

Electronic Thesis and Dissertation Repository

9-18-2023 10:30 AM

Protein Stability in Solution and in the Gas Phase.

Yousef Haidar,

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

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Chemistry

© Yousef Haidar 2023

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

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

Recommended Citation

Haidar, Yousef, "Protein Stability in Solution and in the Gas Phase." (2023). *Electronic Thesis and Dissertation Repository*. 9673.

<https://ir.lib.uwo.ca/etd/9673>

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.

Abstract

Electrospray Ionization mass spectrometry (ESI-MS) is widely used for probing proteins, yet many aspects of this technique remain elusive. Using MS, ion mobility spectrometry (IMS), and circular dichroism (CD) spectroscopy, this thesis sheds light on the stability differences of proteins in the gas phase and solution. After a general introduction (Chapter 1), Chapter 2 scrutinizes some aspects of native ESI. Our data highlight the significance of cone voltage in maintaining a native-like fold and show the advantage of using NH₄Ac in protein experiments. Chapter 3 focuses on hydrogen/deuterium exchange (HDX)-MS. Several studies have reported that D₂O enhances the stability of proteins. We corroborated this effect through thermal unfolding assays. Previous studies tentatively attributed this phenomenon to either strengthened backbone H-bonds or to changes in protein-solvent interactions. To help unravel these contributions, we performed Collision Induced Unfolding experiments (CIU) on gaseous proteins. The indistinguishable CIU profiles of deuterated and unlabeled proteins suggest that D₂O-induced stabilization originates from solvent effects.

Keywords: Proteins, Mass Spectrometry, Ion Mobility Spectrometry, Collision Induced Unfolding, Protein Stability, Hydrogen-Deuterium Exchange.

Summary for Lay Audience:

Proteins are molecules that perform various biological roles important to all living organisms. The function of these molecules depends on their overall structure. Many techniques are available to probe protein structure and conformational dynamics. These include native ESI as well as chemical labelling, in combination with MS. Unfortunately, many fundamental aspects of these techniques remain poorly understood. For every ESI experiment, proteins have to be converted to gaseous ions. This process is carried out by applying high voltage to an ESI capillary where a protein solution is introduced, generating a plume of highly charged droplets that ultimately release protein ions into the gas phase. The extent to which gaseous proteins and non-covalent complexes maintain their native structures depends on how gentle the ESI conditions are. While Native ESI is a useful tool for probing protein structure, it does not provide high-resolution data. Chemical labelling in conjunction with MS fills this role by probing solvent accessible surfaces. The question to what extent these labels are benign has been the subject of debate for decades.

In this work, we investigate ESI parameters that can affect the outcome of protein experiments. The results of chapter 2 demonstrate that cone voltage, and solvent conditions can greatly influence whether proteins retain native-like structures. Moreover, we discussed the chemistry of NH_4Ac (a widely used ESI additive). This chemical has a little buffering capacity at neutral conditions but can stabilize the pH at slightly acidic values. Our data reveal that globular proteins like cytochrome c and lysozyme can tolerate such slight pH drops.

In chapter 3, we shed light on HDX (a widely used chemical labelling method) and scrutinize the assumption that labeling proteins with deuterium is benign. We investigated the purported stability of deuterated proteins in the solution phase using thermal unfolding studies. Similarly, we utilized collision induced unfolding to investigate the stability of deuterated proteins in the gas phase. Our data demonstrates that exposure to D_2O might alter certain aspects of protein structure and dynamics in solution. In contrast, gas-phase experiments revealed that the stability of gaseous protein ions is indistinguishable before and after deuteration.

Co-Authorship Statement

The work in chapter 3 was produced with permission from

Haidar, Y.; Konermann, L. Effects of Hydrogen/Deuterium Exchange on Protein Stability in Solution and in the Gas Phase. *J. Am. Soc. Mass Spectrom.* **2023**, 34, 1447-1458. Copyright ©, American Chemical Society.

Parts of chapter 2 were produced with permission from

Konermann, L.; Haidar, Y. Mechanism of Magic Number NaCl Cluster Formation from Electrosprayed Water Nanodroplets. *Anal.Chem.* **2022**, 94, 16491-16501. Copyright ©, American chemical society.

All experimental work for the first paper was performed and analyzed by the first author, Yousef Haidar . The original draft of the first article was prepared by the first author. Subsequent revisions were done by the author and Dr. Lars Konermann. All experimental work was performed under the supervision of Dr. Lars Konermann.

Acknowledgements

The work of this thesis would not have been possible without the contributions of many people who supported me throughout this journey with endless advice and assurance.

First, I would like to thank Dr. Lars Konermann for allowing me to pursue a master's degree in his lab in September 2021. Under the guidance of Dr. Konermann, I have found myself becoming a better scientist with each day I was under his supervision. I am so grateful for Dr. Konermann unwavering support and guidance throughout my journey. When I first joined the Konermann lab, I was understandably anxious about my progress. Lars made sure that my transition to graduate school was smooth and productive. I will be eternally grateful for his incredible guidance and leadership.

I would also like to thank my committee members Dr. Arghya Paul and Dr. Shawn Li for their useful insights into my work. I would also like to extend my sincere thanks to Darlene DiValentin, Dr. Ken Yeung and Dr. Chris Levy in the chemistry department for their assistance during my graduate studies.

Working in the Konermann Lab with my colleagues was perhaps the best part of my journey. A very special thanks to Pablo Scrosati, who answered every single question I had and made sure that I have the fundamental knowledge for my degree. He always offered help and guidance when asked. Evelyn MacKay-Barr, Kasra Hanafi, Vida Alinezhad, Mahsa Dolatkah, and Kirsty Yuen Ki Ng have always offered help and support when needed. My Fellow Konermann lab members thank you for everything.

I would like to thank my parents Mai and Abd for their unwavering emotional support and supporting me through everything I do in my life. Finally, thank you to Alaa Wardeh for his kindness and reassurance in my abilities. I would not be the person I am without these people and their insurmountable support.

Table of Contents

Abstract.....	i
Summary for Lay Audience	ii
Co-Authorship Statement.....	iii
Acknowledgement.....	iv
List of Tables.....	Vii
List of Figures.....	ix
List of Appendices.....	xi
List of Symbols and Abbreviations.....	xii
Chapter 1. Introduction.....	1
1.1 Native Protein Structures.....	1
1.2 Thermodynamics Aspects of Protein Stability in Solution.....	2
1.2.1 Thermal Protein Unfolding in Solution.....	4
1.2.2 Chemical denaturants.....	6
1.3 Conventional Methods for Protein Structural Analysis.....	8
1.3.1 UV-VIS Absorption Spectroscopy.....	8
1.3.2 Circular Dichroism Spectroscopy.....	9
1.3.3 Nuclear Magnetic Resonance Spectroscopy.....	11
1.3.4 X-ray Crystallography.....	11
1.3.5 Cryogenic Electron Microscopy.....	12
1.4 Protein Chemical Labeling.....	13
1.4.1 Protein Cross Linking.....	14
1.4.2 Hydrogen-Deuterium Exchange.....	15

1.5 Deuterium-Isotope Effects.....	17
1.6 Mass Spectrometry.....	18
1.7 Electrospray Ionization.....	19
1.8 Native ESI of Proteins.....	22
1.9 Protein Structures in the Gas Phase.....	24
1.10 Ion Mobility Spectrometry.....	24
1.11 Collision Induced Unfolding	27
1.12 Scope of this Thesis.....	28
1.13 References.....	29
Chapter 2.....	32
2.1 Introduction.....	32
2.2 Materials and Methods.....	37
2.3 Results and Discussion.....	38
2.3.1 Charge State Distributions and their Information Content.....	38
2.3.2 Experimental Conditions for Native ESI.....	40
2.3.3 Chemistry of Ammonium Acetate.....	43
2.4 Conclusions.....	47
2.5 References.....	48
Chapter 3.....	51
3.1 Introduction.....	51
3.2 Materials and Methods.....	55
3.3 Results and Discussion.....	58
3.3.1 Native ESI Mass Spectra in H ₂ O and D ₂ O.....	58

3.3.2 pH effects on Thermal Protein Unfolding.....	60
3.3.3 D ₂ O-Mediated Protein Stabilization in Solution.....	60
3.3.4 Unfolding of Unlabeled and Deuterated protein in the Gas phase.....	63
3.3.5 Implications of H-Bonds and D-Bonds for Protein Stability in the Gas Phase.	64
3.3.6 Dissecting Isotope Effects on Protein Stability.....	66
3.4 Conclusion	68
3.5 References.....	70
Chapter 4.....	75
4.1 Summary.....	75
4.2 Future Directions.....	76
4.3 References.....	77
Curriculum Vitae.....	79

List of Tables:

Table 3.163

List of Figures

Chapter 1.

Figure 1.1 Crystal structure of lysozyme.....	2
Figure 1.2 Schematic representation of two-state unfolding equilibrium.....	3
Figure 1.3 Fraction of unfolded protein as a function of temperature.....	5
Figure 1.4 Fraction of unfolded protein as a function of denaturant concentration.....	7
Figure 1.5 Absorption spectra of myoglobin.....	9
Figure 1.6 Characteristic CD signals for principal polypeptide structures.....	10
Figure 1.7 Dependence of the chemical rate constant (k_{ch}) on pH.....	16
Figure 1.8 Schematic cartoon of an ESI mass spectrometer.....	20
Figure 1.9 The major steps of ESI processes.....	21
Figure 1.10 The different Ionization mechanism for native and denatured proteins.....	23
Figure 1.11 A stacked ring TWIMS device.....	26

Chapter 2.

Figure 2.1 Schematic free energy difference diagram.....	33
Figure 2.2 The titration curve for NH_4Ac	36
Figure 2.3 ESI mass spectra of native and denatured myoglobin.....	38
Figure 2.4 ESI mass spectra of myoglobin at low and high activation setting.....	40
Figure 2.5 Mass spectra acquired after electrospraying 10 mM NaCl in water.....	42
Figure 2.6 ESI mass spectra acquired for aqueous pH 7 solutions.....	44
Figure 2.7 mass spectra showing the charge state distributions of Ubiquitin, Cytochrome c and Lysozyme in 10mM NH_4Ac at pH 7.00 and pH 5.30	46

Chapter 3.

