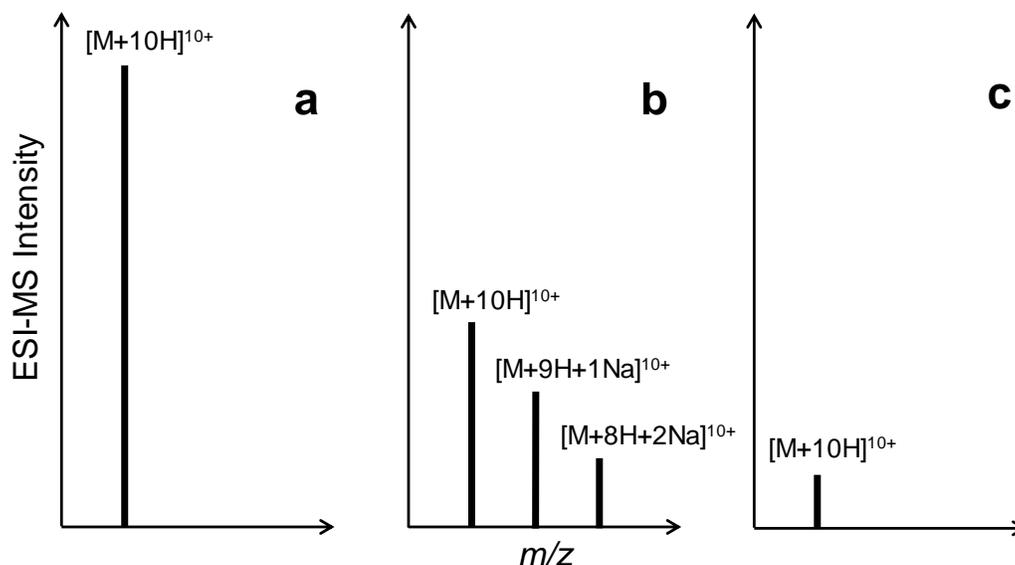
### 2.1. Introduction

Electrospray ionization (ESI) mass spectrometry (MS)<sup>1</sup> has become an indispensable tool for a wide range of bioanalytical applications. The impact of ESI-MS has been particularly pronounced for research in the areas of protein structure, dynamics, and interactions.<sup>2-5</sup> The ESI process commences when analyte solution is passed through a capillary that is held at a high electric potential. Charged droplets are emitted from a Taylor cone at the capillary outlet. These droplets undergo several rounds of solvent evaporation and jet fission,<sup>6</sup> ultimately generating nanodroplets that are close to the Rayleigh limit and from which analyte ions are released into the gas phase.<sup>7</sup>

The mechanism of the final ESI steps remains a matter of debate,<sup>8-11</sup> although recent studies seem to converge towards a consensus view.<sup>7, 12</sup> Accordingly, low molecular weight gaseous ions are produced by field emission from the nanodroplet surface, as envisioned by the ion evaporation model (IEM).<sup>13</sup> Globular proteins are released via nanodroplet evaporation to dryness, in line with the charged residue model (CRM).<sup>14</sup> A third mechanism, the chain ejection model (CEM), has been proposed for unfolded proteins<sup>12</sup> and other disordered polymers.<sup>15</sup> The CEM posits that macromolecular chains get expelled from aqueous droplets by electrostatic and hydrophobic effects.<sup>12, 16</sup> ESI can be conducted under positive and under negative polarity conditions. We will restrict our considerations to positive ion mode, as it is more widely used for protein analyses.

A well-known problem in ESI-MS is the fact that non-volatile salts in the analyte solution can cause a significant degradation of the spectral S/N ratio.<sup>17-24</sup> The presence of Na<sup>+</sup> is particularly problematic, as it represents a ubiquitous contaminant. Also, NaCl is used in many biochemical experiments for ensuring a physiological ionic strength of ~150 mM.<sup>25</sup> Direct infusion of such samples into an ESI source is not a viable option. Numerous desalting

approaches have been devised to address this issue.<sup>17-24</sup> Even with these treated samples, however, the presence of residual salt contaminants often remains problematic.



**Figure 2.1.** Schematic protein ESI mass spectra (only the 10+ region is shown), illustrating two different types of salt effects. (a) Protonated ion signal generated from salt-free solution. (b) Peak splitting due to adduct formation. The integrated signal intensity in panels a and b is identical. (c) Protein ion suppression.

The mechanistic basis of salt interferences in protein ESI-MS continues to be a matter of debate. One fairly trivial aspect is the fact that species such as  $Na_nCl_m^{(n-m)+}$  can obscure protein signals due to spectral crowding.<sup>26-28</sup> Clusters of this type are formed via the CRM as salt-containing nanodroplets evaporate to dryness.<sup>29-31</sup> Of greater concern are issues that are commonly treated under the umbrella of salt-induced “signal suppression”.<sup>17, 23, 32-38</sup> In the literature this term may carry different connotations, a fact that can complicate discussions of the topic. Here, we propose that signal suppression can be conceptually dissected into (i) *peak splitting* due to adduct formation and (ii) *ion suppression*. Both effects will degrade the spectral S/N ratio, as briefly outlined in the following sections.

(1) While salt-free solutions generate clean  $[M + zH]^{z+}$  species, protein ions formed in the presence of nonvolatile salts tend to have a heterogeneous composition such as  $[M + zH +$

$n(\text{Na} - \text{H}) + m(\text{Cl} + \text{H})]^{z+}$  where both  $n$  and  $m$  can cover a wide range.<sup>39-40</sup> This adduct formation is particularly pronounced for native proteins that follow the CRM, because the salt concentration in vanishing droplets increases dramatically during solvent evaporation to dryness.<sup>39, 41</sup> Under these conditions the total ion count of each charge state  $z$  is split into a multitude of peaks, thus reducing the  $[\text{M} + z\text{H}]^{z+}$  intensity.<sup>39, 42</sup> We hypothesize that peak splitting does *not* represent a genuine signal suppression effect, as it may occur without reducing the total protein signal intensity. In other words, the sum of the mixed adduct signals for all values of  $z$ ,  $n$ , and  $m$  may still add up to the same intensity as for  $[\text{M} + z\text{H}]^{z+}$  species produced in the absence of salt (Figure 2.1a, b).

(2) Ion suppression occurs under conditions where the presence of salt reduces the *total* analyte signal (Figure 2.1c). This effect was first examined for singly charged small analytes, assuming that different species compete with each other for excess charge.<sup>32</sup> This framework was subsequently extended into an equilibrium-partitioning model<sup>17</sup> which envisions that ESI droplets consist of an electrically neutral core, and a charged surface layer. Analytes that become part of the surface charge layer are thought to be more easily ejected under IEM conditions. The development of this model was spurred by the observation that surface-active species generally exhibit high ESI-MS signal intensities. Related charge competition models were also proposed to account for the effects of salts on the ESI-MS signals of proteins.<sup>33, 36</sup>

Despite the substantial body of work in this area, the mechanism of salt interferences in protein ESI-MS is still not fully understood. Recent work has put into question the tenet that excess charge carriers always reside on the droplet surface,<sup>12</sup> an assumption that is central to the equilibrium-partitioning model.<sup>17</sup> The purported link between charge competition and ion suppression<sup>17, 32-33, 36</sup> is readily understood for mono-cationic species, but for multiply charged analytes the situation is less clear. For example, charge competition may reduce ESI charge states,<sup>43</sup> but this does not necessarily imply the occurrence of ion suppression because proteins will remain observable as long as  $z > 0$ .<sup>44</sup> Finally, the relationship between adduct formation and ion suppression remains to be elucidated. It is unclear if the phenomena indicated in Figure 2.1b, c can actually occur independently of one another, or if they are always coupled.

This work explores the ESI-MS behavior of cytochrome *c* (Cyt *c*), ubiquitin (Ubq), and lysozyme (Lyz) under native and denaturing solvent conditions. To investigate the effects of salts with different physicochemical properties we tested NaCl, CsCl, and tetra-butyl ammonium chloride (NBu<sub>4</sub>Cl). The three cations cover a wide range of hydration free energies, from highly favorable for Na<sup>+</sup> ( $\Delta_{hyd}G^\circ = -365 \text{ kJ mol}^{-1}$ ) to unfavorable for NBu<sub>4</sub><sup>+</sup> ( $\Delta_{hyd}G^\circ \geq 0$ ). The properties of Cs<sup>+</sup> lie in between these two extremes ( $\Delta_{hyd}G^\circ = -250 \text{ kJ mol}^{-1}$ ).<sup>45</sup> It is found that the effects of NaCl are largely consistent with the “pure” adduction scenario of Figure 2.1b, whereas NBu<sub>4</sub>Cl displays “pure” ion suppression (Figure 2.1c). MD simulations on salt-containing droplets suggest that NBu<sub>4</sub><sup>+</sup> induces ion suppression primarily by interfering with the formation of progeny droplets within the ESI plume.

## 2.2. Materials and Methods

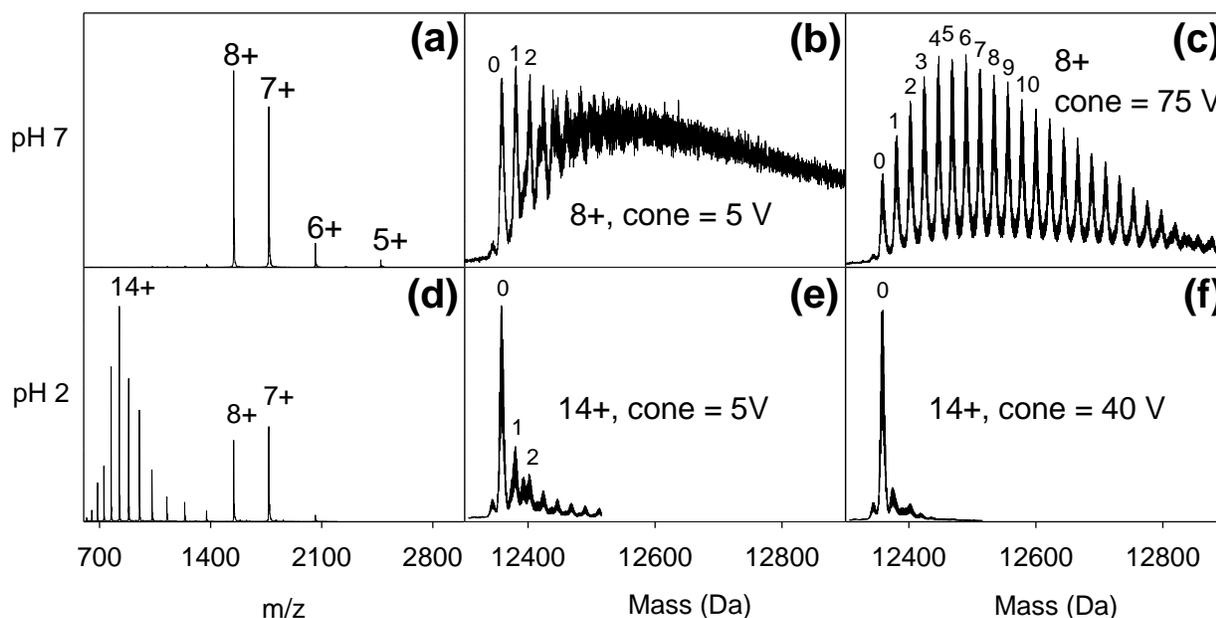
### 2.2.1. Protein Solutions

Horse heart Cyt *c*, hen egg white Lyz, bovine Ubq, 1,4-dithiothreitol (DTT), NBu<sub>4</sub>Cl, ammonium acetate, CsCl, and formic acid were purchased from Sigma (St. Louis, MO). For native ESI-MS, proteins were infused in aqueous solution at pH 7. Denatured (unfolded) samples were prepared by dissolving proteins in water/methanol (50:50 v/v) in the presence of formic acid (pH 2.0). Lysozyme unfolding was further promoted by disulfide reduction.<sup>46</sup> For this purpose, 100  $\mu\text{M}$  Lyz was incubated in 10 mM DTT at 75 °C for 1 h. The reduced protein is referred to as “rLyz”. Protein samples for ESI-MS had a concentration of 5  $\mu\text{M}$  and contained 10 mM ammonium acetate. This volatile electrolyte represents a standard additive in ESI-MS.<sup>7</sup> The solutions were supplemented with NaCl, CsCl, or NBu<sub>4</sub>Cl at concentrations up to 10 mM.

### 2.2.2. Mass Spectrometry (ESI-MS)

ESI mass spectra were acquired in sensitivity mode on a Synapt G2 time-of-flight instrument (Waters, Milford, MA) equipped with a dual ESI source. The ESI capillary was held at +3 kV. Protein solutions were infused using a syringe pump at a flow rate of 5  $\mu\text{L min}^{-1}$ . Cone and desolvation gas flow rates were 50 and 600 L h<sup>-1</sup>, respectively. The desolvation temperature was 200 °C, and the source was kept at 80 °C. For native solutions the cone voltage was set to a

relatively high value of 75 V which promotes the conversion of mixed cation/chloride adducts to  $[M + zH + n(\text{cation} - H)]^{z+}$  species via HCl loss.<sup>39, 41</sup> These settings considerably improve the spectral S/N ratio (Figure 2.2a-f). High charge states formed under denaturing conditions would suffer from rupture of covalent bonds under these conditions. For those measurements the cone voltage was reduced to 40 V (Figure 2.2d-f). Except for the cone voltage, all data were acquired under identical instrument settings as the sum of 500 one-second scans. Intensity units refer to analog-to-digital converter (ADC) count.



**Figure 2.2.** Cone voltage effects on ESI mass spectra of *cyt c*. Panels a-c refer to data acquired at pH 7 in aqueous solution, spectra in panels d-f were recorded at pH 2 in 50:50 (v/v) water/methanol. (a) Complete spectrum of native *Cyt c* electrosprayed in the absence of NaCl. (b) Close-up view of the 8+ charge state in the presence of 1 mM NaCl at a cone voltage of 5 V. (c) Same as in panel b, but for a cone voltage of 75 V. Note that the  $m/z$  axes were converted to *Mass* in panels b, c, e, and f. The “cleaner” appearance of the spectrum in panel c is due to the collisional loss of chloride adducts (as HCl). (d) Complete spectrum of denatured *Cyt c* electrosprayed in the absence of NaCl. (e) Close-up view of the 14+ charge state in the presence of 1 mM NaCl at a cone voltage of 5 V. (f) Same as in panel e, but for a cone voltage of 40 V. Numbers 0, 1, 2, ... indicate how many adducted sodium ions can be resolved in the spectra.

### 2.2.3. Experimental Design and Data Analysis

A potential problem with measurements on salt-contaminated samples is a drift in sensitivity, caused by the deposition of non-volatile material on the ion optics during a series of measurements. An internal standardization procedure was devised to compensate for such drifts, taking advantage of the dual ESI source on the Synapt instrument. The regular (analyte) source was used to deliver actual protein samples at salt concentrations  $C$  between zero and 10 mM. The integrated signal intensity for each of these spectra is referred to as  $I(C)$ . Integration was performed using Microsoft Excel over the spectral range that was deemed to contain adducted protein ions, as identified via comparisons with low salt spectra, up to  $m/z \sim 3000$ . Signals corresponding to protein-free cluster ions (recognizable by their characteristic isotope distributions and low charge states<sup>29, 31</sup>) were eliminated manually prior to integration. After recording each of these spectra, the second (lock spray) source was used to deliver an internal standard comprising the same protein and solvent, but without salt. Salt-containing samples and reference solutions were infused in an alternating fashion, using a mechanical baffle to switch between both sprayers. The flow for the non-operating sprayer was turned off to eliminate any cross-contamination. In each case, it was ensured that the ESI-MS signals had stabilized before data acquisition commenced. From these pairwise measurements a normalized intensity  $N(C)$  can be calculated as  $N(C) = I(C) / I_{ref}$ , where  $I_{ref}$  is the integrated signal intensity of the reference sample. As a final step, these data were normalized according to  $R(C) = N(C) / N(C = 0)$ . These  $R(C)$  values will be referred to as “integrated ion intensity”, and they directly reflect the extent of salt-induced signal attenuation relative to a salt-free sample. To reiterate,  $R(C)$  comprises the contributions of all protein signals regardless of their salt adduction state. All measurements were conducted in triplicate. Error bars represent standard deviations.

## 2.2.4. Molecular Dynamics Simulations

All atom MD simulations of salt-containing droplets were carried out using Gromacs 4.6.5 with CHARMM36 force field<sup>47</sup> and TIP3P water model.<sup>48</sup> Production runs employed trajectory stitching as described previously<sup>31</sup> with Nosé-Hoover temperature coupling. The structure of  $\text{NBu}_4^+$  was obtained from the ZINC database,<sup>49</sup> and it was parameterized using the ParamChem server.<sup>50</sup> All bonds were constrained, thereby allowing for a 2fs integration time step. To take advantage of GPU acceleration the simulations were carried out in a 1  $\mu\text{m}$  box using a potential-shift non-bonded interaction cut-off of 333.3 nm without relying on particle-mesh Ewald summation. Initial system configurations consisted of spherical water droplets with radius 3 nm containing 10 Na/Cl pairs. Each droplet also contained a 12+ excess charge consisting of either 12  $\text{Na}^+$  or 12  $\text{NBu}_4^+$ . These systems were subjected to energy minimization followed by 10ps of equilibration at 330 K using a modified Berendsen thermostat.<sup>51</sup> Individual MD segments consisted of 500 ps windows at 330 K, between which any molecules that had drifted more than 10 nm from the droplet center were removed from the system. The simulated droplets are highly dynamic and undergo occasional distortions into non-spherical shapes. For reporting the droplet size we report an “effective” radius  $r$  corresponding to that of a sphere with the equivalent number of molecules. Calculations involving the surface tension of water employed the experimental value of  $\gamma = 0.06624 \text{ Nm}^{-1}$  at 330 K,<sup>52</sup> although standard MD water models yield values that are somewhat lower.<sup>53-54</sup> Similar to earlier droplet simulations<sup>12, 15, 31, 54-58</sup> the current work uses a non-polarizable force field. It has been noted<sup>59-60</sup> that such an approach may not adequately describe the behavior of “soft” ions such as  $\text{Cs}^+$  which possess a high polarizability (~tenfold higher than that of  $\text{Na}^+$ ).<sup>61</sup> To circumvent polarizability issues, the simulations of this study were therefore limited to droplets containing NaCl and/or  $\text{NBu}_4\text{Cl}$ . The  $\text{NBu}_4^+$  behavior is largely invariant when using polarizable or non-polarizable force fields.<sup>62</sup> All simulations were carried out in triplicate with different starting configurations and random seeds for initial velocity assignments.

## 2.3. Results and Discussion

### 2.3.1. NaCl Effects on Protein Mass Spectra

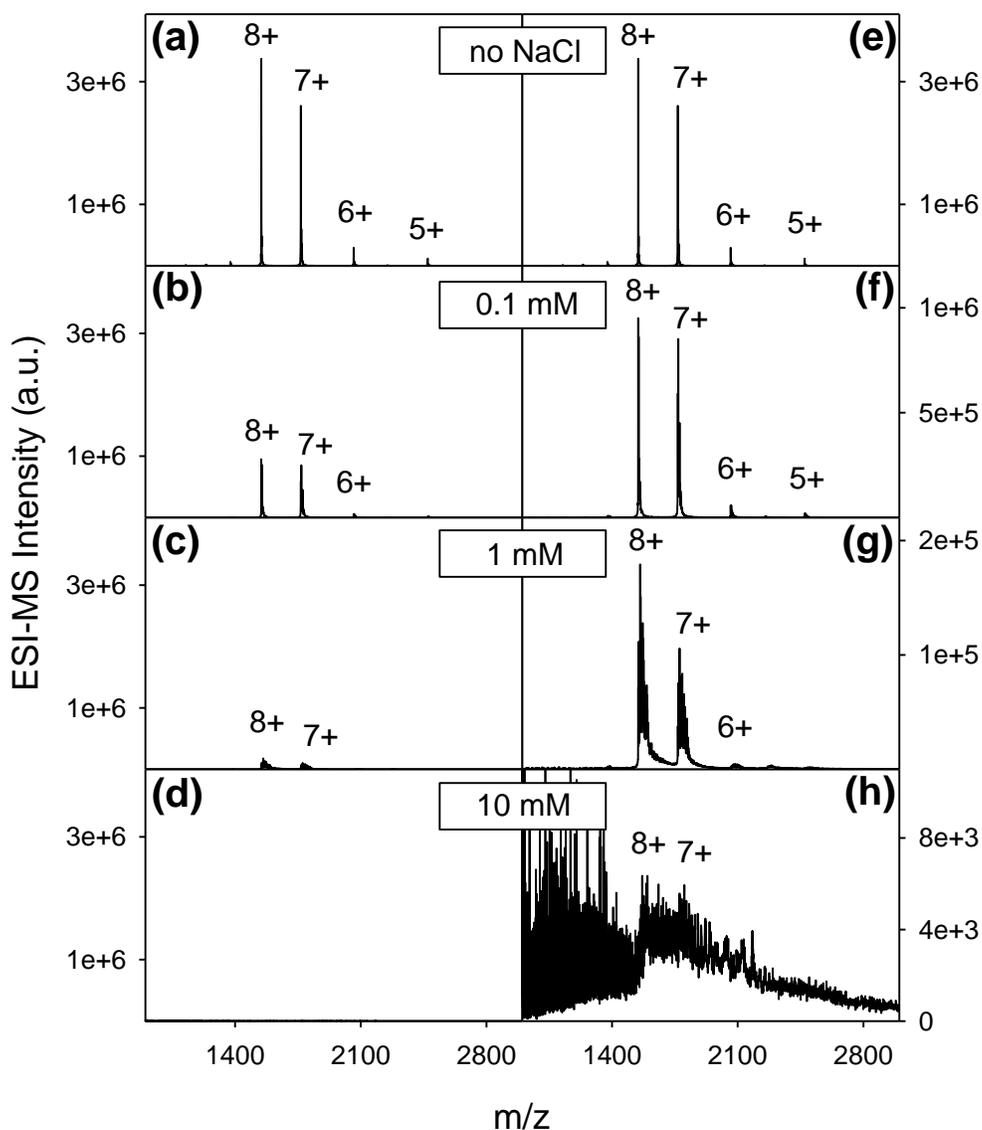
In an initial set of experiments we explored the effects of NaCl on the ESI-MS behavior of native Cyt *c*. When electrosprayed in NaCl-free solution the protein displays a charge state distribution that is dominated by  $[M + 7H]^{7+}$  and  $[M + 8H]^{8+}$  ions (Figure 2.3a).<sup>63-64</sup> Increasing the NaCl concentration from zero to 10 mM dramatically reduces the base peak intensity. Protein signals become almost unobservable for 1 mM NaCl when displaying the data using constant y-axis scaling (Figure 2.3a-d). These observations illustrate the well-known perils of conducting protein ESI-MS in the presence of non-volatile salts.

Figure 2.3e-h provides a closer look at the same Cyt *c* data, using a y-axis range that scales with base peak intensity. The spectrum acquired in 10 mM NaCl (Figure 2.3h) is of particular interest. When considered in isolation, these data would be difficult to interpret. However, comparison with the spectra acquired at lower salt concentrations (Figure 2.3e-g) leaves no doubt that the humps marked 8+ and 7+ in Figure 2.3h represent heavily adducted protein signals with unresolved tails that extend to  $m/z \sim 3000$ . The range of  $m/z < 1500$  in Figure 2.3h is dominated by  $Na_nCl_m^{(n-m)+}$  clusters.<sup>29</sup> When integrating the protein ion intensity for the spectra in Figure 2.3 it is seen that  $R(C)$  of native Cyt *c* is not very strongly affected by NaCl; no intensity loss is evident up to a salt concentration of 1 mM. Even at 10 mM  $R(C)$  remains above 0.6 (Figure 2.4a).

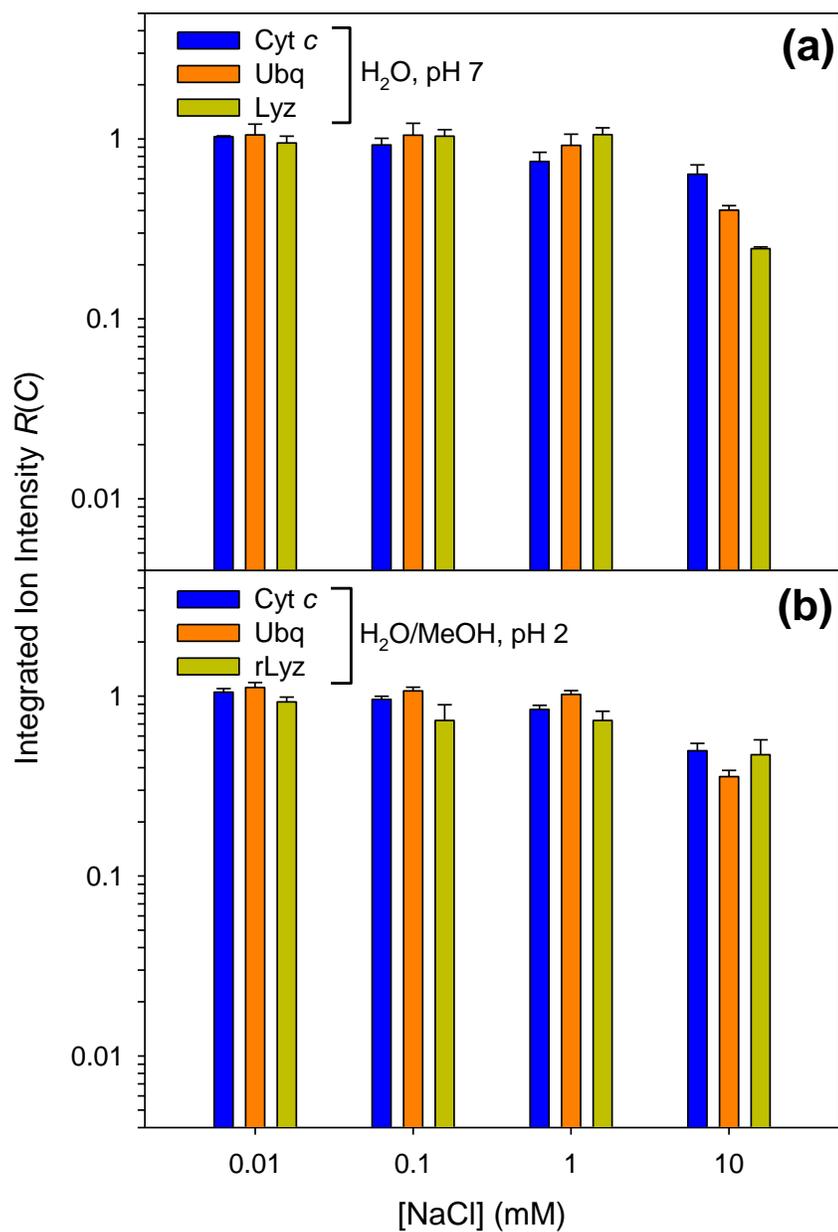
Observations similar to those described above were made when studying NaCl effects for native Ubq and Lyz (Figure 2.5). Those data show a precipitous drop in base peak intensity, whereas the integrated ion intensity  $R(C)$  remains constant up to 1 mM NaCl. At a salt concentration of 10 mM the  $R(C)$  values of Ubq and Lyz exhibit a decrease down to 0.4 and 0.25, respectively (Figure 2.4a).

ESI-MS experiments on the three proteins were also conducted under denaturing conditions. The resulting spectra are shifted to higher charge states<sup>63-64</sup> that show less salt adduction (Figure 2.6). Previous work<sup>41</sup> suggests that adduction is reduced for unfolded proteins

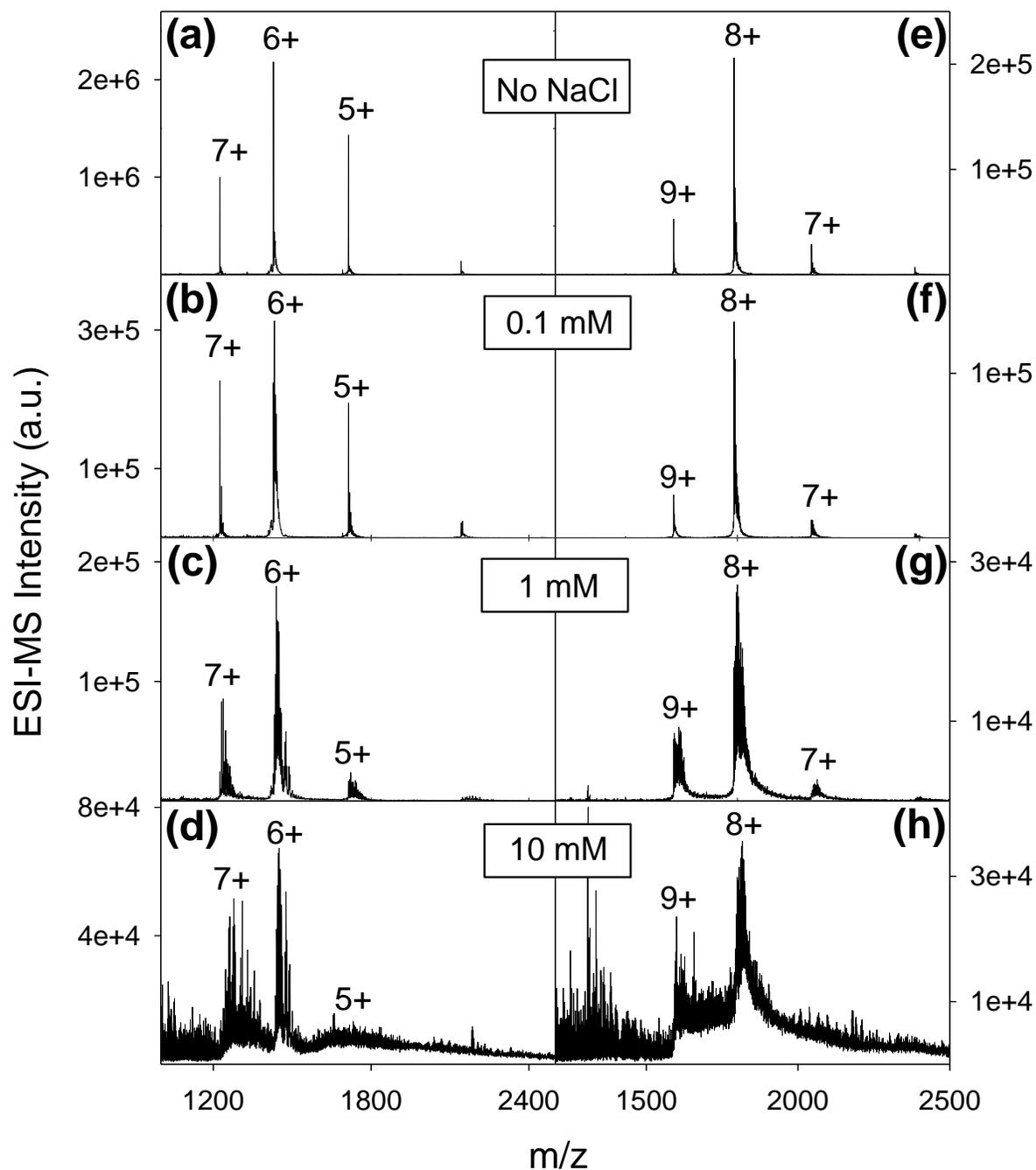
because the chains are ejected from droplets that experienced a relatively low degree of NaCl enrichment. Salt interferences in Figure 2.6 only become prevalent at 10 mM NaCl, at which point the base peak intensity has dropped by one order of magnitude. The integrated ion intensity of the three denatured proteins decreases down to  $R(C) \approx 0.5$  in 10 mM NaCl (Figure 2.4b).



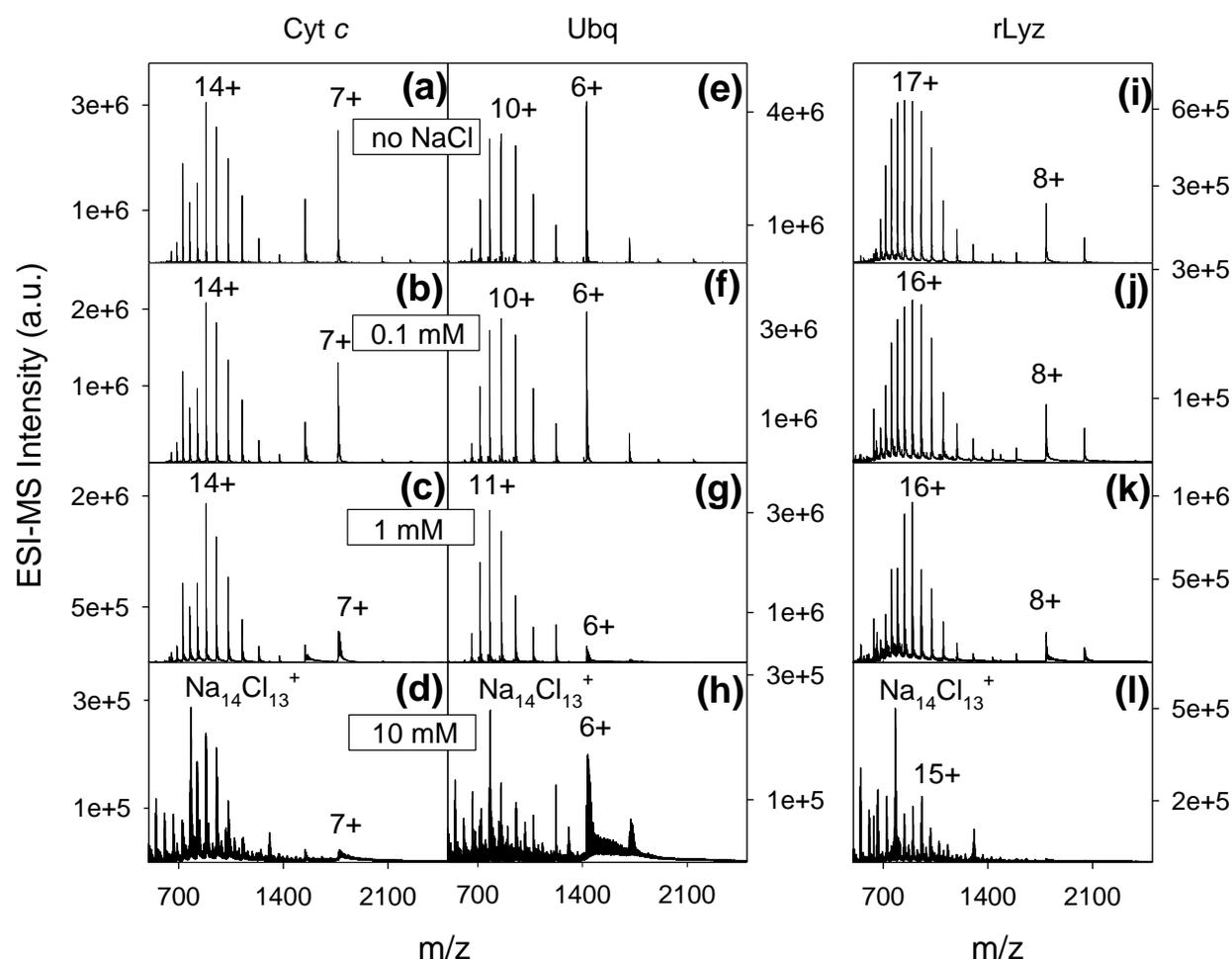
**Figure 2.3.** ESI mass spectra of cyt *c* acquired at pH 7 in the presence of various NaCl concentrations. Panels a/e, b/f, c/g, and d/h display the same data, but with different y (intensity) axis scaling. Panels a-d shares the same y-axis range, whereas panels e-h were scaled according to base peak intensity.



**Figure 2.4.** Salt dependence of the integrated ion intensity  $R(C)$  relative to NaCl-free samples. Data are depicted for three proteins (a) Under native conditions, and (b) In a denaturing environment.



**Figure 2.5.** ESI mass spectra of Ubq (a-d) and Lyz (e-h) acquired at pH 7 in the presence of different NaCl concentrations (zero to 10 mM NaCl, as indicated in the individual panels).



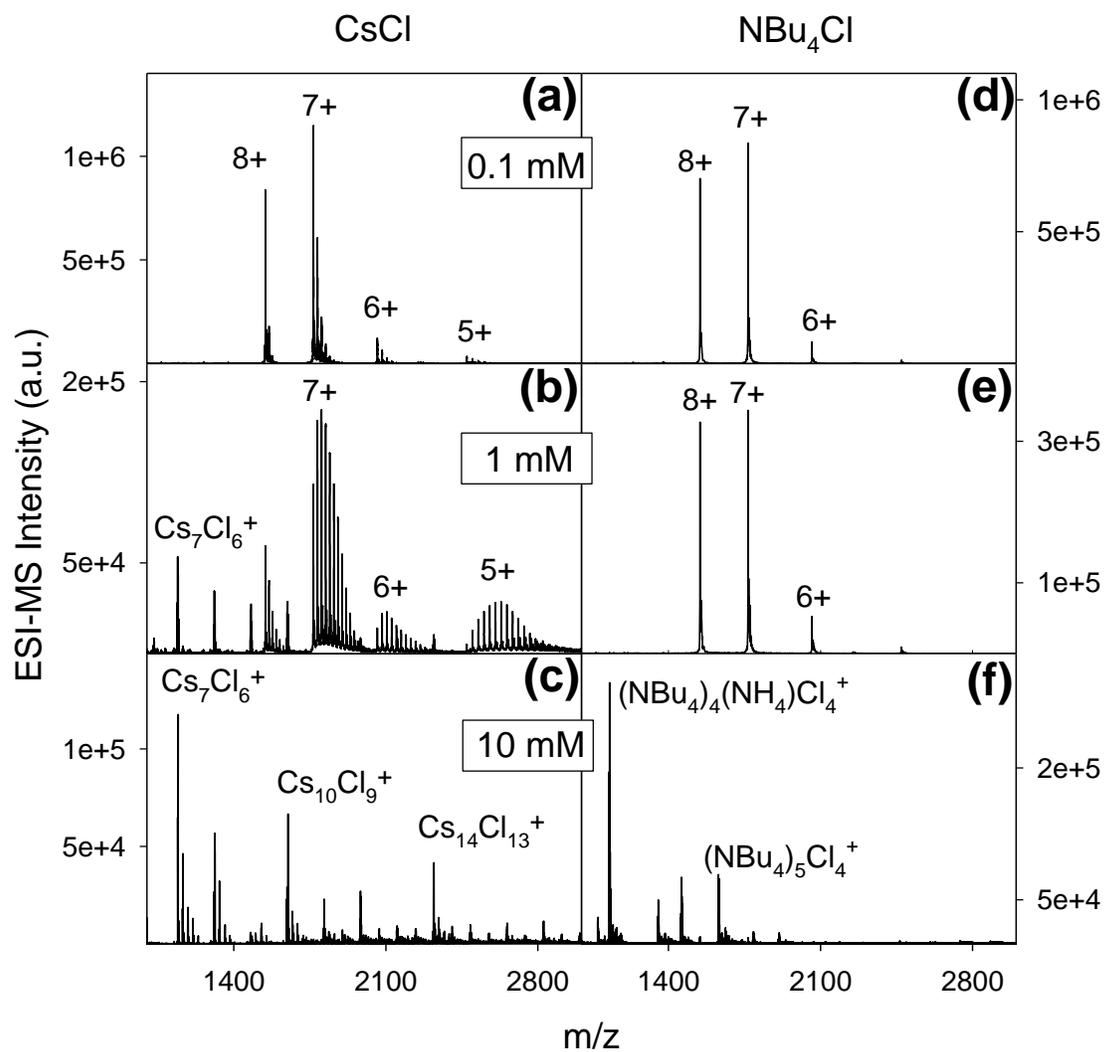
**Figure 2.6.** ESI mass spectra of Cyt c (a-d), Ubq (e-h) and rLyz (i-l) acquired under denaturing solvent conditions (pH 2 in water/methanol) in the presence of different NaCl concentrations (zero to 10 mM NaCl, as indicated in the individual panels).

In summary, NaCl induces a dramatic reduction in the S/N ratio of protein ESI mass spectra by lowering the intensity of  $[M + zH]^{z+}$  ions. However,  $R(C)$  remains surprisingly insensitive to the presence of NaCl. The sum of all adducted protein signals stays relatively close to that of  $[M + zH]^{z+}$  ions observed for NaCl-free samples. Even the  $R(C)$  drop seen for 10 mM NaCl is quite moderate, considering that cursory analysis suggests the near-complete loss of signal (Figure 2.3d). We conclude that NaCl degrades ESI mass spectra primarily via adduct-mediated peak splitting (Figure 2.1b) rather than ion suppression (Figure 2.1c).

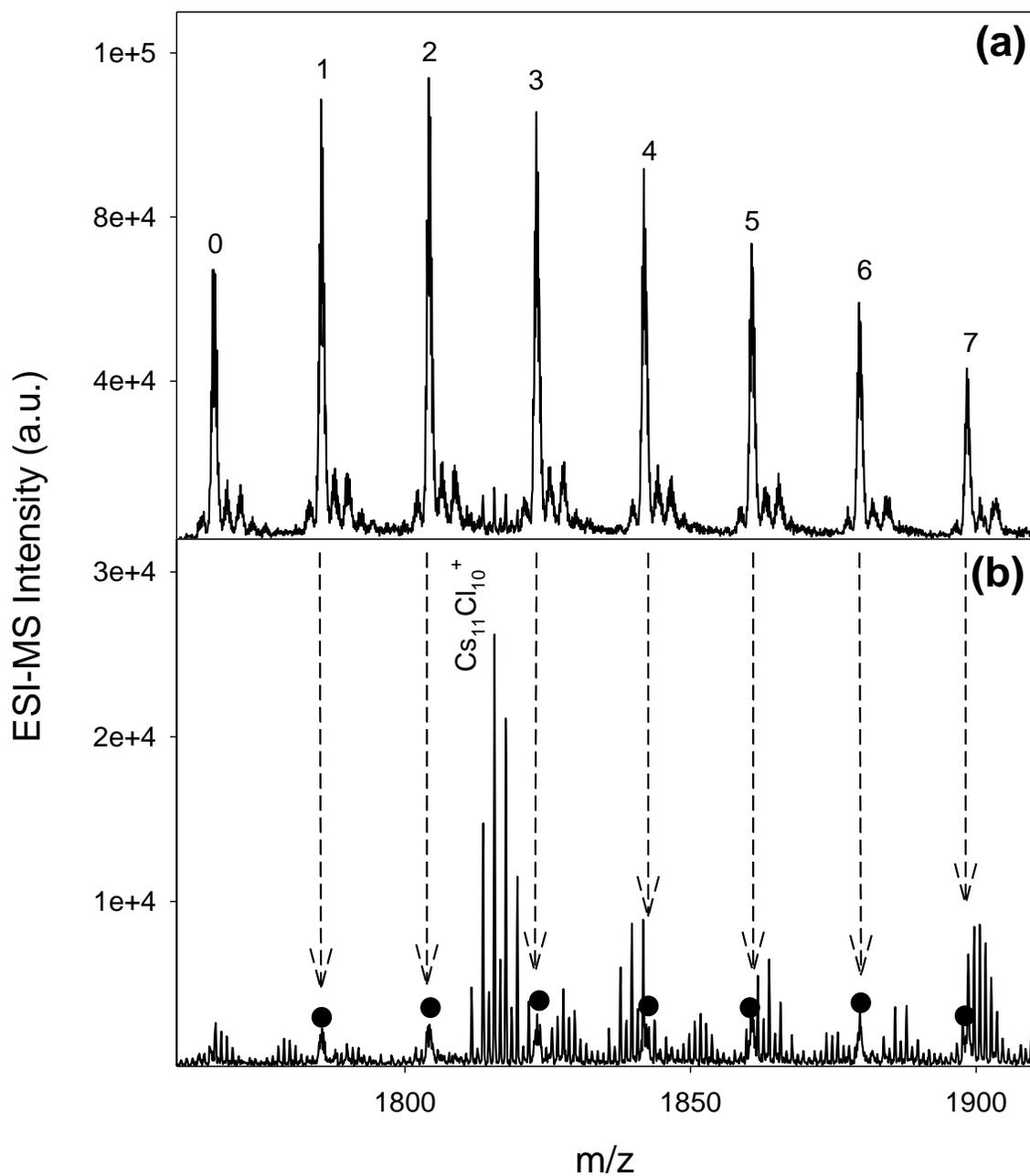
### 2.3.2. Comparison with Other Salts

Although  $\text{Na}^+$  represents the most commonly encountered interfering cation in ESI-MS, additional insights into the mechanism of signal degradation can be obtained by examining the properties of other additives.<sup>33, 36, 65</sup> We chose to study Cyt *c* in the presence of CsCl and  $\text{NBu}_4\text{Cl}$ . The consistent use of chloride salts ensures that any of the observed effects are attributable solely to differences in cation behavior.