Figure 3.1 Schematic illustration of possible deuteration-induced effects.....	53
Figure 3.2 Native mass spectra of different proteins electrosprayed in H ₂ O solution and deuterated samples in D ₂ O solution.....	59
Figure 3.3 Thermodynamic analyses of cyt <i>c</i> and lysozyme thermal unfolding in H ₂ O and in D ₂ O solution.....	62
Figure 3.4 CIU data, displaying collision cross section of cyt <i>c</i> , lysozyme, and Ubiquitin.....	65
Figure 3.5 CIU profiles of deuterated vs. unlabeled proteins.....	66
Figure 3.6 Two-dimensional lattice bead chain model of a protein in water.....	68

List of Appendices

Appendix-I	78
-------------------------	-----------

List of Symbols and Abbreviations

ΔG	Change in Gibbs Free Energy
ΔH	Change in Enthalpy
ΔS	Change in Entropy
[N]	Concentration of Native Protein
[U]	Concentration of Denatured Protein
T	Temperature
ΔC_P	Change in specific Heat Capacity
DSC	Differential Scanning Calorimetry
CD	Circular Dichroism
f	Fraction Unfolded
T_m	Melting Temperature
[D]	Concentration of Denaturant
UV-VIS	Ultraviolet-Visible
E_R	Right-Handed Electric Field
E_L	Left-Handed Electric Field
NMR	Nuclear Magnetic Resonance
B_0	Applied Magnetic Field
M	Magnetization Vector
Cryo-EM	Cryogenic Electron Microscopy
MS	Mass Spectrometry
ESI	Electrospray Ionization

RP-LC	Reverse-Phase Liquid Chromatography
HDX	Hydrogen-Deuterium Exchange
k_{ch}	Intrinsic Exchange Rate
H ₂ O	Water
D ₂ O	Heavy Water
m/z	Mass to Charge ratio
EI	Electron Ionization
GC-MS	Gas Chromatography- Mass Spectrometry
MALDI	Matrix Assisted Laser Desorption Ionization
TOF	Time of Flight
FT-ICR	Fourier Transform- Ion Cyclotron Resonance
E_{kin}	Kinetic energy
E_{pot}	Potential energy
Z	Charge
e	Elementary Charge
CRM	Charge Residue Model
CEM	Chain Ejection Model
IMS	Ion Mobility Spectrometry
TWIMS	Travelling Wave Ion Mobility Spectrometry
DTIMS	Drift Tube Ion Mobility Spectrometry
Ω	Collision Cross Section
CIU	Collision Induced Unfolding
CID	Collision Induced Dissociation

ZPVE Zero Point Vibrational Energy

Cyt-c Cytochrome C

Chapter 1 Introduction

1.1 Native Protein Structures

Proteins are biological macromolecules that carry out various functions, for instance, they regulate cellular division, catalyze metabolic reactions, transcribe DNA, and assist in the synthesis of other biomolecules. The exact function of a protein depends on its three-dimensional structure. These structures are a consequence of various interactions that include covalent bonds, hydrogen bonds, hydrophobic contacts, van der Waals contacts, and Coulombic interactions².

Protein structures can be described at four levels of complexity. The primary structure is the amino acid sequence.³ Secondary structure is defined as the spatial arrangement of backbone atoms into α -helices, β -sheets, and loops that vary in length and topology. These structural elements arise from hydrogen bonds formed between backbone atoms, and they can be identified by their backbone dihedral angles ϕ -and ψ .³ Tertiary structure refers to the three-dimensional arrangement of the backbone and the side chains. Tertiary structures include domains that perform certain functional roles such as ligand binding.³ Finally, oligomeric proteins possess a quaternary structure which defines the association of multiple chains into a closely packed arrangement.³

A protein's three-dimensional structure is encoded in its amino acid sequence⁴. This native structure is the biological active form of the protein. The hydrophobic effect provides much of the thermodynamic driving force for protein folding, through the burial of non-polar side chains⁵. This ensures that proteins are compact, minimizing unfavourable interactions with the solvent. In addition, hydrogen bonds are among the interactions that provide directionality and organization for distinct folds. Proteins contain many hydrogen bond donors and acceptors, most importantly the backbone amide sites, as well as charged and polar functional groups (carboxylates, hydroxyls, amines, and amides)⁵. Backbone CO and NH groups form hydrogen bonds with one other, thereby stabilizing a protein's native fold. As an example, Figure 1.1 shows the crystal structure of lysozyme where backbone hydrogen bonds have been highlighted.

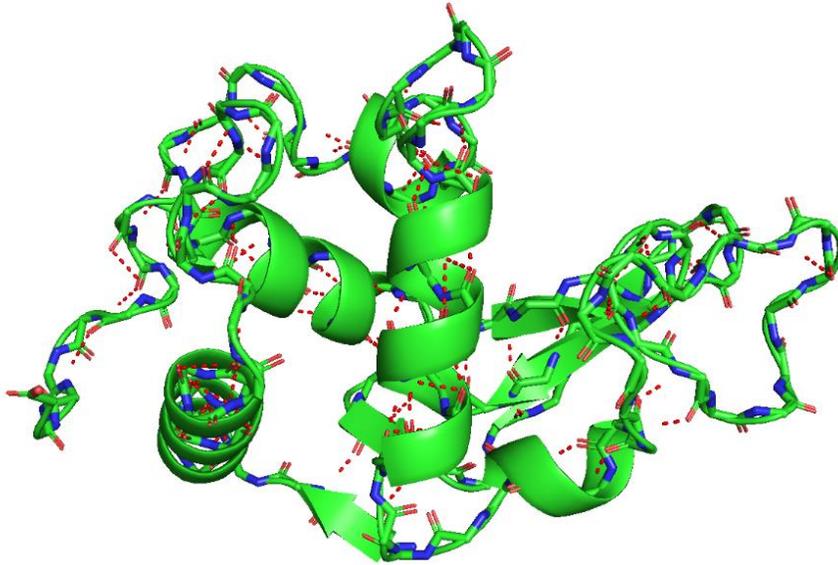


Figure 1.1. Crystal structure of lysozyme, PDB code: 1HEW. The protein secondary structure is shown in cartoon representation. Backbone $\text{NH}\cdots\text{OC}$ hydrogen bonds are highlighted as red dashes.

1.2 Thermodynamic Aspects of Protein Stability in Solution

The native state of a globular protein is characterized by a highly ordered conformation that depends on the amino acid sequence, the solvent environment, and the temperature.⁶ The native conformation contains a tightly packed core dominated by hydrophobic side chains, while hydrophilic residues are located on the outside where they favourably interact with water.⁶ Most native structures are only marginally stable and can unfold as a result of relatively subtle environmental changes.⁶ These unfolding transitions disrupt noncovalent interactions such as intramolecular hydrogen bonds and hydrophobic contacts.⁶ The unfolding transitions of many small proteins are highly cooperative and do not involve any detectable intermediates.⁷ The equilibrium constant of such two-state $\text{N} \rightleftharpoons \text{U}$ reactions is

$$K_{\text{eq}} = [\text{U}]/[\text{N}] \quad (1.1)$$

where $[\text{U}]$ and $[\text{N}]$ denote the equilibrium concentrations of the unfolded and the native states. The extent to which $[\text{U}]$ and $[\text{N}]$ are populated depends on the free energy difference ΔG of unfolding according to

$$K_{\text{eq}} = \exp(-\Delta G/RT) \quad (1.2)$$

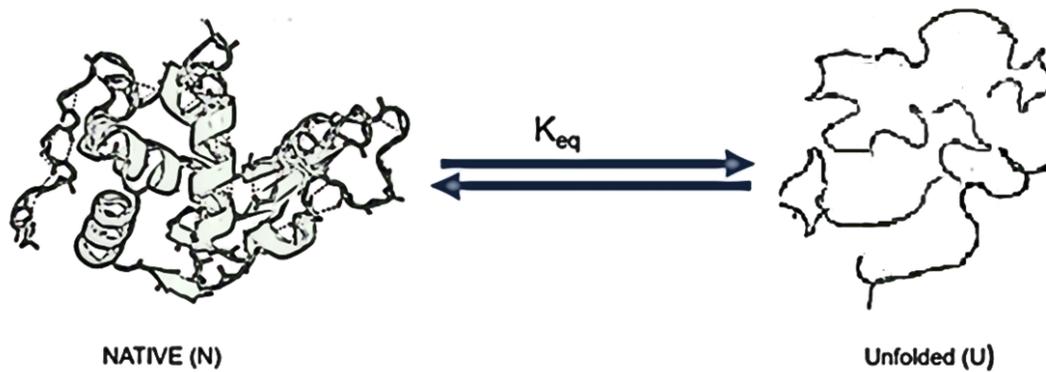


Figure 1.2. Schematic representation of two-state unfolding equilibrium. (N) corresponds to a tightly packed native conformation while (U) represents an unfolded structure.⁶ PDB code: 1HEW

Equations (1.1) and (1.2) define the thermodynamic stability of proteins, where large positive ΔG signifies a highly stable structure. ΔG is composed of enthalpic and entropic contributions both of which comprise stabilizing and destabilizing factors.⁶

$$\Delta G = \Delta H - T\Delta S \quad (1.3)$$

For instance, the large conformational space accessible to the unfolded polypeptide chains results in $\Delta S > 0$, a factor that tends to drive the equilibrium towards [U].⁶ However, under non-denaturing solvent conditions this entropy contribution is counteracted by $\Delta H > 0$, reflecting the presence of numerous non-covalent interactions in the native state, including hydrogen bonds and van der Waals contacts.⁶ In addition, the entropic contributions of surrounding water molecules are unfavourable when proteins are unfolded.⁶ Exposed hydrophobic residues restrict the movement of water molecules, leading to a partially ordered regions (“iceberg water”).^{6,8} This phenomenon is less prominent in the native state as most hydrophobic residues are sequestered in the core. Although intramolecular H-bonds undoubtedly contribute to the stability of N, unfolded proteins form extensive hydrogen bonds with water leading to enthalpic contributions that stabilize U.⁶ Overall, the delicate balance between entropy and enthalpy causes native proteins to be marginally stable with ΔG typically less than 100 kJmol^{-1} .⁶

1.2.1 Thermal Protein Unfolding in Solution

In thermal unfolding studies, proteins are subjected to gradually increasing temperature. The protein will unfold when the positive enthalpy change (ΔH in eq 1.3) can no longer compensate for the $-T\Delta S$ term.⁶ The point where $[U] = [N]$ (and $\Delta G = 0$) corresponds to the melting temperature T_m . Strictly speaking, both ΔH and ΔS are temperature dependent because unfolding is associated with a considerable change in heat capacity ($\Delta C_p > 0$).^{5,8} However, to a first approximation, this dependence can be assumed to be negligible near the transition midpoint.^{10,11,12} In this approximation, plots of ΔG vs temperature are linear according to equation 1.3.