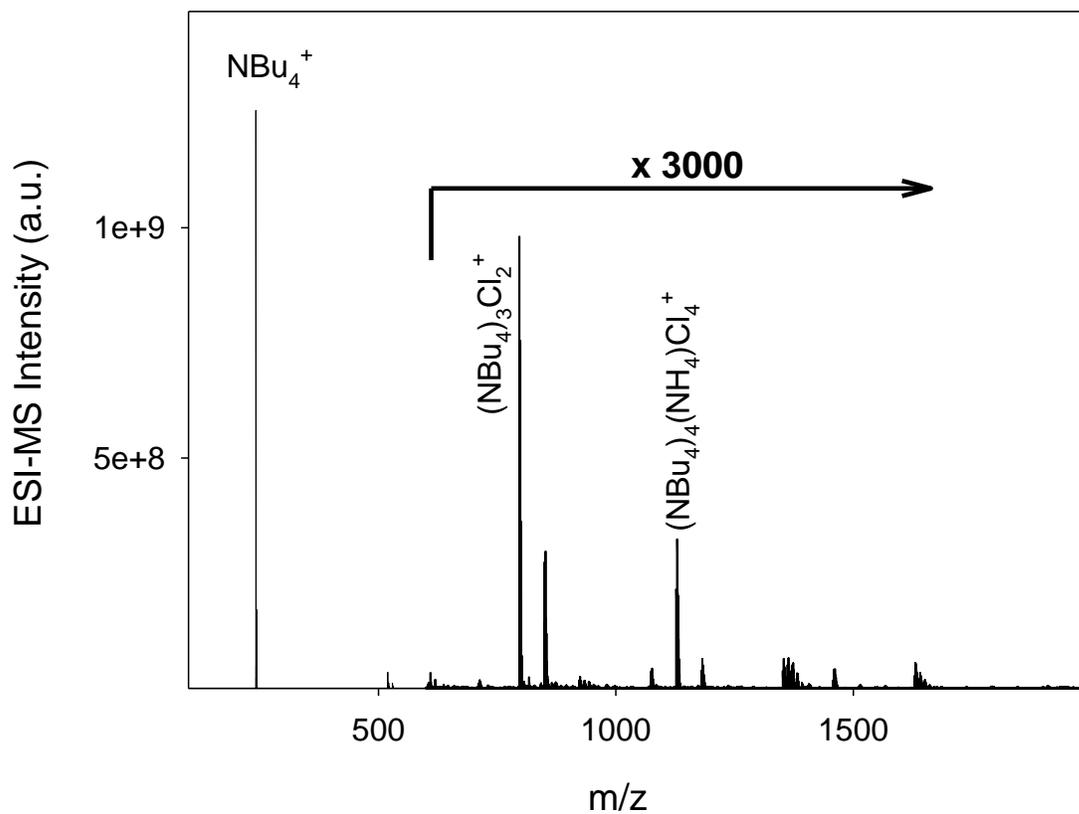
Addition of CsCl to Cyt *c* at pH 7 reduces the base peak intensity and causes extensive adduct formation (Figure 2.7a-c). At a concentration of 10 mM the spectrum is crowded by  $\text{Cs}_{(n+1)}\text{Cl}_n^+$  clusters, but close examination still reveals residual protein signals (Figure 2.8). Experiments conducted with  $\text{NBu}_4\text{Cl}$  reveal a very different behavior, resulting in protein ions that are virtually free of adducts. However, a dramatic signal deterioration is seen for this salt, where protein signals become almost undetectable in 10 mM  $\text{NBu}_4\text{Cl}$  (Figure 2.7d-f). Spectra at this high salt concentration are dominated by a  $\text{NBu}_4^+$  peak that dwarfs all other signals by a factor of several thousand (Figure 2.9). Observations similar to those described here for pH 7 were also made in Cyt *c* measurements under denaturing conditions (Figure 2.10).



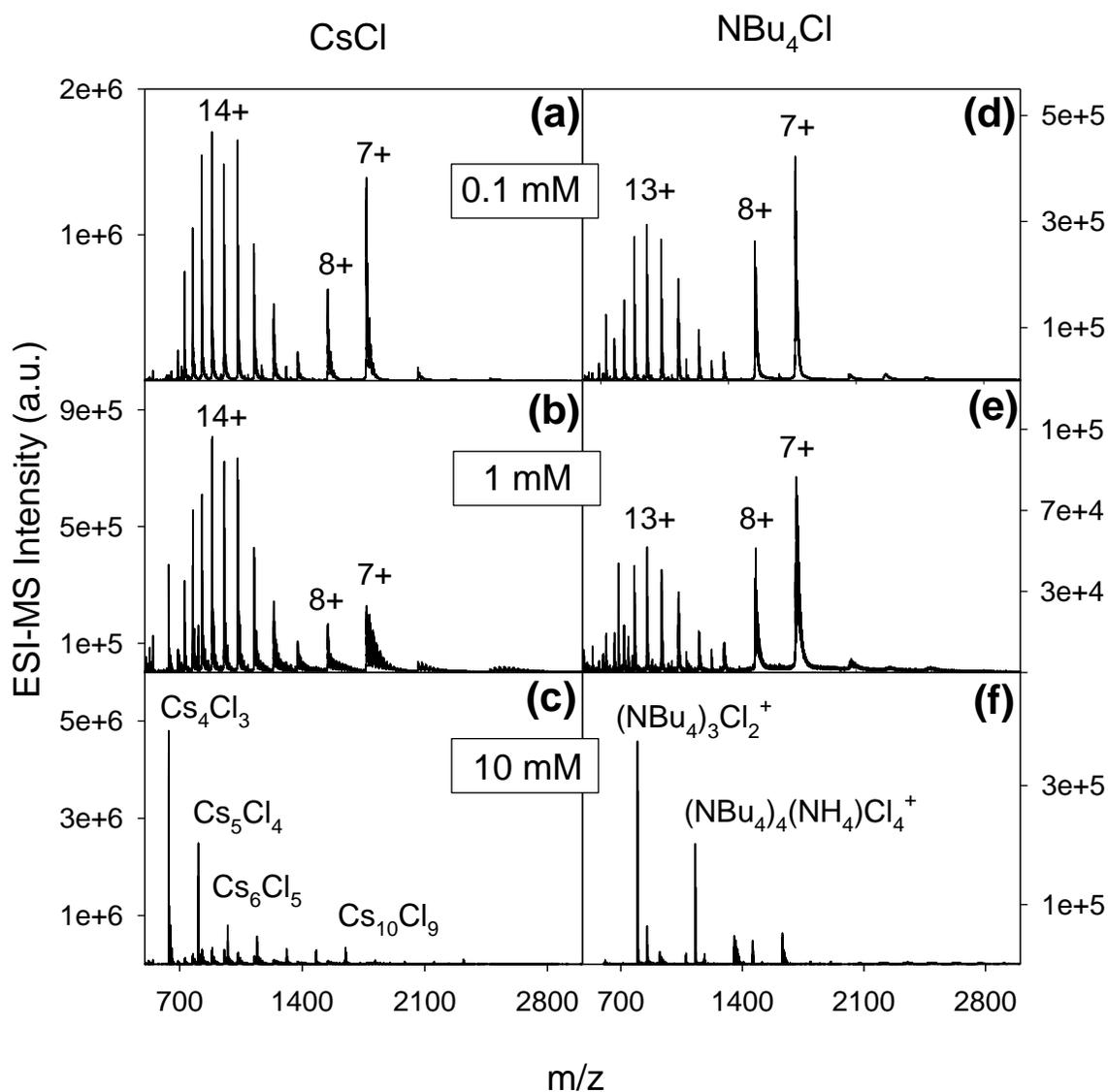
**Figure 2.7.** ESI mass spectra of Cyt *c* acquired at pH 7 in the presence of different CsCl (a-c) and NBu<sub>4</sub>Cl concentrations (d-f). Protein charge states as well as the composition of selected salt cluster signals are indicated.



**Figure 2.8.** Close-up views of the Cyt *c* the 7+ charge state acquired at pH 7 in the presence of (a) 1 mM CsCl and (b) 10 mM CsCl. Numbers  $n = 0, 1, 2, \dots$  in panel A denote how many Cs<sup>+</sup> ions are bound to the  $[\text{M} + 7\text{H} + n(\text{Cs} - \text{H})]^{7+}$  ions. Dashed arrows in panel (b) emphasize that protein signals are still observable in 10 mM CsCl, despite the dominance of salt cluster signals. Protein signals in panel (b) are further highlighted by filled circles.



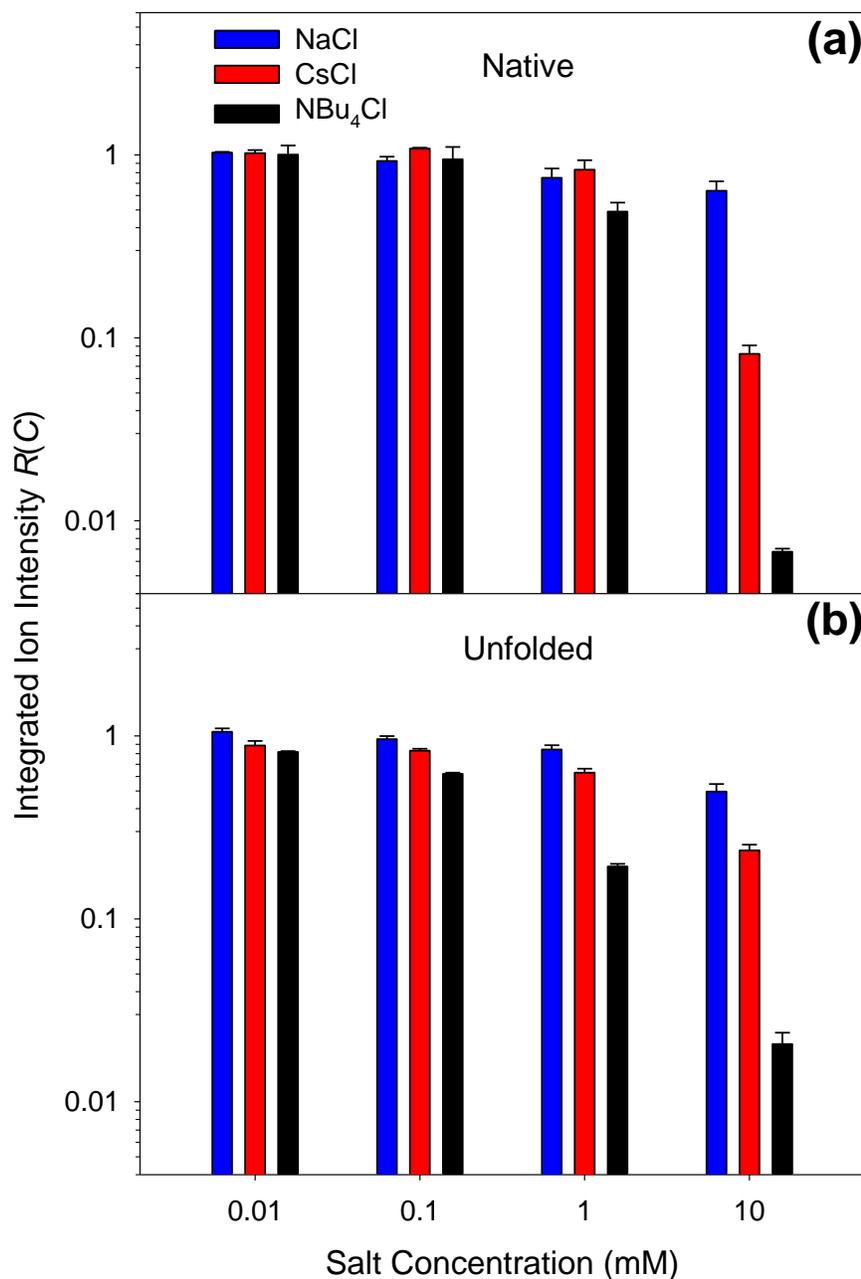
**Figure 2.9.** ESI mass spectrum of Cyt *c* acquired at pH 7 in the presence of 10 mM NBu<sub>4</sub>Cl. Protein signals are almost completely suppressed; instead the spectrum is dominated by NBu<sub>4</sub><sup>+</sup>. Note that the signal intensity has been magnified by a factor of 3000 for  $m/z > 600$ .



**Figure 2.10.** ESI mass spectra of Cyt *c* acquired at pH 2 in water/methanol and in the presence of different concentrations of CsCl (a-c) and NBu<sub>4</sub>Cl (d-f). Protein charge states as well as the composition of selected salt cluster signals are indicated.

Figure 2.11 compares the integrated ion intensities of Cyt *c* in CsCl, and NBu<sub>4</sub>Cl under native and denaturing conditions. The NaCl data discussed above are included in Figure 2.11 for reference purposes. In the case of CsCl the integrated ion intensity remains roughly constant up to a salt concentration of 1 mM. Elevating the CsCl concentration further to 10 mM induces a

notable  $R(C)$  reduction, down to around 0.1 and 0.2 for native and unfolded Cyt  $c$ , respectively. A much more significant drop in  $R(C)$  is seen for  $\text{NBu}_4\text{Cl}$ . This salt begins to lower  $R(C)$  already at a concentration of 1 mM. For 10 mM  $\text{NBu}_4\text{Cl}$  the integrated ion intensity is reduced by two orders of magnitude (Figure 2.11).



**Figure 2.11.** Salt dependence of the integrated ion intensity  $R(C)$  for Cyt  $c$  under (a) Native solvent conditions and (b) In denaturing solution in the presence of NaCl, CsCl, and  $\text{NBu}_4\text{Cl}$ .

The data presented above reveal that the detrimental effects associated with the three salts are quite different. As noted above, NaCl does *not* cause a dramatic  $R(C)$  decrease. Instead, the degradation of the spectral S/N ratio is primarily due to adduct formation which leads to extensive peak splitting. The effects of NaCl can therefore be approximated by the scenario depicted in Figure 2.1b. The opposite applies to NBu<sub>4</sub>Cl, i.e., a pronounced reduction in  $R(C)$  without adduct formation. The effects of NBu<sub>4</sub>Cl are thus consistent with the ion suppression scenario of Figure 2.1c. CsCl shows a behavior in-between these two extremes; it promotes the formation of adducts while at the same time reducing  $R(C)$ . It is clear from Figure 2.11, however, that ion suppression by CsCl is less pronounced than for NBu<sub>4</sub>Cl. In summary, the tendency to form protein/salt adducts follows the sequence NaCl > CsCl > NBu<sub>4</sub>Cl. Conversely, the potential of these species to cause protein ion suppression exhibits the reverse order, NBu<sub>4</sub>Cl > CsCl > NaCl.

### 2.3.3. MD Simulations of Salt-Containing Droplets.

To better understand the observed salt effects, we turned to MD simulations. Considerable recent progress has been made in the application of MD techniques to MS-related phenomena,<sup>12, 15, 31, 54-58</sup> although realistic simulations of the complete ESI process (from micrometer-sized early droplets to bare protein ions) still remain out of reach. The simulations of the current work will be limited to protein-free nanodroplets, comparable in size to “late” droplets in the ESI plume.<sup>7, 32</sup> It will be seen that even this simplified approach provides important insights. We focus on aqueous systems for two reasons. (i) Native ESI-MS studies generally use water-based solvents.<sup>2, 66-67</sup> (ii) ESI of water/organic mixtures induces H<sub>2</sub>O enrichment because the organic component usually has a higher evaporation rate.<sup>68-69</sup> Thus, late ESI droplets tend to be predominantly aqueous, regardless of the initial solvent composition. For reasons outlined in the Methods section we will focus on the behavior of Na<sup>+</sup> and NBu<sub>4</sub><sup>+</sup>, i.e., the types of ions that epitomize the two scenarios of adduction *vs.* ion suppression in Figure 2.1.

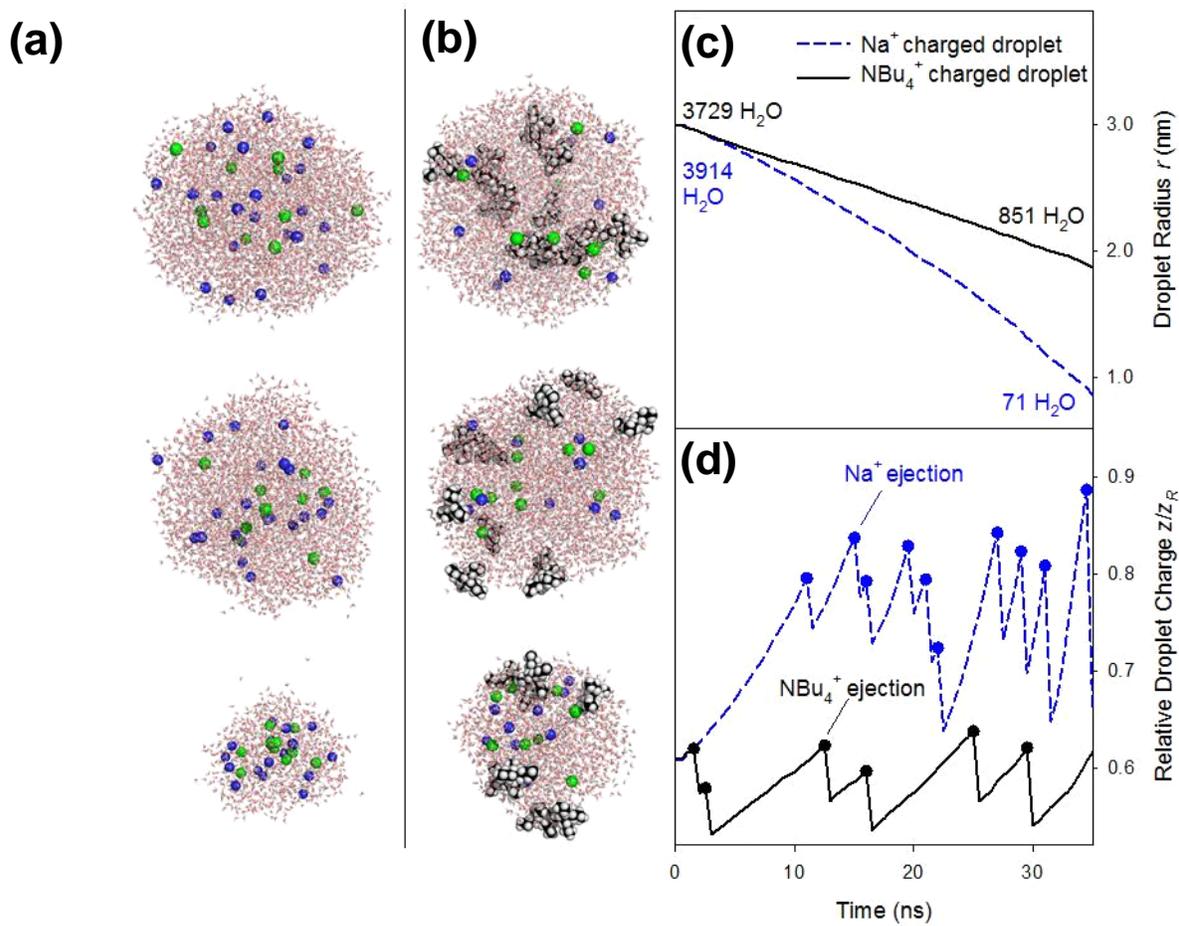
The theoretical number of maximum excess charges  $z_R$  that can be accommodated on a droplet is usually estimated by the Rayleigh equation<sup>7, 70</sup> which states that

$$z_R = \frac{8\pi}{e} \sqrt{\epsilon_0 \gamma r^3} \quad (2.1)$$

where  $e$  is the elementary charge,  $r$  is the radius,  $\epsilon_0$  is the vacuum permittivity, and  $\gamma$  is the surface tension. For the droplets considered here with an initial radius of 3 nm Equation 2.1 yields  $z_R \approx 20$ . The droplet charge at the onset of our simulations was chosen to be  $z = 12$ , well below the Rayleigh limit. Two droplet types were examined, where this charge was implemented via incorporation of either 12  $\text{Na}^+$  or 12  $\text{NBu}_4^+$ . In addition, all droplets contained 10  $\text{Na}^+ / \text{Cl}^-$  pairs, keeping in mind that realistic simulations have to include some counter ions.<sup>31</sup>

MD snapshots for a  $\text{Na}^+$  charged droplet are depicted in Figure 2.12a. As reported earlier,<sup>15, 31, 54-55</sup> droplet shrinkage due to water evaporation is accompanied by the IEM ejection of solvated  $\text{Na}^+$ . Prior to being ejected,  $\text{Na}^+$  ions tend to reside at intermediate radial positions rather than at the surface.<sup>12</sup>  $\text{NBu}_4^+$  containing droplets show a different behavior.  $\text{NBu}_4^+$  ions rapidly move to the surface, with protrusion of alkyl chains into the vacuum environment (Figure 2.12b). As the droplets shrink,  $\text{NBu}_4^+$  ions detach from the surface. Both the low surface affinity of  $\text{Na}^+$  and the high surface affinity of  $\text{NBu}_4^+$  seen here are consistent with previous reports,<sup>12, 17, 33, 36, 59-60, 71</sup> reflecting the fact that  $\text{Na}^+$  interacts very favorably with water, with a hydration free energy  $\Delta_{\text{hyd}}G^\circ = -365 \text{ kJ mol}^{-1}$ . In contrast,  $\text{NBu}_4^+$  hydration is unfavorable ( $\Delta_{\text{hyd}}G^\circ \geq 0$ ).<sup>45</sup>

Figure 2.12c displays changes in droplet radius throughout the 35 ns simulation window. When plotting the relative droplet charge  $z/z_R$  on the same time axis it is seen that IEM events give rise to saw tooth patterns (Figure 2.12d) that are reminiscent of experimental Doppler interferometry data.<sup>72</sup> Notably, the relative droplet charge remains much lower for the  $\text{NBu}_4^+$  systems ( $z/z_R = 0.59 \pm 0.05$ ) than for the  $\text{Na}^+$  charged droplets ( $z/z_R = 0.74 \pm 0.05$ ). These averages and standard deviations are based on three independent runs for each set of conditions. Readers are reminded that both types of droplets contain  $\text{NaCl}$  as background electrolyte. In principle, therefore,  $\text{NBu}_4^+$  droplets should also be able to undergo  $\text{Na}^+$  ejection. However, we did not observe a single  $\text{Na}^+$  ejection event as long as there was any  $\text{NBu}_4^+$  present. In other words, the ejection of  $\text{NBu}_4^+$  is much more facile than that of  $\text{Na}^+$ . This difference is due to the tendency of  $\text{Na}^+$  to stay inside the droplet where it is tightly solvated. In contrast, the hydrophobic alkyl chains of  $\text{NBu}_4^+$  drive this ion to the surface, thereby facilitating its ejection at much lower  $z/z_R$  values.



**Figure 2.12.** MD simulation results for the evaporation of ESI droplets carrying an initial 12+ charge caused by excess Na<sup>+</sup> or NBu<sub>4</sub><sup>+</sup>. All droplets contained an additional 10 Na<sup>+</sup>/Cl<sup>-</sup> pairs as background electrolyte. (a) Na<sup>+</sup> charged droplet at  $t = 0$ , 2.5 ns, and 21.5 ns, top to bottom. (b) NBu<sub>4</sub><sup>+</sup> charged droplet at the same time points as in panel (a). Element coloring: Na<sup>+</sup> blue, Cl<sup>-</sup> green, C black, O red. (c) Droplet size as a function of time. Also indicated is the number of water molecules in the droplet at the beginning and at the end of the simulation window. (d) Droplet charge  $z$  relative to the Rayleigh charge  $z_R$ . Solid circles indicate ion ejection events.

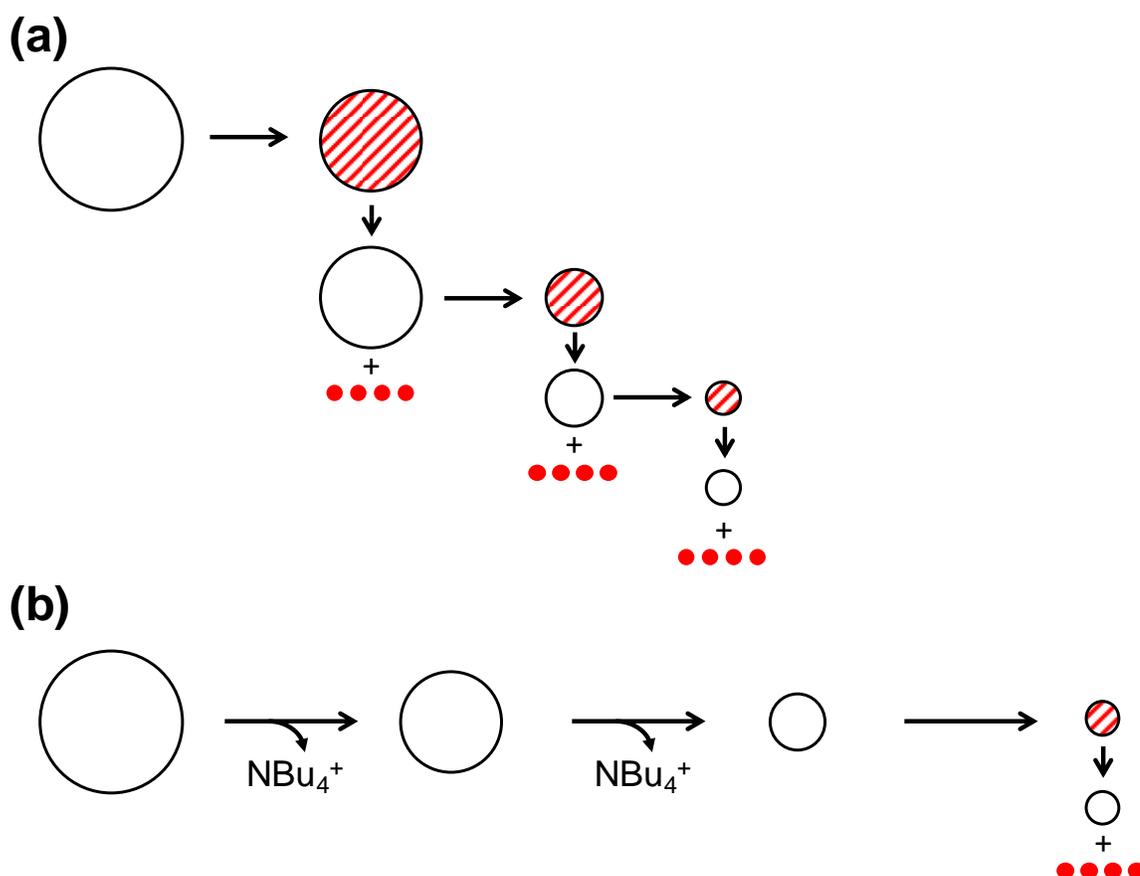
Another interesting observation is that solvent evaporation proceeds faster for the Na<sup>+</sup> charged droplets than for the NBu<sub>4</sub><sup>+</sup> containing systems (Figure 2.12c). The former shrink from an initial value close to 4000 H<sub>2</sub>O down to less than 100 H<sub>2</sub>O in 35 ns. In the presence of NBu<sub>4</sub><sup>+</sup> the number of water molecules remaining after the same time period is roughly ten times greater.

The different evaporation rates likely reflect a destabilization of the hydrogen bonding network within the electrostatically stressed (high  $z/z_R$ )  $\text{Na}^+$  containing droplets. Another contributing factor could be the presence of  $\text{NBU}_4^+$  ions at the surface which keep the droplet partially wrapped in an organic layer, thereby inhibiting water evaporation due to steric effects.

#### 2.3.4. Mechanism of Protein Ion Suppression

On the basis of the data described above one can propose a mechanistic framework to account for the interfering effects that are associated with different salts. Shrinking droplets that approach the Rayleigh limit have two options for relieving electrostatic stress. The first possibility is the emission of small progeny droplets via jet fission.<sup>6</sup> Alternatively, the droplet may eject ions via the IEM.<sup>13</sup> The two pathways are in kinetic competition with each other, governed by the corresponding activation barriers.<sup>13</sup> Under typical conditions fission is the preferred option for larger droplets, whereas for radii below  $\sim 10$  nm the high surface electric field favors ion ejection.<sup>7, 13</sup> Importantly, it is the small progeny droplets that ultimately produce gaseous protein ions.<sup>7, 12, 32</sup> Hence, the analyte ion yield is directly related to the number of nanodroplets generated in the ESI plume.<sup>7</sup>

*Protein ESI in the Absence of Salt Contaminants:* Under favorable conditions the evolution of ESI droplets proceeds as outlined in Figure 2.13.<sup>7, 32</sup> Early droplets with  $z/z_R \approx 0.5$  undergo solvent evaporation until they approach  $z/z_R \approx 0.8$  (indicated by hatch markings in Figure 2.13a).<sup>6, 32</sup> At this point jet fission produces a litter of progeny droplets (indicated in solid red).<sup>6</sup> The charge-depleted parents undergo further evaporation until another round of jet fission produces additional progeny droplets. These evaporation/fission events repeat themselves through several generations. The prevalence of  $\text{H}^+$  or  $\text{NH}_4^+$  as excess charge carriers under typical conditions<sup>12, 73</sup> favors the formation of clean  $[\text{M} + z\text{H}]^{z+}$  ions from the nanodroplets, keeping in mind that  $\text{NH}_4^+$  adducts will be lost as  $\text{NH}_3$  during ion sampling.<sup>7</sup>



**Figure 2.13.** Schematic of droplet evolution within the ESI plume. Time increases from left to right. Droplet shrinkage due to solvent evaporation is indicated by horizontal arrows, fission events that produce progeny droplets (solid red) are represented by vertical arrows. The progeny droplets will subsequently produce gaseous protein ions. Parent droplets close to the Rayleigh limit carry a red hatch pattern. (a) Shrinkage/fission scenario for droplets charged with  $\text{H}^+$ ,  $\text{NH}_4^+$ , or  $\text{Na}^+$ .<sup>7, 32</sup> (b) Scenario encountered in the presence of  $\text{NBu}_4^+$ , where shrinking droplets remain far below  $z_R$  due to the facile loss of  $\text{NBu}_4^+$ . Droplet fission only takes place *after* ejection of all  $\text{NBu}_4^+$ . Key features to note in panel (b): (i) the low number of progeny droplets implies a low protein ion yield; and (ii) droplet evaporation is slowed down compared to panel (a).

*ESI in NaCl Solutions:* Our data demonstrate that the main problem encountered in the presence of NaCl is extensive adduction that spreads the total ion count over numerous  $[M + zH + n(\text{Na} - H) + m(\text{Cl} + H)]^{z+}$  species. Adducts are particularly prevalent under CRM conditions.<sup>41</sup> Figure 2.4 reveals that  $R(C)$  is not strongly affected by the presence of NaCl. In addition, the protein charge state distributions with and without NaCl remain almost unchanged (Figures 2.3, 2.5, 2.6). Taken together, this implies that (i) NaCl does *not* cause major interferences with the droplet evaporation/fission events outlined in Figure 2.13, and (ii) NaCl does *not* strongly interfere with release of protein ions from nanodroplets. The cartoon representation of Figure 2.13a therefore applies to both salt-free and NaCl contaminated protein solutions.

*Effects of NBu<sub>4</sub><sup>+</sup>:* A key feature observed in our simulations is the facile IEM ejection of NBu<sub>4</sub><sup>+</sup> which keeps the corresponding droplets far below the Rayleigh limit (Figure 2.12d). Emission of NBu<sub>4</sub><sup>+</sup> will take place already for droplets much larger than those in Figure 2.12. NBu<sub>4</sub><sup>+</sup> emission thus prevents these “would-be parents” from reaching the threshold of  $z/z_R \approx 0.8$  which is required for progeny droplet formation via jet fission.<sup>6, 32</sup> Since progeny droplets are the precursors of gaseous analyte ions,<sup>7</sup> the reduced value of  $z/z_R$  represents a key reason why the protein ion yield is lowered by NBu<sub>4</sub><sup>+</sup>. Analogous considerations likely apply to other ionic additives with high surface affinity. Jet fission will take place only *after* all NBu<sub>4</sub><sup>+</sup> have drained from the droplet, because it is only then that droplets can reach the critical range of  $z/z_R \approx 0.8$  (Figure 2.13b). The absence of NBu<sub>4</sub><sup>+</sup> from the final nanodroplets is evident from the fact that charge states formed in NBu<sub>4</sub>Cl are similar to those observed for salt-free solution (Figure 2.7, 2.10). The presence of residual NBu<sub>4</sub><sup>+</sup> at the point of protein release would manifest itself as a shift to lower charge states,<sup>7, 12, 14</sup> because the corresponding nanodroplets would be far below the Rayleigh limit (Figure 2.12d). Overall, a key factor responsible for protein ion suppression by NBu<sub>4</sub><sup>+</sup> appears to be a shift in the kinetic competition from jet fission towards charge carrier ejection (Figure 2.13b).

An additional aspect to consider is that solvent evaporation in the presence of NBu<sub>4</sub><sup>+</sup> is quite slow (Figure 2.12c). The prolonged droplet lifetime increases the likelihood of droplet collisions with metal components of the ion sampling interface, thereby further lowering the yield of gaseous protein ions. The cartoon in Figure 2.13b indicates both the lower progeny droplet yield and the longer droplet lifetime in the presence of NBu<sub>4</sub><sup>+</sup>.

*Interferences Caused by Cs<sup>+</sup>*: The propensity of Cs<sup>+</sup> to suppress the protein ion intensity is lower than that of NBu<sub>4</sub><sup>+</sup> (Figure 2.11). This suggests that the ability of Cs<sup>+</sup> to interfere with the ESI process via charge loss from early droplets is less pronounced. Consistent with this interpretation, the hydration free energy of Cs<sup>+</sup> ( $\Delta_{hyd}G^\circ = -250 \text{ kJ mol}^{-1}$ ) is intermediate between those of Na<sup>+</sup> and NBu<sub>4</sub><sup>+</sup>.<sup>45</sup> From an experimental point of view the main problem encountered in the presence of Cs<sup>+</sup> is the formation of adducts (Figures 2.7, 2.10).

## 2.4. Conclusions

The starting point for the current investigation was the hypothesis that salt interferences in protein ESI-MS can be dissected into two unrelated effects, i.e., adduct formation and ion suppression (Figure 2.1). The findings discussed above strongly support the validity of this idea. NaCl induces major adduction, while not strongly interfering with the total ion intensity. Conversely, NBu<sub>4</sub>Cl leads to the near-complete breakdown of protein ion formation, while not causing any adduction. Other additives may show hybrid scenarios, exemplified by CsCl in this work. The key factor that governs the nature of the interfering effects appears to be the surface affinity of the cation. Species with high surface affinity ( $\Delta_{hyd}G^\circ \geq 0$ ) tend to cause ion suppression, whereas those with low surface affinity ( $\Delta_{hyd}G^\circ \ll 0$ ) tend to cause adduction. It is possible that the extent of adduction can be modulated somewhat by factors such as protein size and amino acid composition.

Equation 2.1 plays a key role in mechanistic discussions of the ESI process. Specifically, Rayleigh's theory is commonly used for estimating the amount of charge that can be accommodated on a stable droplet.<sup>7, 32</sup> However, Equation 2.1 was derived on the basis of a continuum model, without any atomistic details.<sup>70</sup> Recent work<sup>11, 74</sup> has already suggested that shrinking droplets may lose charge due to IEM events far below the Rayleigh limit. The feasibility of these premature ion ejection events depends on the nature of the charge carrier, as well as the electric field at the droplet surface. Consistent with those proposals<sup>11, 74</sup> the current work demonstrates that ions with high surface affinity (such as NBu<sub>4</sub><sup>+</sup>) get ejected from ESI droplets at  $z/z_R \ll 1$ . We propose that this phenomenon represents the root cause for the occurrence of protein ion suppression in the presence of low MW ions that possess high surface affinity. Specifically, we propose that the hemorrhaging of charge at  $z/z_R \ll 1$  prevents shrinking

ESI droplets from undergoing jet fission. Reduced jet fission implies that a lower number of progeny droplets will be formed which are the precursors of gaseous protein ions.<sup>7</sup> In addition, the reduced droplet evaporation rate in the presence of  $\text{NBu}_4^+$  may lower the yield of gaseous protein ions as well.

Finally, we return to the concept of charge competition that is frequently mentioned in the context of ESI-MS signal suppression.<sup>17, 32-33, 36</sup> The framework proposed here (Figure 2.13) is consistent with a competition mechanism in the ESI plume, where kinetic partitioning favors either the IEM emission of small ions or the production of progeny droplets. The emission of small ions drains the parent droplet charge in an unproductive fashion, whereas the formation of nanodroplets will ultimately lead to production of gaseous protein ions. The competing nature of these two processes becomes clear when realizing that the total amount of excess charge on the initial ESI droplets is limited; hence, any charge that is lost via ejection of small ions will not be available for the production of protein ions. This competition framework is suitable for understanding the effects of species such as  $\text{NBu}_4^+$  that reduce the total protein ion signal (Figure 2.1c). On the other hand, the behavior of salts such as  $\text{NaCl}$  is primarily based on the rather trivial spreading of the ion count over a wide  $m/z$  range via adduct formation (Figure 2.1b). It is hoped that the ideas put forward here will improve the general understanding of salt interferences in ESI-MS, paving the way towards strategies to better cope with contaminants in biological samples.

## 2.5. References

1. Fenn, J. B., *Angew. Chem. Int. Ed.* **2003**, *42*, 3871-3894.
2. Lössl, P.; Snijder, J.; Heck, A. J. R., *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 906-917.
3. Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R., *Protein Sci.* **2013**, *22*, 530-544.
4. Marciano, D. P.; Dharmarajan, V.; Griffin, P. R., *Curr. Op. Struct. Biol.* **2014**, 105-111.
5. Konermann, L.; Vahidi, S.; Sowole, M. A., *Anal. Chem.* **2014**, *86*, 213-232.
6. Gomez, A.; Tang, K., *Phys. Fluids* **1994**, *6*, 404-414.
7. Kebarle, P.; Verkerk, U. H., *Mass Spectrom. Rev.* **2009**, *28*, 898-917.
8. Marchese, R.; Grandori, R.; Carloni, R.; Raugei, S., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1903-1910.
9. Ogorzalek Loo, R. R.; Lakshmanan, R.; Loo, J. A., *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1675-1693.
10. Nguyen, S.; Fenn, J. B., *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1111-1117.
11. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L., *Anal. Chem.* **2009**, *81*, 369-377.
12. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., *Anal. Chem.* **2013**, *85*, 2-9.
13. Iribarne, J. V.; Thomson, B. A., *J. Chem. Phys.* **1976**, *64*, 2287-2294.
14. de la Mora, F. J., *Anal. Chim. Acta* **2000**, *406*, 93-104.
15. Consta, S.; Oh, M. I.; Soltani, S., *Int. J. Mass spectrom.* **2014**, (in press).
16. Larriba, C.; de la Mora, F.; Clemmer, D. E., *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1332-1345.
17. Cech, N. B.; Enke, C. G., *Mass Spectrom. Rev.* **2001**, *20*, 362-387.
18. Krueve, A.; Kaupmees, K.; Liigand, J.; Oss, M.; Leito, I., *J. Mass Spectrom.* **2013**, *48*, 695-702.
19. Burdette, C. Q.; Marcus, R. K., *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 975-978.
20. Hannis, J. C.; Muddiman, D. C., *Rapid Commun. Mass Spectrom.* **1999**, *13*, 323-330.
21. Cavanagh, J.; Benson, L. M.; Thompson, R.; Naylor, S., *Anal. Chem.* **2003**, *75*, 3281-3286.
22. Flick, T. G.; Cassou, C. A.; Chang, T. M.; Williams, E. R., *Anal. Chem.* **2012**, *84*, 7511-7517.

23. Chen, Y. J.; Mori, M. X.; Pastusek, A. C.; Schug, K. A.; Dasgupta, P. K., *Anal. Chem.* **2011**, *83*, 1015-1021.
24. DeMuth, J. C.; McLuckey, S. A., *Anal. Chem.* **2015**, *87*, 1210-1218.
25. Creighton, T. E., *Proteins*. W. H. Freeman & Co: New York, **1993**.
26. Hop, C. E. C. A., *J. Mass Spectrom.* **1996**, *31*, 1314-1316.
27. Zhang, D. X.; Cooks, R. G., *Int. J. Mass spectrom.* **2000**, *195*, 667-684.
28. Feketeova, L.; O'Hair, R. A. J., *Rapid Commun. Mass Spectrom.* **2009**, *23*, 60-64.
29. Juraschek, R.; Dulcks, T.; Karas, M., *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 300-308.
30. Wang, G. D.; Cole, R. B., *Anal. Chem.* **1998**, *70*, 873-881.
31. Konermann, L.; McAllister, R. G.; Metwally, H., *J. Phys. Chem. B* **2014**, *118*, 12025-12033.
32. Kebarle, P.; Tang, L., *Anal. Chem.* **1993**, *65*, 972A-986A.
33. Pan, P.; McLuckey, S. A., *Anal. Chem.* **2003**, *75*, 5468-5474.
34. Zhou, S.; Cook, K. D., *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 206-214.
35. Chalcraft, K. R.; Lee, R.; C., M.; Britz-McKibbin, P., *Anal. Chem.* **2009**, *81*, 2506-2515.
36. Wang, G. D.; Cole, R. B., *Anal. Chem.* **1994**, *66*, 3702-3708.
37. Koszinowski, K.; Lissy, F., *Int. J. Mass Spectrom.* **2013**, *354*, 219-228.
38. Chang, D.-Y.; Lee, C.-C.; Shiea, J., *Anal. Chem.* **2002**, *74*, 2465-2469.
39. Verkerk, U. H.; Kebarle, P., *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1325-1341.
40. Grewal, R. N.; El Aribi, H.; Smith, J. C.; Rodriquez, C. F.; Hopkinson, A. C.; Siu, K. W. M., *Int. J. Mass spectrom.* **2002**, *219*, 89-99.
41. Yue, X.; Vahidi, S.; Konermann, L., *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1322-1331.
42. Huber, C. G.; Buchmeister, M. R., *Anal. Chem.* **1998**, *70*, 5288-5295.
43. Pan, P.; McLuckey, S. A., *Anal. Chem.* **2003**, *75*, 1491-1499.
44. Hagan, N.; Fabris, D., *Biochemistry* **2003**, *42*, 10736-10745.
45. Marcus, Y., *J. Chem. Soc.-Faraday Trans.* **1991**, *87*, 2995-2999.
46. Loo, J. A.; Edmonds, C. G.; Udseh, H. R.; Smith, R. D., *Anal. Chem.* **1990**, *62*, 693-698.
47. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I., *J. Comp. Chem.* **2010**, *31*, 671-690.
48. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., *J. Chem. Theory Comput.* **2008**, *4*, 435-447.

49. Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G., *J. Chem. Inf. Mod.* **2012**, *52*, 1757-1768.
50. Vanommeslaeghe, K.; Raman, E. P.; MacKerell Jr., A. D., *J. Chem. Inf. Mod.* **2012**, *52*, 3155-3168.
51. Bussi, G.; Donadio, D.; Parrinello, M., *J. Chem. Phys.* **2007**, *126*, 0141011-0141017.
52. Lide, D. R., *CRC Handbook of Chemistry and Physics*. 82nd ed.; CRC Press: Boca Raton, London, New York, Washington, 2001.
53. Vega, C.; de Miguel, E., *J. Chem. Phys.* **2007**, *126*, 154707.
54. Caleman, C.; van der Spoel, D., *Phys. Chem. Chem. Phys.* **2007**, *9*, 5105-5111.
55. Znamenskiy, V.; Marginean, I.; Vertes, A., *J. Phys. Chem. A* **2003**, *107*, 7406-7412.
56. Daub, C. D.; Cann, N. M., *Anal. Chem.* **2011**, *83*, 8372-8376.
57. Steinberg, M. Z.; Breuker, K.; Elber, R.; Gerber, R. B., *Phys. Chem. Chem. Phys.* **2007**, *9*, 4690-4697.
58. Luedtke, W. D.; Landmann, U.; Chiu, Y.-H.; Levandier, D. J.; Dressler, R. A.; Sok, S.; Gordon, M. S., *J. Phys. Chem. A* **2008**, *112*, 9628-9649.
59. Caleman, C.; Hub, J. S.; van Maaren, P. J.; van der Spoel, D., *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6838-6842.
60. Jungwirth, P.; Tobias, D. J., *Chem. Rev.* **2006**, *106*, 1259-1281.
61. Coker, H., *J. Phys. Chem.* **1976**, *80*, 2078-2084.
62. Heyda, J.; Lund, M.; Oncak, M.; Slavicek, P.; Jungwirth, P., *J. Phys. Chem. B* **2010**, *114*, 10843-10852.
63. Grandori, R., *Protein Sci.* **2002**, *11*, 453-458.
64. Kaltashov, I. A.; Eyles, S. J., *Mass Spectrom. Rev.* **2002**, *21*, 37-71.
65. Susa, A. C.; Mortenson, D. N.; Williams, E. R., *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 918-927.
66. Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431-441.
67. Benesch, J. L. P.; Robinson, C. V., *Nature* **2009**, *462*, 576-577.
68. Girod, M.; Dagany, X.; Boutou, V.; Broyer, M.; Antoine, R.; Dugourd, P.; Mordehai, A.; Love, C.; Werlich, M.; Fjeldsted, J.; Stafford, G., *Phys. Chem. Chem. Phys.* **2012**, *14*, 9389-9396.

69. Wang, R.; Zenobi, R., *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 378-385.
70. Rayleigh, L., *Philos. Mag.* **1882**, *14*, 184-186.
71. Hrobarik, T.; Vrbka, L.; Jungwirth, P., *Biophys. Chem.* **2006**, *124*, 238-242.
72. Smith, J. N.; Flagan, R. C.; Beauchamp, J. L., *J. Phys. Chem. A* **2002**, *106*, 9957-9967.
73. Van Berkel, G. J.; Kertesz, V., *Anal. Chem.* **2007**, *79*, 5511-5520.
74. Allen, S. J.; Schwartz, A. M.; Bush, M. F., *Anal. Chem.* **2013**, *85*, 12055–12061.