Protein unfolding studies can utilize several tools to uncover thermodynamic parameters. For instance, differential scanning calorimetry (DSC) measures the heat capacity of a protein as a function of temperature.⁹ Alternatively, changes in secondary structures can be monitored by circular dichroism (CD) spectroscopy, usually at 222 nm. This optical technique is commonly used to construct unfolding curves, that is, the fraction of unfolded protein $f = [U]/([U] + [N])$ as a function of temperature.⁶ Assuming a two-state model, f is given by

$$f = \frac{e^{-\Delta G/RT}}{1 + e^{-\Delta G/RT}} \quad (1.4)$$

with ΔG from equation 1.3. Alternatively, by noting that $\Delta H - T_m \Delta S = 0$ (see equation 1.3) ΔG can be expressed as

$$\Delta G = \Delta H \left(1 - \frac{T}{T_m} \right) \quad (1.5)$$

As an example, Figure 1.3 provides plots that were calculated on the basis of equations 1.4 and 1.5, with $\Delta H = 435$ kJ/mol and $T_m = 359$ K.

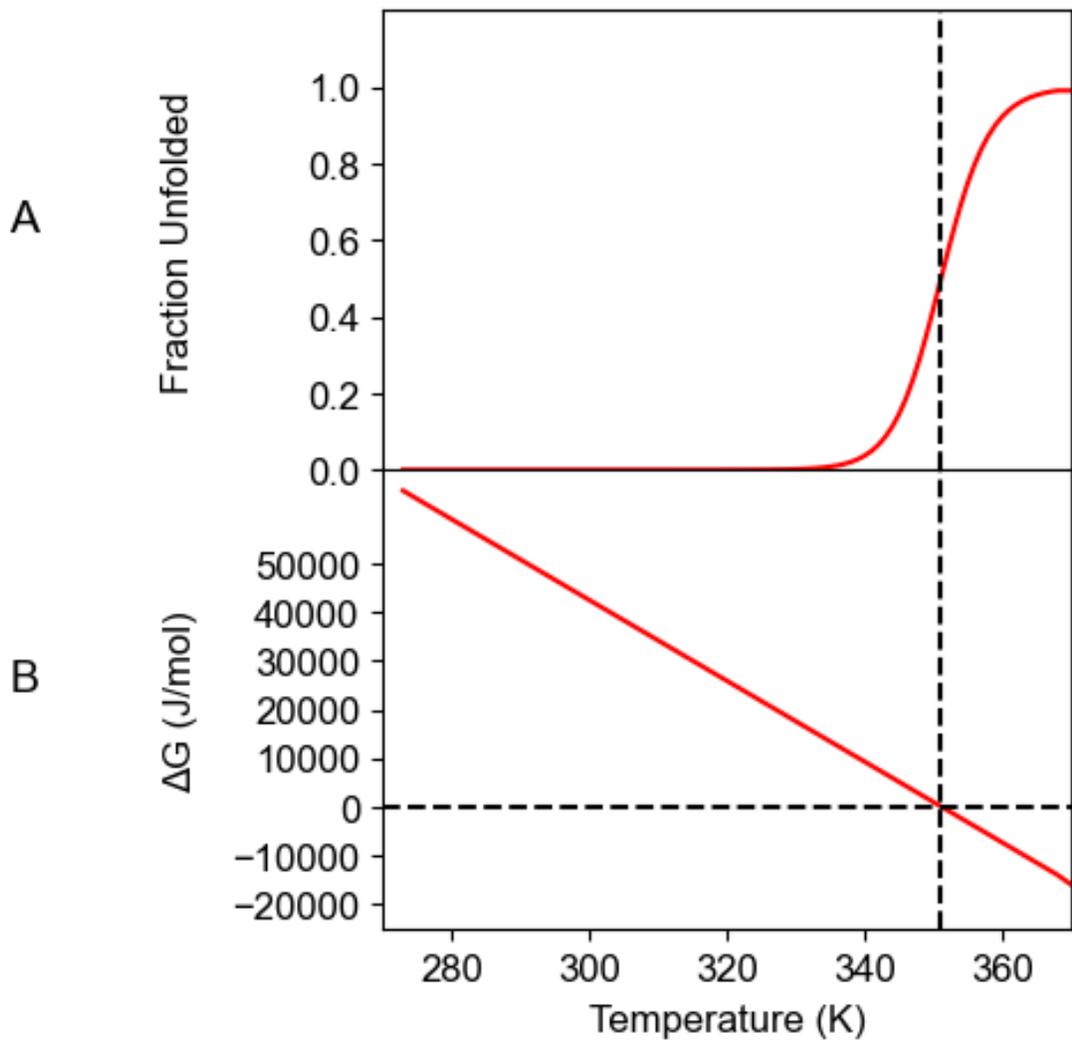


Figure 1.3. (A) Fraction of unfolded protein as a function of temperature.(from eq 4). (B) Free energy profile of unfolding (from eq 5) The x-intercept of $\Delta G(T) = 0$ corresponds to T_m , highlighted by the vertical dashed line.

1.2.2 Chemical Denaturants

In chemical denaturation, proteins are usually exposed to urea or guanidinium chloride to induce unfolding. The concentration required to achieve complete unfolding is typically around 8M for urea and 6M for guanidinium chloride.⁶ Despite numerous studies, the exact mechanism of action for these denaturants remains elusive. However, it is thought that both urea and guanidinium chloride increase the solubility of polar and non-polar side chains, reducing the magnitude of the hydrophobic effect.⁶ In other words, polypeptide chains appear to interact more favourably with denaturants compared to pure water. These interactions are maximized when the protein adopts an unfolded conformation.⁶

Modifying equation (1.1) and (1.2), the stability of proteins can be expressed as

$$\Delta G = -RT \ln K_{eq} = -RT \ln \frac{[U]}{[N]} \quad (1.6)$$

In the case of chemical denaturation, it has been found empirically that the free energy of unfolding is linearly dependent on the concentration of urea or guanidinium chloride⁶

$$\Delta G = \Delta G_{water} - m[D] \quad (1.7)$$

Where ΔG_{water} is the free energy in absence of any denaturants and $[D]$ is the denaturant concentration. Furthermore, m is the dependence of ΔG on $[D]$ and is related to the change in solvent-accessible surface area of the protein upon unfolding. Hence, the most relevant parameter for describing protein stability is ΔG_{water} . This value is not possible to ascertain under native conditions, where the concentration of $[U]$ is extremely small.⁶ Exposing the protein to an increasing concentration of chemical denaturants, spanning from 0M to 6M, allows experimentalists to determine ΔG_{water} from a linear extrapolation of ΔG vs D Figure(1.4).⁶ Also it can be noted that

$$\Delta G_{water} = m[D]_{50} \quad (1.8)$$

Where $[D]_{50}$ is a transition midpoint where $[U]=[N]$ and $\Delta G=0$. Like in thermal unfolding studies, the transition midpoint is not necessarily correlated to the thermodynamic stability of the protein in the absence of denaturant. For instance, large proteins tend to unfold at lower denaturant concentrations.⁶

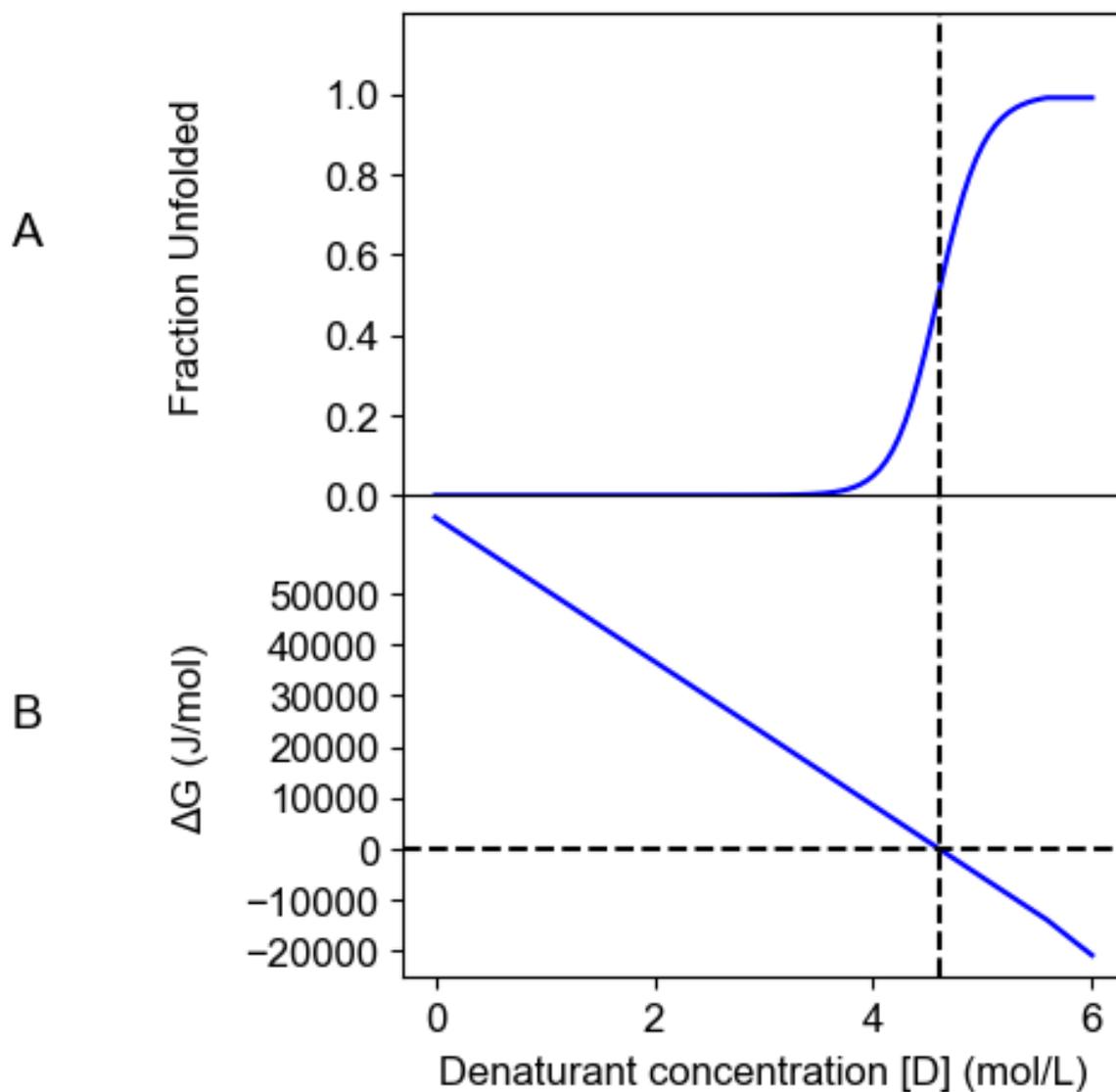


Figure 1.4. (A). Fraction of unfolded protein as a function of denaturant concentration from equation (1.4). (B) Free energy profile of unfolding from equation (1.5). The x-intercept of $\Delta G[D] = 0$ corresponds to $[D]_{50}$, highlighted by the vertical line. ΔG_{water} can be elucidated using equation (1.8) where m is the slope of the free energy profile.

1.3 Conventional Methods for Protein Structural Analyses

There are several methods that can be utilized to study protein structures. These include ultra-violet visible absorption spectroscopy (UV-Vis), circular dichroism spectroscopy (CD), Nuclear magnetic resonance (NMR) spectroscopy, X-ray crystallography, and cryo-electron microscopy. Elucidating protein structure is a crucial aspect of protein chemistry as it is widely accepted that a protein function is dependent on its structure. This section will discuss fundamental aspects of these methods and highlight their pitfalls.

1.3.1 UV-VIS Absorption Spectroscopy

Absorption spectroscopy is a technique used to measure the interaction of protein-associated chromophores with electromagnetic radiation.^{13,14} Typical UV-Vis experiments utilize a spectrophotometer operating in the ultraviolet (150-400nm) and visible (400nm-800nm) range.^{13,14} The light is passed through a monochromator and is focused into a cuvette. The intensity of the transmitted light is then detected by a photomultiplier tube.¹³ Photon absorptions by the chromophore can excite its electrons from their ground state to an excited state. This happens only if the energy of the photons matches the energy gap between those two levels. Hence, the wavelength and the intensity of absorption depends on the chemical nature and the molecular environment of the chromophore.¹³ For instance, proteins typically exhibit an absorption maxima between 275 nm and 280 nm due to aromatic amino acids (tyrosine and tryptophan).¹³ In addition, some proteins such as myoglobin contain heme cofactors that display absorption in the visible region around 400 nm, known as the soret band.

The absorption of a chromophore is linearly dependent on its concentration C ; thus, UV-Vis spectroscopy is well suited for concentration measurements.^{13,15} The relationship between concentration and absorption can be described by the Beer Lambert law

$$A = \log \frac{I_0}{I} = \epsilon lc \quad (1.9)$$

where A is the absorption of a chromophore, I_0 is the intensity of light, at measured wavelength, passing through a reference cell, while I is the intensity of light passing through a chromophore. ϵ is the molar absorptivity constant and is defined as how strong a chromophore absorbs light at a specific wavelength. Finally, l is the length of the cuvette. Equation (1.9) conveys that absorption is directly proportional to concentration.

UV-VIS spectroscopy usually does not provide in-depth structural information.¹⁶ For example, the Soret band of myoglobin is red shifted when the heme Fe is reduced to Fe(2+), reflecting changes in the electronic structure of heme Figure(1.5).

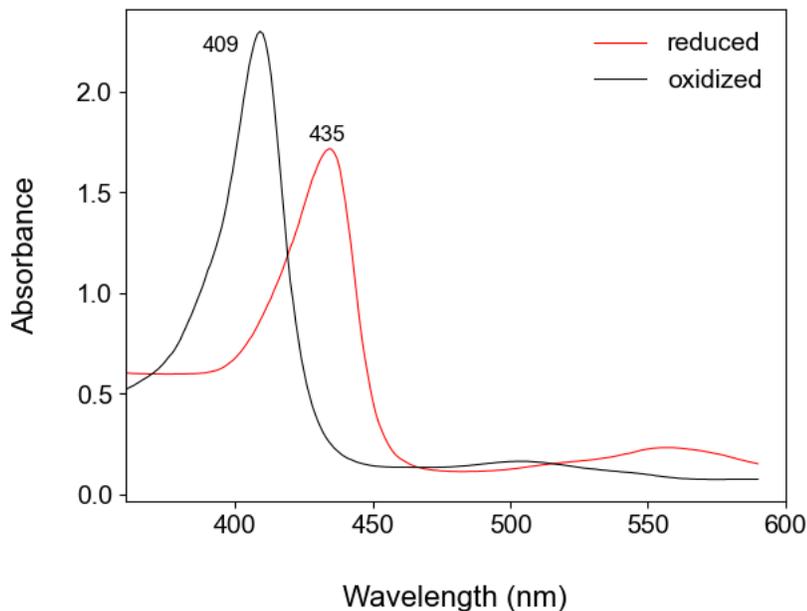


Figure 1.5. The absorption spectra of reduced myoglobin (Fe²⁺) in black and oxidized myoglobin in red (Fe³⁺). Reduced myoglobin has a maximum absorption at 435 nm while oxidized myoglobin displays a maximum peak at 409 nm.

1.3.2 Circular Dichroism (CD) Spectroscopy

CD spectroscopy is one of the most utilized tools for studying protein structures and unfolding transitions. Linearly polarized light has an electric field that oscillates sinusoidally in a single plane.¹⁷ This wave can be visualized as a combination of two vectors of equal length, forming circular traces where one rotates clockwise (E_R) and the other rotates counter-clockwise (E_L).¹⁷ In the CD spectrometer, these two components (E_L and E_R) are separated. When chiral molecules interact with circularly polarized light, they may absorb E_L and E_R differently.^{17,18} In CD experiments, a chiral chromophore is alternatively exposed to equal amounts of E_L and E_R of selected wavelengths, resulting in a CD spectrum that represents the differential absorption (molar ellipticity) of a molecule.

The amide backbone of proteins acts as a chromophore and shows differential absorption in the far-UV range(180-250nm).¹⁷ Importantly, proteins display characteristic CD signals depending on

the type of secondary structures they possess Figure(1.6). For example, α -helical proteins show a strong negative band near 222 nm due to $n \rightarrow \pi^*$ electronic transitions.¹⁹ In contrast, β -sheets display a negative band near 216 nm and a positive band near 200 nm.¹⁹ One application of CD spectroscopy is estimating the fraction of residues involved in α -helix or β -sheets.¹⁹ For instance, the α -helical content of a protein can be correlated with the CD signal at 222nm.¹⁹ This makes CD spectroscopy a reliable tool in protein unfolding studies, allowing practitioners to monitor changes in secondary structures and obtain thermodynamic parameters. However, this method does not provide site-specific structural data and needs to be complemented with other techniques like NMR or X-ray crystallography for high resolution data.¹⁹

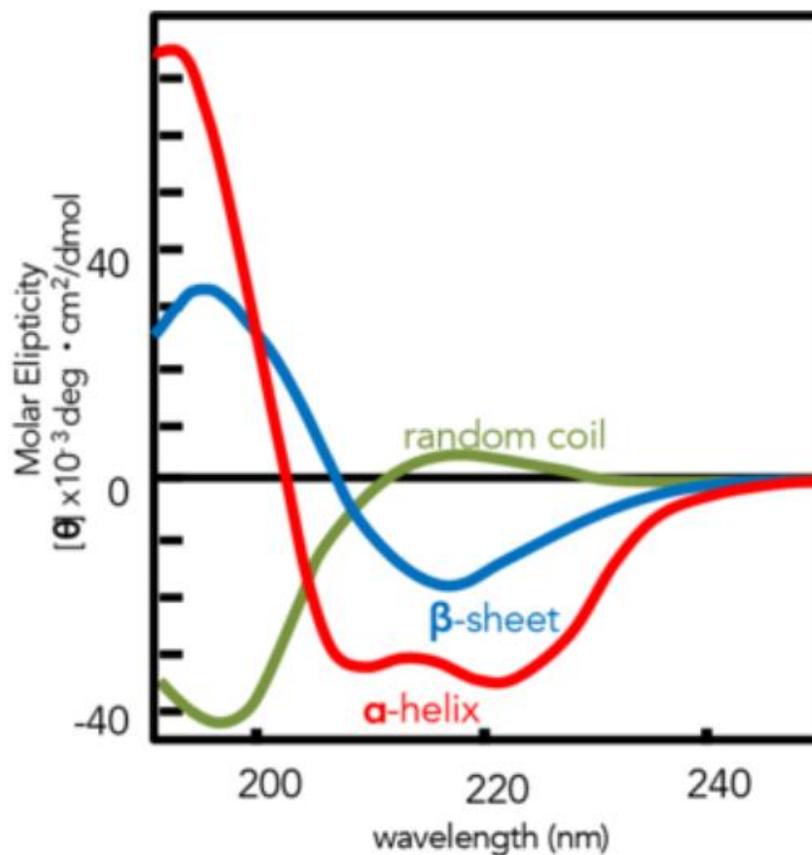


Figure 1.6. Characteristic CD signals for principal polypeptide secondary structures (α -helix, β -sheets, and random coil). This was adapted from reference 20.

1.3.3 Nuclear Magnetic resonance (NMR) Spectroscopy

NMR spectroscopy is a useful technique that can provide detailed structural information at atomic resolution and insights into protein dynamics. In general, NMR utilizes active nuclei or those with $\frac{1}{2}$ spin nuclei, which possess magnetic dipoles capable of aligning with external magnetic fields.²¹ Once aligned with a magnetic field, the nuclei can adopt two conformations. One corresponds to the lowest energy level (parallel to the magnetic field), while the other corresponds to a higher energy level (anti-parallel to the magnetic field). Atomic nuclei under the effect of magnetic field (B_0) have a magnetization vector (M) that precess around an axis parallel to B_0 (z-axis).^{21,22} The frequency of this precession (Larmor frequency) is specific to each nucleus.²¹ Consequently, nuclear excitation is achieved by applying a strong radio frequency field for a short period of time. As a result, the precession will shift away from the z-axis in the direction of the y-axis. After the pulse, only the magnetic field B_0 acts upon M .²¹ The time signal induced in the detector coil through the precession of the x,y component fades away through relaxation.²¹ This time signal is converted to a frequency signal through Fourier transformation.