## Chapter 3. Mechanism of Protein Supercharging by Sulfolane and m-NBA: Molecular Dynamics Simulations of the Electrospray Process

### 3.1. Introduction

The capability to transfer intact proteins and protein complexes from solution into the gas phase by electrospray ionization (ESI)<sup>1</sup> has revolutionized the area of mass spectrometry (MS).<sup>2-5</sup> During the ESI process analyte solution is dispersed into charged droplets. These droplets undergo evaporation and jet fission close to the Rayleigh limit,<sup>6-7</sup> ultimately producing nanodroplets from which gaseous protein ions are released.<sup>8-9</sup> These ions are then analyzed according to their mass-to-charge ratio. Additional information is obtainable by applying various excitation and fragmentation techniques,<sup>10-14</sup> as well as ion mobility spectrometry (IMS).<sup>15-18</sup> Early ESI experiments mostly employed denaturing solvents. Recent years have witnessed the rise of “native” ESI, which uses aqueous solutions at neutral pH. Protein ions generated by native ESI can retain solution-like structures, provided that collisional activation is minimized.<sup>2, 5, 15, 19</sup>

A key factor that governs the behavior of gaseous proteins is their charge state  $z$ . Low  $z$  values are usually associated with compact conformations, while proteins with elevated charge states tend to adopt extended structures.<sup>15-16, 20</sup> High  $z$  values are beneficial for fragmentation experiments,<sup>11, 21</sup> and they increase the performance characteristics of Fourier transform mass analyzers.<sup>14, 22</sup> The following discussion is restricted to the widely used positive ESI mode, where analytes carry excess protons and/or metal cations.<sup>8</sup>

Not surprisingly, there is tremendous interest in methods for manipulating the charge states of gaseous proteins. Native ESI usually results in low  $z$  values, whereas denaturing conditions produce ions that are highly charged.<sup>23-25</sup> This effect has been linked to the mechanisms by which different conformers emerge from aqueous ESI nanodroplets.<sup>9</sup> While several aspects of the ESI process remain contentious,<sup>26-27</sup> there is considerable support for the view that compact structures are released upon solvent evaporation to dryness<sup>28-29</sup> as described by the charged residue mechanism (CRM).<sup>30</sup> The CRM produces ions that have  $z$  values close to

the Rayleigh charge of protein-sized water droplets.<sup>8-9, 31</sup> For unfolded proteins a chain ejection model (CEM)<sup>2, 9</sup> has been proposed that envisions the extrusion of polymer chains from the nanodroplet surface.<sup>32</sup> The CEM attributes the high  $z$  values of unfolded proteins to H<sup>+</sup> migration from the droplet to the protruding chain.<sup>9</sup> Shifts in charge states may also originate from protein structural changes within the droplet.<sup>33-34</sup> For example, exposure of a native ESI plume to heat<sup>35</sup> or acidic vapors<sup>36</sup> can induce unfolding, resulting in elevated  $z$  values. Proton transfer reactions<sup>16, 37-38</sup> and electron capture<sup>39-40</sup> allow modifying charge states after protein release from the droplet.

Supercharging agents (SCAs) offer an alternative way to modulate protein  $z$  values during ESI.<sup>41</sup> SCAs are added to the bulk solution in low concentration where they do not affect protein structure or stability.<sup>42</sup> Yet, ESI of these samples produces greatly elevated charge states.<sup>41, 43</sup> Supercharging can increase  $z$  values for native<sup>42, 44-49</sup> and for denaturing solutions.<sup>50-53</sup> SCAs have large dipole moments.<sup>42-43, 51</sup> Their surface tension values are in-between those of methanol and water.<sup>51</sup> In addition to one or more polar groups, most SCAs possess an aliphatic moiety.<sup>43, 51</sup> The low volatility of SCAs (bp > 180 °C)<sup>42-43, 51</sup> causes differential evaporation. Hence, ESI nanodroplets have a SCA content that is significantly elevated relative to the bulk solution.<sup>27, 51</sup> Two widely used SCAs that embody all these characteristics are sulfolane (C<sub>4</sub>H<sub>8</sub>SO<sub>2</sub>) and *m*-NBA (*m*-nitrobenzyl alcohol, HO-CH<sub>2</sub>-C<sub>6</sub>H<sub>4</sub>-NO<sub>2</sub>, see Figure 3.1 insets).<sup>43, 45-46, 48, 51-52</sup>

The mechanism of supercharging has yet to be elucidated.<sup>27, 51</sup> Despite the common features noted above, there is no clear correlation between the effectiveness of SCAs and their physicochemical properties.<sup>51</sup> Initial work focused on the surface tension  $\gamma$ .<sup>41</sup> The charge of a droplet with radius  $r$  at the Rayleigh limit<sup>8</sup> is  $z_R = 8\pi(\epsilon_0 \gamma r^3)^{1/2}/e$ , where  $\epsilon_0$  is the vacuum permittivity and  $e = 1.602 \times 10^{-19}$  C. If SCA enrichment were to increase surface tension, the droplet should support more charge and produce protein ions with higher  $z$ .<sup>9, 41</sup> Unfortunately, this idea is inconsistent with experimental data.<sup>54</sup> For example, *m*-NBA enrichment in aqueous droplets *lowers* the surface tension ( $\gamma_{m\text{-NBA}} < \gamma_{\text{water}}$ ), but *m*-NBA *increases* protein charging in native ESI.<sup>43</sup> Other proposals envision that SCA-enriched droplets trigger chemical<sup>42</sup> or thermal<sup>55</sup> denaturation. Unfolding within the droplet would then produce high charge states, analogous to conditions where proteins are electrosprayed in denaturing bulk solutions.<sup>23-25</sup> This unfolding mechanism is at odds with reports that some supercharged proteins retain a compact fold.<sup>44, 56</sup>

Other investigations considered direct SCA-protein interactions,<sup>47, 52</sup> as well as the SCA Brønsted acid/base chemistry.<sup>27</sup> None of those earlier studies has yielded compelling evidence for any particular mechanism.

Molecular dynamics (MD) simulations can provide detailed insights into the ESI process.<sup>7, 9, 32, 57-63</sup> Several studies demonstrated the field emission of small charge carriers from the droplet surface.<sup>7, 9, 32, 58, 63</sup> These events are consistent with the ion evaporation model (IEM).<sup>64-65</sup> MD work from our laboratory resulted in the an atomistic view of CRM protein release from aqueous nanodroplets.<sup>57</sup> The IEM ejection of small ions (such as Na<sup>+</sup>) plays an ancillary role during the CRM<sup>57</sup> by keeping the shrinking water droplets close to  $z_R$ .<sup>28-29</sup> Surprisingly, there appear to be no previous MD studies on protein supercharging.

Here we use a combination of experiments and MD techniques to elucidate the supercharging mechanism for two archetypical SCAs, sulfolane and *m*-NBA. Our simulations focus on late generation nanodroplets within the ESI plume from which protein ions are released into the gas phase.<sup>8</sup> Holo-myoglobin (hMb) was chosen as model system, reflecting the use of this heme-protein complex in numerous earlier supercharging experiments.<sup>27, 42-43, 46-47</sup> Also, hMb represents a minimalist example of a biologically relevant complex, keeping in mind the steadily increasing use of SCAs for investigations on noncovalent assemblies.<sup>42-43, 48-49, 56</sup> We specifically focus on the role of SCAs in neutral aqueous solution, where proteins adopt their biologically active conformations as they enter the ESI process.<sup>2, 5, 15, 19</sup> Our data reveal that supercharging proceeds via a previously unrecognized mechanism, where solvent segregation induces charge trapping in the droplet and on the protein.

## 3.2. Materials and Methods

### 3.2.1. Electrospray Ionization Mass Spectrometry and Ion Mobility Spectrometry

Equine hMb, sulfolane, *m*-NBA and ammonium acetate were purchased from Sigma (St. Louis, MO). Protein samples were centrifuged to remove small amounts of insoluble debris, followed by dialysis against 10 mM aqueous ammonium acetate. All solutions had a protein

concentration of 10  $\mu\text{M}$  and contained 10 mM ammonium acetate. For supercharging experiments the samples were supplemented with 1% (v/v) sulfolane or *m*-NBA.

Mass spectra were acquired on a Synapt HDMS instrument (Waters, Milford, MA). The Z-spray ESI source was operated at 2.8 kV. The desolvation and source temperatures were 200 °C and 80 °C, respectively, and the cone was at 20 V. Average charge states were calculated as  $z_{av} = \Sigma(z_i I_i)/I_i$ , where  $I_i$  is the signal intensity of charge state  $i$ . IMS data were recorded using the instrument's travelling-wave cell. Potential gradients were tuned to be as gentle as possible to minimize collisional activation. This includes lowering the cone to 5 V.<sup>66</sup> Drift times were converted to collision cross sections ( $\Omega$ ). Instrument settings and  $\Omega$  calibration procedure have been summarized elsewhere.<sup>66</sup>

### 3.2.2. MD Simulations

All-atom MD simulations were conducted using the CHARMM36 force field<sup>67</sup> in Gromacs 5.0 with GPU acceleration.<sup>68</sup> Our choice of CHARMM is based on its excellent performance for proteins in solution<sup>69</sup> as well as for droplet systems.<sup>57</sup> All bonds were constrained, using LINCS<sup>70</sup> for protein and SCAs, and SETTLE<sup>71</sup> for water, thereby allowing for a 2 fs integration time step. The hMb X-ray coordinates 1WLA<sup>72</sup> served as starting structure. Side chains and termini were set to their default pH 7 protonation states; this includes deprotonation of both heme propionates. TIP4P/2005 water<sup>73</sup> was chosen because it reproduces the experimental surface tension better than other models.<sup>74</sup> The topology and structure files for sulfolane were taken from the GROMACS Molecule & Liquid Database.<sup>75</sup> Parameters for *m*-NBA were obtained using the ParamChem Server<sup>76</sup> and the ZINC database.<sup>77</sup> Initial ESI droplets were generated by surrounding the protein with pre-equilibrated water. For SCA simulations 315 sulfolane or *m*-NBA were inserted at random positions, while eliminating any water molecules that would result in steric clashes. In-house Perl software was used to carve the systems into spherical droplets. The initial radius was  $r_0 = 3.5$  nm, corresponding to 5400 water molecules for purely aqueous droplets, with the protein at its center. 26 random water molecules were replaced with  $\text{Na}^+$ . Together with the hMb 2- net charge the resulting net droplet charge was 24+, corresponding to  $z_R$  for a 3.5 nm aqueous droplet at 370 K.<sup>8</sup> This charge regime is well within

the range of experimentally measured values.<sup>6</sup> A few short runs were also conducted on larger droplets (5.5 nm radius).

Simulations were run by placing the droplets in a vacuum environment, effectively without cutoffs for van der Waals or electrostatic interactions.<sup>57</sup> The droplets were initially subjected to steepest descent energy minimization. Subsequent production runs followed a recently developed trajectory stitching approach.<sup>57</sup> This method breaks up simulations into 250 ps segment during which the temperature was controlled by a Nosé-Hoover<sup>78</sup> thermostat with a coupling constant of 0.5 ps. After each segment the system was run through an in-house Fortran program to remove any evaporated moieties (H<sub>2</sub>O, sulfolane, *m*-NBA, or Na<sup>+</sup>) that had moved more than 7 nm from the center of mass. New velocities were then assigned from a Maxwell-Boltzmann distribution prior to beginning of the next simulation window. This trajectory stitching dramatically reduces wall clock time by gradually decreasing the number of atoms in the simulation.<sup>57</sup> Unless noted otherwise, the simulations were initially run at 370 K for 75 ns. Subsequently the temperature was elevated to 450 K for facilitating the evaporation of any remaining solvent molecules. This temperature profile reflects progressive heating of protein ions as they traverse the ion optics of the mass spectrometer.<sup>79</sup> The temperatures used are in line with experimental studies.<sup>80-81</sup> The overall simulation window was 350 ns. All ESI simulations were repeated five times with different initial atom positions and velocities.

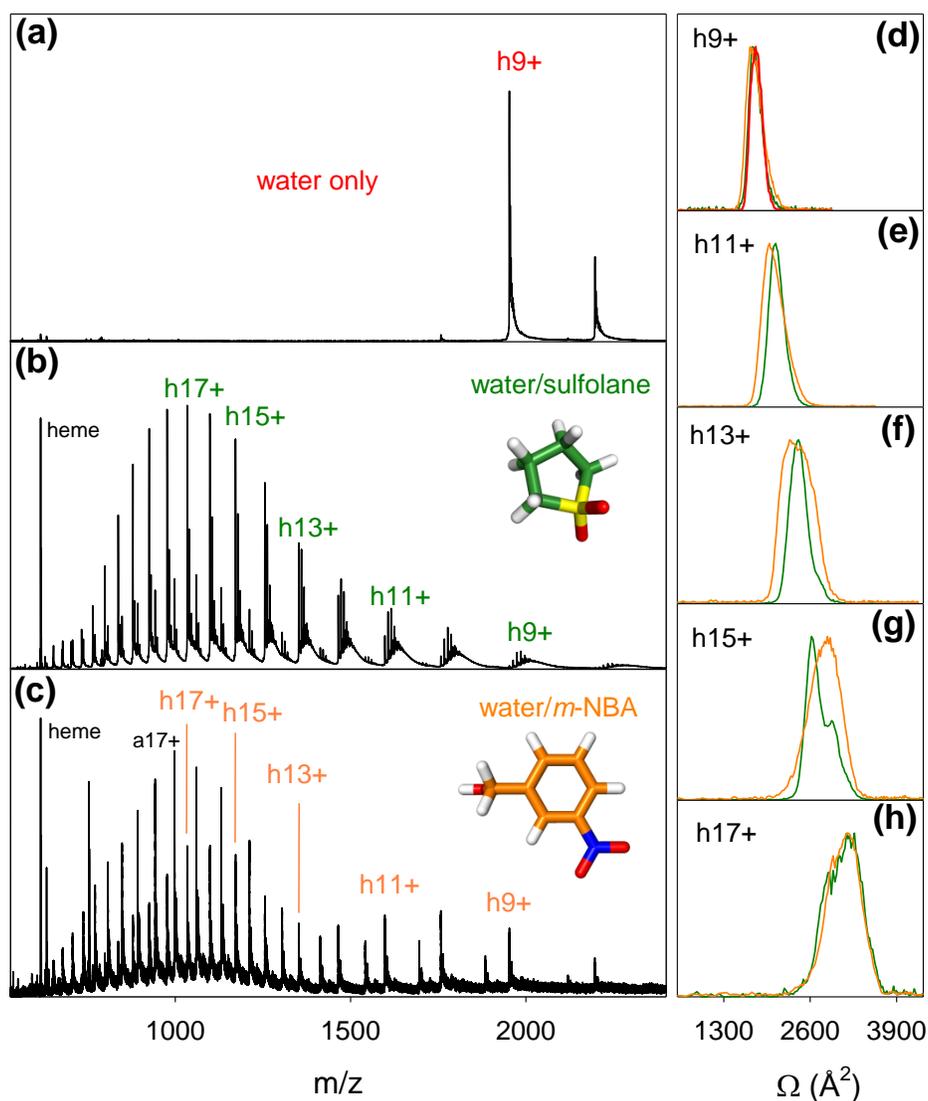
Early tests revealed that the slow SCA evaporation precluded formation of desolvated hMb within reasonable wall clock time. Starting at  $t = 150$  ns, SCA molecules were therefore subjected to “forced evaporation”. Under this scheme the SCA molecule farthest from the droplet center was eliminated after each 250 ps run segment, corresponding to an SCA evaporation rate of 4 ns<sup>-1</sup>. This strategy is reminiscent of biased<sup>82</sup> or steered<sup>83</sup> MD techniques used by others. Forced evaporation reduces the wall clock time from months to roughly one week for each droplet run. It could be argued that this forced evaporation scheme biases the temporal evolution of the system towards a CRM scenario. However, the approach chosen here seems well justified, considering that previous studies strongly argue against IEM or CEM-type behavior under native ESI conditions.

## 3.3. Results and Discussion

### 3.3.1. Supercharging of Holo-Myoglobin

Before discussing MD simulations, it is beneficial to examine the effects of SCAs experimentally. The hMb mass spectrum acquired in neutral aqueous solution (without SCA) is dominated by 9+ ions of the intact heme-protein complex (Figure 3.1a). Addition of 1% sulfolane shifts the spectrum to higher charge states, with hMb<sup>17+</sup> as the most intense ion. In addition, the spectrum shows minor apo-myoglobin (aMb) signals (Figure 3.1b). Similar supercharging effects are seen in the presence of 1% m-NBA, albeit with an elevated aMb abundance (Figure 3.1c).

The structures of the gaseous hMb were probed by IMS (Figure 3.1d-h). Regardless of solvent conditions, hMb<sup>9+</sup> displays a collision cross section around 1740 Å<sup>2</sup>. This value coincides with literature data, matching the calculated  $\Omega$  of the crystal structure to within 2%.<sup>84</sup> Elevated charge states show larger  $\Omega$  values, *e.g.*, 3260 Å<sup>2</sup> for hMb<sup>17+</sup>. Clearly, these highly charged ions are extensively unfolded.<sup>42</sup> Some of the  $\Omega$  distributions in Figure 3.1 are multimodal, reflecting the presence of co-existing conformers.<sup>15-18</sup> The data in Figure 3.1 are consistent with earlier reports,<sup>24, 27, 42-43, 46-47</sup> illustrating the dramatic effects of SCAs on protein charge states.<sup>42-43, 48-49,</sup>



**Figure 3.1.** ESI mass spectra of hMb in aqueous solution at pH 7. (a) Control without SCA. (b) Supercharged, 1% sulfolane, and (c) Supercharged, 1% *m*-NBA. (d) – (h) IMS data for selected charge states. Coloring: Red, purely aqueous; green, 1% sulfolane; orange, 1% *m*-NBA. The heme peak in b, c was scaled by 1/3. Insets in b, c show the structures of sulfolane and *m*-NBA. The signal at  $m/z$  754 in Figure 1c corresponds to a singly charged contaminant ion.

### 3.3.2. Choice of Charge Carrier for Simulations

Electrosprayed proteins usually appear as  $[M + zH]^{z+}$  ions, *i.e.*, they are charged by excess protons (exemplified in Figure 3.1). The inclusion of  $H^+$  in standard MD force fields is problematic.<sup>85</sup> Realistic proton simulations call for QM/MM<sup>86</sup>, *ab initio*<sup>60, 87</sup>, or DFT/MD

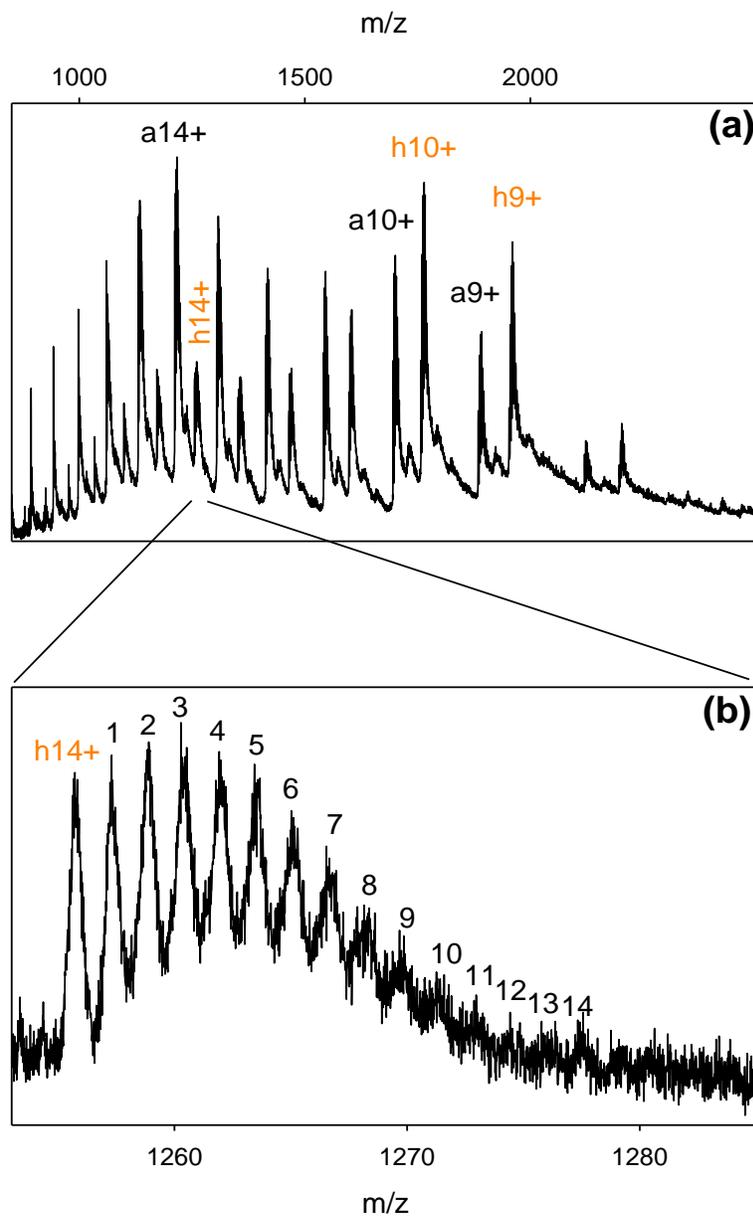
methods<sup>85</sup> that are out of reach for the size and time regime considered here. Luckily, there is a way to sidestep this issue.<sup>57</sup> Proteins electrosprayed in the presence of  $\text{Na}^+$  salts are charged by a combination of protonation and sodiation, all the way to fully sodiated  $[\text{M} + z\text{Na}]^{z+}$ .<sup>8</sup> Figure 3.2 demonstrates that this also applies to supercharged hMb. Similar to our earlier ESI simulations,<sup>57</sup> the present study thus uses  $\text{Na}^+$  as charge carrier, thereby simulating the formation of  $[\text{M} + z\text{Na}]^{z+}$  ions. The inclusion of  $\text{Na}^+$  in MD force fields is straightforward.<sup>67</sup> In neutral solution (and in the initial droplets used here) hMb has a net charge of 2-. Thus, the charge state of hMb ions carrying  $n$   $\text{Na}^+$  in our simulations is  $z = (n - 2)$ .

### 3.3.3. Charge Carrier Behavior in Different Solvents

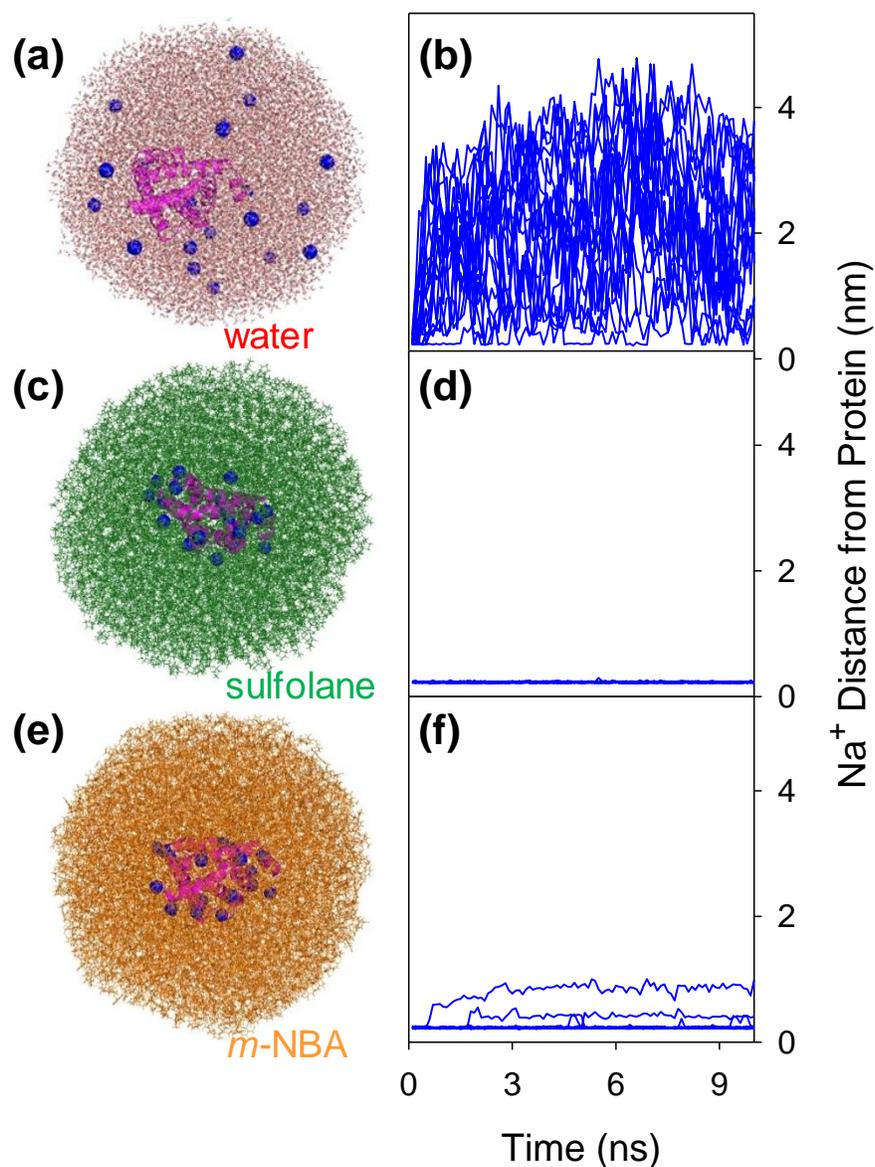
As a final step before conducting ESI simulations we examine the behavior of water, sulfolane, and *m*-NBA droplets under quasi-equilibrium conditions. Pure solvents (not mixtures) were used to generate protein-containing droplets with radius  $r_0 \approx 5.5$  nm. The temperature was kept low to ensure minimum evaporation (see Figure 3.3 for details). 20  $\text{Na}^+$  were included as excess charge carriers. For these tests we chose a starting configuration where all  $\text{Na}^+$  were initially placed in contact with carboxylates and/or carbonyl groups on the protein surface.

The three droplet types exhibit very different behavior. In water each of the  $\text{Na}^+$  immediately separates from the protein and diffuses throughout the droplet (Figure 3.3a, b). In sulfolane all  $\text{Na}^+$  remain associated with hMb (Figure 3.3c, d). Similarly, in *m*-NBA most of the  $\text{Na}^+$  remain attached; only a few charge carriers venture slightly away from the protein (Figure 3.3 e, f).

The unconventional starting configuration used for the simulations of Figure 3.3 (with all  $\text{Na}^+$  located at the protein surface for  $t = 0$ ) helps illustrate an important point. Water is an excellent solvent for  $\text{Na}^+$ ; its favorable hydration<sup>88</sup> readily outcompetes  $\text{Na}^+$  interactions with the protein.  $\text{Na}^+$  solvation by sulfolane and *m*-NBA is much less favorable, as reported for other solvents that carry aliphatic moieties.<sup>89</sup> In conclusion, small charge carriers such as  $\text{Na}^+$  exhibit high affinity for water, lower affinity for the protein surface, and very low affinity for SCA environments. It will be seen that these trends cause partitioning effects that are responsible for ESI supercharging.



**Figure 3.2.** (a) ESI mass spectrum obtained after infusion of holo-myoglobin (hMb) in aqueous solution containing 0.1 mM sodium acetate and 1% *m*-NBA. Some of the hMb undergoes heme loss, generating apo-myoglobin (aMb). Selected charge states are denoted using “h” and “a” for hMb and aMb, respectively. (b) Expanded view of hMb<sup>14+</sup> supercharged ions. Numbers 1, 2, 3, ... indicate how many Na<sup>+</sup> ions are bound to the protein ions as charged adducts. The sodiation progression extends all the way to [M + 14Na]<sup>14+</sup>.



**Figure 3.3.** Quasi-equilibrium MD data for ESI droplets containing hMb (radius  $\sim 5.5$  nm) and 20  $\text{Na}^+$ . Simulation snapshots on the left were taken after 2 ns. Plots show the distance of individual  $\text{Na}^+$  from the closest protein heavy atom. (a, b) Water at 330 K; (c, d) Sulfolane at 370 K; (e, f) *m*-NBA at 370 K. At the onset of the simulations all  $\text{Na}^+$  were placed in direct contact with the protein. The net droplet charge corresponds to  $0.42 z_R$  (for water with  $\gamma = 0.066 \text{ N m}^{-1}$  at 330 K).

### 3.3.4. ESI Simulations

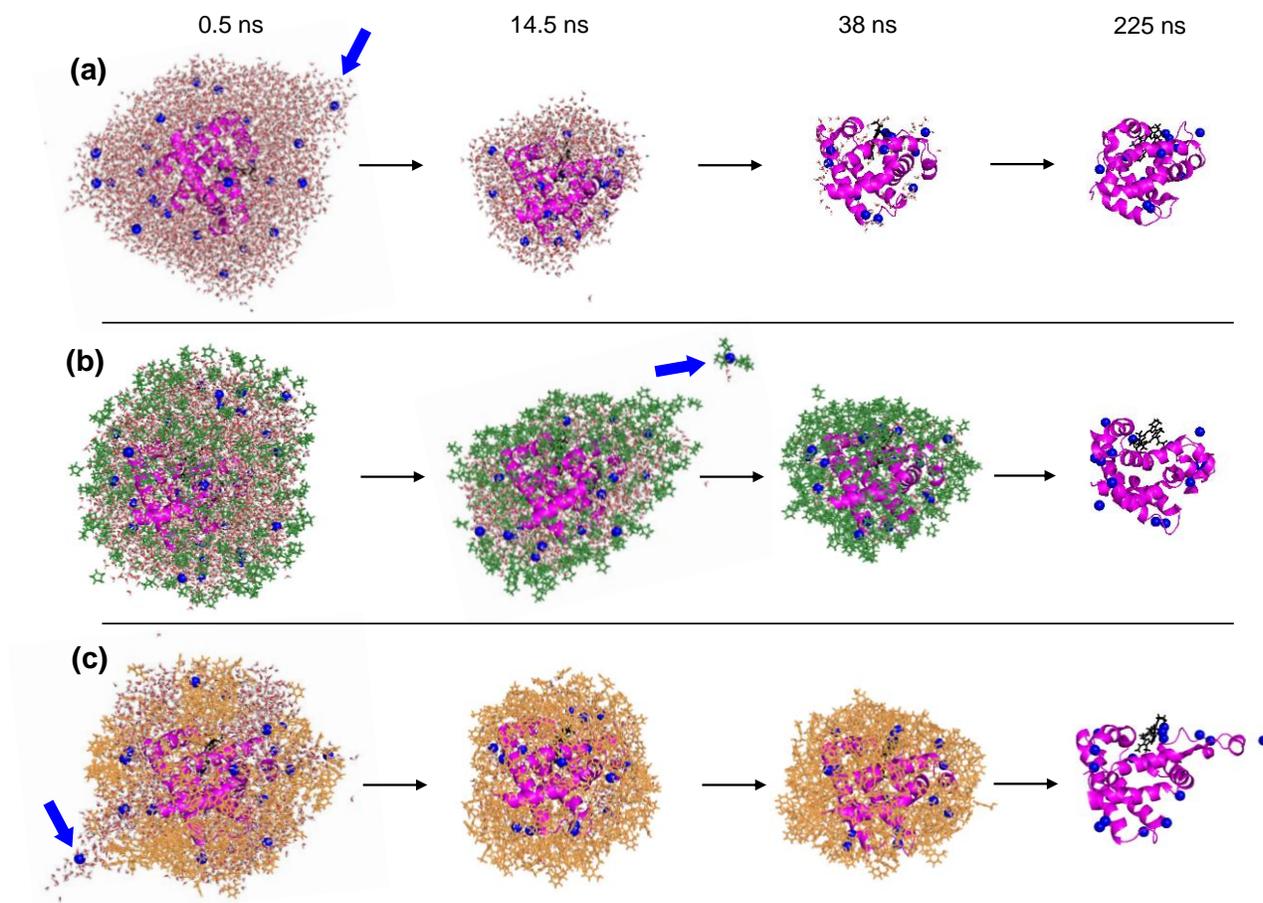
We modeled the release of proteins from nanodroplets consisting of water, water/sulfolane, and water/*m*-NBA. The extended time frame of these runs necessitated the use of slightly smaller droplets than in Figure 3.3 ( $r_0 \approx 3.5$  nm instead of 5.5 nm). Na<sup>+</sup> and SCA molecules were inserted in random locations, thereby ensuring unbiased starting configurations.

Purely aqueous droplets underwent rapid shrinkage due to water evaporation, accompanied by ejection of solvated Na<sup>+</sup>. Any remaining Na<sup>+</sup> associate with the protein (Figure 3.4a).<sup>57</sup> MD data for water/sulfolane are exemplified in Figure 3.4b. Rapid water evaporation results in sulfolane enrichment. Figure 3.4b/panel 2 reveals partial solvent segregation, where an aqueous droplet core harbors the protein, while the outer shell primarily consists of sulfolane. The sulfolane shell is dynamic enough to permit water evaporation, as well as the occasional ejection of small clusters comprising one Na<sup>+</sup> and a few water and sulfolane molecules (Figure 3.4b/panel 2). After ~40 ns most of the water has vanished, leaving behind the protein surrounded by sulfolane (Figure 3.4b/panel 3). Complete sulfolane evaporation eventually releases the protein into the gas phase (Figure 3.4b/panel 4). A very similar progression was seen for water/*m*-NBA droplets (Figure 3.4c).

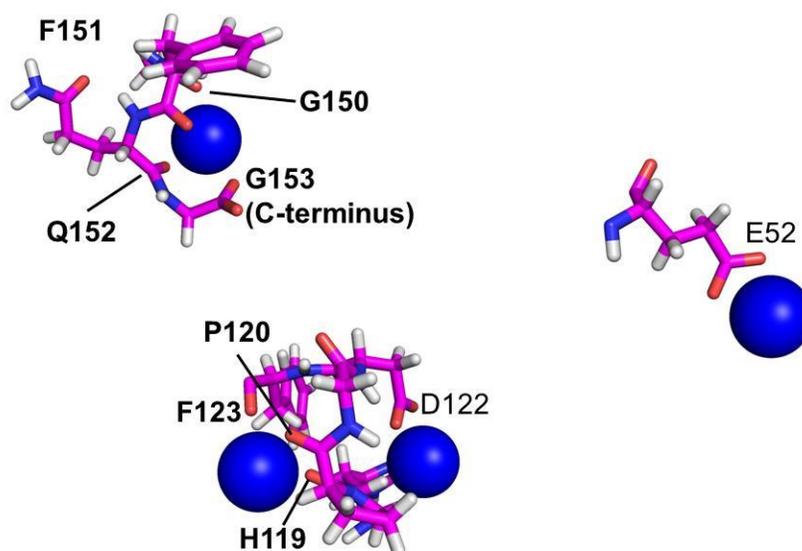
The three ESI scenarios in Figure 3.4 share several features. Droplet shrinkage is accompanied by the field emission of Na<sup>+</sup>, in line with IEM events previously described for other systems.<sup>9, 28-29, 64-65</sup> The droplets produce gaseous proteins via solvent evaporation to dryness. This morphological characteristic implies that ESI is a CRM process for the conditions studied here,<sup>30, 57</sup> regardless whether SCAs are present or not.

Protein charge states predicted by simulations with and without SCAs are summarized in Figure 3.6. Water droplets produce  $z$  values around 9+. This is identical to the result of our earlier simulations,<sup>57</sup> despite the use of larger droplets in the present work. This consistency is reassuring, as it confirms that the MD results are not affected by droplet size. Strikingly, simulations of both water/sulfolane and water/*m*-NBA droplets produce charge states around 15+, *i.e.*, significantly higher than for the purely aqueous systems. The simulated  $z$  values show remarkable agreement with the experimental average charge states for all three solvent systems (Figure 3.6).

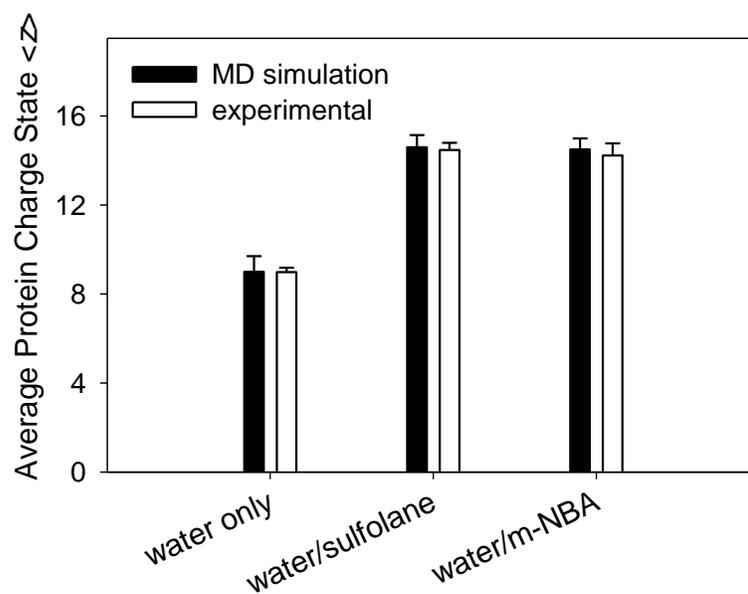
Experimental mass spectra acquired in the presence of SCAs cover a range of charge states, whereas the water/SCA simulations produce  $z$  values that are quite well defined (see Figure 3.6 caption). Charge heterogeneity in the experimental data likely reflects the fact that individual ESI droplets contain different SCA concentrations, with higher concentrations favoring formation of higher  $z$  values.<sup>42</sup> Capturing this heterogeneity in our simulations is difficult, as it would require knowledge of the exact ESI plume composition. Nonetheless, the agreement between experimental and simulated  $z$  values in Figure 3.6 implies that our MD approach successfully captures the key aspects of SCA behavior during protein ESI.



**Figure 3.4.** MD snapshots illustrating the temporal evolution of hMb-containing nanodroplets. (a) Water, (b) Water/sulfolane, and (c) Water/*m*-NBA. Time points are indicated along the top. Blue arrows indicate various stages of  $\text{Na}^+$  IEM ejection. Coloring: Protein, pink; heme, black;  $\text{Na}^+$ , blue; water oxygen, red; sulfolane, dark green; *m*-NBA, orange. All  $\text{Na}^+$  in the final frames are bound to hMb, connecting side chains are not shown to prevent clutter (see also fig. 3.5).



**Figure 3.5.** Details of hMb interactions with selected  $\text{Na}^+$  ions after release from a water/*m*-NBA droplet ( $t = 225$  ns of Figure 3.4 c). Side chain contacts are indicated using regular labels, main chain contacts are highlighted in bold.



**Figure 3.6.** ESI charge states of hMb predicted by MD simulations compared to experimental average charge states. MD results are the average of five independent runs, resulting in  $z$  values 8/9/9/9/10 (water), 14/14/15/15/15 (water/sulfolane), and 14/14/14/15/15 (water/*m*-NBA). Experimental values are averaged over all peaks in the spectra from three measurements. Error bars represent standard deviations.

### 3.3.5. Anatomy of the ESI Process

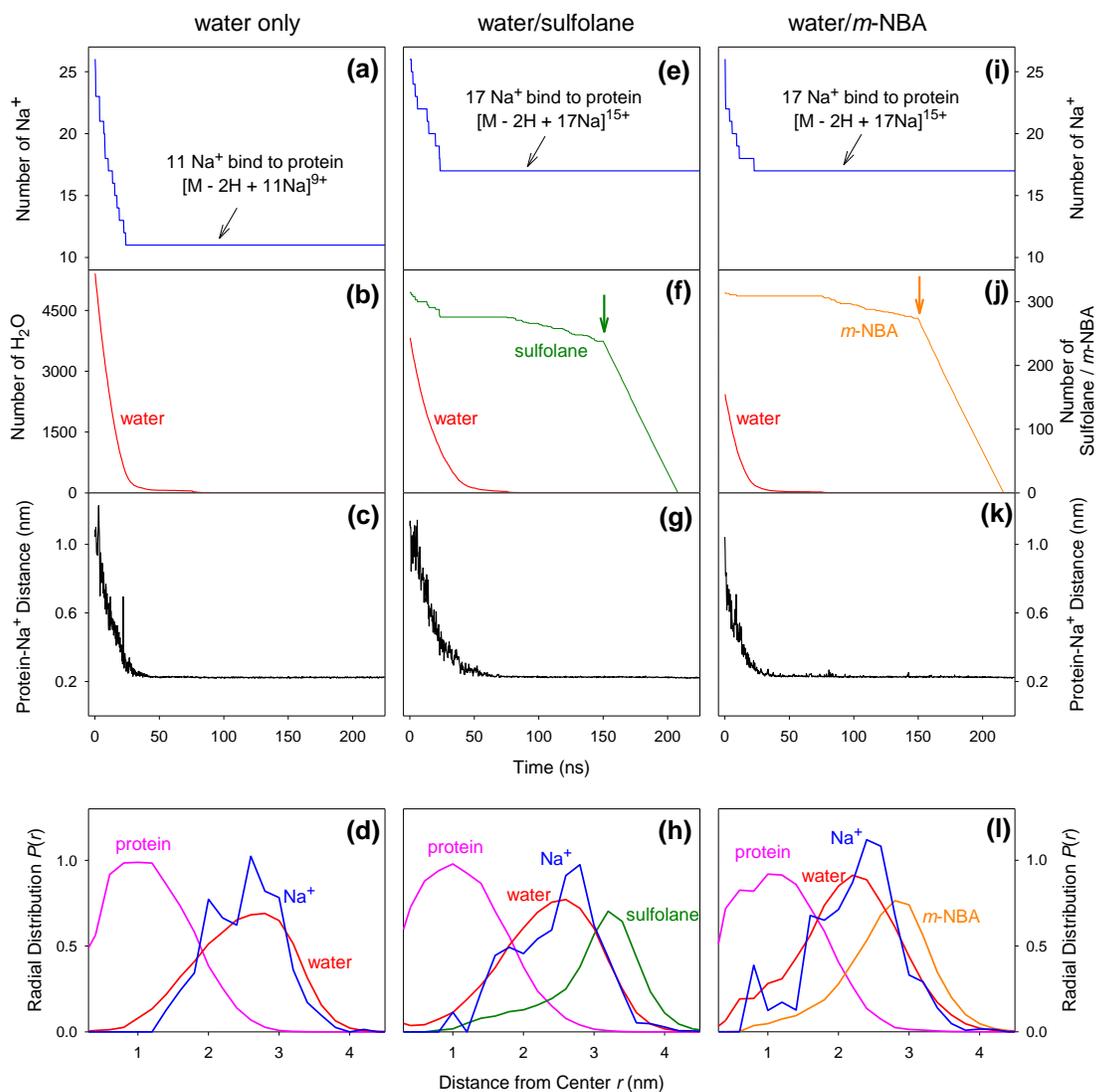
After confirming that the MD results are consistent with experiments, we now dissect the simulation trajectories to scrutinize the ESI events with and without SCAs. Most  $\text{Na}^+$  ejection events take place within the first  $\sim 25$  ns. *More*  $\text{Na}^+$  are retained in water/SCA than in purely aqueous droplets (Figure 3.7a, e, i). Any retained  $\text{Na}^+$  will eventually contribute to the protein charge. Understanding why aqueous droplets shed  $\text{Na}^+$  with higher efficiency thus holds the key to deciphering the supercharging mechanism. We will discuss this central point in more detail below.

$\text{Na}^+$  attachment provides the protein with its net charge. Early on,  $\text{Na}^+$  ions show noisy trajectories, arising from diffusive movement in the space between the protein and the droplet surface. Within  $\sim 50$  ns the profiles settle down at 0.22 nm, reflecting  $\text{Na}^+$  binding to carboxylate and carbonyl oxygens on the protein (Figure 3.7c, g, k, see Figure 3.5 for interaction details). There is a striking correlation between  $\text{Na}^+$  attachment (Figure 3.7c, g, k) and water evaporation (Figure 3.7b, f, j):  $\text{Na}^+$  ions can roam the droplet interior only as long as there is water. Once the water vanishes all remaining  $\text{Na}^+$  attach to the protein. The pure SCA environments encountered for  $t > 50$  ns *cannot* prevent  $\text{Na}^+$  attachment to the protein (see Figures 3.7g, k,  $t = 50 - 150$  ns, where all remaining  $\text{Na}^+$  are bound to hMb while most SCA molecules are still present). This behavior arises from the  $\text{Na}^+$  affinity trends discussed earlier (Figure 3.3):  $\text{Na}^+$  prefers to be solvated in water. In the absence of water,  $\text{Na}^+$  binding to the protein is preferred. Being dissolved in an SCA environment is least favored.

Radial distributions report on the internal droplet structure (Figure 3.7, bottom row,  $t = 3-5$  ns). In all three cases the protein resides close to the droplet center. It is surrounded by water containing the  $\text{Na}^+$  ions. Note that the  $\text{Na}^+$  and water distributions closely coincide. Water-SCA segregation causes sulfolane and *m*-NBA to accumulate in the outermost droplet layers (Figures 3.7h, l). This segregation is consistent with nanoscale de-mixing of other binary systems such as water/methanol.<sup>9, 90</sup> Formation of an aqueous droplet core maximizes enthalpically favorable water-water hydrogen bonding as well as water-protein contacts.<sup>91</sup>

The number of solvent molecules in the droplet is tallied in Figure 3.7b, f, j. Most of the water has evaporated after  $\sim 50$  ns, while only a handful of SCA molecules have left at this point.

This enrichment reflects the low volatility of SCAs, in line with experiments.<sup>27, 51</sup> As noted in the Materials and Methods section, starting at  $t = 150$  ns a “forced evaporation” scheme was applied for speeding up the temporal evolution of SCA-containing droplets (arrows in Figures 3.7f, j).



**Figure 3.7.** MD data, illustrating the evaporation of hMb-containing ESI droplets: (a-d) Water, (e-h) Water/sulfolane, (i-l) Water/*m*-NBA. Top row:  $\text{Na}^+$  ejection from the droplet. Second row: Solvent evaporation (vertical arrows in f, j indicate the onset of forced evaporation). Third row: Average distance of  $\text{Na}^+$  in the droplet from the closest protein heavy atom. Bottom row: Radial distributions, averaged over 3 - 5 ns and five MD runs. These  $P(r)$  data were generated by tallying heavy atoms for each component with  $4\pi r^2$  normalization. Protein  $P(r)$  data are scaled by  $\times 0.33$ .

### 3.3.6. Electrospray Ionization Supercharging via Charge Trapping

For understanding the supercharging mechanism, it is necessary to recall three basic facts. (1) Charge accumulation on the protein competes with IEM ejection of small ions (such as  $\text{Na}^+$ ) from the droplet.<sup>28-29, 57</sup> (2) Strong electrostatic interactions (e.g.,  $\text{R-COO}^- \text{Na}^+$ ) prevent the protein from ejecting charge once the solvent has left.<sup>92</sup> Thus, any charge carriers that did not leave the droplet during solvent evaporation will contribute to the  $z$  value of the protein.<sup>28-29, 57</sup> (3) Charge carriers can undergo IEM ejection only after coming close to the droplet surface.<sup>64</sup> Hence, any factors that restrict charge carrier access to the droplet surface will reduce the IEM ejection rate, thereby boosting the protein charge.