In contrast to optical methods, NMR provides information on the location of secondary structures within a protein sequence. For instance, $^1\text{H}\alpha$ chemical shifts can be used to identify regions of secondary structures where helical conformations produce upfield shifts and β -sheets produce downfield $^1\text{H}\alpha$ chemical shifts.¹⁹ In addition, NMR techniques like 2D NMR and NOESY (nuclear Overhauser effect spectroscopy) have been proven useful in elucidating protein structures.²³ Despite this, NMR spectroscopy has some drawbacks that can hinder protein structural studies. For instance, proteins larger than 30 kDa are usually investigated by using labor-intensive methods like TROSY (transverse relaxation optimized spectroscopy) in conjunction with side chain deuteration, whereas the structure of smaller proteins can be elucidated using simple homonuclear ^1H NMR experiments.^{23,24}

1.3.4 X-ray crystallography

X-ray crystallography is commonly used for precise structural determination. Its great contribution to structural biology can be exemplified by the fact that most protein structures in the protein databank have been generated using this method. The basic principle of x-ray crystallography involves subjecting purified crystals, present at high concentrations, to beams of x-ray.^{25,26} The resulting diffraction patterns are then analyzed to generate a 3D map, representing

time and space-average of the electron density of a molecule.²⁵ The refined model derived from the electron density map contains the cartesian coordinates for each atom. Despite its versatility, x-ray crystallography has many limitations. For instance, hydrogen atoms with lone electrons are not visible in electron density maps.²⁵ Furthermore, x-ray crystallography requires prior knowledge of chemical entities present in the crystal, as it cannot distinguish between carbon, nitrogen, and oxygen.²⁵ Another drawback is its inability to capture highly flexible regions due to poor electron density.²⁵

1.3.5 Cryogenic electron microscopy (cryo-EM)

Recent developments in electron microscopy coupled with advances in modelling softwares have established cryo-EM as a highly effective tool in biology.²⁷ Similar to X-ray crystallography, Cryo-EM enables precise structural determination, generating 3D models at unprecedented resolution.²⁷ Cryo-EM experiments begins with placing a molecule on a specialized grid and flash-freezing it in a thin sheet of amorphous ice. 2D images are obtained by applying a beam of electron that is accelerated down a microscope column and are scattered by the sample. These scattered electrons are focused into an electromagnetic lens.²⁸ The resulting 2D images are incredibly noisy and yield little structural information. To enhance the resolution, a large number of projection images, featuring the molecule of interest in different orientations, are computationally averaged to generate a 3D model.²⁷ This allows the construction of macromolecules with a resolution rivalling those obtained using X-ray crystallography. For instance, one study managed to determine the structure of ferritin (iron-storing protein) to a resolution of 1.24 Å.²⁹ Although cryo-EM has contributed significantly to structural biology, it remains inaccessible for many institutions due to the high cost associated with operating the instrument.²⁷ Furthermore, the inherent structural flexibility of many macromolecules presents a challenge in achieving high-resolution models. However, advances in sample preparation combined with new algorithms may hold a potential for probing the conformational landscapes of such disordered molecules.²⁷

1.4 Protein Chemical Labeling

Many of techniques discussed above are quite time consuming and labor-intensive, and most of them are applicable only to highly ordered native protein structures in a certain size range. There is still a need for complementary techniques that can interrogate specific aspects of structure and dynamics, especially for proteins that are not amenable to the aforementioned high-resolution techniques e.g., because of a high degree of disorder, or their short lifetimes (in the case of folding intermediates).

Chemical labelling of proteins in solution in conjunction with mass spectrometry (MS) fills an important niche in this context by probing solvent-accessible surfaces, conformational changes, and folding pathways.³¹ In other words, chemical labels function as sensors for perturbations in structure or dynamics.³¹ Changes in labelling behavior can be monitored by measuring protein or peptide mass shifts after exposure to the labeling agent.³¹ The three most common labeling methods are covalent labeling, covalent cross-linking, and H/D exchange.³⁰ These three labeling techniques are somewhat complementary to each other, and one can benefit from combining different methods to gain high-resolution data.³⁰

Analysis of labeled proteins can be achieved by utilizing MS, which enables the detection of mass shifts associated with chemical labeling. One simple approach involves taking an intact mass spectrum of a labeled protein. This provides insights into the number of modifications generated by a labelling reagent. However, this simple approach offers no spatial resolution and is seldom used. Instead, chemically labelled proteins are usually enzymatically digested, and the resulting peptides are analyzed by MS.³⁰ This method is typically used in conjunction with reverse-phase liquid chromatography (RP-LC) where digested proteins are loaded onto a trapping column for desalting.³⁰ Subsequently, peptides are separated on an analytical C-18 column and identified using MS. Tandem MS mapping allows for the determination of amino acid residues that have been labeled. Modified amino acids are identified by finding proteolytic peptides that differ in mass from a predicted value.

Covalent labeling is used to tag surface-accessible side chains. Multiple approaches have been used in this context, some of which are highly specific for certain side chains whereas others are rather non-selective.³² One avenue to introduce these labels involves hydroxyl radical ($\cdot\text{OH}$) which can be generated by a variety of methods, including Fenton reactions and laser induced

photochemical H_2O_2 .³³ Exposure to $\cdot\text{OH}$ results in the oxidation of side chains. Common oxidation reactions include the insertion of oxygen in aliphatic side chains, or the formation of a carbonyl groups.²⁹ for example, the reaction of arginine with $\cdot\text{OH}$ would yield a major oxidation product with a mass shift of -43 Da, resulting from the loss of a guanidino group and the formation of a carbonyl group. Furthermore, $\cdot\text{OH}$ is a nonspecific reagent where the extent of labelling is governed by solvent accessibility and the intrinsic reactivity of each amino acid.³³ In contrast, amino acid-specific labels rely on the reactivity of specific functional groups. For instance, vicinal dicarbonyl compounds are used to label arginine, forming cyclic products.³³ Arginine labeling is a very valuable probe in studying protein interactions as it is thought to be a key residue in protein interfaces.³³ This can be attributed to the fact that arginine is involved in multiple types of favourable interactions like hydrogen bonds, salt bridges, and hydrophobic interactions. Applications of covalent labeling involve probing local domain structures for changes in solvent accessibility.³³ For instance, protein-ligand binding can be probed using covalent labeling. In this approach, reactivity of side chains is compared in the ligand-bound and free protein where the former is expected to have side chains that are protected from chemical modification due to solvent inaccessibility.³⁴

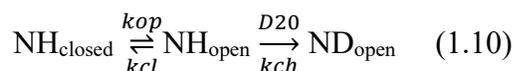
1.4.1 Protein Cross Linking

Another labeling method involves covalently cross-linking spatially proximate side chain pairs. This method can provide distance constraints and reveal information on subunit topology.³⁵ Cross-linkers of various lengths and selectivity for functional groups have been designed to link different reactive groups over defined distances.³⁵

While covalent labelling and cross-linking can provide valuable insights into protein structure and dynamics, there is a general concern that these modifications can perturb protein structures.^{32,36} Hence, there is great interest in developing labelling methods that are more 'benign'.

1.4.2 Hydrogen-Deuterium Exchange

Hydrogen-deuterium exchange (HDX) is one of the most popular protein labeling methods, it is widely believed that this method causes only minimal structural perturbations.³⁷ For HDX experiments, proteins are incubated in D₂O, and labile hydrogens (those bonded to oxygen, nitrogen, or sulfur) are exchanged with deuterium.³⁸ HDX is typically used to monitor structural dynamics, binding events, or conformational changes.³⁸ Backbone amide hydrogens in the protein backbone act as reporters, as they exchange for deuterium over various time intervals.²¹ The exchange rate can be monitored by mass shifts that manifest themselves as changes in isotopic distributions.³⁹ Alternatively, HDX can be monitored using NMR where the amide proton signal disappears as it is exchanged with deuterium.⁴⁰ Side chain hydrogens are not usually monitored as they undergo rapid back-exchange prior to detection.³⁸ In native proteins, most backbone sites are engaged in hydrogen bonds. Thus, HDX occurs via transient opening/closing fluctuations that disrupt hydrogen bonds and provide access to NH sites.³⁹ This can be represented by the widely accepted Linderstrøm-Lang model



Here k_{op} and k_{cl} are the opening and closing rate constants, while k_{ch} is the intrinsic exchange rate constant, which is determined by steric and electronegativity effects of neighboring chains. In addition, temperature is shown to affect k_{ch} and should be controlled in HDX experiments.⁴¹ The ionization constant of deuterium oxide ($K_{\text{D}_2\text{O}}$) is correlated with the solution temperature, affecting the concentration of D_3O^+ and OD^- .⁴¹ A theoretical value for k_{ch} can be computed using the Arrhenius equation

$$k_{ch}(T) = k_{ch}(293) \exp\left(\frac{-E_a}{R} \left(\frac{1}{T} - \frac{1}{293}\right)\right) \quad (1.11)$$

In this equation, T refers to the experimental temperature, $k_{ch}(293)$ is the intrinsic exchange rate at room temperature, E_a is the activation energy, and R is the molar gas constant. Equation 1.10 implies that k_{ch} increases 10-fold with every 22 °C increase in temperature.⁴¹ While this relationship appears to be maintained through a wide range of temperatures, it is important to consider that temperature likely affects the structure and dynamics of proteins. Another factor that governs k_{ch} is pH. Controlling solution pH (or pD) is fundamental to HDX experiments where the sensitivity of amide exchange rate is often utilized for analysis by MS.⁴¹ In typical HDX

experiments, proteins are labelled under near physiological conditions (pH 7).^{39,41} Under these conditions, labeling is fast which enables the probing of diverse protein motions. In contrast, MS analyses are typically carried out under acidic condition where the rate of HDX is quenched. This suppresses the loss of deuterium label caused by the presence of H₂O in subsequent steps. The dependence of k_{ch} on pH is approximated in Figure (1.7), which implies that HDX is quenched at pH 2.5. Other factors that affect chemical rate constant include pressure and ionic strength.⁴¹