In purely aqueous systems, favorable solvation allows charge carriers to roam the entire liquid space within the droplet, including positions close to the droplet surface (Figures 3.3a, 3.7d). The IEM ejection of charge carriers from aqueous droplets is therefore a facile process, provided that the system net charge is sufficiently high.<sup>64-65</sup> This facile charge loss causes protein ions generated from water droplets via the CRM to have low  $z$  values.<sup>8-9, 31</sup>

Water/SCA droplets undergo solvent segregation. SCA molecules are enriched in the outermost layers, while protein and water reside close to the center. Charge carrier partitioning proceeds according to solvation preferences, *i.e.*, small ions reside in the aqueous core (Figures 3.7h, i). The resulting charge carrier depletion at the droplet surface *disfavors* IEM ejection,<sup>64</sup> a view that is supported by experiments on *m*-NBA droplets.<sup>6</sup> As the final water molecules escape (leaving the SCA behind), all remaining charge carriers bind to the protein. Being a poor solvent for small ions, the SCA is incapable of dissolving charge carriers away from the protein (Figure 3.3b, c), such that IEM ejection is shut down. Charge carriers remain associated with the protein, causing it to emerge as highly charged CRM product once the SCA has evaporated.

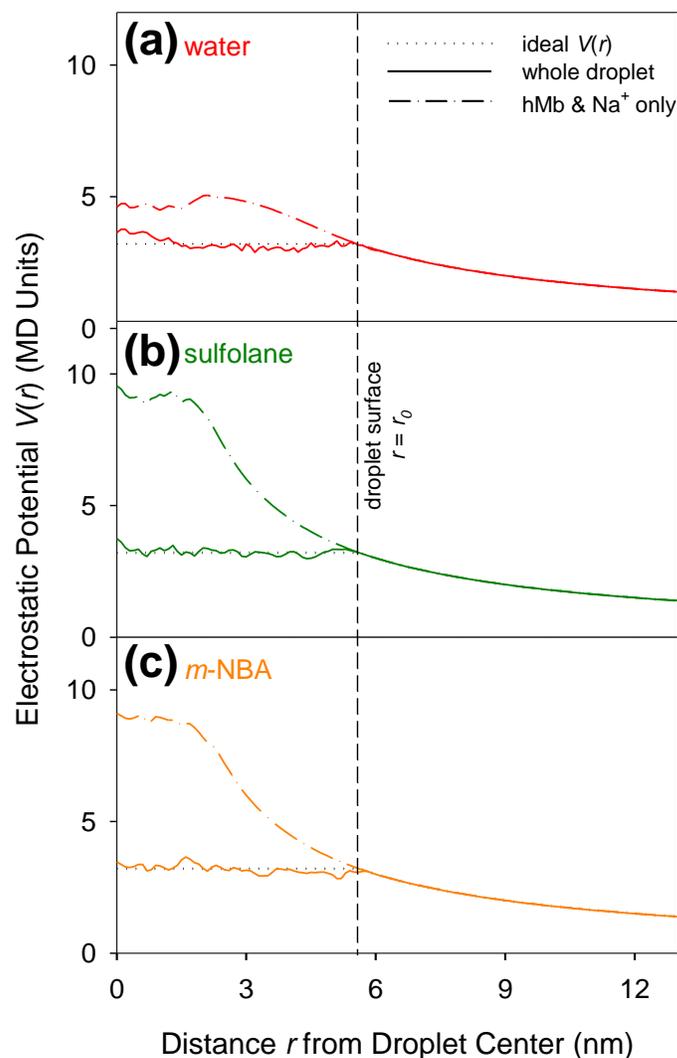
In summary, protein supercharging is based on charge trapping within ESI nanodroplets. This trapping results from two effects that are related to one another. (i) Early during nanodroplet evaporation, SCA accumulation in the outermost layers restricts charge carrier access to the droplet surface. (ii) After all the water has evaporated, charge carriers are bound to the protein deep within the droplet. Unfavorable solvation characteristics prevent these charge carriers from venturing into the surrounding SCA. Both (i) and (ii) interfere with the capability of the

nanodroplets to shed charge via the IEM. As a result, more charge is imparted to the protein than in the absence of SCAs.

### 3.3.7. Surface Charge of ESI Droplets

The trapping mechanism outlined above involves the binding of numerous charge carriers to the protein, surrounded by an SCA shell, after all water has left. One might argue that such a charge carrier accumulation should be electrostatically disfavored. This apparent conundrum is resolved when solvation effects are taken into account.

We previously noted that water droplets containing excess  $\text{Na}^+$  undergo dipole ordering.<sup>9</sup> As a result, the droplets carry their entire net charge on the surface (as widely assumed in the ESI literature<sup>8</sup>) while excess charge carriers (e.g.,  $\text{Na}^+$ ) are located in the interior where solvation is more favorable.<sup>88</sup> This surface charge is a consequence of Gauss' Law,<sup>93</sup> which states that a conducting sphere with radius  $r_0$  will carry its entire charge at  $r = r_0$ . In MD units<sup>68</sup> the electrostatic potential under such "ideal" condition is  $V(r) = z/r$  for  $r > r_0$ , and  $V(r) = z/r_0 = \text{const}$  for  $r \leq r_0$ .<sup>93</sup> The  $V(r)$  profile predicted by Gauss' Law is shown in Figure 3.8 (dotted lines in each panel).



**Figure 3.8.** Electrostatic potential  $V(r)$  of nanodroplets containing hMb and 20  $\text{Na}^+$ , generated from the trajectories of Figure 3.3. (a) Water, (b) Pure sulfolane, (c) Pure *m*-NBA. Each panel contains three data sets. 1. Gauss' Law for an ideal conductor (black dotted lines). 2. Whole droplet with hMb, 20  $\text{Na}^+$ , and solvent (solid lines). 3. hMb / 20  $\text{Na}^+$  without solvent contributions (dash-dotted lines). Each curve represents the average of 140 scans between 2 ns and 3 ns.

It has never been confirmed that the surface charge projection seen for water/ $\text{Na}^+$  droplets<sup>9</sup> also applies in the presence of proteins and/or other solvents. We examined this aspect by mapping  $V(r)$  for the three droplets depicted in Figure 3.3. a virtual test particle was scanned from the droplet midpoint in radial direction. For each position  $r$  the electrostatic potential was

calculated as  $V(r) = \sum(q_i/d_i)$  where the sum includes all atoms  $i$ ,  $q_i$  is the atom charge  $i$  defined in the force field,<sup>67</sup> and  $d_i$  is the distance between the test particle and atom  $i$ .

Droplet profiles were first generated by only scanning protein and  $\text{Na}^+$ , omitting the solvent.  $V(r)$  profiles generated under these conditions show dramatic deviations from Gauss' Law (dash-dotted lines, Figure 3.8a-c). The mismatch is most pronounced for sulfolane and *m*-NBA, reflecting  $\text{Na}^+$  accumulation close to the center (cf. Figure 3.3). In contrast,  $V(r)$  scans of the whole droplets (protein,  $\text{Na}^+$ , and solvent) agree closely with the ideal profiles (solid lines, Figure 3.8a-c).

The data of Figure 3.8 confirm the expectation<sup>9</sup> that ESI droplets containing a dipolar solvent will project their net charge to the surface. The droplet interior is free of static fields because  $V(r) \approx \text{const}$  for  $r \leq r_0$ . As a result, large-scale electrostatic repulsion among charges within the droplet is absent. This effect greatly reduces the energetic penalty associated with charge carrier accumulation on the protein close to the center. Also, bound charge carriers will not trigger electrostatic protein unfolding while the polypeptide chain is surrounded by a solvent shell, because the charge is projected away from the protein. Instead, the net charge destabilizes the whole droplet by forcing surface solvent molecules into orientations that satisfy Gauss' Law<sup>93</sup> at the expense of intermolecular contacts.<sup>9</sup> The net charge of SCA droplets under these conditions may exceed  $z_R$ , in line with experiments on various non-aqueous systems.<sup>6</sup>

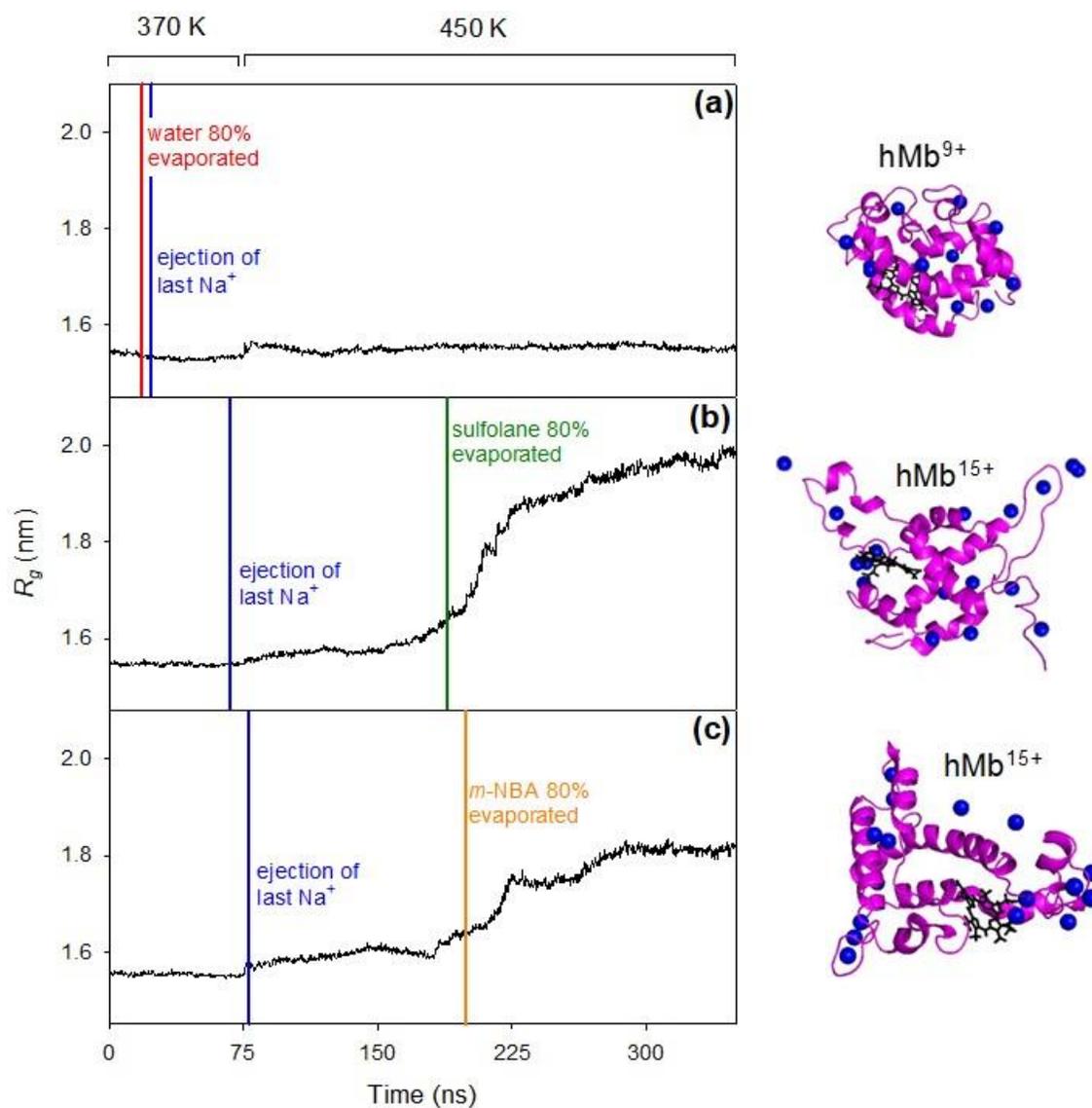
### 3.3.8. Supercharging and Protein Unfolding

Highly charged hMb ions generated in the presence of SCAs are unfolded (Figure 3.1f-h). Similar observations prompted Williams et al. to propose that chemical<sup>42</sup> or thermal<sup>55</sup> unfolding within ESI droplets is the primary origin of supercharging. Bulk solution studies confirmed that high SCA percentages can reduce protein stability.<sup>34, 42</sup> However, it is unclear if this effect can trigger unfolding on the very short time scale of the final ESI steps,<sup>34</sup> keeping in mind that protein/SCA contact is limited by segregation (Figure 3.7h, l). These questions necessitate a closer look at the relationship between supercharging and unfolding.

Figure 3.9 tracks the hMb radius of gyration ( $R_g$ ) over 350 ns. For purely aqueous ESI the protein retains a native-like compactness with  $R_g \approx 1.5$  nm throughout the entire time window

(Figure 3.9a). For water/sulfolane and water/*m*-NBA compact conformations are retained only for *ca.* 180 ns (Figure 3.9b, c). The final protein  $z$  values are determined after the last  $\text{Na}^+$  ejection event (blue vertical lines at  $\sim 70$  ns, Figure 3.9). The  $R_g$  profiles therefore imply that hMb supercharging is complete prior to unfolding. Significant conformational changes start to occur later, during evaporation of the final solvent molecules (Figure 3.9b, c).

In summary, our data imply that unfolding within ESI droplets is not the origin of supercharging by sulfolane or *m*-NBA. Electrostatically driven unfolding<sup>15-16, 20</sup> takes place *after* protein charging is complete. Major structural changes start during the final stages of desolvation, and they likely continue for the bare proteins beyond the 350 ns window of Figure 3.9. Unfolding under these conditions is facilitated by electrostatic repulsion once the stabilizing effects caused by solvent-mediated charge projection have disappeared (cf. Figure 3.8).



**Figure 3.9.** Radius of gyration ( $R_g$ ) as a function of time for ESI droplets consisting of water (a), Water/sulfolane (b) Water/*m*-NBA (c) Averaged over five MD runs. Blue vertical lines denote ejection of the last  $\text{Na}^+$  from the droplet. Also indicated are the time points where 80% of water (a), sulfolane (b), and *m*-NBA (c) have evaporated. Representative structures populated at the end of the simulation window ( $t = 350$  ns) are shown along the right.

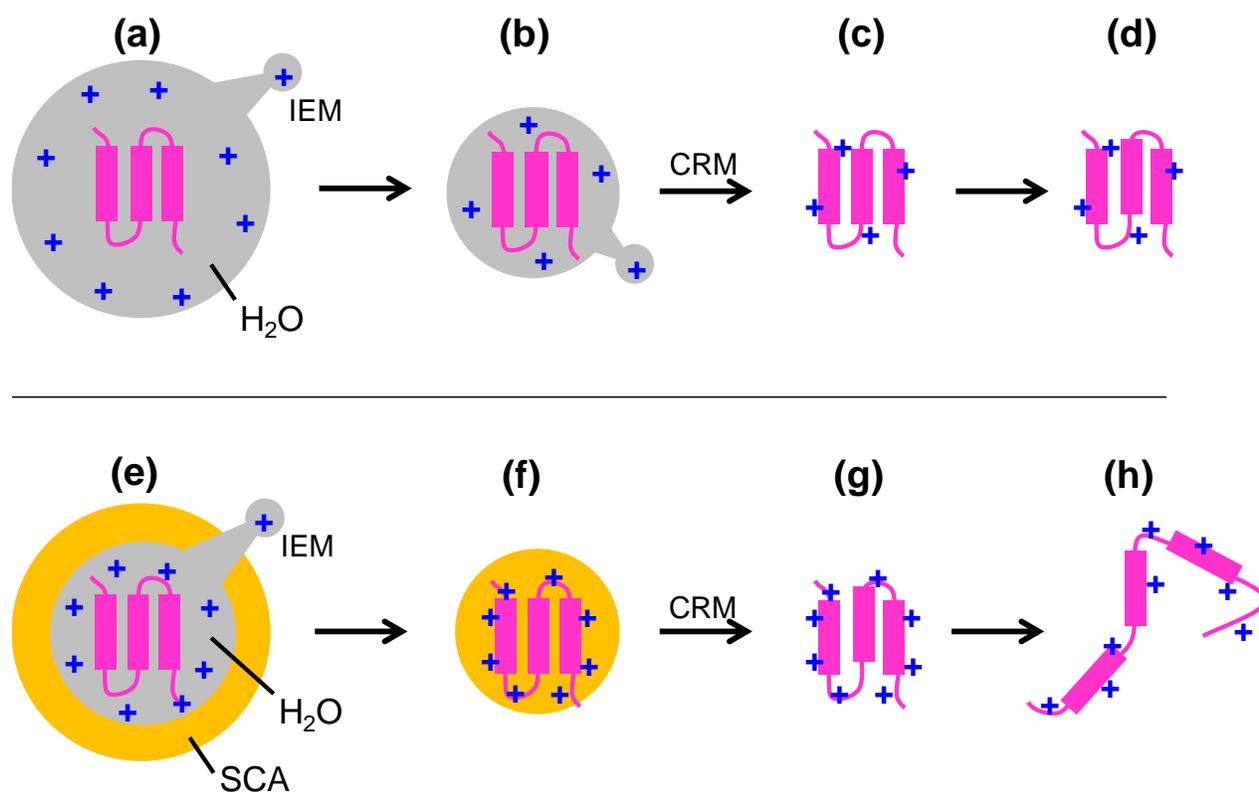
### 3.4. Conclusions

The ESI charge states predicted by our MD simulations are in remarkable agreement with experimental data (Figure 3.6). This high level of consistency suggests that the MD trajectories describe the physics of the ESI process quite well, despite the limitations of classical force fields. We focused on the fate of native hMb in charged nanodroplets consisting of water, water/sulfolane, and water/*m*-NBA. In all cases the protein is released via solvent evaporation to dryness, as envisioned by the CRM.<sup>30</sup> ESI charge states generated in water are close to the Rayleigh charge of protein-sized aqueous droplets.<sup>31, 57</sup> In contrast, the elevated  $z$  values formed in the presence of SCAs are inconsistent with the predicted  $z_R(\gamma)$  trend. This behavior implies that supercharging is not a simple surface tension effect.<sup>43, 54</sup> Our results also indicate that neither unfolding within ESI droplets,<sup>42, 55</sup> nor the Brønsted acid/base properties of SCAs<sup>27</sup> represent key components of the supercharging mechanism.

Figure 3.10 summarizes the ESI process for water and water/SCA droplets. Without SCAs, highly charged aqueous droplets generated by native ESI undergo rapid evaporation. The aqueous environment favors relatively unrestricted movement of charge carriers, allowing them to approach the droplet surface where IEM ejection takes place with high efficiency (Figure 3.10 a, b). The few remaining charge carriers bind to the protein as the final water layers evaporate (Figure 3.10c). The low  $z$  values render proteins electrosprayed in water quite resistant to gas phase unfolding, favoring the retention of solution-like conformations (Figure 3.10d).<sup>2, 5, 15, 19</sup>

Supercharging is caused by a charge trapping mechanism. Mixed water/SCA nanodroplets undergo solvent segregation, resulting in an aqueous core that contains the protein and an outer shell of SCA molecules. Charge carriers partition into the aqueous core where solvation is more favorable than in the outer SCA layers. As a consequence, the IEM ejection of charge carriers from the droplet surface proceeds at a lower rate (Figure 3.10e). Poor solvation afforded by the SCA causes binding of all remaining charge carriers to the protein once the water has evaporated. IEM ejection is no longer feasible after this point because charge carrier affinity to the protein exceeds their affinity for the SCA environment; the charge carriers are trapped (Figure 3.10f). SCA evaporation releases the protein into the gas phase without any additional

charge loss (Figure 3.10g). As a result of their high  $z$  values, supercharged protein ions are prone to electrostatically driven unfolding in the gas phase (Figure 3.10h).



**Figure 3.10.** ESI in water without (a-d) and with supercharging (e-h). (a, b) Water droplet undergoing solvent evaporation. Charge carriers ( $\text{Na}^+$  or  $\text{H}^+$ , blue) can move to the surface and undergo IEM ejection. (c) Native-like nascent gas phase protein. (d) Low charge favors retention of compact structures. (e) Water/SCA droplet after solvent segregation.  $\text{Na}^+$  or  $\text{H}^+$  accumulate in the aqueous core, reducing IEM ejection. (f)  $\text{Na}^+$  or  $\text{H}^+$  are trapped on the protein after water evaporation. (g) Supercharged nascent gas phase protein. (h) High charge states favor unfolding.

The supercharging mechanism outlined in Figure 3.10e-h suggest that an “ideal” SCA will be dipolar and moderately miscible with water, while its evaporation rate will be much lower than that of water. In addition, interactions of the SCA with the protein surface and with excess charge carriers have to be less favorable than in the case of water. These features are consistent with typical SCA characteristics previously noted in the literature.<sup>43, 51</sup>

The supercharging mechanism outlined in Figure 3.10 applies to proteins that are electrosprayed from non-denaturing bulk solution in positive ions mode. Supercharging also takes place for proteins that enter the ESI process as unfolded conformers, e.g. in the presence of acids.<sup>50-53</sup> Meaningful MD simulations under those conditions will require the use of mobile proton methods, as the process likely involves protein release via the CEM<sup>9</sup>. It will also be of interest to examine the effects of SCAs in negative ion mode,<sup>47</sup> and for analytes such as nucleic acids.<sup>94-95</sup> Work in this direction is currently ongoing in our laboratory, and the results of those endeavors will be reported elsewhere.

### 3.5. References

1. Fenn, J. B., *Angew. Chem. Int. Ed.* **2003**, *42*, 3871-3894.
2. Mehmood, S.; Allison, T. M.; Robinson, C. V., *Annu. Rev. Phys. Chem.* **2015**, *66*, 453-474.
3. Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R., *Protein Sci.* **2013**, *22*, 530-544.
4. Ashcroft, A. E., *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1087-1096.
5. Uetrecht, C.; Rose, R. J.; van Duijn, E.; Lorenzen, K.; Heck, A. J. R., *Chem. Soc. Rev.* **2010**, *39*, 1633-1655.
6. Grimm, R. L.; Beauchamp, J. L., *J. Phys. Chem. A* **2010**, *114*, 1411-1419.
7. Nemes, P.; Marginean, I.; Vertes, A., *Anal. Chem.* **2007**, *79*, 3105-3116.
8. Kebarle, P.; Verkerk, U. H., *Mass Spectrom. Rev.* **2009**, *28*, 898-917.
9. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., *Anal. Chem.* **2013**, *85*, 2-9.
10. Liu, L.; Michelsen, K.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S., *J. Am. Chem. Soc.* **2012**, *134*, 3054-3060.
11. Han, X.; Jin, M.; Breuker, K.; McLafferty, F. W., *Science* **2006**, *314*, 109-112.
12. Ly, T.; Julian, R. R., *J. Am. Chem. Soc.* **2010**, *132*, 8602-8609.
13. Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S., *J. Am. Chem. Soc.* **2013**, *135*, 12646-12651.
14. Zubarev, A. R.; Makarov, A., *Anal. Chem.* **2013**, *85*, 5288-5296.
15. Wyttenbach, T.; Pierson, N. A.; Clemmer, D. E.; Bowers, M. T., *Annu. Rev. Phys. Chem.* **2014**, *65*, 175-196.
16. Shelimov, K. B.; Jarrold, M. F., *J. Am. Chem. Soc.* **1997**, *119*, 2987-2994.
17. Chen, S. H.; Chen, L. X.; Russell, D. H., *J. Am. Chem. Soc.* **2014**, *136*, 9499-9508.
18. Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V., *Science* **2005**, *310*, 1658-1661.
19. Benesch, J. L. P.; Ruotolo, B. T., *Curr. Op. Struct. Biol.* **2011**, *21*, 641-649.
20. Borysik, A. J.; Kovacs, D.; Guharoy, M.; Tornpa, P., *J. Am. Chem. Soc.* **2015**, *137*, 13807-13817.
21. Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L., *Anal. Chem.* **2011**, *83*, 6868-6874.
22. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S., *Mass Spectrom. Rev.* **1998**, *17*, 1-35.

23. Chowdhury, S. K.; Katta, V.; Chait, B. T., *J. Am. Chem. Soc.* **1990**, *112*, 9012-9013.
24. Dobo, A.; Kaltashov, I. A., *Anal. Chem.* **2001**, *73*, 4763-4773.
25. Grandori, R., *Protein Sci.* **2002**, *11*, 453-458.
26. Li, J.; Santambrogio, C.; Brocca, S.; Rossetti, G.; Carloni, P.; Grandori, R., *Mass Spectrom. Rev.* **2016**, *35*, 111-122.
27. Ogorzalek Loo, R. R.; Lakshmanan, R.; Loo, J. A., *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1675-1693.
28. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L., *Anal. Chem.* **2009**, *81*, 369-377.
29. Allen, S. J.; Schwartz, A. M.; Bush, M. F., *Anal. Chem.* **2013**, *85*, 12055-12061.
30. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B., *J. Chem. Phys.* **1968**, *49*, 2240-2249.
31. de la Mora, F. J., *Anal. Chim. Acta* **2000**, *406*, 93-104.
32. Consta, S.; Oh, M. I.; Soltani, S., *Int. J. Mass Spectrom.* **2015**, *377*, 557-567.
33. Lin, H.; Kitova, E. N.; Johnson, M. A.; Eugenio, L.; Ng, K. K. S.; Klassen, J. S., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 2122-31.
34. Hamdy, O. M.; Julian, R. R., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1-6.
35. Mirza, U. A.; Chait, B. T., *Int. J. Mass. Spectrom. Ion Proc.* **1997**, *162*, 173-181.
36. Kharlamova, A.; Prentice, B. M.; Huang, T.-Y.; McLuckey, S. A., *Anal. Chem.* **2010**, *82*, 7422-7429.
37. McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L., *J. Am. Chem. Soc.* **1990**, *112*, 5668-5670.
38. Laszlo, K. J.; Bush, M. F., *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 2152-2161.
39. Abzalimov, R. R.; Kaltashov, I. A., *Anal. Chem.* **2010**, *82*, 7523-7526.
40. Lermyte, F.; Williams, J. P.; Brown, J. M.; Martin, E. M.; Sobott, F., *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1068-1076.
41. Iavarone, A. T.; Williams, E. R., *J. Am. Chem. Soc.* **2003**, *125*, 2319-2327.
42. Sterling, H. J.; Daly, M. P.; Feld, G. K.; Thoren, K. L.; Kintzer, A. F.; Krantz, B. A.; Williams, E. R., *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1762-1774.
43. Lomeli, S. H.; Peng, I. X.; Yin, S.; Ogorzalek Loo, R. R.; Loo, J. A., *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 127-131.

44. Hogan, C. J.; Loo, R. R. O.; Loo, J. A.; de la Moraa, J. F., *Phys. Chem. Chem. Phys.* **2010**, *12*, 13476-13483.
45. Hall, Z.; Politis, A.; Bush, M. F.; Smith, L. J.; Robinson, C. V., *J. Am. Chem. Soc.* **2012**, *134*, 3429-3438.
46. Fisher, C. M.; Kharlamova, A.; McLuckey, S. A., *Anal. Chem.* **2014**, *86*, 4581-4588.
47. Chingin, K.; Xu, N.; Chen, H., *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 928-934.
48. Zhou, M.; Dagan, S.; Wysocki, V. H., *Analyst* **2013**, *138*, 1353-1362.
49. Ferguson, C. N.; Benchaar, S. A.; Miao, Z. X.; Loo, J. A.; Chen, H., *Anal. Chem.* **2011**, *83*, 6468-6473.
50. Samalikova, M.; Grandori, R., *J. Mass Spectrom.* **2005**, *40*, 503-510.
51. Teo, C. A.; Donald, W. A., *Anal. Chem.* **2014**, *86*, 4455-4462.
52. Douglass, K. A.; Venter, A. R., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 489-497.
53. Zhang, J.; Loo, R. R. O.; Loo, J. A., *Int. J. Mass Spectrom.* **2015**, *377*, 546-556.
54. Samalikova, M.; Grandori, R., *J. Am. Chem. Soc.* **2003**, *125*, 13352-13353.
55. Sterling, H. J.; Williams, E. R., *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1933-1943.
56. Hall, Z.; Robinson, C. V., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1161-1168.
57. McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L., *J. Am. Chem. Soc.* **2015**, *137*, 12667-12676.
58. Higashi, H.; Tokumi, T.; Hogan, C. J.; Suda, H.; Seto, T.; Otani, Y., *Phys. Chem. Chem. Phys.* **2015**, *17*, 15746-15755.
59. Steinberg, M. Z.; Breuker, K.; Elber, R.; Gerber, R. B., *Phys. Chem. Chem. Phys.* **2007**, *9*, 4690-4697.
60. Iyengar, S. S.; Day, T. J. F.; Voth, G. A., *Int. J. Mass Spectrom.* **2005**, *241*, 197-204.
61. Daub, C. D.; Cann, N. M., *Anal. Chem.* **2011**, *83*, 8372-8376.
62. Patriksson, A.; Marklund, E.; van der Spoel, D., *Biochemistry* **2007**, *46*, 933-945.
63. Caleman, C.; van der Spoel, D., *Phys. Chem. Chem. Phys.* **2007**, *9*, 5105-5111.
64. Iribarne, J. V.; Thomson, B. A., *J. Chem. Phys.* **1976**, *64*, 2287-2294.
65. Labowsky, M.; Fenn, J. B.; Fernandez de la Mora, J., *Anal. Chim. Acta* **2000**, *406*, 105-118.
66. Sun, Y.; Vahidi, S.; Sowole, M. A.; Konermann, L., *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 31-40.

67. Huang, J.; MacKerell, A. D., *J. Comput. Chem.* **2013**, *34*, 2135-2145.
68. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E., *Bioinformatics* **2013**, *29*, 845-854.
69. Piana, S.; Lindorff-Larsen, K.; Shaw, D. E., *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 5915-5920.
70. Hess, B.; Henk, B.; Berendsen, H. J. C.; Fraaije, J. G. E. M., *J. Comput. Chem.* **1997**, *18*, 1463-1472.
71. Miyamoto, S.; Kollman, P. A., *J. Comput. Chem.* **1992**, *13*, 952-962.
72. Maurus, R.; Overall, C. M.; Bogumil, R.; Luo, Y.; Mauk, A. G.; Smith, M.; Brayer, G. D., *Biochim. Biophys. Acta* **1997**, *1341*, 1-13.
73. Abascal, J. L. F.; Vega, C., *J. Chem. Phys.* **2005**, *123*, 234505.
74. Vega, C.; de Miguel, E., *J. Chem. Phys.* **2007**, *126*, 154707.
75. van der Spoel, D.; van Maaren, P. J.; Caleman, C., *Bioinformatics* **2012**, *28*, 752-753.
76. Fischer, N. M.; van Maaren, P. J.; Ditz, J. C.; Yildirim, A.; van der Spoel, D., *J. Chem. Theor. Comp.* **2015**, *11*, 2938-2944.
77. Sterling, T.; Irwin, J. J., *J. Chem. Inf. Model.* **2015**, *55*, 2324-2337.
78. Hoover, W. G., *Phys. Rev. A* **1985**, *31*, 1695-1697.
79. Covey, T. R.; Thomson, B. A.; Schneider, B. B., *Mass Spectrom. Rev.* **2009**, *28*, 870-897.
80. Gabelica, V.; De Pauw, E., *Mass Spectrom. Rev.* **2005**, *24*, 566-587.
81. Merenbloom, S. I.; Flick, T. G.; Williams, E. R., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 553-562.
82. Nam, K.; Pu, J. Z.; Karplus, M., *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 17851-17856.
83. Isralewitz, B.; Gao, M.; Schulten, K., *Curr. Opin. Struct. Biol.* **2001**, *11*, 224-230.
84. Hopper, J. T. S.; Oldham, N. J., *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1851-1858.
85. Nakai, H.; Sakti, A. W.; Nishimura, Y., *J. Phys. Chem. B* **2016**, *120*, 217-221.
86. Goyal, P.; Qian, H. J.; Irle, S.; Lu, X. Y.; Roston, D.; Mori, T.; Elstner, M.; Cui, Q., *J. Phys. Chem. B* **2014**, *118*, 11007-11027.
87. Marx, D.; Chandra, A.; Tuckerman, M. E., *Chem. Rev.* **2010**, *110*, 2174-2216.
88. Mahler, J.; Persson, I., *Inorg. Chem.* **2012**, *51*, 425-438.
89. Pinho, S. P.; Macedo, E. A., *J. Chem. Eng. Data* **2005**, *50*, 29-32.

90. Dixit, S.; Crain, J.; Poon, W. C. K.; Finney, J. L.; Soper, A. K., *Nature* **2002**, *416*, 829-832.
91. Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D., *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 5129-5134.
92. Yin, S.; Xie, Y.; Loo, J. A., *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1199-1208.
93. Halliday, D.; Resnick, R.; Krane, K. S., *Physics*. 4 ed.; Wiley: New York, **1992**.
94. Brahim, B.; Alves, S.; Cole, R. B.; Tabet, J.-C., *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1988-1996.
95. Xu, N.; Chingin, K.; Chen, H. W., *J. Mass Spectrom.* **2014**, *49*, 103-107.

## Chapter 4. Crown Ethers Modulate the Location of Charge Carriers in Electrospray Droplets: Implications for the Mechanism of Protein Charging and Supercharging

### 4.1. Introduction

Electrospray ionization (ESI)<sup>1</sup> transforms solution phase proteins into multiply charged gaseous ions for analysis by mass spectrometry (MS). Of particular interest are “native” ESI-MS experiments<sup>2-4</sup> that aim to preserve solution structures and interactions in the gas phase. These studies employ non-denaturing aqueous solutions and gentle ion sampling conditions. Native ESI-MS reports on protein binding stoichiometries.<sup>2-5</sup> Complementary information is obtainable from dissociation experiments<sup>2, 6-7</sup> and ion mobility spectrometry.<sup>8-10</sup>

In positive ESI the protein solution is dispersed into droplets that carry excess  $H^+$ ,  $NH_4^+$ , or  $Na^+$ .<sup>11-12</sup> Evaporation and fission events close to the Rayleigh limit produce progressively smaller droplets.<sup>11, 13</sup> The mechanisms of analyte ion release from these nanodroplets were shrouded in controversy for many years.<sup>1, 11-12, 14-17</sup> Recent molecular dynamics (MD) simulations helped address some of the questions in this area.<sup>18-24</sup> For example, it is now widely accepted that globular proteins are released via droplet evaporation to dryness during native ESI,<sup>20-24</sup> as envisioned by the charged residue model (CRM).<sup>11, 25</sup> Charge carriers bind to the protein during the final stages of evaporation, generating ions such as  $[M + zH]^{z+}$  or  $[M + zNa]^{z+}$ .<sup>11</sup>

The dissociation behavior<sup>26-32</sup> and conformations of gaseous proteins<sup>33-35</sup> are governed by their charge state  $z$ . As a result, there is considerable interest in ways to manipulate these charge states.<sup>33-34, 36-39</sup> Native ESI generates low  $z$  values close to the Rayleigh charge of protein-sized water droplets,<sup>11, 25</sup> in accordance with the CRM.<sup>20-24</sup> Much higher charge states are generated from proteins that are unfolded in bulk solution.<sup>40-41</sup> According to the chain ejection model (CEM) these highly charged ions form during protein expulsion from the droplet surface.<sup>42-44</sup>

A common strategy for modulating protein charge states is the use of supercharging agents (SCAs).<sup>45-46</sup> SCAs are added to the sample at low concentrations that do not significantly affect the protein structure in bulk solution.<sup>44, 47</sup> Yet, SCAs significantly enhance charging during ESI. Typical SCAs (such as sulfolane, C<sub>4</sub>H<sub>8</sub>SO<sub>2</sub>) possess a nonpolar hydrocarbon moiety and one or more polar groups.<sup>32, 45, 47</sup> Their low volatility makes them evaporate more slowly than water, such that late ESI nanodroplets are SCA-enriched.<sup>32, 45, 47-48</sup> Supercharging takes place for native<sup>30, 38, 45, 47, 49-53</sup> and for denaturing solutions.<sup>32, 44, 54</sup> Here, we focus on the role of SCAs in native ESI, i.e., the conversion of folded solution phase proteins to highly charged gaseous ions.<sup>30, 38, 45, 47, 49-52, 55</sup>

The mechanism of native ESI supercharging remains controversial.<sup>45-46, 48, 50, 54, 56</sup> According to one proposal supercharging is caused by thermal or chemical unfolding in SCA-enriched droplets.<sup>47</sup> Within this model, SCAs cause proteins to switch from the CRM to the CEM.<sup>44</sup> Although this “unfolding model” offers an intuitive explanation for supercharging, its validity is under dispute.<sup>30, 45, 55, 57-59</sup> The elevated CCSs observed for some supercharged proteins<sup>44, 47, 58</sup> do not prove that unfolding takes place within the droplet; alternatively, unfolding could be caused by Coulombic repulsion *after* release into the gas phase.<sup>58, 60</sup> A number of supercharged proteins retain a native-like compactness,<sup>30, 60</sup> making it unlikely that unfolding constitutes the root cause of supercharging. Also, weakly bound complexes can be supercharged without undergoing dissociation,<sup>30, 45, 55</sup> prompting, Robinson et al.<sup>59</sup> to conclude that “*supercharging does not appear to perturb the structure in that unfolding is not detected*”.

Our group recently addressed this issue by applying MD simulations in which Na<sup>+</sup> served as excess charge carrier for probing the supercharging mechanism, focusing on the SCAs sulfolane and *m*-nitrobenzyl alcohol (*m*-NBA).<sup>58</sup> The MD-generated [M + *z*Na]<sup>*z*+</sup> ions closely matched the *z* values of experimentally observed sodiated and protonated protein ions, both with and without SCAs. In the simulations proteins were released via droplet evaporation to dryness. Droplet shrinkage was accompanied by charge carrier ejection. The remaining charge carriers underwent binding to the protein during the final stage of evaporation. The simulations indicated (i) SCA enrichment at the droplet surface, followed by (ii) formation of an SCA layer around the protein after complete water evaporation. Both factors inhibited charge carrier ejection from the droplet because SCAs are ionophobic<sup>58</sup> (e.g., the NaCl solubility in sulfolane is four orders of

magnitude lower than in water<sup>61</sup>). We thus proposed<sup>58</sup> that supercharging is caused by charge trapping, not by unfolding. This “charge trapping model”<sup>58</sup> and the aforementioned unfolding model<sup>44, 47</sup> represent two very different mechanistic views of protein supercharging in native ESI.

The current work examines the supercharging mechanism from a new perspective. Crown ethers can bind small cations, thereby enhancing the solubility of these charge carriers in nonaqueous solvents.<sup>62-63</sup> We will test the following hypothesis: The charge trapping model<sup>58</sup> envisions that supercharging is caused by the low solubility of charge carriers in the SCA-enriched droplet layers. Under such conditions, the capability of crown ethers to act as phase transfer catalysts<sup>62-63</sup> should facilitate the shuttling of charge carriers to the droplet surface, thereby favoring charge ejection. We predict that these conditions will lower the extent of protein charging in the presence of SCAs. 18-crown-6 (18C6) is of particular interest due to its ability to accommodate ESI-relevant species ( $\text{Na}^+$ ,  $\text{NH}_4^+$  and  $\text{H}_3\text{O}^+$ ) in solution and in the gas phase.<sup>62, 64-67</sup> Previous studies explored 18C6 binding to  $\text{Lys}^+$  and  $\text{N}^+$ -termini of gaseous peptides or proteins,<sup>68-72</sup> but the consequences of crown ethers for the ESI process remain largely unexplored.<sup>39</sup> The MD simulations of this work, as well as experiments on proteins and dendrimers, support the proposed hypothesis. We report for the first time that 18C6 acts as a powerful supercharging antidote. These findings support the view that native ESI supercharging is caused by charge trapping.

## 4.2. Materials and Methods

### 4.2.1. Proteins and Reagents

Horse holo-myoglobin (17568 Da) bovine ubiquitin (8565 Da), 18C6 (264 Da) and polyamidoamine dendrimer (generation 5, PAMAM succinamic acid dendrimer with 1,4-diaminobutane core, theoretical mass 41669 Da) were from Sigma (St. Louis, MO). Neutral solutions were prepared at a protein concentration of 5  $\mu\text{M}$ , with 1 mM ammonium acetate or 1 mM NaCl. Dendrimers were electrosprayed in water containing 100 mM ammonium acetate. As needed, the samples were supplemented with 1% (v/v) sulfolane and/or 1 mM 18C6.

## 4.2.2. Mass Spectrometry

Spectra were acquired on a Synapt ESI mass spectrometer (Waters, Milford, MA). Proteins were electrosprayed at 1.5 kV using gold-coated nanoESI emitters borosilicate glass emitters at a flow rate of  $\sim 40 \text{ nL min}^{-1}$ .<sup>73</sup> Standard experiments were conducted under gentle conditions with cone voltage = 20 V, trap collision energy = 4 V, and source temperature = 80 °C. If required, collisional heating was applied by raising the cone voltage to 120 V, or by raising the trap collision energy (CE) to 40 V. Average charge states were calculated as  $z_{av} = \sum(z_i I_i)/I_i$ , where  $I_i$  is the integrated signal intensity of charge state  $i$ .

## 4.2.3. MD Simulations

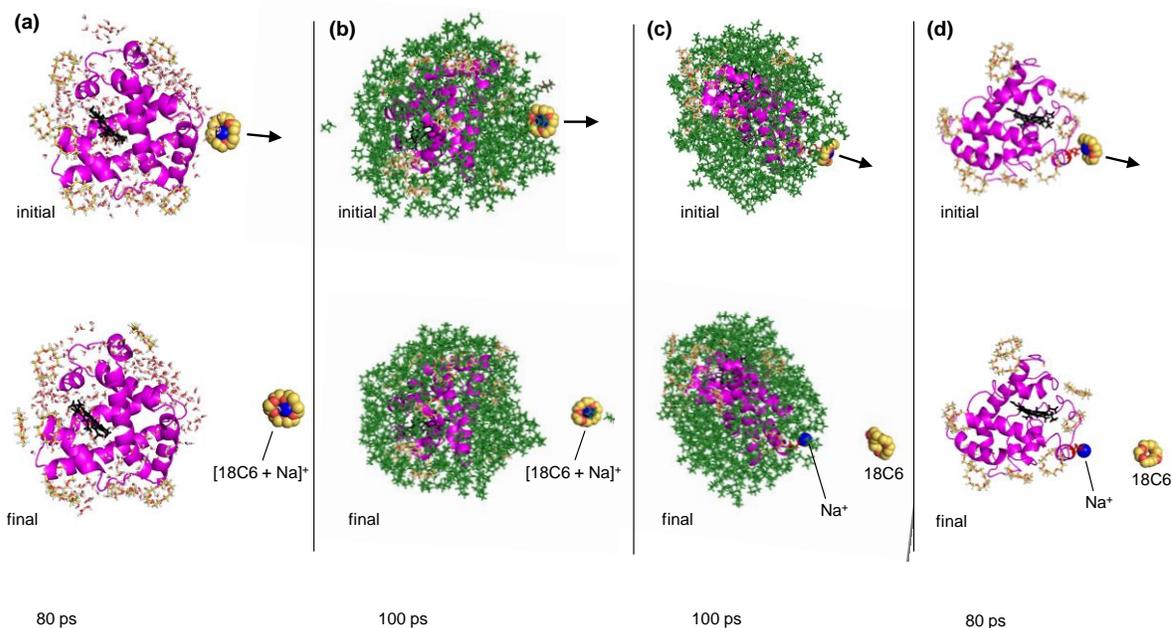
ESI droplet simulations followed a strategy similar to that described earlier,<sup>24, 58</sup> using Gromacs 5<sup>74</sup> and the CHARMM36 force field.<sup>75</sup> The TIP4P/2005 H<sub>2</sub>O model<sup>76</sup> was used because it closely mimics the experimental surface tension of water.<sup>77</sup> 18C6 parameters were obtained from the ParamChem Server<sup>78</sup> and the ZINC database.<sup>79</sup> Initial protein coordinates were taken from the crystal structure 1WLA, with default charges (NT<sup>+</sup>, Arg<sup>+</sup>, Lys<sup>+</sup>, His<sup>0</sup>, Asp<sup>-</sup>, Glu<sup>-</sup>, heme<sup>2-</sup>, CT<sup>-</sup>), for an intrinsic hMb charge of 2-. The initial droplet radius was  $r = 4 \text{ nm}$ . The droplet compositions tested were (a) 8000 waters, (b) 7500 waters and 32 18C6, (c) 5300 waters and 460 sulfolane, (d) 4800 waters, 460 sulfolane, and 32 18C6. The initial water:sulfolane ratio was chosen in accordance with our earlier supercharging simulations.<sup>58</sup> 32 18C6 molecules were included because these conditions generated hMb adducted with  $(10 \pm 1)$  18C6 prior to the onset of forced evaporation in water/18C6 runs. This number is consistent with the experimental data of Figure 4.2b (gentle source conditions), where the most intense signal corresponded to hMb bound to 10 18C6. In all cases, 32 Na<sup>+</sup> provided a net charge of 30+ which corresponds to the Rayleigh limit of a 4 nm aqueous droplet.<sup>11</sup> Initially the protein was centered, and all other constituents had random positions within the droplet. After equilibration, production runs were conducted at 370 K for 75 ns, followed by 200 ns at 450 K. Starting at  $t = 150 \text{ ns}$ , sulfolane and 18C6 were subject to forced evaporation (see subsequent section).<sup>58</sup> Runs were repeated five times with different initial positions and velocities.

Under experimental conditions ESI usually produces  $[M + zH]^{z+}$  ions.<sup>11</sup> Modeling the formation of such species would require QM or DFT methods for properly describing proton transfer events. Unfortunately, the system size and time scale studied here exceed the capabilities of such high level methods.<sup>80</sup> As in previous work, our simulations thus focused on droplets carrying  $Na^+$  as charge carriers, which produce  $[M + zNa]^{z+}$  ions instead of  $[M + zH]^{z+}$ .<sup>11</sup> In contrast to protons, ions such as  $Na^+$  are well described by classical force fields.<sup>75</sup> The final simulated charge states reflect the sum of protein-bound  $Na^+$ , minus the intrinsic 2- hMb charge.

#### 4.2.4. Accelerated MD with Forced Evaporation

The ESI modeling strategy outlined in the main text readily produces free  $[M + zNa]^{z+}$  ions from aqueous droplets.<sup>24</sup> Unfortunately the slow evaporation of sulfolane and 18C6 cause an unacceptable increase in wall clock time.<sup>58</sup> This problem was overcome by subjecting sulfolane and 18C6 to forced evaporation<sup>58</sup> after 150 ns of regular MD. i.e., after the droplet had shrunk to a fraction of its initial size due to complete water evaporation. Under this scheme, the sulfolane or 18C6 with the largest distance from the droplet center was removed at a rate of  $4 \text{ ns}^{-1}$  (or  $0.17 \text{ ns}^{-1}$  for 18C6 after all other solvent molecules had left). This approach resembles biased MD techniques that are widely used for various applications.<sup>81-83</sup>

18C6 forced evaporation causes a slight complication because some crown ethers existed as  $[18C6 + Na]^+$  complexes. It has to be decided if forced evaporation of these 18C6 should include the bound  $Na^+$ . Steered MD<sup>82</sup> revealed that departing  $[18C6 + Na]^+$  lost their  $Na^+$  *only* if the  $Na^+$  was bound to a protein carboxylate (Figure 4.1). For reasons of simplicity,  $[18C6 + Na]^+$  forced evaporation was therefore always conducted by removing the crown ether along with its  $Na^+$ , which represents the appropriate scenario for most of the  $[18C6 + Na]^+$  (those that were not protein-bound). The simulated average protein  $z$  values were corrected for instances that involved carboxylate-bound  $[18C6 + Na]^+$  ( $\Delta z = + 1.7$  for water/18C6, and  $\Delta z = + 3.6$  for water/sulfolane/18C6). The magnitude of these corrections is relatively small, and all of the arguments made in the main text remain valid regardless whether the correction is applied or not.



**Figure 4.1.** Illustrative results of COM (center of mass) pulling simulations for hMb-containing ESI droplets. The aim of these steered MD runs was to determine the proper forced evaporation strategy for  $[18C6 + Na]^+$ . In other words, we determined whether  $[18C6 + Na]^+$  that leave the droplet depart with their  $Na^+$ , or if the  $Na^+$  stays behind. Data were generated by adapting a literature method.<sup>82</sup> An external force (indicated by black arrows) was applied to pull on the 18C6 scaffold of  $[18C6 + Na]^+$ , one crown ether at a time. This was achieved by placing the 18C6 COM in a harmonic potential with  $k = 1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ , and by moving the potential minimum away from the droplet center at a velocity of  $0.02 \text{ nm ps}^{-1}$ . Droplet displacement was eliminated by harmonically restraining all protein backbone atoms with  $k = 1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ . No external force was applied to  $Na^+$  or any other ion/molecule. The pulled 18C6 (orange/red) with their  $Na^+$  (blue) are shown in spacefill representation. All other solvent molecules are shown as sticks. The “initial” configurations in this figure represent MD-generated droplets at intermediate stages of the ESI process. The “final” configurations represent snapshots taken  $\sim 90$  ps later. (a) 18C6 pulling in the presence of water, culminating in departure of the entire  $[18C6 + Na]^+$  moiety. (b) 18C6 pulling in the presence sulfolane (green), culminating in departure of the entire  $[18C6 + Na]^+$  moiety. (c) 18C6 pulling out of sulfolane under conditions where the  $Na^+$  is bound to a side chain carboxylate. This causes departure of an empty 18C6, while the  $Na^+$  remains bound to the side chain. (d) 18C6 pulling for a largely desolvated protein that only retains a few crown ethers. The  $Na^+$  is bound to a side chain carboxylate. Pulling causes the departure of an empty 18C6, while the  $Na^+$  remains bound to the side chain. In total, 50 COM pulling simulations consistently revealed that departing  $[18C6 + Na]^+$  leave their  $Na^+$  behind *only if* the metal is bound to a protein carboxylate (protein- $COO^- \cdots Na^+ \cdots 18C6 \rightarrow$  protein- $COO^- \cdots Na^+ + 18C6$ ). In all other cases the entire  $[18C6 + Na]^+$  complex departed from the droplet.

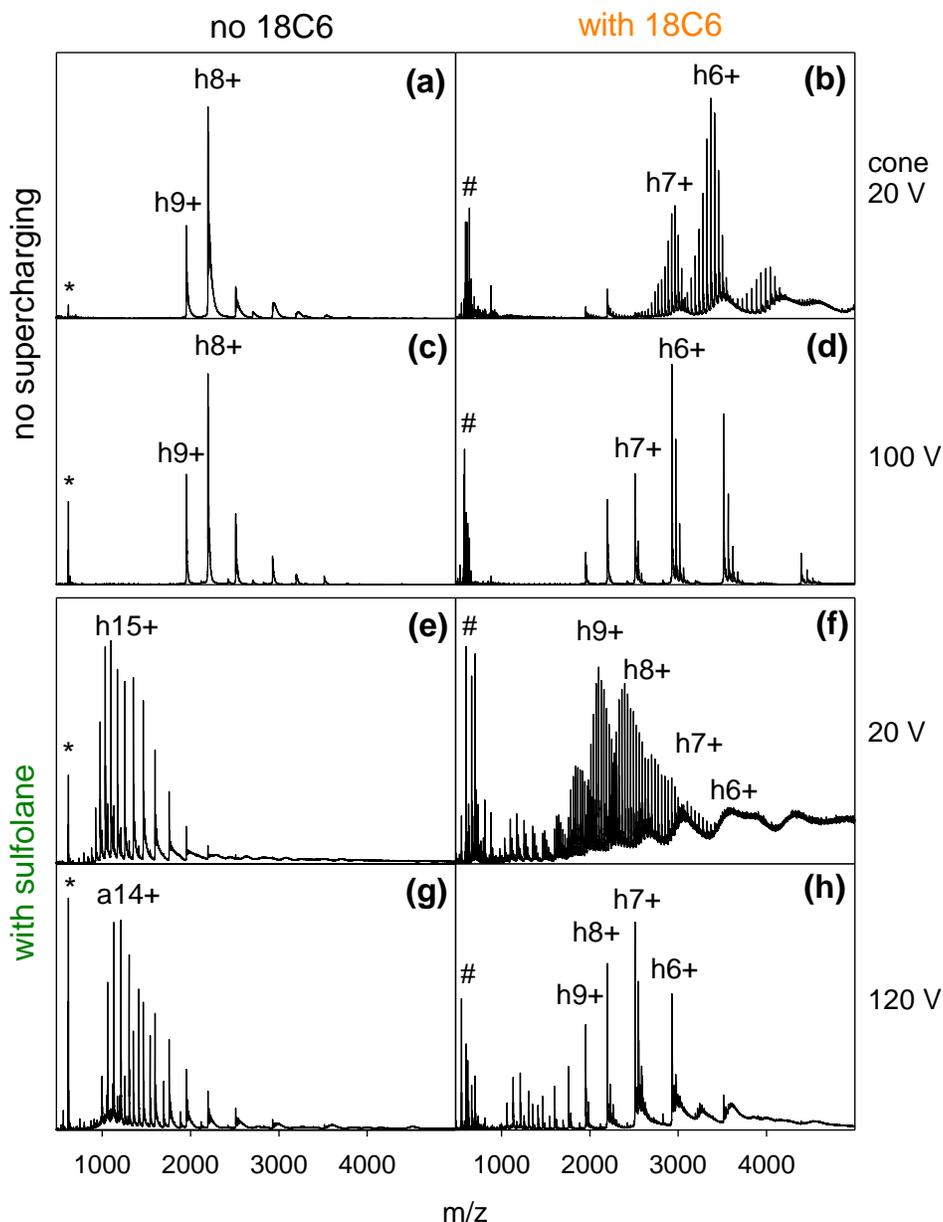
## 4.3. Results and Discussion

### 4.3.1. Effects of 18C6 on ESI Charge States

This work focused on holo-myoglobin (hMb), a heme-protein complex that served as model system for many earlier mechanistic studies.<sup>38, 45, 47, 53, 58</sup> Native ESI in aqueous ammonium acetate solution generated hMb ions in the 8+ and 9+ charge states (Figure 4.2a), very similar to earlier data recorded on different instruments and with different ESI sources.<sup>40, 47, 58</sup> Addition of 1 mM 18C6 resulted in crown ether adduction,<sup>68</sup> and a shift to slightly lower charge states (from 8+/9+ to 6+/7+, Figure 4.2b). Collisional activation caused loss of the 18C6 adducts, while the charge state distribution remained virtually unchanged (Figure 4.2d).

### 4.3.2. 18C6 Suppresses Supercharging

Sulfolane is a typical SCA. As expected from earlier reports,<sup>30, 32, 38, 44-45, 47, 49-52, 54</sup> high charge states (around 15+, Figure 4.2e) were observed when electrospraying hMb from sulfolane-containing aqueous ammonium acetate. The supercharged protein ions largely retained their heme group,<sup>45</sup> whereas denaturation in solution usually disrupts heme-protein interactions.<sup>40</sup> Remarkably, the addition of 1 mM 18C6 to the sulfolane-containing solution suppressed supercharging, i.e., the hMb charge state distribution shifted from around 15+ into the 6+ to 9+ range (Figure 4.2f). The experiments were repeated with collisional activation (cone 120 V). Similar to the sulfolane-free solutions (Figure 4.2b, d), these harsher conditions removed 18C6 adducts without major changes of the charge state distribution (Figure 4.2f, h). It appears that the capability of 18C6 to act as supercharging antidote has not been reported before.

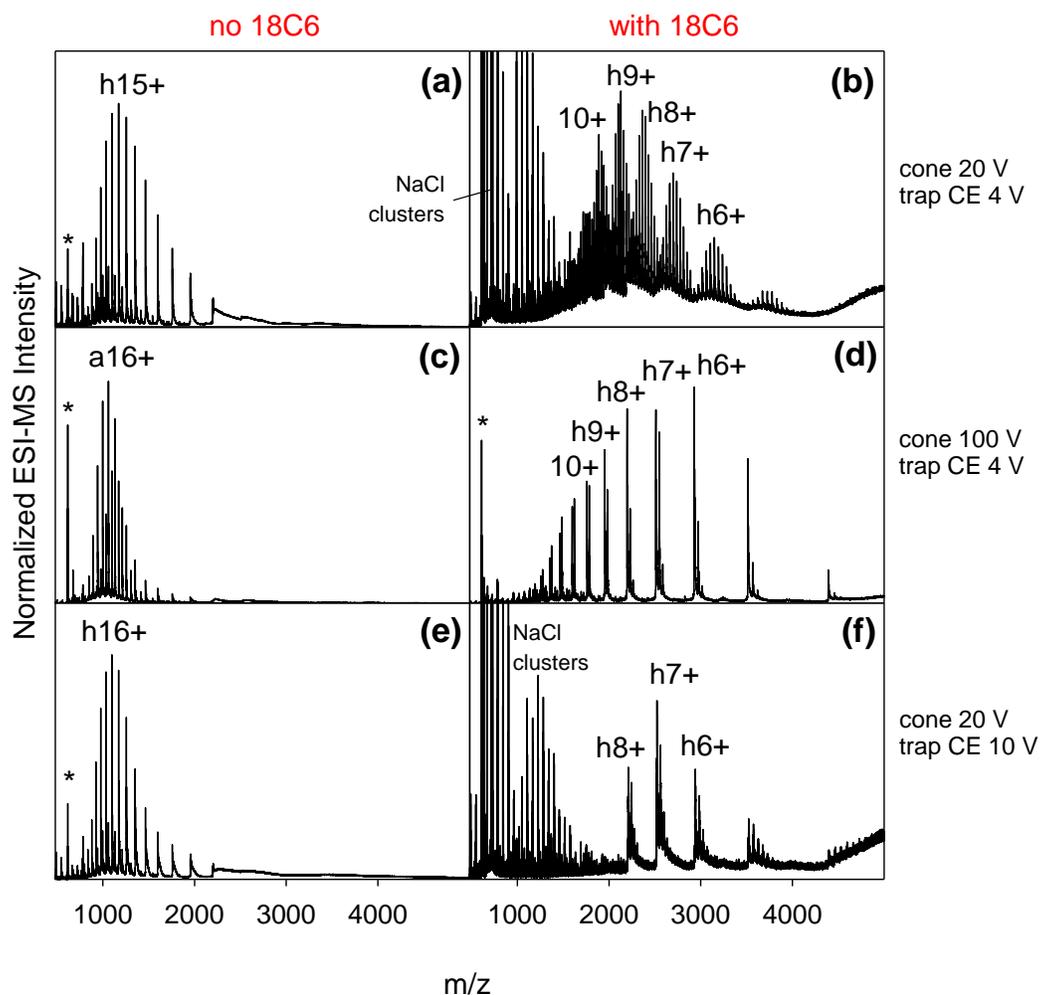


**Figure 4.2.** Mass spectra acquired after electrospraying holo-myoglobin without (panels on left) and with 1 mM 18C6 (panels on right) in neutral aqueous solution containing 1 mM ammonium acetate. (a, b) Data recorded without sulfolane under gentle conditions, i.e., cone 20 V. Extensive adduct formation in (b) is due to noncovalent attachment of up to ~14 18C6. (c, d) Same as in (a) and (b), but with in-source activation (cone 100 V). (e, f) Spectra acquired after addition of 1% sulfolane under gentle conditions (cone 20 V). (g, h) Same as in (e) and (f), but with in-source activation (cone 120 V). h8+, a14+, etc. denote hMb and aMb charge states. \* indicates free heme; # refers to an unidentified cluster.

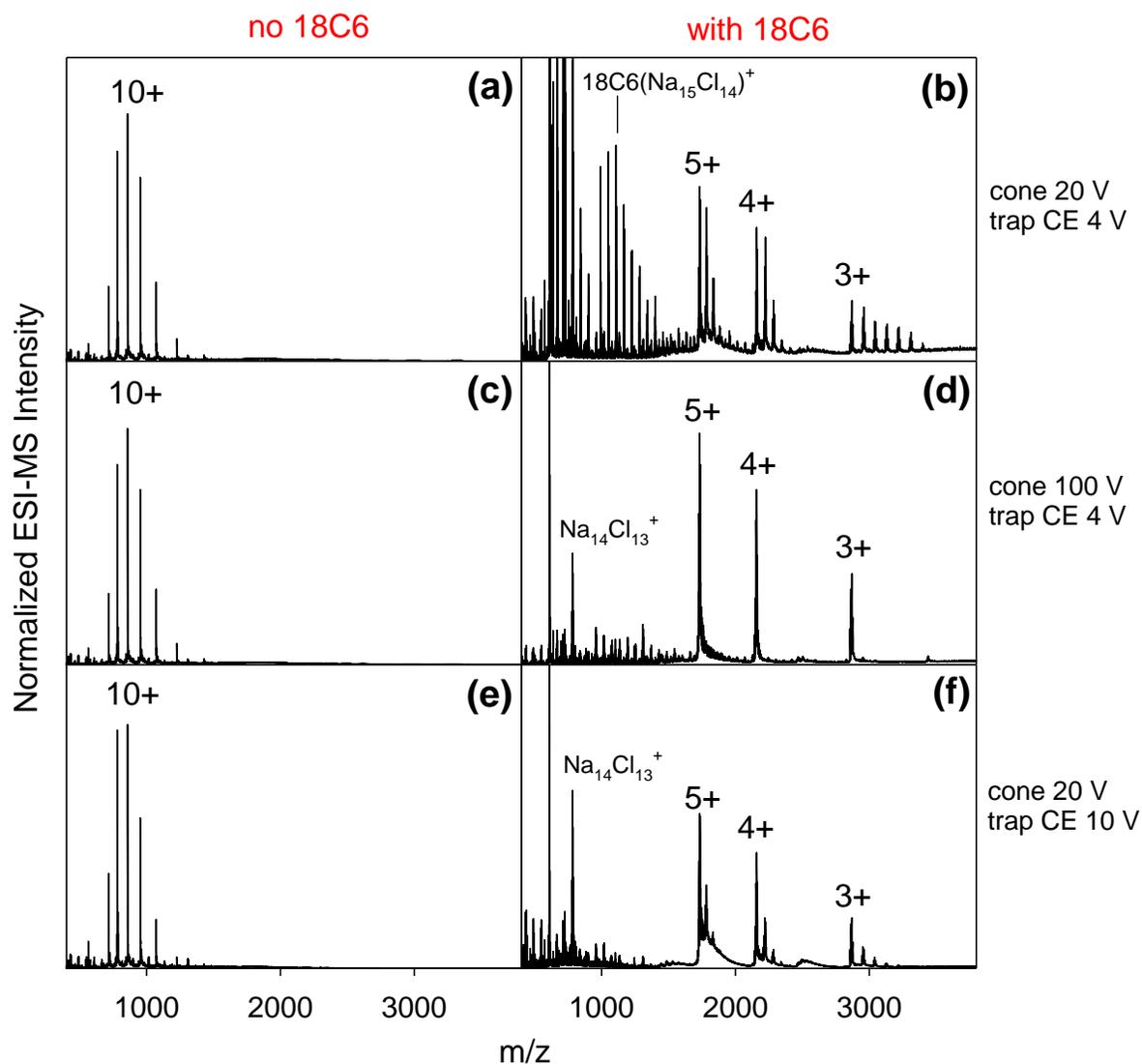
### 4.3.3. 18C6 Effects in NaCl-Containing Solutions

Native ESI-MS experiments on hMb were repeated in solutions containing NaCl instead of ammonium acetate, giving rise to the formation of  $[M + (z-n)H + nNa]^{z+}$  ions, all the way to fully sodiated  $[M + zNa]^{z+1}$ . These conditions resemble those used for the subsequent MD simulations, where  $Na^+$  served as charge carrier. In addition to gentle ESI conditions, we tested the effects of source activation by raising the cone voltage. Alternatively, collisional activation was applied by raising the trap CE (Figure 4.3). All experiments were also repeated with ubiquitin, another common test protein (Figure 4.4).

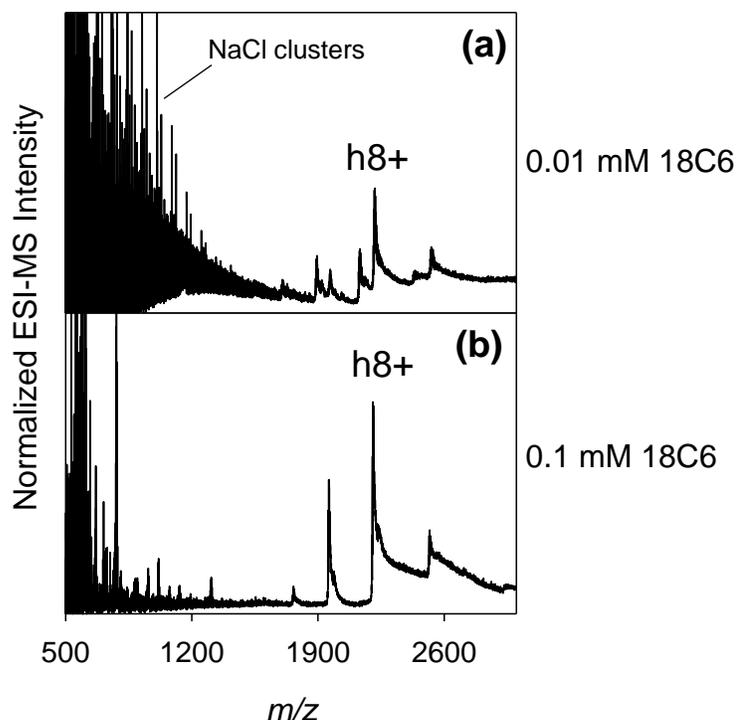
The data obtained for the NaCl-containing samples resemble those of Figure 4.2 and can be summarized as follows: Sulfolane causes supercharging. Addition of 1 mM 18C6 to the sulfolane-containing solution shifts the spectra back to low charge states. This supercharging suppression was observed even when lowering the 18C6 concentration from 1 mM to 0.1 mM or 0.01 mM (Figure 4.5). Collisional activation removes 18C6 adducts without major charge state alterations. Thus, 18C6-induced shifts to lower charge states are *not* primarily caused by the loss of 18C6-bound charge carriers from the gaseous protein. This conclusion is consistent with previous work,<sup>71</sup> where it was noted that collisional charge loss (such as  $\text{protein-NH}_3^+ \cdots 18\text{C6} \rightarrow \text{protein-NH}_2 + [18\text{C6+H}]^+$ ) is enthalpically unfavorable, in agreement with the data of Figure 4.1. Therefore, the capability of 18C6 to act as supercharging antidote must have a different origin. We resorted to MD simulations for uncovering the basis of this phenomenon.



**Figure 4.3.** Mass spectra acquired after electrospraying holo-myoglobin without (panels on left) and with 18C6 (panels on right) in neutral aqueous solution containing 1% sulfolane and 1 mM NaCl. (a, b) Data recorded under gentle conditions. Protein ions in (b) show extensive sodium and 18C6 adduction. (c, d) Same as in (a) and (b), but with in-source activation (cone 100 V). (e, f) Same as in (a) and (b), but with quadrupole activation (trap collision energy 10 V). h15+, a16+, etc. denote hMb and aMb charge states. \* indicates free heme, scaled by 1/3 in panel (c). The low  $m/z$  range in panels (b) and (f) is dominated by various NaCl/18C6 cluster ions.



**Figure 4.4.** Mass spectra acquired after electrospraying ubiquitin without (panels on left) and with 18C6 (panels on right) in neutral aqueous solution containing 1% sulfolane and 1 mM NaCl. (a, b) Data recorded under gentle conditions. Protein ions in (b) show extensive sodium and 18C6 adduction. (c, d) Same as in (a) and (b), but with in-source activation (cone 100 V). Same as in (a) and (b), but with quadrupole activation (trap collision energy 10 V). The low  $m/z$  range in panels (b) and (f) is dominated by various NaCl/18C6 cluster ions.



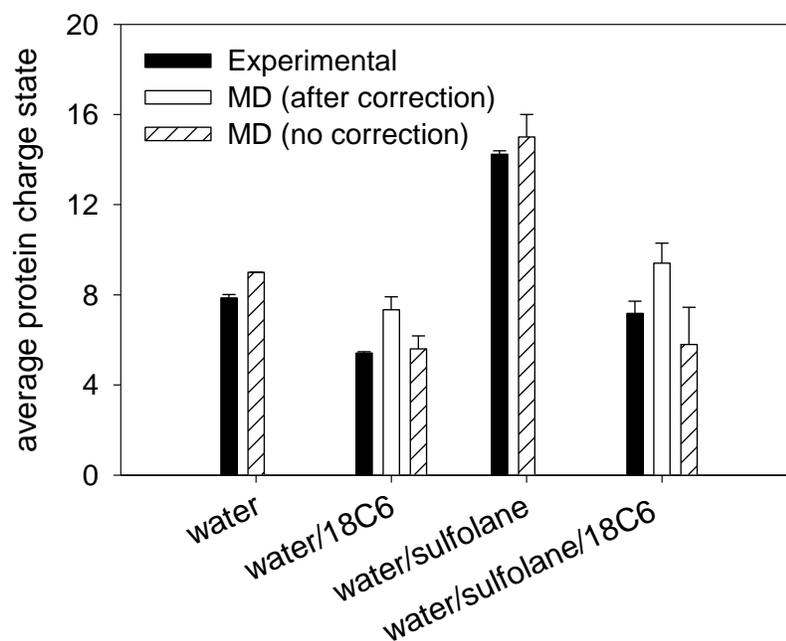
**Figure 4.5.** Mass spectra acquired after electrospraying holo-myoglobin in neutral aqueous solution containing 1% sulfolane, 1 mM NaCl and (a) 0.01 mM 18C6, and (b) 0.1 mM 18C6. Cone 20 V, trap CE 10 V. Even these reduced 18C6 concentrations still suppress supercharging.

#### 4.3.4. Comparing MD and Experimental Results

The ESI droplets modeled here (initial radius 4 nm) were significantly larger than in earlier studies,<sup>18-24</sup> resembling the size regime encountered in experiments.<sup>11</sup> Charge states predicted by MD simulations on four types of hMb-containing droplets are compiled in Figure 4.6, along with the corresponding experimental data. Gratifyingly, the MD data reproduced the experimental trends. Water droplets produced charge states around 8+/9+. Slightly lower charge states were seen for water/18C6. Supercharged proteins with  $z$  values around 14+/15+ were produced from water/sulfolane droplets. The addition of 18C6 to the water/sulfolane droplets dramatically reduced the extent of protein charging. The subsequent MD trajectory analyses reveal the physical reasons underlying the sulfolane and 18C6 effects on protein ESI charge states.

The forced evaporation tool applied during the final stages of droplet shrinkage necessitated small corrections for MD charge states generated in the presence of 18C6 (see

Methods). Figure 4.6 includes results obtained with and without this correction. The MD data reproduced the experimental trends, regardless whether the correction was applied or not.

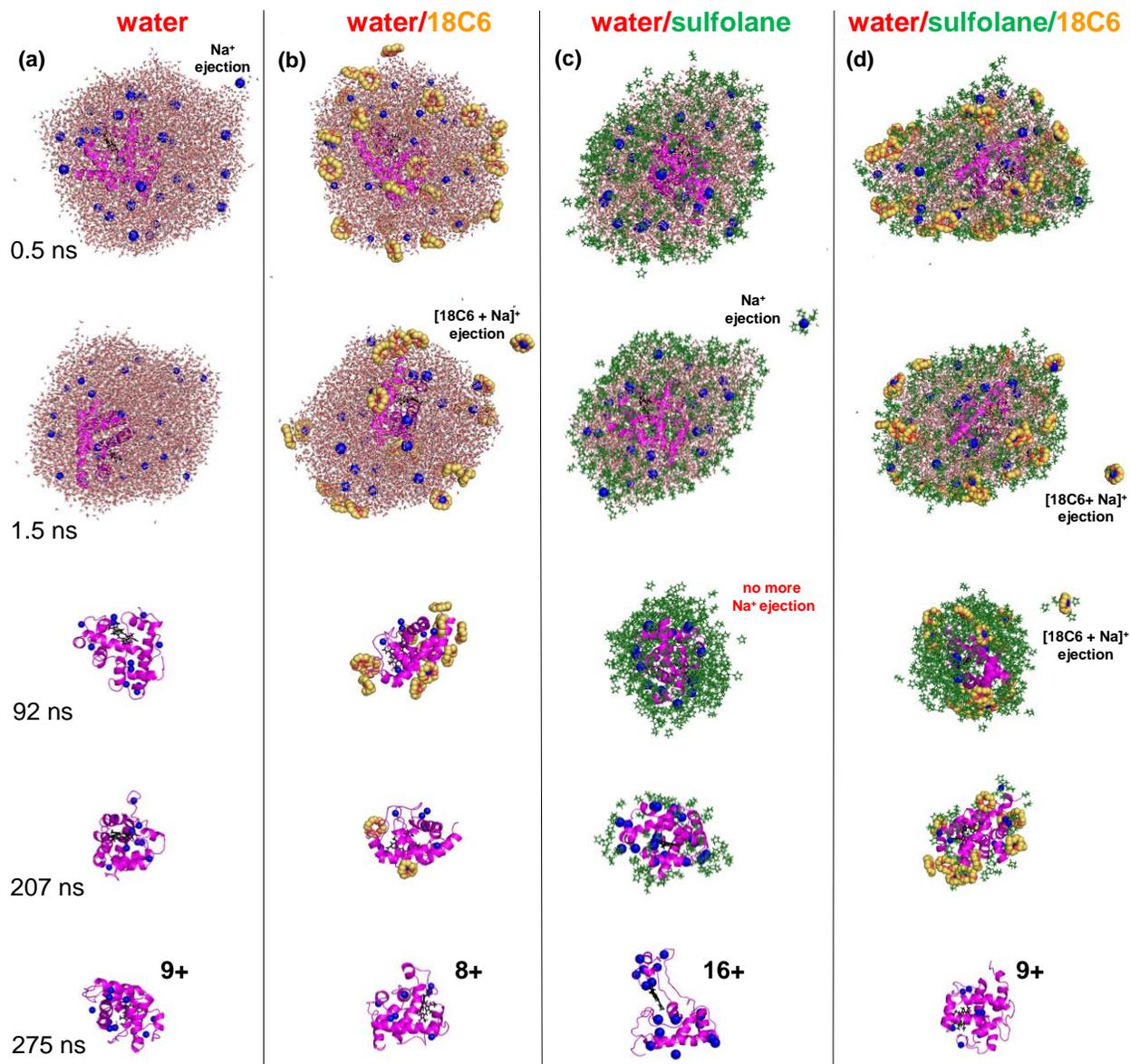


**Figure 4.6.** ESI charge states of hMb obtained experimentally and from MD simulations. Experimental values are averages of three measurements acquired under the conditions of Figure 4.2 (cone 100 V, trap CE 4 V). MD data are based on five replicate runs for each condition. MD data for 18C6-containing droplets are shown with and without forced evaporation correction. Error bars represent standard deviations.

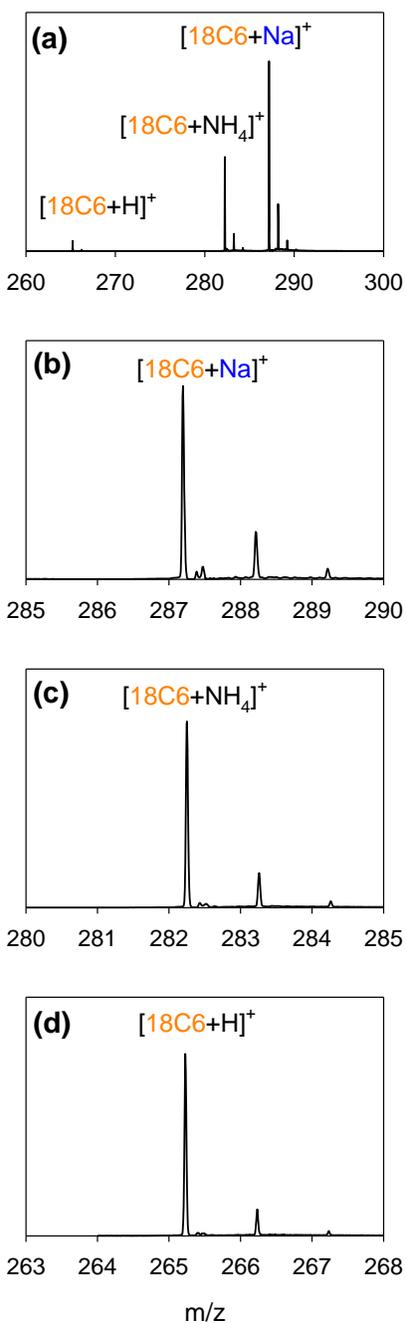
#### 4.3.5. Common Features of MD Trajectories

Snapshots taken from representative MD runs for the four conditions are summarized in Figure 4.7. All trajectories shared several features: The evaporating droplets retained an approximately spherical shape, with the protein in the interior. Multiply charged gaseous hMb was produced via solvent evaporation to dryness, as envisioned by the CRM.<sup>20-24, 58</sup> Droplet shrinkage was accompanied by Na<sup>+</sup> ejection from the droplet surface. None of these IEM events<sup>14, 18, 42</sup> involved completely desolvated Na<sup>+</sup>, instead the departing charge carriers were bound to 18C6 and/or several water or sulfolane molecules. Examples of such IEM events are highlighted in Figure 4.7. Ejected 18C6-bound charge carriers can be observed experimentally. Aqueous ammonium

acetate/18C6 solution produced intense signals for  $[18C6 + NH_4]^+$  and  $[18C6 + Na]^+$  (Figure 4.8), underscoring the role of  $Na^+$  as ubiquitous contaminant in analyte solutions,<sup>11</sup> and giving credence to the use of  $Na^+$  in our simulations.<sup>24, 58</sup>



**Figure 4.7.** Snapshots taken from MD trajectories that culminate in the production of desolvated hMb ions from ESI nanodroplets. Four solvent conditions were tested: (a) Water, (b) Water/18C6, (c) Water/sulfolane, (d) Water/sulfolane/18C6. Time points are indicated along the left hand side. Charge states of protein ions at the end of the simulation runs are shown (at  $t = 275$  ns). IEM ejection events of solvated and/or 18C6-complexed  $Na^+$  are highlighted. Coloring: Protein, pink; heme, black;  $Na^+$ , blue; water oxygen, red; sulfolane, dark green; 18C6, orange/red.



**Figure 4.8.** (a) Mass spectrum obtained upon electrospraying 1 mM aqueous ammonium acetate containing 1 mM 18C6. Close-up views of the [18C6 + Na]<sup>+</sup>, [18C6 + NH<sub>4</sub>]<sup>+</sup>, and [18C6 + H]<sup>+</sup> signals are shown in panels (b) – (d). The intense [18C6 + Na]<sup>+</sup> signal is consistent with the presence of Na<sup>+</sup> ions as ubiquitous contaminants in ESI-MS analyte solutions,<sup>11</sup> supporting the validity of our MD strategy that is based on nanodroplets charged with excess Na<sup>+</sup>. [18C6 + H]<sup>+</sup> in panel (d) may be ejected from the droplet as protonated crown ether, or it could arise from loss of ammonia from [18C6 + NH<sub>4</sub>]<sup>+</sup>.<sup>7, 1</sup>

### 4.3.6. Principles that Govern the ESI Charge States of Proteins

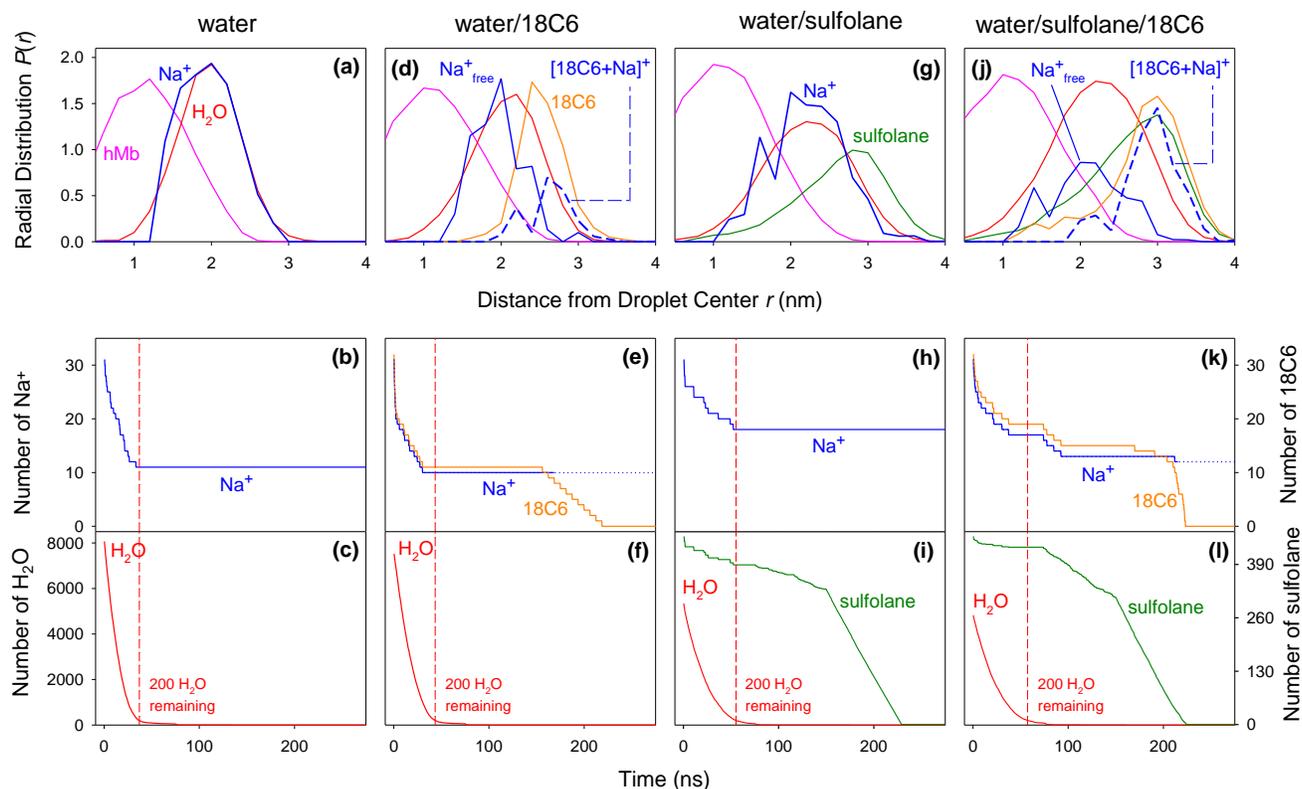
Before proceeding, it is helpful to summarize the rules that govern the charge carrier behavior in native ESI.<sup>24, 58</sup> (i) Charge carriers can experience only two fates, ejection from the droplet or binding to the protein. (ii) The ejection of bare charge carriers from desolvated proteins is not feasible, at least not for the charge state range considered here.<sup>84</sup> (iii) Any charge carriers that are not ejected from the droplet will become part of the protein charge. (iv) Prerequisite for each charge ejection from the droplet is that the charge carrier can (at least transiently) reside close to the droplet surface; any factor that tends to exclude charge carriers from the surface will lower the charge ejection efficiency. From (i) - (iv) it follows that any factor that hinders charge carrier access to the droplet surface will increase the protein charge  $z$ . The repercussions of these considerations will become clear in the next section.

### 4.3.7. MD Trajectories Reveal the Basis of Charging and Supercharging

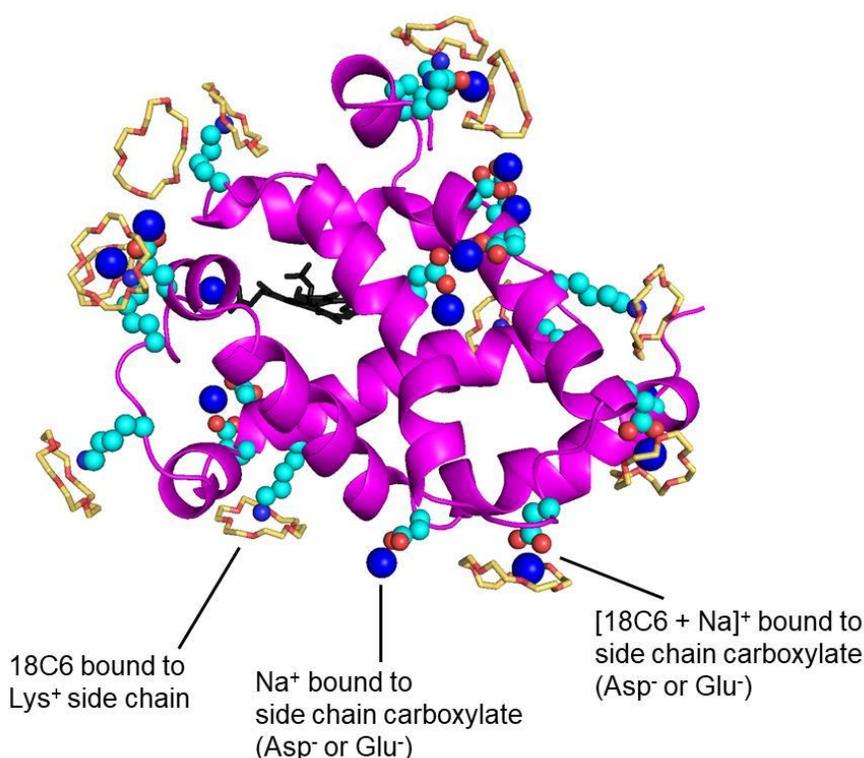
In *aqueous droplets* water and  $\text{Na}^+$  shared the same radial distributions (Figure 4.9a), reflecting the favorable solvation of  $\text{Na}^+$  in  $\text{H}_2\text{O}$ .<sup>85</sup> Under these conditions  $\text{Na}^+$  can roam the entire solvent-occupied volume, including positions at the droplet surface from where charge ejection readily takes place (Figure 4.9b). As the final water molecules evaporated, charge ejection came to a halt (Figure 4.9b, c). The remaining  $\text{Na}^+$  underwent irreversible binding to hMb (Figure 4.7a,  $t \approx 92$  ns) at protein carboxylates (Figure 4.10).

For *water/18C6 droplets* (Figure 4.9b) all 18C6 were located close to the droplet surface, consistent with their amphiphilic nature.<sup>62-63, 86-88</sup>  $\text{Na}^+$  exhibited a bimodal radial distribution, comprising water-solvated  $\text{Na}^+$  in the droplet interior and  $[\text{18C6} + \text{Na}]^+$  at the droplet surface (Figure 4.9d). The positioning of  $[\text{18C6} + \text{Na}]^+$  at the liquid/vapor interface facilitated ejection of these complexes from the droplet. Thus, the high surface affinity of  $[\text{18C6} + \text{Na}]^+$  was responsible for the slightly lower protein charge states in water/18C6 compared to pure water (Figure 4.6). The behavior seen here for  $[\text{18C6} + \text{Na}]^+$  mirrors the facile IEM ejection of other cationic solutes that carry nonpolar groups.<sup>12</sup> Charge ejection ceased just prior to evaporation of the last water molecules (Figure 4.9e,f). At this point, residual  $\text{Na}^+$  and  $[\text{18C6} + \text{Na}]^+$  underwent ion pairing with protein carboxylates, while non-sodiated 18C6 bound to  $\text{Lys}^+$

(Figure 4.10).<sup>68-72</sup> These MD-predicted 18C6 adducts were experimentally observable under gentle ESI conditions (Figure 4.2b). The number of MD-adducted 18C6 ( $10 \pm 1$ ) was consistent with the experimental data of Figure 4.2b, where the most intense signal corresponded to hMb<sup>6+</sup> attached to 10 crown ethers. 18C6 removal by MD forced evaporation (Figures 4.7b, 4.9e) or by collisional heating (Figure 4.2d) produced fully desolvated hMb.



**Figure 4.9.** MD data for four types of hMb-containing ESI droplets, as noted along the top. Top row: Spatial distribution of solutes, averaged over  $t = 25$  to  $t = 32$  ns in all repeat runs. For 18C6-containing droplets two separate  $\text{Na}^+$  distributions are shown, reflecting the behavior of sodium in  $[\text{18C6} + \text{Na}]^+$  vs. all other sodium ions (denoted as  $\text{Na}^+_{\text{free}}$ ). Middle and bottom rows: Time-dependent changes in droplet composition for four typical trajectories, reflecting the occurrence of solvent evaporation and charge ejection. Dashed vertical lines indicate the point where 200 water molecules remain in the droplet; after this point only panel (k) shows  $\text{Na}^+$  ejection (as  $[\text{18C6} + \text{Na}]^+$ ). Dotted blue lines indicate  $\text{Na}^+$  data after correction for  $[\text{18C6} + \text{Na}]^+$  forced evaporation (see Methods). The coloring of droplet components matches that of Figure 4.7.



**Figure 4.10.** Protein interactions with Na<sup>+</sup>, 18C6, and [18C6 + Na]<sup>+</sup>, exemplified for a  $t = 205$  ns MD frame taken from a water/sulfolane/18C6 run. Coloring is as in Figure 4.7. Large blue spheres represent Na<sup>+</sup>; small blue spheres represent N $\epsilon$  atoms of Lys<sup>+</sup>; hydrogens have been omitted.

*Water/sulfolane droplets* (Figure 4.7c) showed sulfolane enrichment in the outermost layers. This nanoscale segregation is consistent with phenomena reported for other binary mixtures, such as methanol/water or ethanol/water, where the nonaqueous component undergoes surface enrichment despite being “fully miscible” in bulk solution.<sup>86-89</sup> A key driving force for the sulfolane/water segregation seen here is the maximization of enthalpically favorable water-water contacts in the droplet core.<sup>90</sup> (Analogous arguments explain why 18C6 is driven to the surface of the water/18C6 droplets discussed in the preceding paragraph.<sup>86-90</sup>) For the ESI water/sulfolane droplets of Figure 4.7c these segregation phenomena produced a sulfolane-enriched surface layer that surrounded the protein-containing aqueous core (Figure 4.9g). The ionophobic nature of sulfolane<sup>58, 61</sup> (along with the high water affinity of Na<sup>+</sup>)<sup>85</sup> largely confined the charge carriers to the droplet interior. The resulting Na<sup>+</sup> depletion at the droplet surface

reduced the charge ejection probability. Because of this reduced IEM efficiency, a larger number of  $\text{Na}^+$  remained trapped within the droplet compared to the water or water/18C6 droplets (Figures 4.9h vs. 4.9b, 4.9e). Complete  $\text{H}_2\text{O}$  loss from the water/sulfolane droplets subsequently generated sulfolane-encapsulated hMb, concomitant with irreversible binding of all remaining  $\text{Na}^+$  to protein carboxylates (Figure 4.7c,  $t = 92$  ns). Further charge loss would require the occurrence of highly unfavorable events, i.e., dissociation of protein- $\text{COO}^- \cdots \text{Na}^+$  ion pairs, followed by  $\text{Na}^+$  diffusion through the ionophobic sulfolane environment. Hence, the protein attained its final “supercharged”  $z$  value after all water had left the droplet, but long before evaporation of the sulfolane shell had gone to completion.

These data for water/sulfolane reflect the charge trapping model outlined in the Introduction.<sup>58</sup> This mechanism causes supercharging via two interrelated factors, i.e., (i) formation of a SCA layer at the droplet surface that tends to confine charges to the aqueous interior, thereby impeding charge ejection. (ii) After complete water loss, the protein becomes surrounded by an SCA shell and the remaining charge carriers are forced to associate with protein carboxylates. Charge partitioning during droplet shrinkage is governed by the *low* affinity of charge carriers for the SCA, their *greater* affinity for protein carboxylates, and their *high* affinity for water.<sup>58, 85</sup> For the three droplet types discussed so far (water, water/18C6, water sulfolane), only  $\text{Na}^+$  solvation by water prevented irreversible charge carrier binding to the protein. Charge ejection came to a halt once the number of  $\text{H}_2\text{O}$  dropped below  $\sim 200$  (vertical lines in Figure 4.9b, e, h), forcing freely diffusible  $\text{Na}^+_{\text{aq}}$  to transition into carboxylate-bound environments.