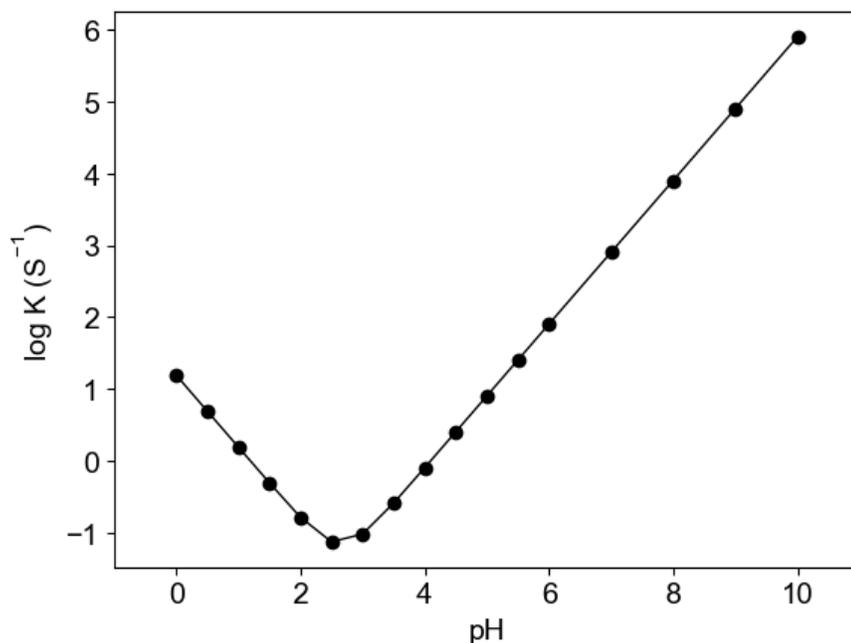


Figure 1.7. Dependence of the chemical rate constant (k_{ch}) on pH. This figure is based on reference 41.

In neutral solutions, HDX is base-catalyzed. In the commonly encountered EX2 regime, and proceeds with a rate constant k_{ex} described by

$$k_{ex} = \frac{k_{op}}{k_{cl}} \times k_{ch} \quad (12)$$

where $k_{op} \ll k_{cl}$. Amide hydrogens can exchange with $k_{ex} = k_{ch}$ only if they are solvent-accessible and free from hydrogen bonding.³⁹ Hence, the presence of secondary structure results in $k_{ex} \ll k_{ch}$. The extent of this reduction in the exchange rate can be expressed as a protection factor $P = k_{ch}/k_{ex}$.³⁹

HDX is often thought to cause minimal changes to the properties of native proteins, particularly when compared to the covalent labeling techniques discussed in the preceding section. Nonetheless, there is evidence that the substitution of H₂O for D₂O can affect some aspects such as the thermal unfolding behaviour many proteins.³⁷ (see Next section)

1.5 Deuterium-Isotope Effects

Water plays an important role during protein folding and for stabilizing the native state.⁶ After all, one of the major factors that stabilizes native structures is the hydrophobic effect i.e., the tendency of nonpolar amino acids to get buried in the core where they are shielded from water.⁶ Conversely, hydrophilic side chains are on the protein surface where they can favourably interact with water.⁶ Naturally occurring water, (H₂O) contains protium (¹H). In contrast, heavy water (D₂O) contains deuterium (²H). Various experimental techniques like NMR and infrared spectroscopy utilize D₂O as solvent since H₂O interferes with the protein signal.³⁹ In addition, HDX studies label proteins by incubating them in D₂O.³⁸ An important point that tends to be overlooked is that the physiochemical properties of D₂O differ from those of H₂O, and these differences may introduce a bias in structural studies. For instance, the substitution of H₂O for D₂O was shown to affect the melting temperature of several proteins, stabilizing their native fold by few degrees.³⁷ The origins of this effect are non-trivial, but stabilization is typically attributed to two phenomena: (1) solvent effects caused by the different properties of D₂O or (2) an increase in the hydrogen bond strength in D₂O vs H₂O. The following paragraph examines these factors in more detail.

Several studies reported an increase in melting temperature T_m for proteins in D₂O.^{37,43,44} Some authors have proposed that heavy water stabilizes native protein structures by increasing the strength of the hydrophobic effect. This idea goes back to a study by Scheraga et al. in 1965.⁴⁵ The authors analyzed the solubility of non-polar amino acids in both solvents and found that the free energy of transfer from H₂O to D₂O was positive (unfavourable). This indicated that heavy water is a poorer solvent for non-polar amino acids. Several calorimetry studies reported an increase in enthalpy of unfolding for proteins in heavy water.⁴³ The molecular basis behind this change in enthalpy is difficult to pinpoint, as both hydrogen bonds and the hydrophobic effect have enthalpic contributions.^{44,45}

Other studies attributed the enhanced stability of proteins in heavy water to an increased strength of deuterium bonds.^{46,47,48} For instance, calculations have shown that the zero-point energy of D₂O is lower than that of H₂O. This difference can be traced back to bending vibrations that displace the bridging atom away from the hydrogen bonding axis.⁴³ The heavier mass of deuterium lowers the frequency of these vibrations, decreasing their zero-point energy. Furthermore, a study by Cuma et al. found that replacing a bridging protium atom with deuterium can increase the binding energy by 0.92 kJ/mol.⁴⁶ This is consistent with observations that D₂O reduces protein flexibility.⁴⁶ Overall, the origin of protein stability enhancements in D₂O is poorly understood. This is highlighted by a quote from a recent paper that stated “*To the best of our knowledge, a reliable rationalization of the D₂O stabilizing effect has not yet been provided, even though it has been proposed that heavy water stabilizes the N-state because it increases the strength of the hydrophobic effect*”.⁵⁰

1.6 Mass spectrometry

Mass spectrometry (MS) has become one of the most widely used tools in biochemistry. Its speed, sensitivity, and ability to discern coexisting species is advantageous for analyzing complex mixtures.³⁰ The three main components of a mass spectrometer are the ion source, the mass analyzer, and the detector. The ion source is the compartment where gaseous ions are produced. These ions are separated by a mass analyzer based on their mass-to-charge values (m/z) and are quantified by a detector.⁵¹

Various ionization techniques are available for MS. For instance, electron ionization (EI) employs energetic electron beams that interact with gas phase molecules. The collisions between gaseous molecules and electrically accelerated electrons cause analytes to expel electrons, resulting in the formation of positively charged cations.⁵¹ While EI is a popular ionization method for GC-MS, it is not suitable for large non-volatile and thermolabile analytes such as proteins.⁵¹ Matrix assisted laser desorption ionization (MALDI) or electrospray ionization (ESI) is usually the standard for large biomolecules. In MALDI, ions are desorbed from the solid phase. The sample is first dissolved in a solvent and embedded in a matrix of UV or IR absorbing molecules.⁵¹ Subsequently, the sample is crystallized with the matrix and brought to the gas phase via a pulse of laser light that hits the sample-matrix crystal, leading to the desorption and ionization of the analytes. In contrast, ESI produces gaseous ions by spraying a sample solution

under a high electric potential, resulting in the formation of highly charged droplets (See next section).^{51,52} Both MALDI and ESI are “soft” ionization methods that transfer analytes into the gas phase without inducing unwanted fragmentations.⁵¹

A mass analyzer separates ions based on their m/z values. There are various mass analyzers that are commonly used in MS instruments. These include quadrupole (Q), orbitrap, time of flight (TOF), and Fourier transform ion cyclotron resonance (FT-ICR).⁵¹ These analyzers vary in term of resolution, price, and mass range. TOF analyzers will be the focus of this paragraph as it is used for this work. TOF mass analyzers rely on the flight of ionized ions in a tube of 1-2 meters in length where a short pulse of ions is accelerated by a fixed potential generated by a high-voltage pusher.^{51,53} The potential energy E_{pot} of each ion is converted into kinetic energy according to equation 1.13 and 1.14

$$E_{\text{pot}} = E_{\text{kin}} \quad (1.13)$$

$$ze\Delta U = \frac{1}{2}mv^2 \quad (1.14)$$

where z is the charge of the ion, e is the elementary charge, ΔU is the applied voltage, m is the mass of the ion and v is the speed of the ion. Equation 1.14 can be rearranged to

$$v = \sqrt{\frac{2\Delta Ue}{m/z}} \quad (1.15)$$

Equation 1.15 indicates that ions with lower m/z will reach the detector first. However, ions with similar m/z values might have slightly different kinetic energies, subsequently diminishing the resolution and broadening the arrival time of ions. This drawback is overcome by utilizing a reflectron which is a device that contains multi-stage electrostatic mirrors. The reflectron changes the paths of ions within the TOF such that ions with higher kinetic energy will travel further into the ion mirror (Figure 1.8).^{51,54} Hence, ions with the same m/z but different kinetic energies will reach the detector at the same time.⁵⁴

TOF mass analyzers usually employ a microchannel plate detector. A microchannel plate is made of semi-conductive materials and functions as a signal amplifier. Ions impacting the plate start a cascade of electrons that propagate throughout the channel. Hence, microchannel plates function as a dynode electron multiplier.⁵⁵ The electrical signal produced by the impact of

the ion is then converted to a digital signal by an Analog-to-digital-converter, where intensity vs m/z distributions can be obtained.⁵⁵

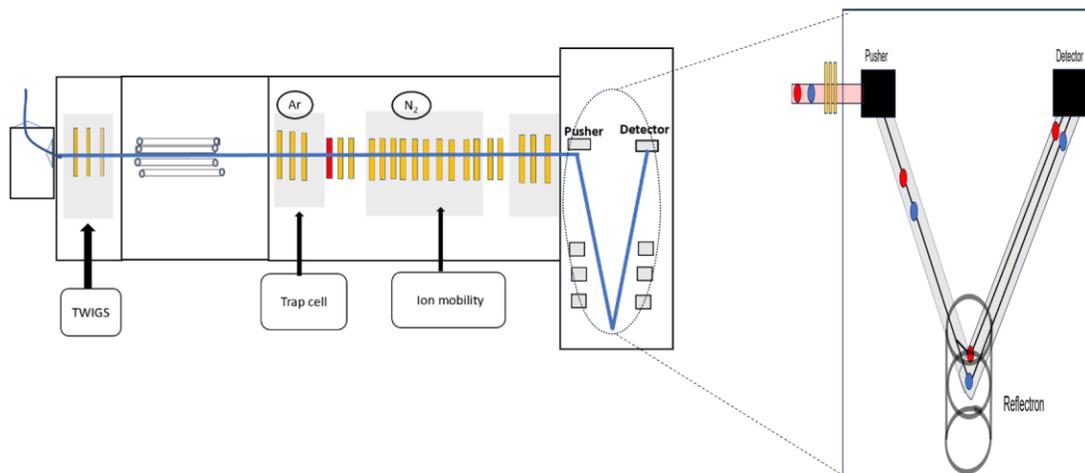


Figure 1.8. Schematic cartoon of an ESI mass spectrometer. The TOF reflectron is magnified for visual aid.

1.7 Electro spray Ionization

The advent of soft ionization techniques such as electrospray ionization (ESI) has enabled the characterization of proteins and other large molecules involved in biological processes. In ESI, there are three major steps in the production of gaseous ions from an analyte in solution (Figure 1.9).⁵² The ESI process occurs at atmospheric pressure, and the resulting ions enter the vacuum of the mass spectrometer.^{51,53} The ionization process starts with applying a voltage of +2-3 kV to a spray capillary, resulting in a high electric field at the capillary tip.⁵² This will cause water electrolysis, and the accumulation of a net positive charge as a result of excess H⁺ and other cations. All positive charges will be enriched near the meniscus, while anions will move away from the capillary tip.⁵² The resulting electrostatic forces distort the meniscus into a Taylor cone, and a fine jet emerges from the cone tip.⁵² The repulsion between charges in the jet causes it to break up into small, positively charged droplets. In the second stage, the charged droplets drift towards the mass spectrometer which represents the counter-electrode.⁵² Solvent evaporation leads to an increase in the electric field at the droplet surface. When the repulsion between charges overcomes surface tension, Coulomb fission yields smaller droplets. This evaporation/fission cycle can repeat itself through several generations. The final steps involve the release of gaseous ions

from nanometer-sized droplets.⁵² The exact mechanisms of these final steps are still under debate. Several pathways exist, discussed below.

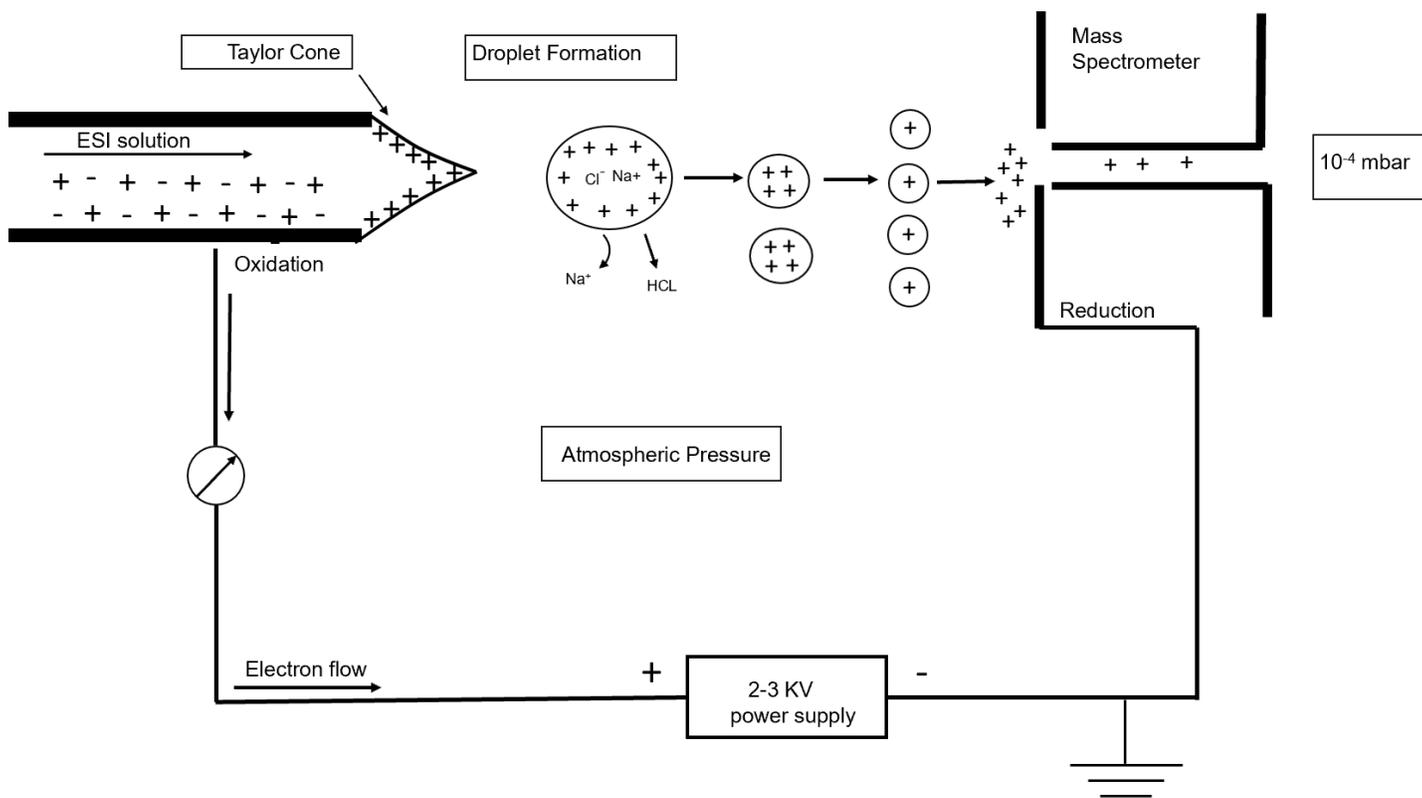
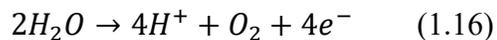


Figure 1.9. The major steps of ESI processes. This figure is based on reference 52.

1.8 Native ESI of Proteins

Two common ESI conditions can be distinguished, referred to as “denaturing” or “native”. These conditions reflect the solution environment experienced by proteins within the ESI capillary, and on their way into the vacuum of the mass analyzer.⁵⁶ Denaturing ESI is implemented by electrospraying proteins that are unfolded in bulk solution, usually by employing acidic pH and/or organic cosolvents.⁵⁶ Such denaturing conditions generate $[M+zH]^{z+}$ protein ions with high charge states z . These protein ions are electrostatically ejected from the shrinking droplets according to the chain ejection model (CEM). Conversely, native ESI utilizes aqueous solutions at pH 7, generating protein ions in much lower z -values, close to the Rayleigh charge of protein-sized water droplets. Protein charge originates from small cations like Na^+ , H^+ , which are present in excess compared to their counterions.⁵⁶ Protons are generated from water electrolysis at the metal/liquid interface of the ESI emitter. In contrast, metals such as Na^+ are ubiquitous in biological samples.⁵⁶ The charge state distributions can provide structural information on proteins in solution.^{56,56,57} Protein ions under native ESI conditions likely form according to the charged residue model (CRM). This model postulates that ESI-generated droplets will undergo cycles of fission and evaporation until there is one analyte per droplet.³⁰ Protein CRM is accompanied by the occasional ejection of charge carriers (such as Na^+ ions), keeping the shrinking droplets close to the Rayleigh limit. Furthermore, folded proteins remain in the droplet interior throughout the evaporation process. This reflects the tendency of polar side chains to maximize solvation by aqueous environment.⁵⁶ Evaporation of the final water layers eventually releases protein ions into the gas phase, and charge carriers that were still present during the final stages of evaporation associate with acidic side chains.⁵⁶ Hence, Native ESI experiments are typically conducted in the absence of non-volatile salts to avoid adducts that may associate with titratable side chains.⁵⁶ Ammonium acetate is usually the additive of choice for native ESI, allowing some background electrolyte to mimic physiological conditions. The compatibility of ammonium acetate with ESI experiments is attributed to its volatility as it evaporates during the final ESI stages without forming any adducts. NH_4^+ that initially binds to the surface of a nascent protein droplet is thought evaporate as NH_3 , leaving a proton behind. Although ammonium acetate produces a pH 7 solution, it has very little buffering capacity in neutral solution.

Inadvertent pH alterations during the ESI process can negatively affect the outcome of the experiment, causing proteins to unfold in solution prior to ESI. pH changes during the ESI process can be attributed to the water electrolysis



Water oxidation is the main charge balancing process in positive ion mode, resulting in acidification of the droplets at the ESI emitter.⁵⁷ The magnitude of this pH drop has been demonstrated to be significant from near-neutral range to pH 3.⁵⁸ Additional pH changes can take place within the shrinking droplet. These pH changes are consistent with the CRM where initial droplets undergo cycles of fission and evaporation, and the pH decreases as droplets shrink.⁵⁷ These pH alterations can induce unwanted structural perturbation. For instance, several studies have reported discrepancies between solution and gas-phase behaviour for metal complexes, which was attributed to ESI-induced pH artifacts.⁵⁹⁻⁶¹ Native ESI experiments and some of its considerations will be the premise of chapter 2.