As noted, a key element of the charge trapping model is that SCAs exhibit a low charge carrier affinity. This aspect is consistent with earlier work that focused on protein charging by  $\text{H}^+$ , where it was noted that SCAs generally exhibit a low  $\text{H}^+$  affinity in solution (i.e., a weak Brønsted basicity).<sup>48, 91</sup> These parallels support the view that the charge trapping model is not limited to droplets containing  $\text{Na}^+$ , but that it also applies to  $\text{H}^+$  and other ESI-relevant charge carriers.

In *water/sulfolane/18C6 droplets* the charge carriers exhibited a bimodal distribution, with  $\text{H}_2\text{O}$ -solvated  $\text{Na}^+$  in the interior and abundant sulfolane-solvated  $[\text{18C6} + \text{Na}]^+$  in the

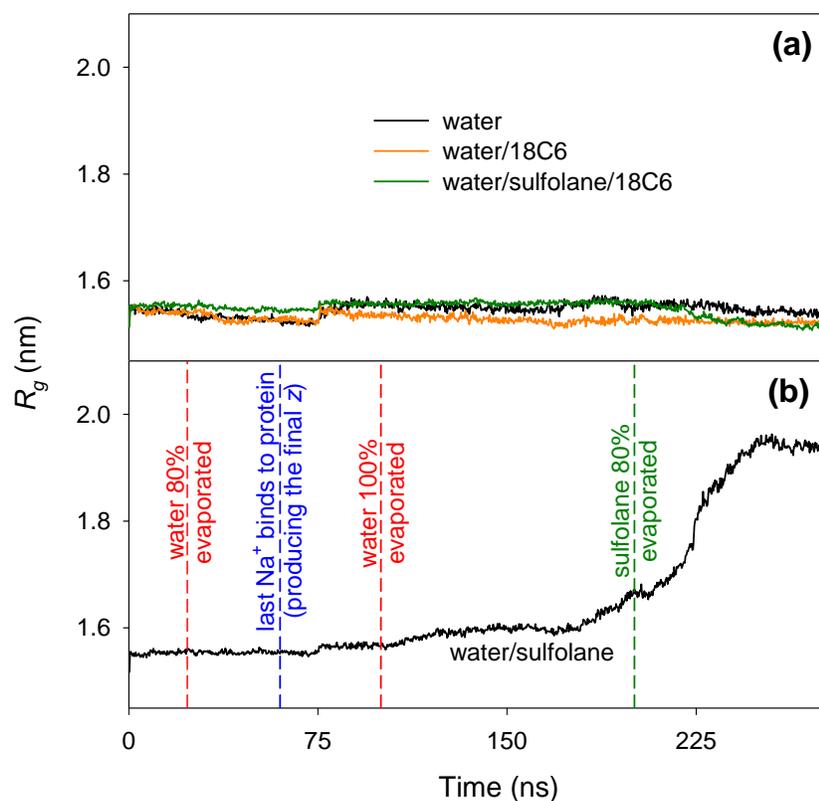
outermost droplet layers (Figure 4.9j). The latter reflect the capability of crown ethers to solubilize cations in unfavorable solvents.<sup>62-63</sup> [18C6 + Na]<sup>+</sup> enrichment at the droplet surface facilitated the ejection of these charged complexes. Numerous [18C6 + Na]<sup>+</sup> remained at the surface of the droplet even after complete water evaporation, thereby ensuring *continued* charge ejection and suppression of charge carrier binding to the protein (Figure 4.7d at  $t = 92$  ns, Figure 4.9k, l).

In summary, the MD data reveal that 18C6 nullifies the charge-enhancing effects of the SCA by eliminating charge trapping. Our results confirm the hypothesis stated in the Introduction, i.e., the proposal that 18C6 binds charge carriers (Na<sup>+</sup>, NH<sub>4</sub><sup>+</sup>, H<sub>3</sub>O<sup>+</sup>)<sup>62, 64-67</sup> and shuttles them through the SCA trapping layer.

#### 4.3.8. Relationship between Supercharging and Unfolding

Our MD runs showed that hMb in water/sulfolane retained a compact structure until the droplet had almost completely dried out (Figure 4.7c), long after the final  $z$  value had been attained via Na<sup>+</sup> binding. Coulombically driven unfolding of the supercharged protein started to take place during the final solvent evaporation steps (Figure 4.11). Hence, for the conditions examined here, unfolding is caused by supercharging and not *vice versa*. In other words, our data argue against the idea that native ESI supercharging is caused by protein unfolding within the droplet, with subsequent protein extrusion from the droplet surface.<sup>44, 47</sup> Instead, our findings support the view that supercharging is caused by charge trapping. While supercharged hMb undergoes unfolding in the gas phase,<sup>47, 58</sup> other supercharged proteins are more resilient and retain native-like properties.<sup>30, 45, 55, 60</sup> The experimental observation of such compact supercharged proteins<sup>30, 45, 55, 60</sup> would be difficult to reconcile with the unfolding model,<sup>44, 47</sup> while our charge trapping model readily explains how such species can form.

Notwithstanding the aforementioned conclusions, we do not rule out that unfolding may contribute to supercharging under some conditions, likely in combination with the SCA-mediated charge trapping outlined above. The latter scenario is supported by data on disulfide intact and reduced proteins.<sup>44</sup> The possible occurrence of protein unfolding within the droplet will depend on various factors, including the droplet lifetime,<sup>92</sup> confinement effects,<sup>93</sup> and interactions with gas/liquid interfaces.<sup>94</sup>



**Figure 4.11.** Protein compactness during ESI, reported as radius of gyration ( $R_g$ ) vs. MD simulation time, averaged over all runs. (a) Behavior of hMb in water, water/18C6, and water/sulfolane/18C6 droplets, all of which culminate in low charge states. In all three cases the protein retains a tightly folded conformation. (b)  $R_g$  of hMb in water/sulfolane droplets, which produce high charge states (supercharging). Vertical dashed lines in panel (b) indicate the behavior of water, sulfolane, and  $\text{Na}^+$  in the evaporating droplets. Coulombically driven unfolding starts to take place during the final stages of droplet evaporation. The protein reaches its final ESI charge state  $z$  long before unfolding commences (the last  $\text{Na}^+$  binding event is indicated in blue), implying that supercharging is not caused by unfolding. Instead, supercharging is the cause of unfolding.

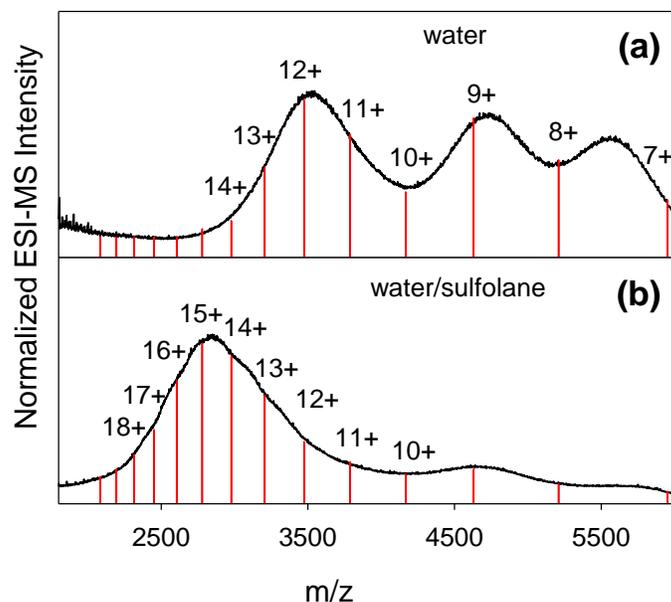
#### 4.3.9. Dendrimer Supercharging

Dendrimers are hyperbranched globular macromolecules that are incapable of large-scale unfolding.<sup>95-96</sup> These analytes are well suited for further scrutinizing the competing supercharging models. The charge trapping model predicts that dendrimers will undergo supercharging, while according to the unfolding model dendrimers should be immune to SCAs.<sup>44</sup>

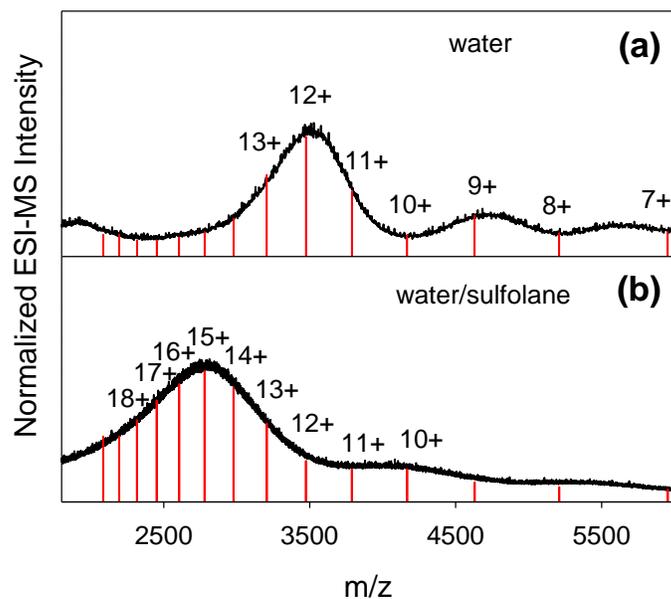
<sup>47</sup> For meaningful test experiments it is important to use dendrimers with "protein-like" properties, i.e., with both acidic and basic sites, and with a MW similar to that of typical proteins.

Earlier work <sup>46</sup> examined the behavior of DAB-16 (1687 Da) and DAB-64 dendrimers (7168 Da). The former did not undergo supercharging. For the latter, the presence of SCA in water caused a broadened charge state distribution that suggested supercharging for a sub-population of the analytes. The implications of those data<sup>46</sup> for proteins are inconclusive. The absence of carboxylates in DAB dendrimers prevents analyte charging via R-COO<sup>-</sup> neutralization, unlike for proteins where carboxylates represent the main charge carrier binding sites.<sup>24</sup> In addition, the small size of DAB-16 raises questions as to whether this species exhibits true CRM behavior.<sup>1</sup>

Here we tested the behavior of G5 PAMAM succinamic acid dendrimer. Its theoretical MW (41669 Da) is comparable to proteins that have previously been subjected to native ESI supercharging.<sup>30, 45, 55, 59-60</sup> This dendrimer possesses both basic sites (tertiary amines) and acidic moieties (succinamic acid groups). A slight complication is the fact that PAMAM dendrimers generally exhibit mass heterogeneity arising from defects in their branched structures, particularly for large species that are comparable in size to proteins.<sup>96</sup> ESI mass spectra recorded in aqueous solution showed several broad maxima that can be attributed to the charge state range of 7+ to 12+ (Figure 4.12a). Upon addition of sulfolane the spectra underwent a dramatic shift towards lower  $m/z$ , corresponding to charge states around 15+ and higher (Figure 4.12b). Similar data were obtained when repeating the experiments in NaCl-containing solution (Figure 4.13). These spectra demonstrate the occurrence of sulfolane-induced supercharging for the PAMAM dendrimer, confirming the prediction of the charge trapping model.



**Figure 4.12.** ESI mass spectra of G5 PAMAM dendrimer in aqueous solution containing ammonium acetate (a) Without sulfolane, (b) With sulfolane. Mass heterogeneity obscures individual charge states. Red lines indicate expected peak positions for the calculated theoretical mass.



**Figure 4.13.** ESI mass spectra of G5 PAMAM dendrimer in aqueous solution containing ammonium acetate (a) Without sulfolane, (b) With sulfolane. These data are analogous to those of Figure 4.12, but here the experiments were conducted in the presence of 1 mM NaCl to promote the formation of sodiated ions. Mass heterogeneity precludes the identification of individual charge states. Red lines indicate peak positions for the calculated theoretical mass of the analyte. Supercharging of the dendrimer in the presence of sulfolane is clearly apparent.

## 4.4. Conclusions

The current work demonstrates that MD simulations represent a powerful tool for probing the behavior of ESI droplets containing multiple interacting components. We applied crown ethers as a mechanistic probe of the ESI process. The ability of 18C6 to act as phase transfer catalyst<sup>62-63</sup> alters the location of charge carriers inside ESI droplets, with dramatic consequences for protein charge states. Crown ethers may also influence protein structures in solution<sup>97-98</sup> and in the gas phase,<sup>70</sup> but our data do not support the view that such conformational factors are responsible for the charge state shifts reported here. The mechanistic insights obtained in the current work can be summarized in cartoon form (Figure 4.14), where blue “+” symbols indicate charge carriers. While the current work focused on Na<sup>+</sup>, Figures 4.2, 4.3, 4.4 suggest that similar considerations also apply to other ESI-relevant charge carriers such as H<sub>3</sub>O<sup>+</sup> and NH<sub>4</sub><sup>+</sup>.<sup>11, 91</sup>

Evaporating ESI nanodroplets experience internal Coulomb repulsion which tends to trigger charge ejection via the IEM.<sup>17</sup> The efficiency of these events is modulated by the capability of charge carriers to migrate to the droplet surface (because only *surface* charges can undergo IEM ejection<sup>12, 14</sup>). In aqueous solution (Figure 4.14a) the favorable solvation characteristics afforded by water allow charge carriers to adopt positions throughout the droplet, including locations close to the surface from where they can be ejected. The relatively few remaining charge carriers in the vanishing droplet bind to the protein, producing low CRM charge states.<sup>11, 24-25</sup>

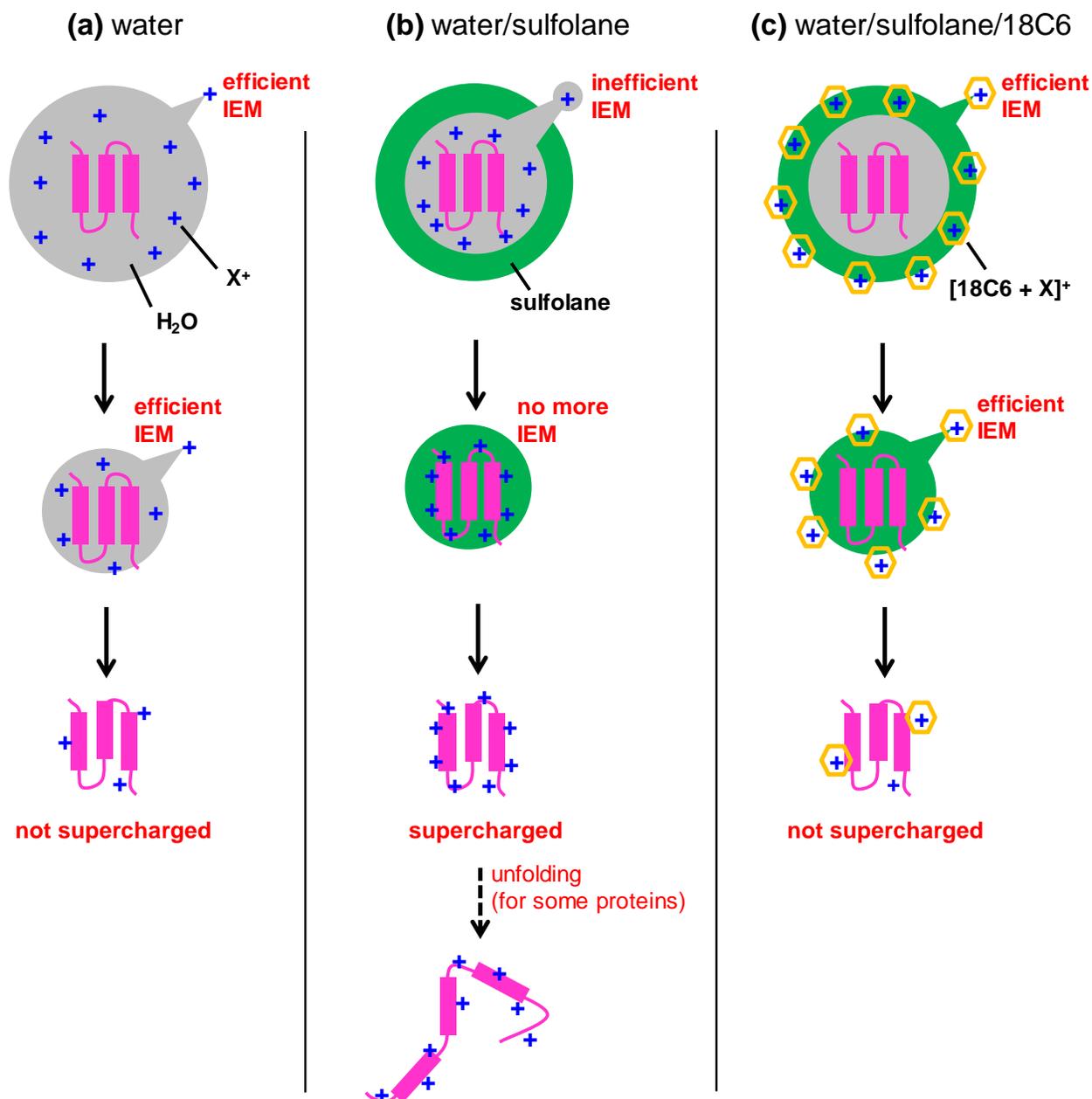
The addition of SCA leads to supercharging via charge trapping (Figure 4.14b)<sup>58</sup>. The SCA initially forms an ionophobic surface layer. As a result, charge carriers preferentially reside in the droplet interior such that their IEM efficiency is reduced; thus, a larger number of charge carriers remain in the droplet compared to the purely aqueous droplets. Once all the water has evaporated the SCA encapsulates the protein. Unfavorable interactions with the SCA force the remaining charge carriers to associate with the protein. All these (many) charge carriers remain bound until the SCA layer has evaporated – producing a desolvated supercharged protein ion. Depending on their structural resilience, supercharged proteins may undergo electrostatically driven unfolding (as in the case of hMb<sup>47, 58</sup>), or they can retain compact conformations (as demonstrated for several other proteins<sup>30, 45, 55, 59-60</sup>). Our MD simulations and dendrimer

supercharging data do not support the view that unfolding within the ESI droplet is the root cause of supercharging.<sup>44, 47</sup>

Figure 4.14c illustrates how 18C6 acts as supercharging antidote. 18C6 solubilizes charge carriers in the SCA, allowing the charge carriers to reside at the droplet surface such that IEM ejection proceeds with high efficiency. Compared to the supercharging conditions of Figure 4.14b the droplets lose more charge, such that the dried-out protein at the end of the process has a lower  $z$ .

Experimental supercharging spectra exhibit wide charge state distributions (Figure 4.2e, g), whereas our simulations produced fairly well defined  $z$  values (Figure 4.6). The experimentally observed charge state range is attributed to the heterogeneous nature of evaporation/fission events in the ESI plume, which will yield nanodroplets with different SCA concentrations.<sup>11</sup> Nanodroplets containing more SCA will produce higher  $z$  values.<sup>47, 58</sup> Modeling this heterogeneity is difficult, as it would require knowledge of the exact ESI plume composition. Regardless of these nuances, it is remarkable how well the current MD data capture the experimental trends, i.e., a dramatic shift to higher charge states in the presence of sulfolane and the suppression of supercharging by 18C6 (Figure 4.6).

Our use of 18C6 as a mechanistic ESI probe expands on previous studies, where this remarkable molecule was applied to examine protein structures in solution<sup>68</sup> and in the gas phase.<sup>69-70</sup> In future work we hope to apply strategies similar to those used here for uncovering the mechanistic basis of supercharging under denaturing conditions, where the charge states formed are even higher than under the native ESI conditions examined here.<sup>32, 44, 54</sup>



**Figure 4.14.** Cartoon summary of MD and experimental results. (a) Native ESI in aqueous solution, producing low charge states. A similar scenario is encountered for water/18C6 droplets (not shown). (b) Supercharging via the charge trapping. A highly charged protein is formed because charge ejection from the droplet is hindered. The supercharged protein may undergo gas phase unfolding. (c) 18C6 acts as supercharging antidote; it prevents charge trapping by promoting the ejection of charge carriers from the droplet surface. “X<sup>+</sup>” represents charge carriers such as Na<sup>+</sup>, H<sub>3</sub>O<sup>+</sup> or NH<sub>4</sub><sup>+</sup>. “IEM” indicates charge carrier ejection (field emission).

## 4.5. References

1. Nguyen, S.; Fenn, J. B., *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1111-1117.
2. Marty, M. T.; Zhang, H.; Cui, W. D.; Blankenship, R. E.; Gross, M. L.; Sligar, S. G., *Anal. Chem.* **2012**, *84*, 8957-8960.
3. Leney, A. C.; Heck, A. J. R., *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 5-13.
4. Marcoux, J.; Robinson, C. V., *Structure* **2013**, *21*, 1541-1550.
5. Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R., *Int. J. Mass Spectrom.* **2002**, *216*, 1-27.
6. Bonner, J. G.; Lyon, Y. A.; Nellessen, C.; Julian, R. R., *J. Am. Chem. Soc.* **2017**, *139*, 10286-10293.
7. Han, L.; Hyung, S.-J.; Mayers, J. J. S.; Ruotolo, B. T., *J. Am. Chem. Soc.* **2011**, *133*, 11358-11367.
8. El-Baba, T. J.; Woodall, D. W.; Raab, S. A.; Fuller, D. R.; Laganowsky, A.; Russell, D. H.; Clemmer, D. E., *J. Am. Chem. Soc.* **2017**, *139*, 6306-6309.
9. Wytenbach, T.; Bleiholder, C.; Bowers, M. T., *Anal. Chem.* **2013**, *85*, 2191-2199.
10. Marklund, E. G.; Degiacomi, M. T.; Robinson, C. V.; Baldwin, A. J.; Benesch, J. L. P., *Structure* **2015**, *23*, 791-799.
11. Kebarle, P.; Verkerk, U. H., *Mass Spectrom. Rev.* **2009**, *28*, 898-917.
12. Cech, N. B.; Enke, C. G., *Mass Spectrom. Rev.* **2001**, *20*, 362-387.
13. Grimm, R. L.; Beauchamp, J. L., *J. Phys. Chem. A* **2010**, *114*, 1411-1419.
14. Iribarne, J. V.; Thomson, B. A., *J. Chem. Phys.* **1976**, *64*, 2287-2294.
15. Cole, R. B., *J. Mass. Spectrom.* **2000**, *35*, 763-772.
16. Loscertales, I. G.; de la Mora, J. F., *J. Chem. Phys.* **1995**, *103*, 5041-5060.
17. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L., *Anal. Chem.* **2009**, *81*, 369-377.
18. Znamenskiy, V.; Marginean, I.; Vertes, A., *J. Phys. Chem. A* **2003**, *107*, 7406-7412.
19. Steinberg, M. Z.; Breuker, K.; Elber, R.; Gerber, R. B., *Phys. Chem. Chem. Phys.* **2007**, *9*, 4690-4697.
20. Patriksson, A.; Marklund, E.; van der Spoel, D., *Biochemistry* **2007**, *46*, 933-945.
21. Oh, M. I.; Consta, S., *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 2262-2279.

22. Kim, D.; Wagner, N.; Wooding, K.; Clemmer, D. E.; Russell, D. H., *J. Am. Chem. Soc.* **2017**, *139*, 2981-2988.
23. Porrini, M.; Rosu, F.; Rabin, C.; Darre, L.; Gomez, H.; Orozco, M.; Gabelica, V., *ACS Central Sci.* **2017**, *3*, 454-461.
24. McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L., *J. Am. Chem. Soc.* **2015**, *137*, 12667-12676.
25. de la Mora, F. J., *Anal. Chim. Acta* **2000**, *406*, 93-104.
26. Siuti, N.; Kelleher, N. L., *Nat. Methods* **2007**, *4*, 817-821.
27. Zubarev, R. A., *Mass Spectrom. Rev.* **2003**, *22*, 57-77.
28. Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528-9533.
29. Quintyn, R. S.; Zhou, M.; Yan, J.; Wysocki, V. H., *Anal. Chem.* **2015**, *87*, 11879-11886.
30. Hall, Z.; Politis, A.; Bush, M. F.; Smith, L. J.; Robinson, C. V., *J. Am. Chem. Soc.* **2012**, *134*, 3429-3438.
31. Zhang, J.; Loo, R. R. O.; Loo, J. A., *Int. J. Mass Spectrom.* **2015**, *377*, 546-556.
32. Teo, C. A.; Donald, W. A., *Anal. Chem.* **2014**, *86*, 4455-4462.
33. Shelimov, K. B.; Jarrold, M. F., *J. Am. Chem. Soc.* **1997**, *119*, 2987-2994.
34. Laszlo, K. J.; Munger, E. B.; Bush, M. F., *J. Am. Chem. Soc.* **2016**, *138*, 9581-9588.
35. Florez, A. I. G.; Mucha, E.; Ahn, D. S.; Gewinner, S.; Schollkopf, W.; Pagel, K.; von Helden, G., *Angew. Chem.-Int. Edit.* **2016**, *55*, 3295-3299.
36. Jhingree, J. R.; Bellina, B.; Pacholarz, K. J.; Barran, P. E., *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 1450-1461.
37. Lermyte, F.; Williams, J. P.; Brown, J. M.; Martin, E. M.; Sobott, F., *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1068-1076.
38. Fisher, C. M.; Kharlamova, A.; McLuckey, S. A., *Anal. Chem.* **2014**, *86*, 4581-4588.
39. Pagel, K.; Hyung, S.-J.; Ruotolo, B. T.; Robinson, C. V., *Anal. Chem.* **2010**, *82*, 5363-5372.
40. Dobo, A.; Kaltashov, I. A., *Anal. Chem.* **2001**, *73*, 4763-4773.
41. Natalello, A.; Santambrogio, C.; Grandori, R., *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 21-28.
42. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., *Anal. Chem.* **2013**, *85*, 2-9.

43. Mehmood, S.; Allison, T. M.; Robinson, C. V., *Annu. Rev. Phys. Chem.* **2015**, *66*, 453-474.
44. Donor, M. T.; Ewing, S. A.; Zenaidee, M. A.; Donald, W. A.; Prell, J. S., *Anal. Chem.* **2017**, *89*, 5107-5114.
45. Lomeli, S. H.; Peng, I. X.; Yin, S.; Ogorzalek Loo, R. R.; Loo, J. A., *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 127-131.
46. Iavarone, A. T.; Williams, E. R., *J. Am. Chem. Soc.* **2003**, *125*, 2319-2327.
47. Sterling, H. J.; Daly, M. P.; Feld, G. K.; Thoren, K. L.; Kintzer, A. F.; Krantz, B. A.; Williams, E. R., *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1762-1774.
48. Ogorzalek Loo, R. R.; Lakshmanan, R.; Loo, J. A., *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1675-1693.
49. Hogan, C. J.; Ogorzalek Loo, R. R.; Loo, J. A.; de la Mora, J. F., *Phys. Chem. Chem. Phys.* **2010**, *12*, 13476-13483.
50. Chingin, K.; Xu, N.; Chen, H., *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 928-934.
51. Zhou, M.; Dagan, S.; Wysocki, V. H., *Analyst* **2013**, *138*, 1353-1362.
52. Ferguson, C. N.; Benchaar, S. A.; Miao, Z. X.; Loo, J. A.; Chen, H., *Anal. Chem.* **2011**, *83*, 6468-6473.
53. Wang, H.; Yong, G.; Brown, S. L.; Lee, H. E.; Zenaidee, M. A.; Supuran, C. T.; Donald, W. A., *Anal. Chim. Acta* **2018**, *1003*, 1-9.
54. Douglass, K. A.; Venter, A. R., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 489-497.
55. Lomeli, S. H.; Yin, S.; Loo, R. R. O.; Loo, J. A., *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 593-596.
56. Samalikova, M.; Grandori, R., *J. Mass Spectrom.* **2005**, *40*, 503-510.
57. Yao, Y.; Richards, M. R.; Kitova, E. N.; Klassen, J. S., *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 498-506.
58. Metwally, H.; McAllister, R. G.; Popa, V.; Konermann, L., *Anal. Chem.* **2016**, *88*, 5345-5354.
59. Hall, Z.; Robinson, C. V., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1161-1168.
60. Sterling, H. J.; Kintzer, A. F.; Feld, G. K.; Cassou, C. A.; Krantz, B. A.; Williams, E. R., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 191-200.
61. Burgess, J., *Metal Ions in Solution*. Ellis Horwood: New York, **1978**.

62. Pedersen, C. J., *J. Am. Chem. Soc.* **1967**, *89*, 7017.
63. Landini, D.; Montanar, F.; Pirisi, F. M., *J. Chem. Soc.-Chem. Commun.* **1974**, , 879-880.
64. Armentrout, P. B., *Int. J. Mass Spectrom.* **1999**, *193*, 227-240.
65. Cram, D. J.; Cram, J. M., *Accounts Chem. Res.* **1978**, *11*, 8-14.
66. Hurtado, P.; Gamez, F.; Hamad, S.; Martinez-Haya, B.; Steill, J. D.; Oomens, J., *J. Phys. Chem. A* **2011**, *115*, 7275-7282.
67. Igumnov, S. N.; Mamontov, M. N.; Uspenskaya, I. A., *J. Chem. Eng. Data* **2012**, *57*, 456-461.
68. Ly, T.; Julian, R. R., *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1209-1215.
69. Bonner, J. G.; Hendricks, N. G.; Julian, R. R., *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1661-1669.
70. Warnke, S.; von Helden, G.; Pagel, K., *J. Am. Chem. Soc.* **2013**, *135*, 1177-1180.
71. Chen, Y.; Rodgers, M. T., *J. Am. Chem. Soc.* **2012**, *134*, 5863-5875.
72. Stedwell, C. N.; Galindo, J. F.; Gulyuz, K.; Roitberg, A. E.; Polfer, N. C., *J. Phys. Chem. A* **2013**, *117*, 1181-1188.
73. Sun, Y.; Vahidi, S.; Sowole, M. A.; Konermann, L., *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 31-40.
74. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., *SoftwareX* **2015**, *1-2*, 19-25.
75. Huang, J.; MacKerell, A. D., *J. Comput. Chem.* **2013**, *34*, 2135-2145.
76. Abascal, J. L. F.; Vega, C., *J. Chem. Phys.* **2005**, *123*, 234505.
77. Vega, C.; de Miguel, E., *J. Chem. Phys.* **2007**, *126*, 154707.
78. Fischer, N. M.; van Maaren, P. J.; Ditz, J. C.; Yildirim, A.; van der Spoel, D., *J. Chem. Theor. Comp.* **2015**, *11*, 2938-2944.
79. Sterling, T.; Irwin, J. J., *J. Chem. Inf. Model.* **2015**, *55*, 2324-2337.
80. Chen, M.; Ko, H. Y.; Remsing, R. C.; Andrade, M. F. C.; Santra, B.; Sun, Z. R.; Selloni, A.; Car, R.; Klein, M. L.; Perdew, J. P.; Wu, X. F., *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 10846-10851.
81. Nam, K.; Pu, J. Z.; Karplus, M., *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* , 17851-17856.
82. Lemkul, J. A.; Bevan, D. R., *J. Phys. Chem. B* **2010**, *114*, 1652-1660.
83. Han, W.; Schulten, K., *J. Am. Chem. Soc.* **2014**, *136*, 12450-12460.

84. Zenaidee, M. A.; Leeming, M. G.; Zhang, F. T.; Funston, T. T.; Donald, W. A., *Angew. Chem.-Int. Edit.* **2017**, *56*, 8522-8526.
85. Mahler, J.; Persson, I., *Inorg. Chem.* **2012**, *51*, 425-438.
86. Li, Z. X.; Lu, J. R.; Styrkas, D. A.; Thomas, R. K.; Rennie, A. R.; Penfold, J., *Mol. Phys.* **1993**, *80*, 925-939.
87. Tarek, M.; Tobias, D. J.; Klein, M. L., *J. Chem. Soc.-Faraday Trans.* **1996**, *92*, 559-563.
88. Raina, G.; Kulkarni, G. U.; Rao, C. N. R., *J. Phys. Chem. A* **2001**, *105*, 10204-10207.
89. Dixit, S.; Crain, J.; Poon, W. C. K.; Finney, J. L.; Soper, A. K., *Nature* **2002**, *416*, 829-832.
90. Brini, E.; Fennell, C. J.; Fernandez-Serra, M.; Hribar-Lee, B.; Luksic, M.; Dill, K. A., *Chem. Rev.* **2017**, *117*, 12385-12414.
91. Nshanian, M.; Lakshmanan, R.; Chen, H.; Ogorzalek Loo, R. R.; Loo, J. A., *Int. J. Mass Spectrom.* **2018**, *427*, 157-164.
92. Hamdy, O. M.; Julian, R. R., *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1-6.
93. Zhou, H. X.; Rivas, G. N.; Minton, A. P., *Annu. Rev. Biophys.* **2008**, *37*, 375-397.
94. Hedges, J. B.; Vahidi, S.; Yue, X.; Konermann, L., *Anal. Chem.* **2013**, *85*, 6469-6476.
95. Boas, U.; Heegaard, P. M. H., *Chem. Soc. Rev.* **2004**, *33*, 43-63.
96. Mintzer, M. A.; Grinstaff, M. W., *Chem. Soc. Rev.* **2011**, *40*, 173-190.
97. Suzumura, A.; Paul, D.; Sugimoto, H.; Shinoda, S.; Julian, R. R.; Beauchamp, J. L.; Teraoka, J.; Tsukube, H., *Inorganic Chemistry* **2005**, *44*, 904-910.
98. Lee, C. C.; Maestre-Reyna, M.; Hsu, K. C.; Wang, H. C.; Liu, C. I.; Jeng, W. Y.; Lin, L. L.; Wood, R.; Chou, C. C.; Yang, J. M.; Wang, A. H. J., *Angew. Chem.-Int. Edit.* **2014**, *53*, 13054-13058.

## Chapter 5. Chain Ejection Model for Electrospray Ionization of Unfolded Proteins: Evidence from Atomistic Simulations and Ion Mobility Spectrometry

### 5.1. Introduction

Electrospray ionization (ESI) has revolutionized mass spectrometry (MS) by allowing the facile transfer of proteins and other analytes from solution into the gas phase.<sup>1</sup> During ESI charged droplets emanate from a high voltage capillary. Solvent evaporation and jet fission decrease the droplet size to the nanometer range, while maintaining a charge close to the Rayleigh limit.<sup>2-7</sup> The mechanisms whereby analyte ions emerge from ESI nanodroplets remain controversial.<sup>4, 8-18</sup> Recent progress in this area has been fueled by molecular dynamics (MD) simulations. MD studies revealed that small ions such as Na<sup>+</sup> undergo field emission from the droplet surface,<sup>15, 19-21</sup> consistent with the ion evaporation model (IEM).<sup>8, 22</sup> Simulations on peptides,<sup>23</sup> globular proteins,<sup>24</sup> and nucleic acid duplexes<sup>25</sup> indicated that these larger species are liberated by droplet evaporation to dryness, as envisioned by the charged residue model (CRM).<sup>3-4</sup>

“Native” protein ESI experiments aim to preserve solution-like structures and interactions in the gas phase by employing non-denaturing aqueous solutions and gentle ion sampling conditions.<sup>26-29</sup> Structural retention under these conditions is promoted by the low CRM charge states of protein ions, which are close to the Rayleigh charge of protein-sized water droplets.<sup>3-4</sup>

Electrosprayed protein ions can be further interrogated by ion mobility spectrometry (IMS)<sup>30-35</sup> and by various activation methods.<sup>36-38</sup> Collision-induced dissociation (CID) of multisubunit complexes usually causes ejection of one highly charged chain.<sup>39</sup> This behavior has been attributed to gradual unfolding of one subunit, H<sup>+</sup> migration onto the unraveling chain, and subsequent separation of this chain from the complex (Figure 5.1a).<sup>38, 40-41</sup> This CID model is consistent with IMS data,<sup>32, 37</sup> the mobile nature of H<sup>+</sup> in gaseous proteins,<sup>42-45</sup> and the fact that H<sup>+</sup> migration onto the unraveling chain will minimize electrostatic repulsion.<sup>46-47</sup>

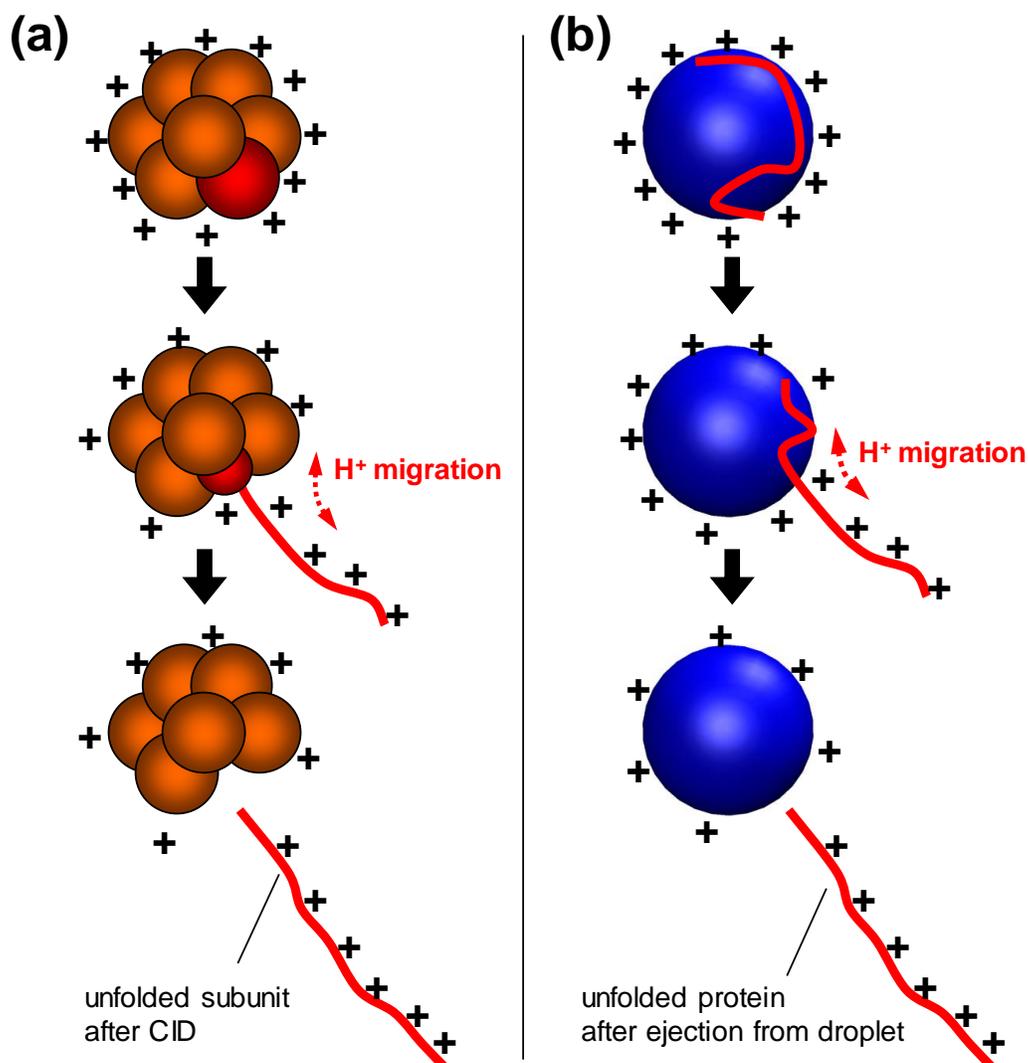
While the CRM is widely accepted for globular proteins in native ESI,<sup>3-4, 11, 24, 48</sup> the behavior of unfolded proteins is more controversial.<sup>12, 48-49</sup>  $[M + zH]^{z+}$  ions formed from unfolded proteins exhibit wide charge state distributions centered at much higher  $z$  values than in native ESI. This effect is encountered after unfolding by acid,<sup>50-51</sup> base,<sup>52</sup> heat,<sup>53</sup> disulfide cleavage,<sup>54</sup> mutations,<sup>55</sup> and cofactor removal.<sup>51</sup> The high charge states of unfolded proteins can boost mass analyzer performance<sup>56-57</sup> and enhance top-down fragmentation.<sup>58-59</sup>

Various attempts have been made to explain the dramatic shift to higher charge states seen for unfolded proteins. Early work proposed that ESI charge states reflect the titration behavior in solution,<sup>60</sup> but subsequent studies showed this not to be the case.<sup>52, 61-62</sup> Other ideas focused on the accessibility of titratable sites,<sup>50, 54, 63</sup> but even in native proteins most titratable sites are accessible at the surface.<sup>64</sup> It has also been proposed that  $\text{COO}^-$  groups may neutralize positive sites in folded gaseous proteins.<sup>12</sup> Although such zwitterionic contacts are well documented,<sup>65</sup> it is unclear if they can account for conformation-dependent charge state changes.<sup>48</sup> Other studies pointed out that the gas-phase basicity of biomolecular ions could result in  $\text{H}^+$  transfer from the solvent vapor.<sup>12, 66</sup> While all these ideas are interesting, they do not directly address the mechanism whereby unfolded proteins emerge from ESI nanodroplets.

Unfolded proteins in solution adopt disordered conformations, similar to certain synthetic polymers such as polyethylene glycol (PEG).<sup>67</sup> MD simulations of PEG-containing aqueous ESI droplets by Consta et al. revealed that PEG binds  $\text{Na}^+$  from the solvent, followed by polymer extrusion from the droplet surface.<sup>68</sup> Our laboratory proposed that the ESI process for unfolded proteins follows similar avenues.<sup>15</sup> According to this “chain ejection model” (CEM, Figure 5.1b), the protein is driven to the droplet surface by electrostatic and hydrophobic factors. The protein then undergoes gradual ejection via “tadpole-like” structures where the droplet carries an extended protein tail. This CEM scenario<sup>15</sup> bears close parallels to the CID process of multi-subunit systems (Figure 5.1a). Specifically, a central aspect of the protein CEM is the migration of mobile  $\text{H}^+$  between the droplet and the protruding polypeptide tail. This  $\text{H}^+$  transfer causes the unfolded protein to depart as a highly charged ion, analogous to  $\text{H}^+$  transfer during CID which causes the departing subunit to be highly charged.<sup>15</sup>  $\text{H}^+$  migration is absent for sodiated PEG,<sup>68</sup> i.e., the CEM scenarios of unfolded proteins and PEG chains are not equivalent.

Several studies have endorsed the CEM.<sup>15, 26, 69-70</sup> However, it is unsettling that protein CEM processes have never been verified by atomistic MD simulations. Instead, the idea relies on simple coarse-grained models, Monte-Carlo methods,<sup>15</sup> and salt adduction studies.<sup>71</sup> Atomistic ESI simulations on unfolded proteins face two challenges. (1) Mobile H<sup>+</sup> are difficult to treat computationally. MD methods are available for H<sup>+</sup> transfer in the gas phase,<sup>46-47</sup> but H<sup>+</sup> migration in water (and between water and protein, Figure 5.1b) requires *ab initio* tools. The computational cost of these tools makes them unsuitable for ESI droplets.<sup>72-74</sup> (2) ESI droplets have to accommodate the analyte at the onset of the simulation. This is not a problem for native proteins, where a few thousand water molecules are sufficient for building a droplet that completely engulfs the analyte.<sup>24-25</sup> In contrast, the increased dimensions of unfolded proteins<sup>67</sup> require larger droplets, driving up computational cost which scales as  $N^2$  with the number of atoms.<sup>75</sup>

The current work scrutinizes the viability of the CEM by conducting atomistic MD simulations on ESI droplets containing unfolded proteins. Myoglobin has been used in numerous earlier ESI mechanistic investigations,<sup>24, 30, 50-51, 63, 69, 71</sup> and therefore it was chosen as model protein for our work as well. Difficulties associated with the treatment of mobile H<sup>+</sup> were circumvented by focusing on carefully selected pH environments. Graphics processing unit (GPU)-accelerated algorithms<sup>76</sup> allowed us to overcome challenges related to droplet size. MD simulations were complemented by ESI-MS/IMS experiments. Our results support the view that the ESI process of unfolded proteins in aqueous solution proceeds via the CEM.



**Figure 5.1.** Cartoon depiction of two analogous gas phase processes. (a) CID of a noncovalent protein complex. Subunits are depicted as spheres, excess H<sup>+</sup> are represented by “+” signs. One subunit (red) undergoes unfolding, and the protruding tail accumulates charge due to H<sup>+</sup> migration from the residual complex. The subunit leaves as a highly charged unfolded ion. (b) Proposed chain ejection model (CEM) for the release of an unfolded protein from an ESI droplet. As the protein gets gradually ejected, the protruding tail undergoes charge equilibration with the droplet via H<sup>+</sup> migration. The protein leaves as a highly charged unfolded ion. Modified from ref.<sup>15</sup>

## 5.2. Materials and Methods

### 5.2.1. Protein, Reagents and Mass Spectrometry

Apo-myoglobin (aMb, Sigma, St. Louis, MO) was prepared by butanone extraction. The protein was dialyzed against 10 mM neutral aqueous ammonium acetate, followed by acidification with formic acid to pH 4 or pH 2. ESI-MS and IMS data were acquired on a Synapt G2 instrument (Waters, Milford, MA). Aqueous aMb (5  $\mu$ M) was infused at 5  $\mu$ L min<sup>-1</sup> using a standard Z-spray ion source at +2.8 kV. The source and desolvation temperatures were kept low (25 °C and 40 °C) and the cone was set to 5 V to minimize the in-source activation. IMS calibration using a set of reference proteins yielded effective He collision cross sections ( $\Omega$ ).<sup>77</sup>

### 5.2.2. MD Simulations

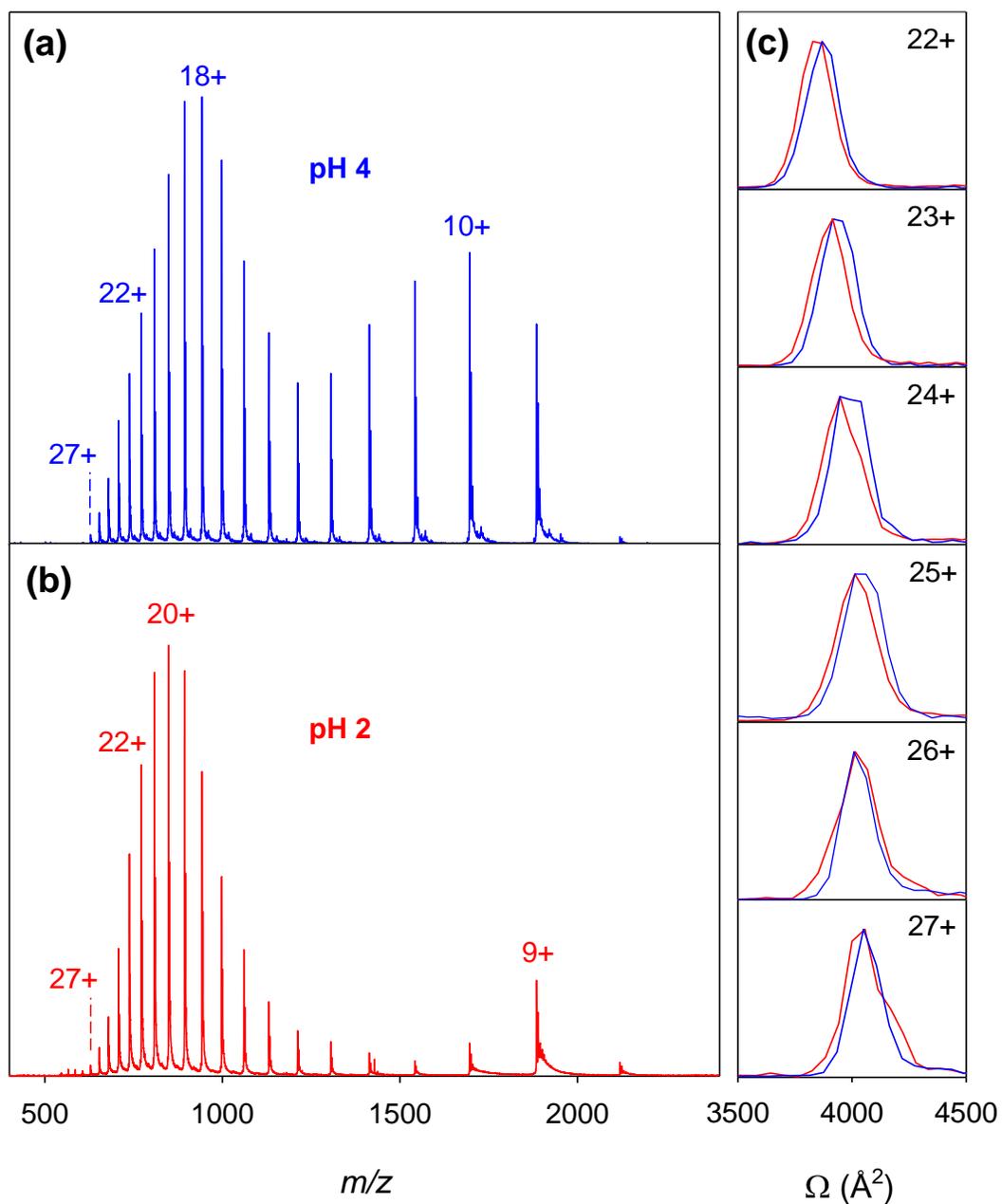
MD simulations were carried out using Gromacs 2016 with GPU acceleration,<sup>76</sup> the Charmm36 force field,<sup>78</sup> and TIP4P/2005 water<sup>79</sup> in trajectory stitching mode.<sup>24</sup> A temperature of 370 K was chosen to mimic the presence of heating elements in typical ESI sources. Unfolded aMb starting conformations were produced by heating aMb (1WLA without heme) from 320 K to 450 K in vacuum using canonical charge states over 20 ns. Spherical water droplets with 5.5 nm radius (~22,500 water molecules) were built around the protein, and the aMb charge was set to 22+, 27+, or 33+. All runs started with a droplet charge of 47+ which corresponds to the Rayleigh charge  $z_R$  of a 5.5 nm aqueous droplet,<sup>3-4, 24</sup> calculated as  $z_R = 8\pi/e \times (\epsilon_0 \gamma r^3)^{1/2}$ . To attain this regime, charges contributed by aMb were supplemented by Na<sup>+</sup> ions in random positions. All runs were repeated three to five times with different initial aMb structures, Na<sup>+</sup> positions, and starting velocities. The simulation time window was 75 ns. After release into the gas phase the desolvated protein was allowed to run for an additional 500 ns at 320 K; this lower temperature was chosen to reflect gentle ion sampling conditions.<sup>26-28, 80</sup> He collision cross sections were calculated using the trajectory method in Collidoscope.<sup>81</sup> These  $\Omega$  values were determined by extracting MD structures from the 500 ns trajectories in 100 ns intervals.

## 5.3. Results and Discussion

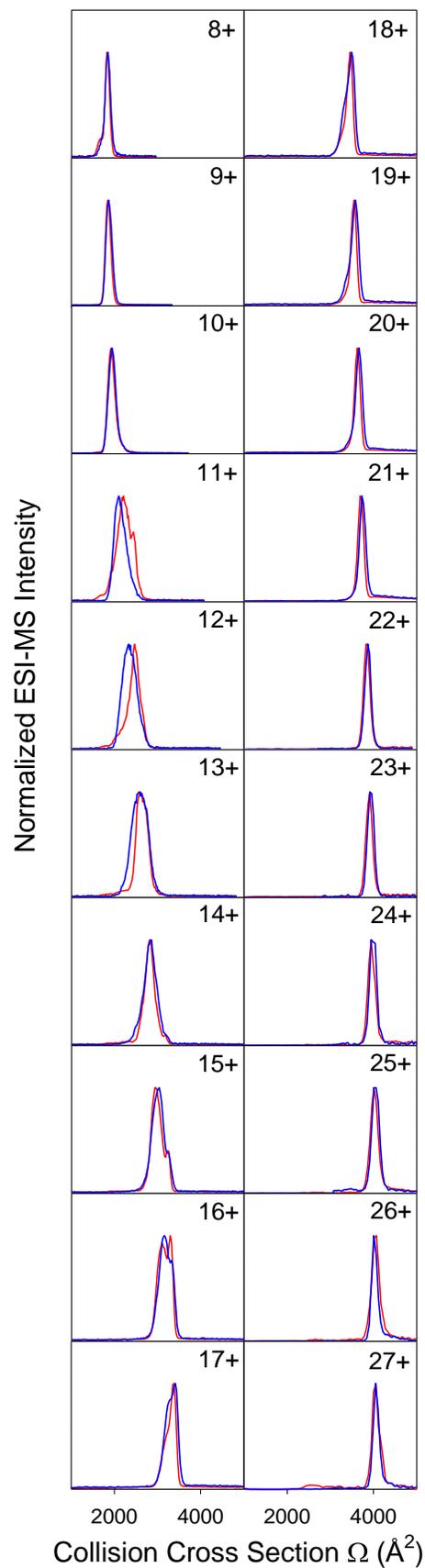
### 5.3.1. ESI-MS and IMS Experiments

ESI mass spectra of aMb acquired at pH 4 showed a bimodal charge state distribution peaking at 9+ and 18+ (Figure 5.2a). The protein is known to be structurally heterogeneous in solution at pH 4, comprising compact conformers and disordered species.<sup>51, 82</sup> This heterogeneity is reflected in the spectrum of Figure 5.2a, where compact conformers gave rise to lower charge states (around 9+), while the more unfolded chains formed charge states around 18+.<sup>51</sup> Acidification to pH 2 caused further unfolding,<sup>82</sup> consistent with a shift to higher ESI charge states (around 20+, Figure 5.2b).<sup>51</sup> The highest detectable charge state was 27+.

IMS data for charge states 22+ to 27+ are shown in Figure 5.2c (the complete dataset is shown in Figure 5.3). 22+ ions had collision cross sections of  $(3870 \pm 30) \text{ \AA}^2$ , whereas higher charge states showed larger  $\Omega$  values, e.g.,  $(4190 \pm 100) \text{ \AA}^2$  for 27+. This trend reflects the internal Coulomb repulsion experienced by the gaseous ions.<sup>30-31</sup> For the highly charged ions considered in Figure 5.2c, collision cross sections measured at pH 4 and pH 2 were virtually indistinguishable.



**Figure 5.2.** Experimental ESI mass spectra for aMb recorded in aqueous solution at pH 4 (a) and pH 2 (b). Selected charge states are indicated. Panel (c) shows IMS collision cross section ( $\Omega$ ) distributions for highly charged ions acquired at pH 4 (blue) and pH 2 (red).



**Figure 5.3.**  $\Omega$  distributions measured by IMS for the complete set of aMb charge states at pH 4 (blue) and pH 2 (red).

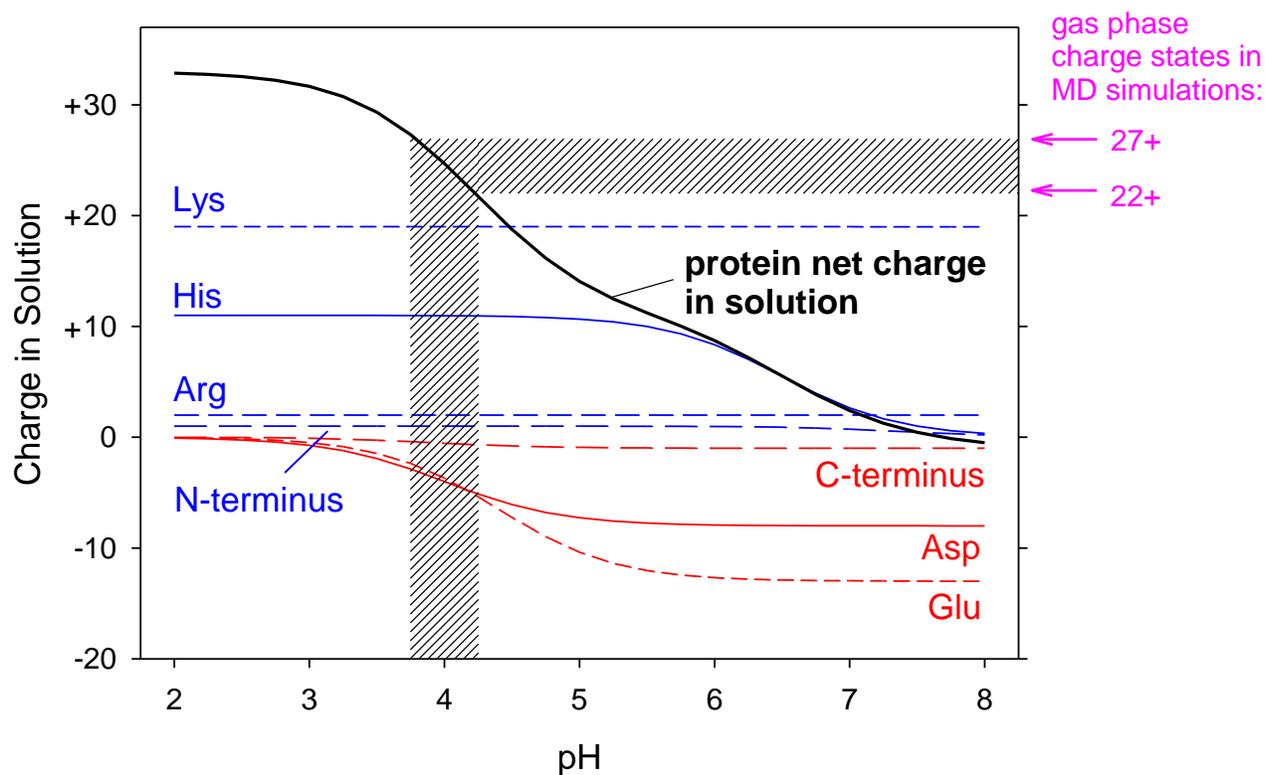
### 5.3.2. ESI Modeling Strategy

We already noted how difficult it is to model  $H^+$  migration in a droplet that carries a polypeptide tail (as envisioned within the CEM, Figure 5.1b).<sup>15</sup> Such  $H^+$  transfer can cause the gas phase protein charge to be very different from that in solution.<sup>52, 61-62</sup> However, the magnitude of this disparity depends on the conditions. The solution charge is dictated by the  $pK_a$  values of titratable sites and by pH.<sup>64</sup> Hence, a judicious choice of pH can result in a scenario where the protein solution charge resembles the charge state of the resulting gaseous ions. Figure 5.4 shows that at pH  $\sim 4$  the aMb solution charge is around 22+ (pH 4.25) to 27+ (pH 3.75). Gas phase ions with charge states 22+ to 27+ can be produced by electrospraying aMb at pH 4 (Figure 5.2a). CEM-related  $H^+$  migration between droplet and protein will be minimal under these conditions because solution charge  $\approx$  gas phase charge. In this specific case it is reasonable to model the ESI process *without* mobile  $H^+$ , thereby greatly simplifying the computational approach.

Prior to applying the strategy outlined above, one has to consider that the droplet pH will likely deviate from that of the bulk solution. Solvent evaporation tends to increase the  $H^+$  concentration. On the other hand, formic acid in the in the aMb solutions used here has buffering capacity around pH 4 ( $pK_a = 3.75$ ). Therefore, the assumptions (1) solution pH  $\approx$  droplet pH and (2) [protein charge in solution]  $\approx$  [protein charge after ESI] provide a reasonable foundation for modeling the ESI behavior of unfolded aMb in the 22+ to 27+ charge states at pH 4.

Another aspect that has to be addressed is the intramolecular  $H^+$  distribution. Most charge states can be implemented via a multitude of protonation patterns.<sup>12</sup> To test whether the aMb behavior is sensitive to this aspect we performed simulations on various protonation patterns (Figure 5.5). Three 22+ patterns were tested, referred to as 22+[A] (all sites protonated, except His), 22+[B] (N-terminus and all Arg/His/Lys protonated, all Asp and some Glu deprotonated), 22+[C] (all sites protonated, except for some Lys). 27+ simulations were conducted where all sites were protonated, except for some Lys. Simulations on fully protonated aMb (33+) were included as well, although this value is beyond the range observed in our experiments.

To ensure that unfolded aMb chains were fully contained within the initial ESI droplet we employed a droplet radius of 5.5 nm. To the best of our knowledge, these are the largest protein-containing ESI droplets modeled to date. GPU-acceleration<sup>76</sup> helped overcome the computational cost associated with this system size. TIP4P/2005 water<sup>79</sup> was chosen because it reproduces the water surface tension, thus ensuring that the simulations yielded realistic data.<sup>24</sup> The droplets were initially charged to the Rayleigh limit<sup>2-4</sup> by supplementing the aMb charge with Na<sup>+</sup> ions, keeping in mind that Na<sup>+</sup> represents a typical ESI charge carrier.<sup>3</sup> Likely, other ions such as NH<sub>4</sub><sup>+</sup> or H<sub>3</sub>O<sup>+</sup> would yield qualitatively similar results to those discussed below for Na<sup>+</sup>.



**Figure 5.4.** Calculated aMb net charge in solution vs. pH (black). Also shown are the contributions of titratable sites (basic = blue; acidic = red), weighted by their abundance in the protein sequence. The number of sites and their  $pK_a$  values<sup>64</sup> are: 1 N-terminus<sup>+</sup> (7.4); 2 Arg<sup>+</sup> (12.0); 19 Lys<sup>+</sup> (10.8); 11 His<sup>+</sup> (6.5); 8 Asp (4.0), 13 Glu (4.4), 1 C-terminus (3.9). Hatch marks highlight the range around pH 4, where the solution charge is between 22+ (pH 4.25) and 27+ (pH 3.75). CEM production of gaseous ions in these charge states from pH ~4 droplets will involve minimal H<sup>+</sup> migration because *gas phase charge*  $\approx$  *solution charge*

**22+[A]**

GLSDGEWQQVLNVWGKVEADIAGHGQEVLI<sup>R</sup>LFTGHPETLE<sup>K</sup>FDK<sup>K</sup>HLKTEAEMKASEDLKKHGT  
 VVLTALGGIIL<sup>KKK</sup>GHHEAEL<sup>K</sup>PLAQSHAT<sup>KHK</sup>PIKYLEFISDAI<sup>I</sup>HVLHSHKHPGDFGADAQQGAMTKAL  
 ELFRNDIAAK<sup>Y</sup>KELGFQG

**22+[B]**

GLSDGEWQQVLNVWGKVEAD<sup>I</sup>AGHGQEVLI<sup>R</sup>LFTGHPETLE<sup>K</sup>FDK<sup>K</sup>HLKTEAEMKASEDLKKHGT  
 VVLTALGGIIL<sup>KKK</sup>GHHEAEL<sup>K</sup>PLAQSHAT<sup>KHK</sup>PIKYLEFISDAI<sup>I</sup>HVLHSHKHPGDFGADAQQGAMTKAL  
 ELFRNDIAAK<sup>Y</sup>KE<sup>L</sup>GFQG

**22+[C]**

GLSDGEWQQVLNVWGKVEADIAGHGQEVLI<sup>R</sup>LFTGHPETLEKFDK<sup>K</sup>HLKTEAEMKASEDL<sup>K</sup>HGT  
 VVLTALGGIIL<sup>KKK</sup>GHHEAEL<sup>K</sup>PLAQSHAT<sup>KHK</sup>PIKYLEFISDAI<sup>I</sup>HVLHSHKHPGDFGADAQQGAMTKAL  
 ELFRNDIAAK<sup>Y</sup>KELGFQG

**27+**

GLSDGEWQQVLNVWGKVEADIAGHGQEVLI<sup>R</sup>LFTGHPETLEKFDK<sup>K</sup>FKH<sup>L</sup>KTEAEMKASEDL<sup>K</sup>HGT  
 VVLTALGGIIL<sup>KKK</sup>GHHEAEL<sup>K</sup>PLAQSHAT<sup>KHK</sup>PIKYLEFISDAI<sup>I</sup>HVLHSHKHPGDFGADAQQGAMTKAL  
 ELFRNDIAAK<sup>Y</sup>KELGFQG

**33+**

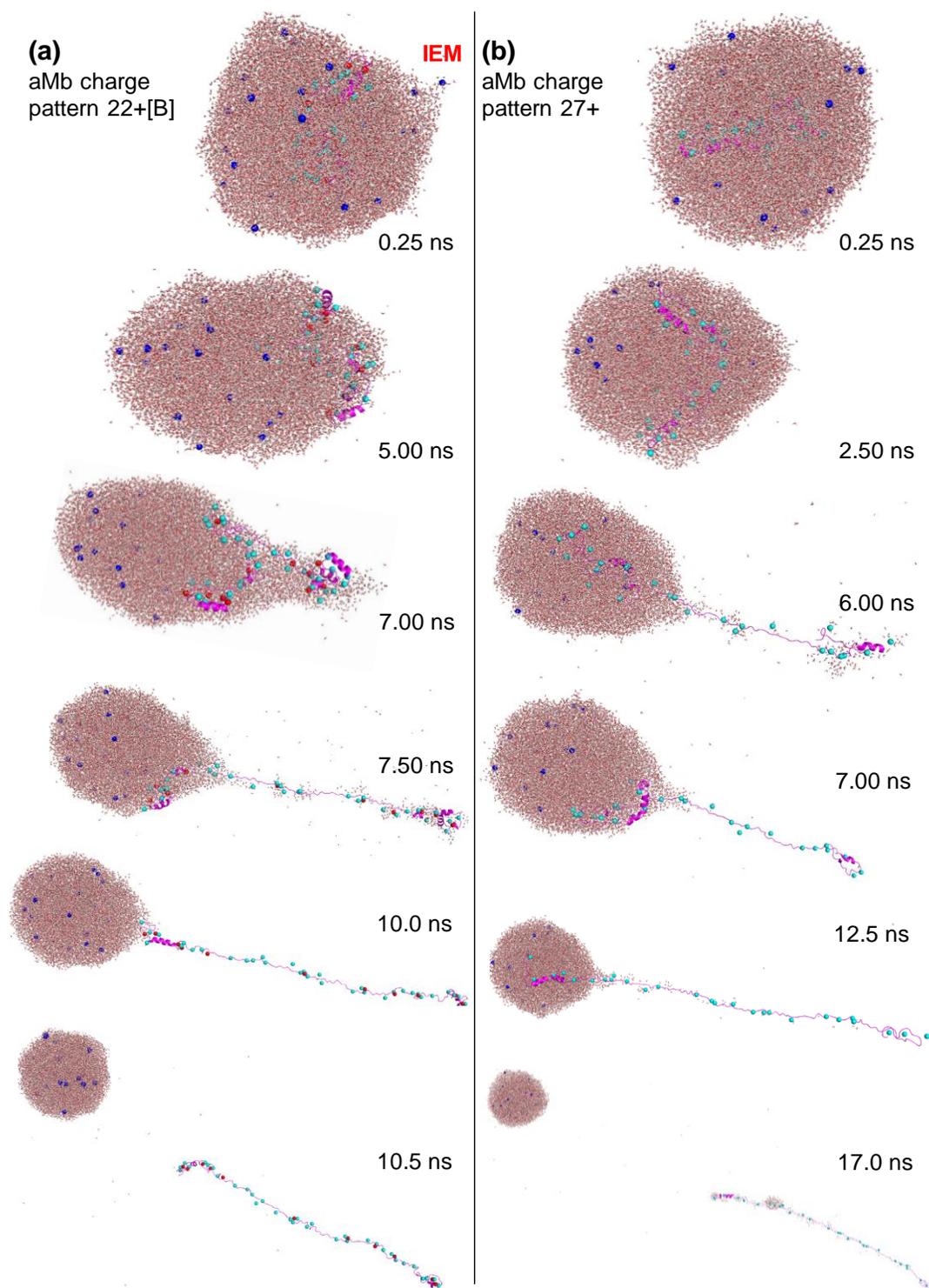
GLSDGEWQQVLNVWGKVEADIAGHGQEVLI<sup>R</sup>LFTGHPETLEKFDK<sup>K</sup>FKH<sup>L</sup>KTEAEMKASEDL<sup>KK</sup>HGT  
 VVLTALGGIIL<sup>KKK</sup>GHHEAEL<sup>K</sup>PLAQSHAT<sup>KHK</sup>PIKYLEFISDAI<sup>I</sup>HVLHSHKHPGDFGADAQQGAMTKAL  
 ELFRNDIAAK<sup>Y</sup>KELGFQG

**Figure 5.5.** aMb charge patterns for MD simulations. Positive sites are indicated in blue, while negative charges are indicated in red. The proton distribution in 22+[A] reflects the view that Arg and Lys are key protonation sites in the gas phase,<sup>49</sup> as well as the most basic residues in solution.<sup>64</sup> 22+[B] allows for negatively charged sites in  $[M + zH]^{z+}$  ions, *i.e.*, zwitterionic and/or salt bridge motifs.<sup>12, 65</sup> This pattern also reflects the partial deprotonation of Asp and Glu in solution at pH 4 (Figure 5.4). 22+[C] was chosen in accordance with the view that protons in the gas phase primarily reside on Arg/His/Lys<sup>46, 66, 83</sup> Some Lys were left unprotonated because Lys has a lower gas phase basicity than His.<sup>84</sup>

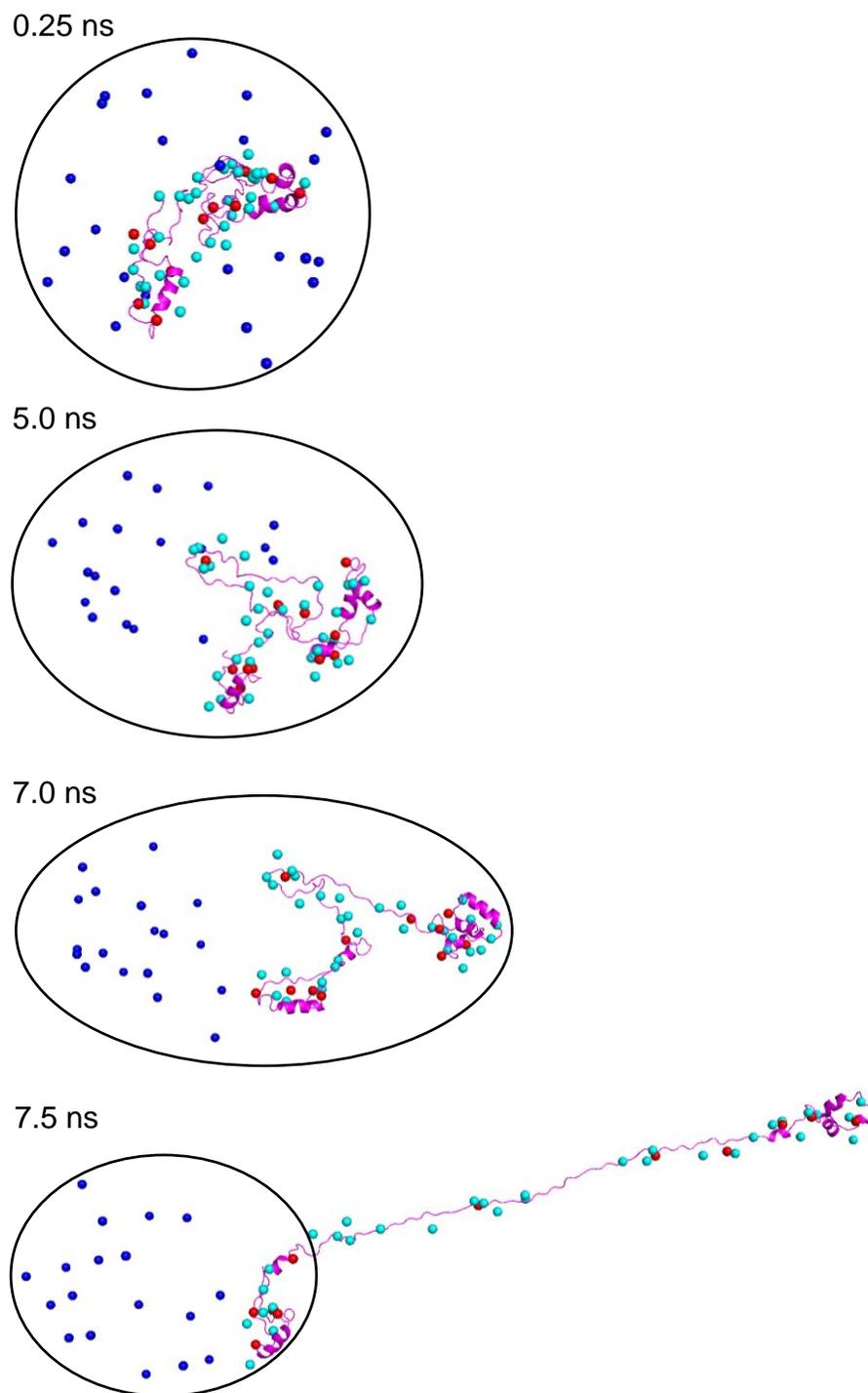
### 5.3.3. MD Simulations Confirm CEM Behavior

Aqueous ESI droplets containing unfolded aMb were subjected to MD simulations. Typical data for protonation patterns 22+[B] and 27+ are shown in Figure 5.6. Within a few ns the protein migrated from the droplet interior close to the surface (Figure 5.6a, 5 ns; Figure 5.6b, 2.5 ns). This was followed by partial aMb eruption as a hydrated bulge (Figure 5.6a, 7 ns). At this particular point  $\text{Na}^+$  and  $\text{aMb}^{22+}$  had partitioned into opposite regions of the droplet, highlighting the electrostatic forces within the system (see also Figure 5.7). Protein expulsion subsequently produced an electrostatically stretched tail that protruded into the vapor phase (Figure 5.6a, 10 ns; Figure 5.6b, 12.5 ns). Further expulsion then caused aMb separation from the droplet. Nascent gaseous proteins retained some water which evaporated within 75 ns. Data very similar to those of Figure 5.6 were also seen for the other protonation patterns (22+[A], 22+[C], 33+, Figure 5.8), confirming that the behavior reported here is robust and reproducible. Figure 5.6 embodies the central result of this work: for the first time atomistic MD simulations confirm the formation of gaseous ions from unfolded proteins via the CEM.

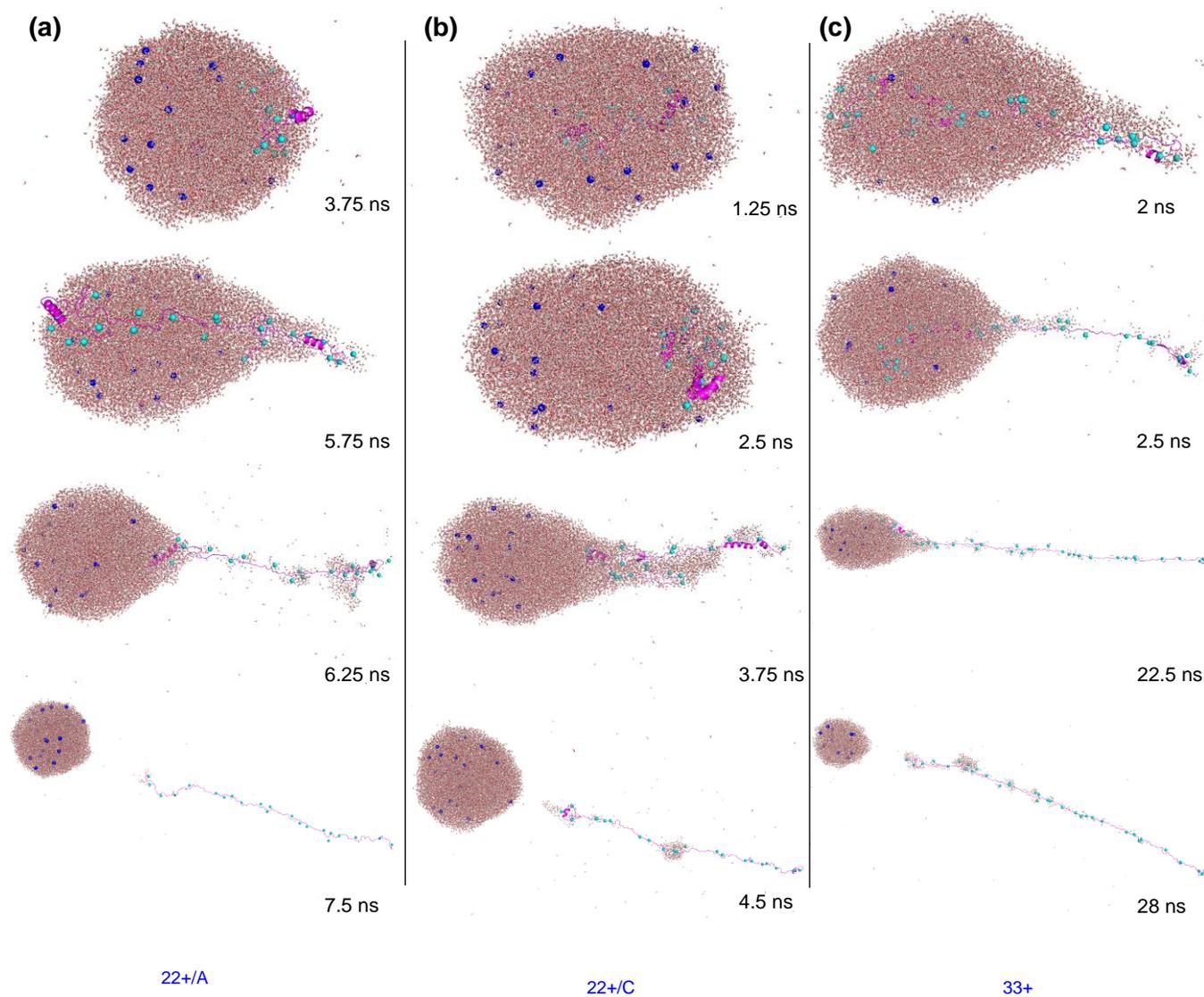
The CEM trajectories showed slight variations. In some instances, aMb ejected with a small droplet attached to one terminus. These small droplets then evaporated without separating from the chain (Figure 5.9). Another variation involved ejection in a hairpin conformation. The hairpins either opened up as they departed from the droplet (22+[A], Figure 5.10a), or they retained looped conformations (22+[B], Figure 5.10b). In one instance aMb ejected without extensive stretching (22+[B], Figure 5.10c). All other (16 out of 17) runs showed the hallmark of the CEM, i.e., aMb ejection via droplets that carried an electrostatically stretched protein tail, consistent with the mechanism of Figure 5.1b.<sup>15</sup> None of the unfolded aMb chains showed CRM behavior, i.e., protein release via droplet evaporation to dryness.



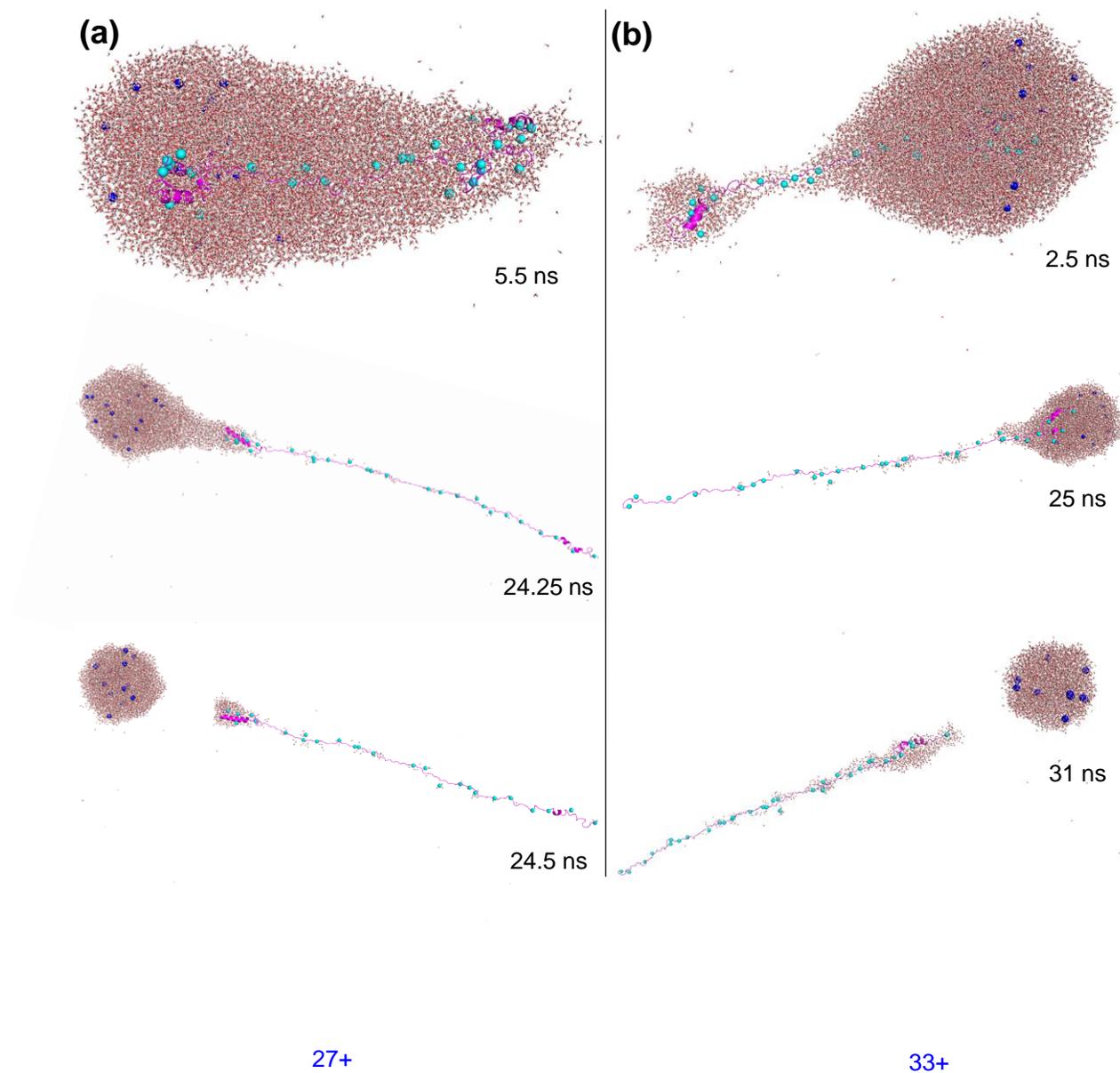
**Figure 5.6.** Typical snapshots from CEM simulation runs on Rayleigh-charged aqueous droplets containing unfolded aMb. The protein net charge was (a) 22+ (pattern 22+[B]) and (b) 27+.  $\text{Na}^+$  is blue, the protein backbone is magenta, positive/negative charges on the protein are highlighted as cyan/red spheres, respectively. Water oxygen is shown in red. The zoom level decreases from top to bottom. “IEM” in (a) highlights the field emission of a  $\text{Na}^+$  ion.



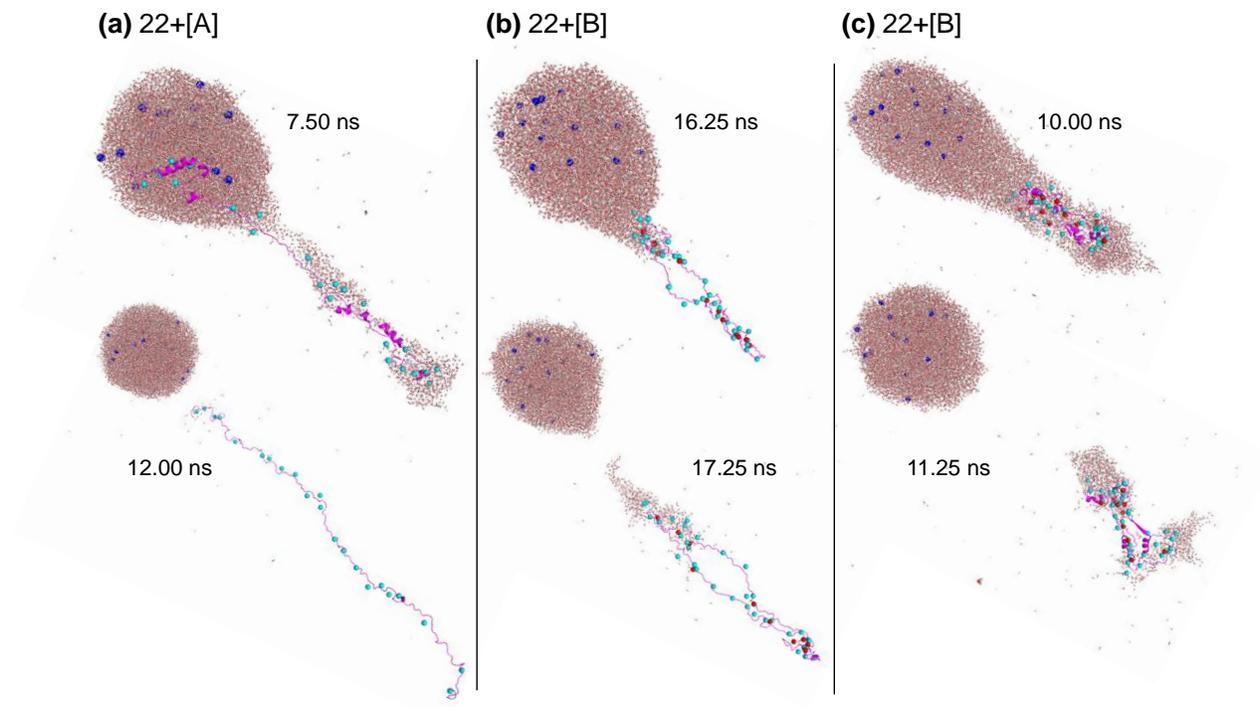
**Figure 5.7.** Snapshots taken from a CEM simulation run on aMb in protonation pattern 22+[B] (same data set as in Figure 5.6). Water was omitted to more clearly reveal the behavior of the protein (magenta) and  $\text{Na}^+$  ions (blue). Positive/negative charges on the protein are highlighted as cyan/red spheres, respectively. Black lines indicate the approximate outline of the droplets.



**Figure 5.8.** CEM ejection aMb from aqueous ESI droplets at the Rayleigh limit. The protonation patterns used for these simulations were (a) 22+[A], (b) 22+[C], and (c) 33+.  $\text{Na}^+$  is indicated in blue, positive/negative charges on the protein are highlighted as cyan/red spheres, respectively. Water oxygen is shown in red.



**Figure 5.9.** CEM ejection of aMb in charge states (a) 27+ and (b) 33+ from the ESI droplet. For these two runs the nascent aMb ions retained a small progeny droplet at one terminus which subsequently evaporated without detaching from the protein.



**Figure 5.10.** Snapshots from simulation runs on three protonation patterns. In contrast to Figure 5.6, these data illustrate less common scenarios, highlighting the variability of the ESI process.

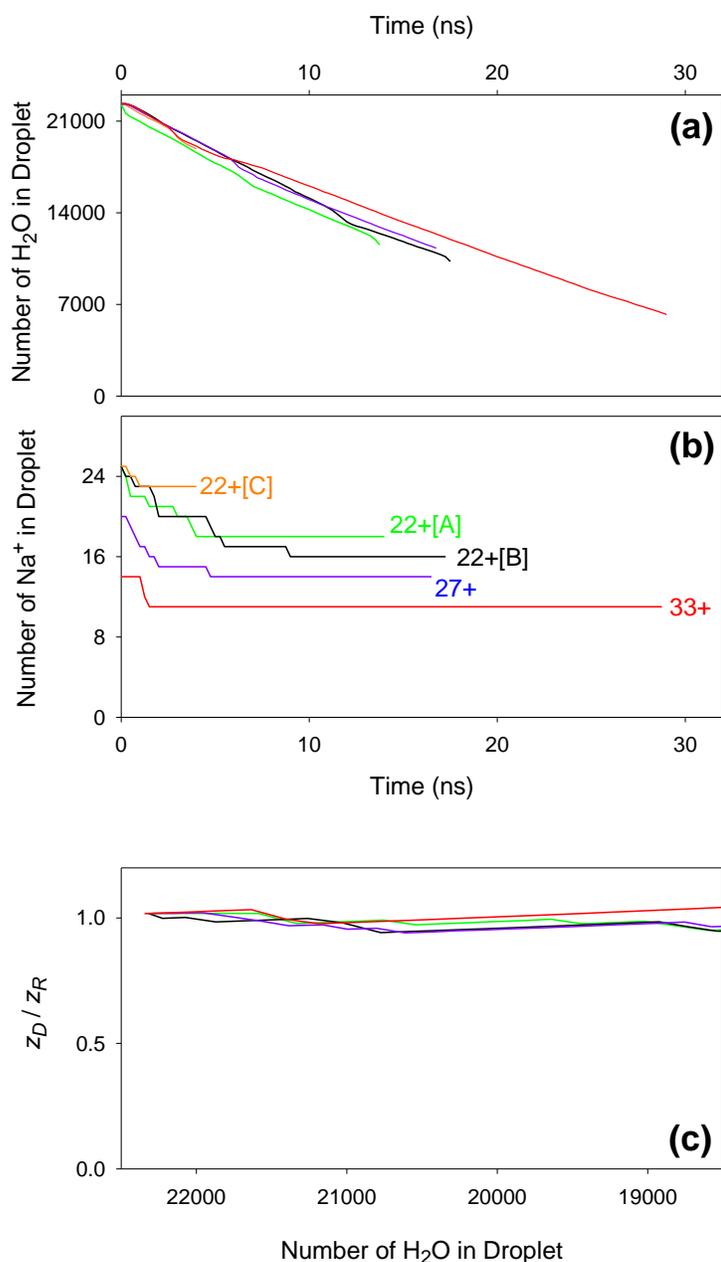
#### 5.3.4. Charge Loss at the Rayleigh Limit

Our simulations started with a droplet charge  $z_D$  close to the Rayleigh limit  $z_R$ .<sup>2-4</sup> Droplet shrinkage due to water evaporation (Figure 5.11a) had a tendency to increase the electrostatic repulsion further, forcing the droplets to shed charge. Three competing charge loss processes were encountered. The first was the IEM ejection of  $\text{Na}^+$  (Figure 5.6a). The second was CEM ejection of the protein. Multiple  $\text{Na}^+$  ejections occurred early during each run, followed by long plateaus where  $\text{Na}^+$  loss had come to a halt (Figure 5.11b). These plateaus represent the regime where formation of a protruding protein tail had decreased the electrostatic repulsion within the droplet to such an extent that  $\text{Na}^+$  ejection became kinetically unfavorable.

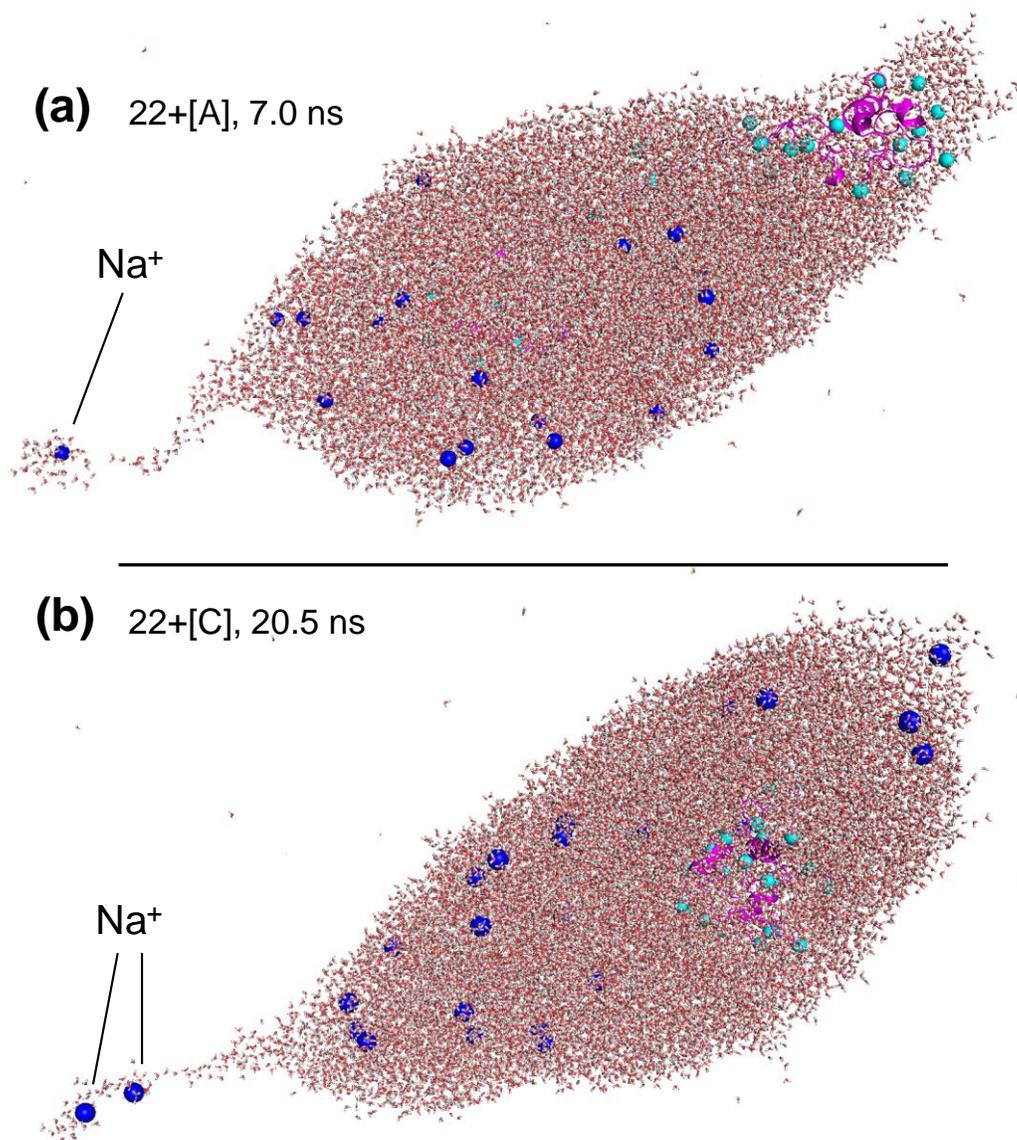
The third charge loss process was the formation of  $\text{Na}^+$ -containing progeny droplets via jet fission, a phenomenon well known from imaging studies on larger ESI droplets.<sup>5-6</sup> Fission events reminiscent of those experimental data,<sup>5-6</sup> involving water filaments, took place in several

MD runs and generated progeny droplets comprising ~50 water molecules and one or two  $\text{Na}^+$  (Figure 5.12). Similar to IEM ejection of  $\text{Na}^+$ , jet fission was limited to time points prior to formation of a protein tail. Plots of  $z_D / z_R$  confirmed that the droplets stayed close to the Rayleigh limit during these early stages (Figure 5.11c). We limited the time frame of Figure 5.11c to the initial ~4 ns during which protein ejection had not started yet in most runs, keeping in mind that the  $z_R$  expression used here<sup>2-4</sup> applies only to spherical systems.

All three of the aforementioned charge loss processes were kinetically viable, but ultimately protein ejection was the main avenue by which the droplets relieved electrostatic stress. In addition to its electrostatic driving force, the CEM behavior is favored by the fact that unfolded proteins possess many solvent-exposed hydrophobic side chains. The tendency of hydrophobic moieties to migrate to the liquid/vapor interface, rather than stay in the droplet interior, is well established.<sup>15, 18</sup> The case is completely different for folded proteins, where most hydrophobic sites are buried, and where water interacts favorably with solvent-exposed hydrophilic/charged residues.<sup>64</sup> The latter conditions cause folded proteins to remain within the ESI droplets until evaporation to dryness, resulting in CRM behavior.<sup>15</sup> Solvent evaporation to dryness causes CRM-produced protein ions to carry adducts arising from nonvolatile solutes, including  $\text{Na}^+$  and other charge carriers.<sup>3, 15</sup> In contrast, the CEM simulations of the current work did not reveal a single instance of  $\text{Na}^+$  adduction to aMb. This behavior is in line with the experimental finding that acid-unfolded proteins are much less prone to adduction than their natively folded counterparts.<sup>71</sup>



**Figure 5.11.** MD simulation data for typical CEM runs on aMb protonation patterns 22+[A] (green), 22+[B] (black), 22+[C] (orange), 27+ (blue), and 33+ (red). (a) Number of water molecules and (b) number of Na<sup>+</sup> in the droplet vs. time. The end of each profile marks the point where the protein chain separates from the droplet. (c) Droplet charge  $z_D$  relative to the Rayleigh charge  $z_R$ , focusing on the initial regime where protein ejection had not started yet.



**Figure 5.12.** Formation of small progeny droplets via jet fission in two ESI simulation runs.

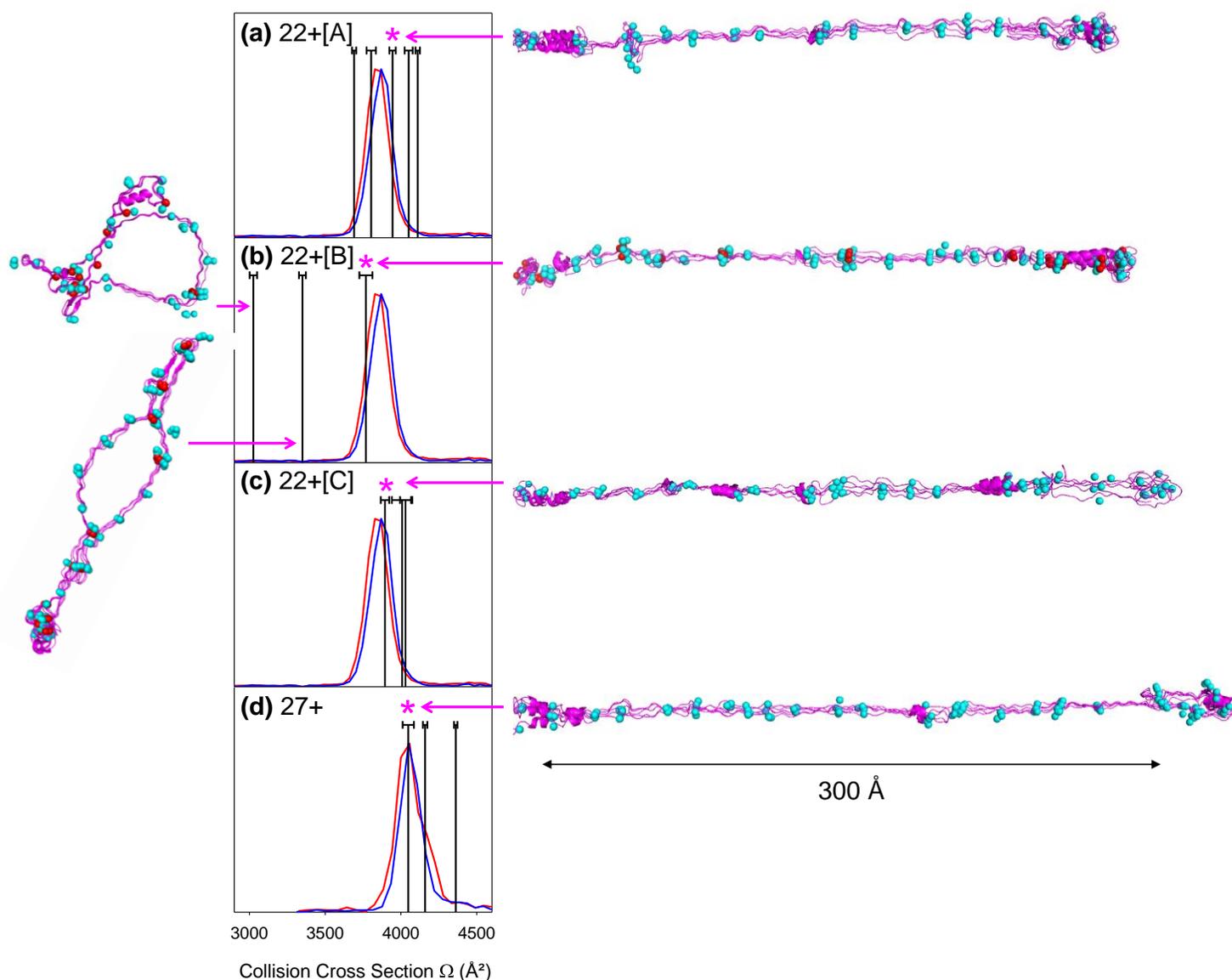
### 5.3.5. Protein Conformations after CEM Ejection

For each ESI simulation the behavior of desolvated aMb after ejection from the droplet was explored in 500 ns MD runs. Representative structures from these vacuum simulations were extracted, and their  $\Omega$  values were compared with experimental IMS data (Figure 5.13). Averaging the results of all aMb 22+ simulations yielded  $\Omega = (3790 \pm 300) \text{ \AA}^2$ , in good agreement with the experimental result of  $(3870 \pm 30) \text{ \AA}^2$ . Simulated 27+ ions had  $\Omega = (4190 \pm 100) \text{ \AA}^2$ , close to the measured value of  $(4050 \pm 30) \text{ \AA}^2$ . The MD structures that most closely matched the experiments are shown along the right-hand side of Figure 5.13. These electrostatically stretched proteins had a stick-like appearance, with local  $\alpha$ -helices, and some coil formation at the termini. Electrostatic repulsion caused 27+ aMb to be slightly longer (overall length  $\sim 350 \text{ \AA}$ ) than the 22+ ions ( $\sim 310 \text{ \AA}$ ).

Overall, the results of MD runs on protonation patterns 22+[A], 22+[C], and 27+ agreed well with the experimental data (Figure 5.13). The zwitterionic pattern 22+[B] deserves a closer look. Two 22+[B] runs produced relatively compact structures that were incompatible with experiments (Figure 5.13b, left). Non-local salt bridges in these runs limited the extent of electrostatic stretching. A third 22+[B] run did yield a stretched structure that matched the measured  $\Omega$  value (Figure 5.13b). Salt bridges in this third run only involved sites in close sequence proximity. Our data thus do not exclude the existence of local zwitterionic motifs in unfolded gaseous proteins, but non-local salt bridges are unlikely. We note that previous evidence for zwitterionic motifs in the gas phase was obtained primarily for tightly folded proteins,<sup>12, 65</sup> rather than unfolded species.

The gas phase conformations generated in our 500 ns vacuum simulations were governed by the morphology of the protein during ejection. Proteins that left the droplet as stretched chains retained extended conformations (Figures 5.3, 5.4a); proteins that ejected in more compact structures (Figure 5.10b, c) gave rise to less extended ions (Figure 5.13b, left). The near-absence of conformational changes in the 500 ns runs suggests that this conformational memory persists for time periods much longer than the simulation window explored here (overlays in Figure 5.13). The agreement between experimental data and MD-derived  $\Omega$  values

in Figure 5.13 thus supports the fidelity of our MD data, and it bolsters the view that the observed ions are CEM products.



**Figure 5.13.** Experimental IMS data measured at pH 4 (blue) and pH 2 (red) for charge states 22+ (a-c) and 27+ (d). Vertical black lines represent average  $\Omega$  values from individual MD runs, standard deviations are indicated as horizontal bars. The four panels represent data for different protonation patterns, as indicated. Shown on the right are MD structures of those trajectories that best matched the experimental data (pink asterisks, overlays of five structures between 100 ns and 500 ns, all at the same zoom level, N-termini pointing to the left). Also included in (b) are MD structures from two runs that yielded more compact structures than those observed experimentally. Positive/negative charges on aMb are shown as cyan/red spheres, respectively.

## 5.4. Conclusions

This work marks the first time that the production of gaseous ions from unfolded proteins via the CEM has been verified in MD simulations using an atomistic force field for protein and solvent. The MD data are supported by IMS experiments, bolstering the view that unfolded proteins are ejected from ESI droplets as electrostatically stretched chains (Figure 5.1b).<sup>15</sup> This CEM behavior can be contrasted to the CRM, which is operational in native ESI where protein are folded.<sup>24</sup>

Computational challenges precluded the inclusion of  $H^+$  migration between droplet and protein (Figure 5.1b) in our simulations. Depending on the titration behavior of the protein,  $H^+$  migration may increase or decrease the protein net charge during ejection. Here we focused on aMb around pH 4, where  $z \approx \text{solution charge}$ , such that  $H^+$  migration could be neglected. How likely is it that more advanced modeling strategies involving mobile  $H^+$  would overthrow the viability of the CEM? Our IMS experiments demonstrated that the gas phase conformations of aMb 22+ to 27+ produced at pH 4 were indistinguishable from those generated at pH 2. This is despite the fact CEM-related  $H^+$  migration at pH 4 is negligible ( $z \approx \text{solution charge}$ ), whereas pH 2 will be associated with significant  $H^+$  transfer ( $z < \text{solution charge}$ ). Hence,  $H^+$  migration does not seem to affect the properties of these ions. In addition, our MD simulations consistently produced CEM behavior for various protonation patterns and charge states, all the way to aMb<sup>33+</sup>. It thus appears that CEM behavior is a robust property of acid-unfolded proteins. It is unlikely that this conclusion would change when applying computational strategies that allow for  $H^+$  migration.

Notwithstanding the aforementioned comments, it is hoped that future studies on the ESI behavior of unfolded proteins will employ more sophisticated models that include  $H^+$  migration. Such strategies will be required for quantitatively explaining the wide range of ESI charge states seen for unfolded proteins (Figure 5.2). We hypothesize<sup>15</sup> that this charge heterogeneity can be attributed to protein ejection from differently sized droplets, in conjunction with the fact that denatured proteins comprise various conformers that may accumulate different numbers of  $H^+$  during ejection.<sup>12, 51</sup> It is possible that the CEM is not the only mechanism that is

operative during ESI of denatured protein. Only the highly charged ions in the spectra were attributed to the CEM, while less abundant ions in low charge states (such as 9+/10+, Figure 5.2b) may represent CRM products. We hope that the current work will stimulate additional studies aimed at deciphering all these intricacies associated with the ESI process, as well as the possible involvement of the CEM in “supercharging” experiments.<sup>69</sup>

## 5.5. References

1. Fenn, J. B., *Angew. Chem. Int. Ed.* **2003**, *42*, 3871-3894.
2. Grimm, R. L.; Beauchamp, J. L., *J. Phys. Chem. A* **2010**, *114*, 1411-1419.
3. Kebarle, P.; Verkerk, U. H., *Mass Spectrom. Rev.* **2009**, *28*, 898-917.
4. de la Mora, F. J., *Anal. Chim. Acta* **2000**, *406*, 93-104.
5. Gomez, A.; Tang, K., *Phys. Fluids* **1994**, *6*, 404-414.
6. Nemes, P.; Marginean, I.; Vertes, A., *Anal. Chem.* **2007**, *79*, 3105-3116.
7. Wang, R.; Zenobi, R., *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 378-385.
8. Iribarne, J. V.; Thomson, B. A., *J. Chem. Phys.* **1976**, *64*, 2287-2294.
9. Cole, R. B., *J. Mass. Spectrom.* **2000**, *35*, 763-772.
10. Wang, G.; Cole, R. B., *Anal. Chim. Acta* **2000**, *406*, 53-65.
11. Iavarone, A. T.; Williams, E. R., *J. Am. Chem. Soc.* **2003**, *125*, 2319-2327.
12. Li, J.; Santambrogio, C.; Brocca, S.; Rossetti, G.; Carloni, P.; Grandori, R., *Mass Spectrom. Rev.* **2016**, *35*, 111-122.
13. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L., *Anal. Chem.* **2009**, *81*, 369-377.
14. Breuker, K.; McLafferty, F. W., *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18145-18152.
15. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., *Anal. Chem.* **2013**, *85*, 2-9.
16. Zhao, F. F.; Matt, S. M.; Bu, J.; Rehrauer, O. G.; Ben-Amotz, D.; McLuckey, S. A., *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 2001-2010.
17. Nshanian, M.; Lakshmanan, R.; Chen, H.; Ogorzalek Loo, R. R.; Loo, J. A., *Int. J. Mass Spectrom.* **2018**, *427*, 157-164.
18. Cech, N. B.; Enke, C. G., *Mass Spectrom. Rev.* **2001**, *20*, 362-387.
19. Higashi, H.; Tokumi, T.; Hogan, C. J.; Suda, H.; Seto, T.; Otani, Y., *Phys. Chem. Chem. Phys.* **2015**, *17*, 15746-15755.
20. Znamenskiy, V.; Marginean, I.; Vertes, A., *J. Phys. Chem. A* **2003**, *107*, 7406-7412.
21. Consta, S.; Mainer, K. R.; Novak, W., *J. Chem. Phys.* **2003**, *119*, 10125-10132.
22. Loscertales, I. G.; de la Mora, J. F., *J. Chem. Phys.* **1995**, *103*, 5041-5060.
23. Kim, D.; Wagner, N.; Wooding, K.; Clemmer, D. E.; Russell, D. H., *J. Am. Chem. Soc.* **2017**, *139*, 2981-2988.

24. McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L., *J. Am. Chem. Soc.* **2015**, *137*, 12667-12676.
25. Porrini, M.; Rosu, F.; Rabin, C.; Darre, L.; Gomez, H.; Orozco, M.; Gabelica, V., *ACS Central Sci.* **2017**, *3*, 454-461.
26. Mehmood, S.; Allison, T. M.; Robinson, C. V., *Annu. Rev. Phys. Chem.* **2015**, *66*, 453-474.
27. Leney, A. C.; Heck, A. J. R., *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 5-13.
28. Dyachenko, A.; Gruber, R.; Shimon, L.; Horovitz, A.; Sharon, M., *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 7235-7239.
29. Gavriilidou, A. F. M.; Holding, F. P.; Mayer, D.; Coyle, J. E.; Veprintse, D. B.; Zenobi, R., *Biochemistry* **2018**, *57*, 1685-1689.
30. Shelimov, K. B.; Jarrold, M. F., *J. Am. Chem. Soc.* **1997**, *119*, 2987-2994.
31. Wyttenbach, T.; Pierson, N. A.; Clemmer, D. E.; Bowers, M. T., *Annu. Rev. Phys. Chem.* **2014**, *65*, 175-196.
32. Ruotolo, B. T.; Hyung, S.-J.; Robinson, P. M.; Giles, K.; Bateman, R. H.; Robinson, C. V., *Angew. Chem. Int. Ed.* **2007**, *46*, 8001-8004.
33. Jurneczko, E.; Barran, P. E., *Analyst* **2011**, *136*, 20-28.
34. Silveira, J. A.; Servage, K. A.; Gamage, C. M.; Russell, D. H., *J. Phys. Chem. A* **2013**, *117*, 953-961.
35. Warnke, S.; von Helden, G.; Pagel, K., *J. Am. Chem. Soc.* **2013**, *135*, 1177-1180.
36. Cammarata, M. B.; Thyer, R.; Rosenberg, J.; Ellington, A.; Brodbelt, J. S., *J. Am. Chem. Soc.* **2015**, *137*, 9128-9135.
37. Quintyn, R. S.; Zhou, M.; Yan, J.; Wysocki, V. H., *Anal. Chem.* **2015**, *87*, 11879-11886.
38. Felitsyn, N.; Kitova, E. N.; Klassen, J. S., *Anal. Chem.* **2001**, *73*, 4647-4661.
39. Schwartz, B. L.; Bruce, J. E.; Anderson, G. A.; Hofstadler, S. A.; Rockwood, A. L.; Smith, R. D.; Chilkoti, A.; Stayton, P. S., *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 459-465.
40. Jurchen, J. C.; Williams, E. R., *J. Am. Chem. Soc.* **2003**, *125*, 2817-2826.
41. Benesch, J. L. P., *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 341-348.
42. Boyd, R. K.; Somogyi, Á., *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1275-1278.
43. Dongré, A. R.; Jones, J. L.; Somogyi, Á.; Wysocki, V. H., *J. Am. Chem. Soc.* **1996**, *118*, 8365-8374.

44. Jørgensen, T. J. D.; Gårdsvoll, H.; Ploug, M.; Roepstorff, P., *J. Am. Chem. Soc.* **2005**, *127*, 2785-2793.
45. Li, J. Y.; Lyu, W. P.; Rossetti, G.; Konijnenberg, A.; Natalello, A.; Ippoliti, E.; Orozco, M.; Sobott, F.; Grandori, R.; Carloni, P., *J. Phys. Chem. Lett.* **2017**, *8*, 1105-1112.
46. Fegan, S. K.; Thachuk, M., *J. Chem. Theory Comput.* **2013**, *9*, 2531-2539.
47. Popa, V.; Trecroce, D. A.; McAllister, R. G.; Konermann, L., *J. Phys. Chem. B* **2016**, *120*, 5114-5124.
48. Nesatyy, V. J.; Suter, M. J.-F., *J. Mass Spectrom.* **2004**, *39*, 93-97.
49. Fenn, J. B., *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 524-535.
50. Katta, V.; Chait, B. T., *J. Am. Chem. Soc.* **1991**, *113*, 8534-8535.
51. Dobo, A.; Kaltashov, I. A., *Anal. Chem.* **2001**, *73*, 4763-4773.
52. Konermann, L.; Douglas, D. J., *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1248-1254.
53. Wang, G.; Abzalimov, R. R.; Kaltashov, I. A., *Anal. Chem.* **2011**, *83*, 2870-2876.
54. Loo, J. A.; Edmonds, C. G.; Udseh, H. R.; Smith, R. D., *Anal. Chem.* **1990**, *62*, 693-698.
55. Borysic, A. J. H.; Radford, S. E.; Ashcroft, A. E., *J. Biol. Chem.* **2004**, *279*, 27069-27077.
56. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S., *Mass Spectrom. Rev.* **1998**, *17*, 1-35.
57. Zubarev, A. R.; Makarov, A., *Anal. Chem.* **2013**, *85*, 5288-5296.
58. Han, X.; Jin, M.; Breuker, K.; McLafferty, F. W., *Science* **2006**, *314*, 109-112.
59. Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L., *Anal. Chem.* **2011**, *83*, 6868-6874.
60. Guevremont, R.; Siu, K. W. M.; Le Blanc, J. C. Y.; Berman, S. S., *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 216-224.
61. Wang, G.; Cole, R. B., *Org. Mass Spectrom.* **1994**, *29*, 419-427.
62. Kelly, M. A.; Vestling, M. M.; Fenselau, C. C.; Smith, P. B., *Org. Mass Spectrom.* **1992**, *27*, 1143-1147.
63. Prakash, H.; Mazumdar, S., *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1409-1421.
64. Creighton, T. E., *Proteins*. W. H. Freeman & Co: New York, **1993**.
65. Bonner, J. G.; Lyon, Y. A.; Nellesen, C.; Julian, R. R., *J. Am. Chem. Soc.* **2017**, *139*, 10286-10293.
66. Schnier, P. D.; Gross, D. S.; Williams, E. R., *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1086-1097.

67. Jha, A. K.; Colubri, A.; Freed, K. F.; Sosnick, T. R., *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13099-13104.
68. Consta, S.; Malevanets, A., *Phys. Rev. Lett.* **2012**, *109*, 144301-144305
69. Donor, M. T.; Ewing, S. A.; Zenaidee, M. A.; Donald, W. A.; Prell, J. S., *Anal. Chem.* **2017**, *89*, 5107-5114.
70. Beveridge, R.; Phillips, A. S.; Denbigh, L.; Saleem, H. M.; MacPhee, C. E.; Barran, P. E., *Proteomics* **2015**, *15*, 2872-2883.
71. Yue, X.; Vahidi, S.; Konermann, L., *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1322-1331.
72. Brini, E.; Fennell, C. J.; Fernandez-Serra, M.; Hribar-Lee, B.; Luksic, M.; Dill, K. A., *Chem. Rev.* **2017**, *117*, 12385-12414.
73. Iyengar, S. S.; Day, T. J. F.; Voth, G. A., *Int. J. Mass Spectrom.* **2005**, *241*, 197-204.
74. Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M., *Nature* **1999**, *397*, 601-604.
75. Caleman, C.; van der Spoel, D., *J. Chem. Phys.* **2006**, *125*, 1545081-1545089.
76. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., *SoftwareX* **2015**, *1-2*, 19-25.
77. Sun, Y.; Vahidi, S.; Sowole, M. A.; Konermann, L., *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 31-40.
78. Huang, J.; MacKerell, A. D., *J. Comput. Chem.* **2013**, *34*, 2135-2145.
79. Abascal, J. L. F.; Vega, C., *J. Chem. Phys.* **2005**, *123*, 234505.
80. Fatunmbi, O.; Abzalimov, R. R.; Savinov, S. N.; Gershenson, A.; Kaltashov, I. A., *Biochemistry* **2016**, *55*, 1918-1928.
81. Ewing, S. A.; Donor, M. T.; Wilson, J. W.; Prell, J. S., *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 587-596.
82. Eliezer, D.; Yao, J.; Dyson, H. J.; Wright, P. E., *Nat. Struct. Biol.* **1998**, *5*, 148-155.
83. Mao, Y.; Woenckhaus, J.; Kolafa, J.; Ratner, M. A.; Jarrold, M. F., *J. Am. Chem. Soc.* **1999**, *121*, 2712-2721.
84. Moser, A.; Range, K.; York, D. M., *J. Phys. Chem. B* **2010**, *114*, 13911-13921.

## Chapter 6: Conclusions and Future Work

### 6.1 Conclusions

Despite the many advances in ESI-MS in various research areas, many fundamental aspects of ESI remain intractable. The photographs of the droplet jet fission and progeny droplet formation by Gomez and Tang<sup>1</sup> have inspired quite a number of researchers to develop techniques that give an atomistic view of the ESI process. In order to achieve a reliable and robust ESI analysis, it is essential to understand the underlying mechanisms, in the words of Richard Cole, we don't want to be '*blind men*'.<sup>2</sup> MD simulations are one of the tools that can provide a detailed information regarding the ESI process.<sup>3-6</sup> There is a general consensus that small ions are ejected from the droplet via the IEM, while globular analytes follow the CRM.<sup>7-9</sup> In this thesis, we have applied MD simulations as a complementary method along with experimental investigations. Our work has broken new ground in terms of applying MD simulations to IEM, CRM and CEM events, and we have significantly contributed to understanding the mechanism of protein supercharging.

The main goal behind the work in Chapter 2 was to investigate the role played by cations during protein charging in the positive ion mode. Analytes that have basic groups can be charged by cation binding which enhances their ionization and detection. Generally, the presence of non-volatile salts or buffers can be beneficial as most biological samples require the existence of cations such as  $\text{Na}^+$  and  $\text{K}^+$  to preserve native structure and non-covalent complexes. However, these non-volatile salts can be incompatible with the ESI-MS source. In literature, these salts can cause interferences and degrade the protein ion signal. The mechanistic understanding of salt interferences is yet to be explained. Instead of depending on experimental work only, we employed all-atom molecular dynamics simulation to gain a more detailed image of the behavior of nanodroplets in the presence of cations such as  $\text{Na}^+$  and  $\text{NBu}_4^+$ . Our findings showed that salts such as  $\text{NaCl}$  will cause a major adduction, while not strongly affecting the protein ion signal. Conversely,  $\text{NBu}_4\text{Cl}$  completely degrades the protein ion formation. The main factor behind their differential response is their free energy of hydration or their surface activity. The surface accessibility of cation was described by the charge competition model, where species like  $\text{NBu}_4^+$  can be lost via facile ejection from the surface of the parent droplet. Species like  $\text{Na}^+$

cannot be described by the charge competition model; rather it is a simple spread of ion count over a wide  $m/z$  range via adduction.

Supercharging is a relatively new method that represents an addition to the field of tandem MS. The efficiency of fragmentation in techniques such as electron capture dissociation (ECD) and CID depends on the extent of charging. To properly apply these dissociation techniques, it will be beneficial to understand the mechanism behind protein supercharging. By using experiments alone, it is impossible to obtain detailed insights into the temporal evolution of the protein/droplet system during supercharging. The mechanism of supercharging has been always controversial. In Chapter 3 we used a combination of MD simulations and ESI-MS/IMS to examine the supercharging mechanism of sulfolane and *m*-NBA. Our data showed that native supercharging takes place due to charge trapping; where supercharging agent (SCA) segregation lead to an aqueous core that contains the protein and  $\text{Na}^+/\text{H}^+$  as a charge carrier with an outer shell of SCA. This layering will allow  $\text{Na}^+/\text{H}^+$  to eject from the droplet via IEM with water. Once water completely evaporates, the outer shell decreases the chances of IEM and the remaining charges will irreversibly bind to the protein. The protein will be released via CRM and may undergo gas phase unfolding due to charge repulsion. In Chapter 4, we tested this proposal by using 18C6. As discussed above, based on our charge trapping hypothesis the solvation characteristics of SCA determine the fate of charge carriers. 18C6 selectively binds  $\text{Na}^+$ , thereby enhancing the solubility of charge carriers in the SCA layer and reversing the supercharging effects. Both MD simulations and experiments show agreement which reflects the utility of MD simulations in interrogating the behavior of ESI-droplets during supercharging.

In Chapter 5, we studied the ESI mechanism for unfolded proteins. Several years ago, our laboratory proposed that the CEM can explain the release of proteins under non-native conditions. There has been a long gap between those earlier proposals and the actual MD simulations of the current work due to difficulties associated with  $\text{H}^+$  migration and droplet size. In the CEM, unfolded proteins migrates to the surface of the droplet due to electrostatic and hydrophobic factors and get ejected via formation of intermediates where the droplets carry extended protein tails. Our MD simulations along with ESI-MS/IMS have provided a clear evidence that unfolded proteins are released via the CEM. To our knowledge, this is the first study that shows the release of unfolded proteins from the droplet in an atomistic fashion.

## 6.2 Future Directions

### 6.2.1 Studying the Effects of Anions on Protein Ion Signals in Negative Ion Polarity

Chapter 2 explored the behavior of cations under positive polarity conditions. Even though positive ion mode is the widely used for protein analysis, negative ion mode is important for studying DNA and RNA duplexes.<sup>10</sup> It will be exciting to explore the effects of anions such as acetate, formate, chloride, and phosphate in proteins, protein complexes, and DNA in negative ESI. Using MD simulations and ESI-MS/IMS can give more insights into the behavior of all the previously mentioned anions especially acetate anion. As ammonium acetate is commonly used as a “buffer” in direct infusion studies, many experiments have shown that the buffering capacity of ammonium acetate is not as effective as other buffers, besides some protein-ligand interactions are unstable in ammonium acetate.<sup>11-12</sup>

### 6.2.2 Investigating the Effects of Salts on Protein Supercharging

Chapters 3 and 4 shed light on the mechanism of supercharging. Exploring the effects of salts on supercharging is important as metal ions such as ( $K^+$ ,  $Li^+$ ) are ligands for some proteins or protein complexes. MD simulations for those cations in SCAs-containing droplets along with observing the changes in charge state distributions in ESI-MS can be used to interrogate the effects of cation size. This can be informative in designing experiments employing supercharging.

### 6.2.3 Additional Supercharging Agents

We focused on the most commonly used SCAs sulfolane and *m*-NBA but studies for additional SCAs can facilitate a better understanding of protein supercharging.<sup>13-14</sup>

## 6.2.4 Examining the Effects of Supercharging Agents in Negative Ion Mode

There is a lot of discussion about whether SCAs increase the charge in the negative ion polarity. Some reports suggest that *m*-NBA can efficiently increase charging in the negative ion mode<sup>15</sup> while sulfolane fails to boost charges in the negative polarity.<sup>16</sup> An expansion of the work presented will be to investigate different SCAs in negative ion mode using ESI-MS/IMS along with the MD protocols of Chapters 3 and 4. In negative ion mode, proteins are charged by proton stripping so MD simulations using mobile proton algorithms have to be applied that can provide more details about supercharging in the negative ion mode by investigating how H<sup>+</sup> is transferred.<sup>17</sup>

## 6.2.5 Uncovering the Mechanism of Supercharging Under Denaturing Conditions

Unfolded proteins have unique features as they show high charge state distributions. Addition of SCAs shift the ions to even lower *m/z*.<sup>18-19</sup> Using our MD techniques of Chapter 3 which involve ‘forced evaporation’ of SCAs along with placing the unfolded protein in a big droplet and judicious choice of pH as in Chapter 5 may be helpful in deciphering the behavior of unfolded proteins during supercharging.

## 6.2.6 Can the Chain Ejection Model (CEM) be applied to Intrinsically Disordered Proteins (IDPs)?

Generally, functional protein should have a compact conformation. IDPs are unique class of proteins, characterized by loose structures at physiological pH. Despite lacking any stable structures, they are vital to many cellular processes. This class of proteins is also known to be associated with human diseases like cancer and Parkinson’s.<sup>20-21</sup> Employing our modeling strategy and ESI-MS/IMS in Chapter 5 will be able to uncover how these proteins are released from the droplet to improve their characterization using ESI-MS.

## 6.3 References

1. Gomez, A., Tang K, *Physics of Fluids* **1994**, *6*, 404-414.
2. B., C. R., *Journal of Mass Spectrometry* **2000**, *35*, 763-772.
3. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., *Analytical Chemistry* **2013**, *85*, 2-9.
4. Kim, D.; Wagner, N.; Wooding, K.; Clemmer, D. E.; Russell, D. H., *Journal of the American Chemical Society* **2017**, *139*, 2981-2988.
5. Consta, S.; Oh, M. I.; Malevanets, A., *Chemical Physics Letters* **2016**, *663*, 1-12.
6. Consta, S.; Malevanets, A., *Physical Review Letters* **2012**, *109*, 148301.
7. Kebarle, P.; Tang, L., *Analytical Chemistry* **1993**, *65*, 972A-986A.
8. Cech, N. B.; Enke, C. G., *Mass Spectrometry Reviews* **2001**, *20*, 362-387.
9. Meyer, T.; Gabelica, V.; Grubmüller, H.; Orozco, M., *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2013**, *3*, 408-425.
10. Porrini, M.; Rosu, F.; Rabin, C.; Darré, L.; Gómez, H.; Orozco, M.; Gabelica, V., *ACS Central Science* **2017**, *3*, 454-461.
11. Konermann, L., *Journal of The American Society for Mass Spectrometry* **2017**, *28*, 1827-1835.
12. Gavriilidou, A. F. M.; Gülbakan, B.; Zenobi, R., *Analytical Chemistry* **2015**, *87*, 10378-10384.
13. Zenaidee, M. A.; Donald, W. A., *Analyst* **2015**, *140*, 1894-1905.
14. Wang, H.; Yong, G.; Brown, S. L.; Lee, H. E.; Zenaidee, M. A.; Supuran, C. T.; Donald, W. A., *Analytica Chimica Acta* **2018**, *1003*, 1-9.
15. Ogorzalek Loo, R. R.; Lakshmanan, R.; Loo, J. A., *Journal of The American Society for Mass Spectrometry* **2014**, *25*, 1675-1693.
16. Douglass, K. A.; Venter, A. R., *Journal of The American Society for Mass Spectrometry* **2012**, *23*, 489-497.
17. Konermann, L., *The Journal of Physical Chemistry B* **2017**, *121*, 8102-8112.
18. Donor, M. T.; Ewing, S. A.; Zenaidee, M. A.; Donald, W. A.; Prell, J. S., *Analytical Chemistry* **2017**, *89*, 5107-5114.
19. Teo, C. A.; Donald, W. A., *Analytical Chemistry* **2014**, *86*, 4455-4462.

20. Beveridge, R.; Phillips, A. S.; Denbigh, L.; Saleem, H. M.; MacPhee, C. E.; Barran, P. E., *Proteomics* **2015**, *15*, 2872-2883.
21. Natalello, A.; Santambrogio, C.; Grandori, R., *Journal of The American Society for Mass Spectrometry* **2016**, *28*, 21-28.

## Appendix I-Permissions



RightsLink®

Home

Create Account

Help



ACS Publications  
Most Trusted. Most Cited. Most Read.

**Title:** Exploring the Mechanism of Salt-Induced Signal Suppression in Protein Electrospray Mass Spectrometry Using Experiments and Molecular Dynamics Simulations

**Author:** Haidy Metwally, Robert G. McAllister, Lars Konermann

**Publication:** Analytical Chemistry

**Publisher:** American Chemical Society

**Date:** Feb 1, 2015

Copyright © 2015, American Chemical Society

<a href="#">LOGIN</a>
If you're a <b>copyright.com user</b> , you can login to RightsLink using your copyright.com credentials. Already a <b>RightsLink user</b> or want to <a href="#">learn more?</a>

### PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK

CLOSE WINDOW



# RightsLink<sup>®</sup>

[Home](#)
[Create Account](#)
[Help](#)


**ACS Publications**  
Most Trusted. Most Cited. Most Read.

**Title:** Mechanism of Protein Supercharging by Sulfolane and m-Nitrobenzyl Alcohol: Molecular Dynamics Simulations of the Electrospray Process

**Author:** Haidy Metwally, Robert G. McAllister, Vlad Popa, et al

**Publication:** Analytical Chemistry

**Publisher:** American Chemical Society

**Date:** May 1, 2016

Copyright © 2016, American Chemical Society

<a href="#">LOGIN</a>
<p><b>If you're a copyright.com user</b>, you can login to RightsLink using your copyright.com credentials.</p> <p>Already a <b>RightsLink user</b> or want to <a href="#">learn more?</a></p>

## PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)
[CLOSE WINDOW](#)



RightsLink®

Home

Create Account

Help



ACS Publications  
Most Trusted. Most Cited. Most Read.

**Title:** Crown Ether Effects on the Location of Charge Carriers in Electrospray Droplets: Implications for the Mechanism of Protein Charging and Supercharging

**Author:** Haidy Metwally, Lars Konermann

**Publication:** Analytical Chemistry

**Publisher:** American Chemical Society

**Date:** Mar 1, 2018

<a href="#">LOGIN</a>
<b>If you're a copyright.com user</b> , you can login to RightsLink using your copyright.com credentials. Already a <b>RightsLink user</b> or want to <a href="#">learn more?</a>

Copyright © 2018, American Chemical Society

### PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK

CLOSE WINDOW

Copyright © 2018 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#). Comments? We would like to hear from you. E-mail us at [customercare@copyright.com](mailto:customercare@copyright.com)



# RightsLink®

[Home](#)
[Create Account](#)
[Help](#)


**ACS Publications**  
Most Trusted. Most Cited. Most Read.

**Title:** Chain Ejection Model for Electrospray Ionization of Unfolded Proteins: Evidence from Atomistic Simulations and Ion Mobility Spectrometry

**Author:** Haidy Metwally, Quentin Duez, Lars Konermann

**Publication:** Analytical Chemistry

**Publisher:** American Chemical Society

**Date:** Jul 1, 2018

Copyright © 2018, American Chemical Society

[LOGIN](#)

If you're a **copyright.com user**, you can login to RightsLink using your [copyright.com](#) credentials.

Already a **RightsLink user** or want to [learn more?](#)

## PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)
[CLOSE WINDOW](#)

Copyright © 2018 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement.](#) [Terms and Conditions.](#) Comments? We would like to hear from you. E-mail us at [customercare@copyright.com](mailto:customercare@copyright.com)



# RightsLink<sup>®</sup>

[Home](#)
[Create Account](#)
[Help](#)


**ACS Publications**  
Most Trusted. Most Cited. Most Read.

**Title:** Unraveling the Mechanism of Electrospray Ionization  
**Author:** Lars Konermann, Elias Ahadi, Antony D. Rodriguez, et al  
**Publication:** Analytical Chemistry  
**Publisher:** American Chemical Society  
**Date:** Jan 1, 2013  
 Copyright © 2013, American Chemical Society

<a href="#">LOGIN</a>
<p><b>If you're a copyright.com user,</b> you can login to RightsLink using your copyright.com credentials.</p> <p>Already a <b>RightsLink user</b> or want to <a href="#">learn more?</a></p>

## PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.

[BACK](#)
[CLOSE WINDOW](#)

## Curriculum Vitae

Haidy Metwally

Department of Chemistry  
Western University

---

### Education

**-Western University- London-Ontario- Canada (September 2013- present)**

- Ph.D. Candidate in Physical and Analytical Chemistry

Principal advisor: Professor/ Lars Konermann.

**-Ain Shams University-Faculty of Pharmacy –Cairo-Egypt (2007-2012)**

- M.Sc. Honors in Pharmaceutical Sciences-Pharmacognosy

Thesis: Phytochemical and Biological Studies of Some *Artocarpus* Species, Family *Moraceae*- Cultivated in Egypt.

**-Ain Shams University –Cairo-Egypt (2000-2005)**

- B.Sc.Honors Specialization in Pharmaceutical Sciences.

### Experience

**-Western University-London-Ontario-Canada (September 2013- present)**

- Teaching and Research Assistant.

**-National Organization of Drug Control and Research, Giza, Egypt, (2006-2012)**

- Quality Control Specialist.

### Honors and Awards

**-Canadian Society for Mass Spectrometry (CSMS) travel award (2016)**

- The 29th Tandem Mass Spectrometry Workshop.

**-Best Oral Presentation (2016)**

- Enabling Technology (ETP) Symposium in Ottawa.

## Peer-Reviewed Publications

- **H. Metwally**, Q. Duez, and L. Konermann “Chain Ejection Model for Electrospray Ionization of Unfolded Proteins: Evidence from Atomistic Simulations and Ion Mobility Spectrometry” *Anal. Chem.* *90*, 10069 (2018).
- Q. Duez, **H. Metwally**, and L. Konermann “Electrospray Ionization of Polypropylene Glycol: Rayleigh-Charged Droplets, Competing Pathways, and Charge State-Dependent Conformations” *Anal. Chem.* *90*, 9912 (2018).
- L. Konermann, **H. Metwally**, R.G. McAllister, and V. Popa, “How to Run Molecular Dynamics Simulations on Electrospray Droplets and Gas Phase Proteins: Basic Guidelines and Selected Applications” *Methods* *144*, 104 (2018).
- **H. Metwally** and L. Konermann “Crown Ethers Modulate the Location of Charge Carriers in Electrospray Droplets: Implications for the Mechanism of Charging and Supercharging” *Anal. Chem.* *90*, 4126 (2018).
- T. D. Schachel, **H. Metwally**, V. Popa, and L. Konermann “Collision-Induced Dissociation of Electrosprayed NaCl Clusters: Using Molecular Dynamics Simulations to Visualize Reaction Cascades in the Gas Phase” *J. Am. Soc. Mass Spectrom.* *27*, 1846 (2016).
- C.E. Bartman, **H. Metwally**, and L. Konermann “Effects of Multidentate Metal Interactions on the Structure of Collisionally Activated Proteins: Insights from Ion Mobility and Molecular Dynamics Simulations” *Anal. Chem.* *28*, 6905 (2016).
- **H. Metwally**, R.G. McAllister, V. Popa, and L. Konermann “Mechanism of Protein Supercharging by Sulfolane and m-NBA: Molecular Dynamics Simulations of the Electrospray Process” *Anal. Chem.* *88*, 5345 (2016).
- R.G. McAllister, **H. Metwally**, Y. Sun, and L. Konermann “Release of Native-Like Gaseous Proteins from Electrospray Droplets via The Charged Residue model: Insights from Molecular Dynamics Simulations” *J. Am. Chem. Soc.* *137*, 12667 (2015).
- **H. Metwally**, R.G. McAllister, and L. Konermann “Exploring the Mechanism of Salt Induced Signal Suppression in Protein Electrospray Mass Spectrometry Using Experiments and Molecular Dynamics Simulations” *Anal. Chem.* *87*, 2434 (2015).
- L. Konermann, R.G. McAllister, and **H. Metwally** “Molecular Dynamics Simulations of The Electrospray Process: Formation of NaCl Clusters via the Charged residue Mechanism” *J. Phy. chem. B.* *118*, 12025 (2014).

## Presentations and Posters

- **H. Metwally**, and L. Konermann “Crown Ethers Modulate the Location of Charge Carriers in Electrospray Droplets: Implications for the Mechanism of Protein Charging and Supercharging” *66<sup>th</sup> ASMS Conference on Mass Spectrometry and Allied Topics* - San Diego- California-USA (2018).
- **H. Metwally**, and L. Konermann “Crown Ethers Suppress Protein Supercharging in Native ESI-MS: Evidence for a Charge Trapping Mechanism” *30<sup>th</sup> Lake Louise Workshop on Tandem Mass Spectrometry*- Lake Louise-Alberta (2017).
- **H. Metwally**, and L. Konermann “Molecular Dynamics Simulation of Electrospray Process: Why Do Crown Ethers Suppress Protein Supercharging?” *9<sup>th</sup> International Symposium on Enabling Technologies*-Ottawa (2017).
- **H. Metwally**, and L. Konermann “Mechanism of Protein Supercharging by Sulfolane and m-NBA” *29<sup>th</sup> Lake Louise Workshop on Tandem Mass Spectrometry*- Lake Louise-Alberta (2016).
- **H. Metwally**, and L. Konermann “Mechanism of Protein Supercharging by Sulfolane and m-NBA: Molecular Dynamics Simulations of the Electrospray Process” *33<sup>rd</sup> Annual Trent Conference workshop on Mass Spectrometry*- Toronto (2016).
- **H. Metwally** and L. Konermann “Mechanism of Protein Supercharging by Sulfolane and m-NBA: Molecular Dynamics Simulations of the Electrospray Process” *21<sup>st</sup> International Mass Spectrometry Conference*-Toronto (2016).
- **H. Metwally** and L. Konermann “Exploring the Mechanism of Salt-Induced Signal Suppression in Protein Electrospray Mass Spectrometry Using Experiments and Molecular Dynamics Simulations” *98<sup>th</sup> Canadian Chemistry Conference and Exhibition*-Ottawa (2015).
- **H. Metwally**, and L. Konermann “Mechanism of Salt-Induced Signal Suppression in Protein ESI-MS” *31<sup>th</sup> Annual Trent Conference on Mass Spectrometry*- Orillia (2014).