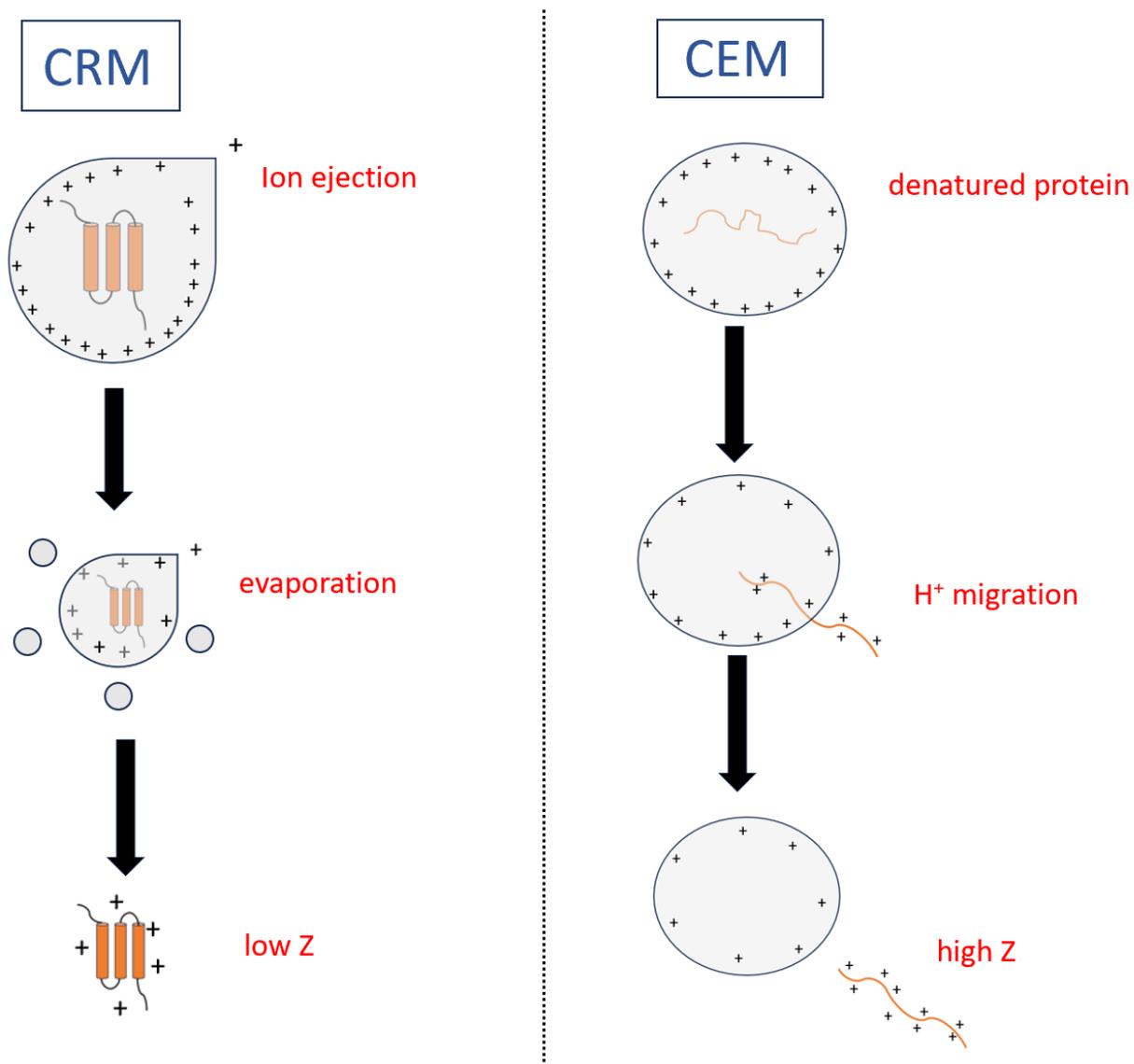


Figure 1.10. The different Ionization mechanism for native and denatured. Native protein undergoes CRM where nascent droplets endure cycles of fission and evaporation with the occasional ejection of charge carriers like Na^+ and H^+ . Evaporation of the final water layers release gaseous Ions with Low Z. In contrast, denatured proteins undergo CEM. This is based on reference 56

1.9 Protein Structures in the Gas Phase

Protein ions produced by native ESI retain conformations close to their native solution structures.^{30,62,63} This might be surprising because the aqueous environment represents a key stabilizing factor for native proteins.⁶ The lack of water might suggest that hydrophobicity is not a major contributor to protein stability in the gas phase.⁵⁶ In general, there can be substantial differences between solution and solvent-free equilibrated structures. For instance, MD simulations have revealed that equilibrated gas-phase proteins can convert to “inside-out” structures, where hydrophilic residues assemble in the core, while hydrophobic residues are exposed to the surface.⁶⁴ Such large-scale rearrangements involve breaking of multiple intramolecular interactions, which is energetically less favorable in vacuum.^{64,65} (Figure 1.11) In contrast, breaking of noncovalent contacts in solution is partially compensated for by formation of intermolecular contacts with water.⁶⁴ Hence, transition states in the gas phase are quite high in energy, and the corresponding transitions tend to be slow.⁶⁴ Experimental evidence suggests that gas phase equilibration where new noncovalent bond formation takes place may take up to minutes or hours.⁶⁵ This shows that in the absence of collisional activation the structures of ESI generated ions are metastable and are more likely to resemble solution like structures.^{64,66}

Furthermore, the fact that electrosprayed protein ions carry a net charge z suggests that their internal electrostatics are dominated by repulsive forces.⁶⁶ The change in dielectric constant from water ($\epsilon \approx 80$) to vacuum ($\epsilon \approx 1$) strengthens electrostatic interactions.⁶⁶ This implies that hydrogen bonds between backbone atoms are largely retained. Additionally, salt bridges (zwitterionic BH^+/A^- contacts) should be much more stable in vacuo and may be a key contributor to the retention of protein compact structures.⁶⁵ However, the mobile nature of H^+ can annihilate some of these zwitterionic contacts, such that H^+ transfer can convert salt bridges into neutral and weaker B^0/HA^0 contacts.⁶⁷ Furthermore, comparisons of protein crystal structure with MD generated structures of gaseous proteins revealed that titratable side chains formed a tightly connected network at the protein surface, involving numerous hydrogen bonds among neutral moieties.⁶⁷ This suggests that hydrogen bonds and electrostatic interactions likely play a major role in enhancing the kinetic stability of protein native-like structures after ESI. For instance, computational studies on cytochrome c have shown the formation of new electrostatic interactions between charged side chains such as protonated lysine or arginine and deprotonated glutamic or

aspartic acid.⁶⁸⁻⁶⁹ These electrostatic interactions result in the collapse of charged side chains onto the protein surface, an effect that may transiently stabilize the protein native fold in the gas phase.³¹

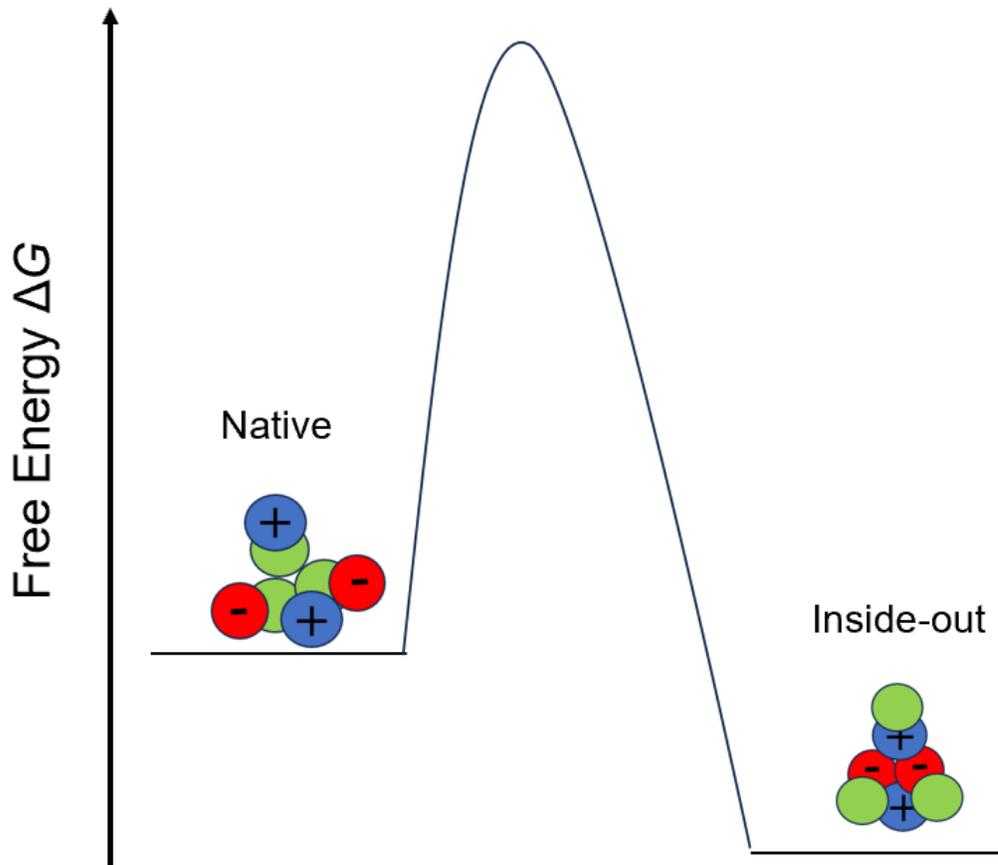


Figure 1.11 Schematic free energy difference diagram in the gas phase. Protein chains are shown with hydrophobic residues (green) and hydrophilic residues (blue/red). Native-like structures are not thermodynamically stable. Instead, inside-out structures are intrinsically favored in the gas phase, but there are large activation barriers that prevent such transitions. This phenomenon is known as kinetic trapping.

1.10 Ion Mobility Spectrometry

MS is often coupled with ion mobility spectrometry (IMS) to characterize biomolecular conformations in more detail.³⁰ IMS separates gaseous ions based on shape and charge. In its simplest setup (drift-tube IMS) ions are introduced into a drift tube that is filled with an inert buffer gas, typically helium. Ions are then exposed to a uniform electric field, propelling them along the drift tube.³⁰ Smaller compact ions undergo fewer collisions with helium, causing them to travel faster than bigger ions. Hence, the time t_d of an ion depends on its collision cross section (Ω).⁷⁰ The recorded t_d can be converted to Ω using the relationship.

$$\Omega = z \times C \times t_d \quad (1.17)$$

Where C is given by Mason-Schamp equation

$$C = \frac{eE}{16NL} \left(\frac{18\pi}{\mu k_B T} \right)^{\frac{1}{2}} \frac{760 \text{ Torr}}{P} \frac{T}{273.2 \text{ K}} \quad (1.18)$$

with the reduced mass μ , Boltzmann constant k_B , gas pressure P, and elementary charge e .⁷⁰ This method of determining collision cross sections provides structural insights into the structure of electrosprayed proteins in the gas phase. IMS offers a way to separate different protein conformations that cannot be resolved by MS alone.⁷⁰ Thus, IMS has become commonplace in research with hundreds of publications every year. Much of the rapid growth of IMS is linked to the commercialization of a slightly different approach by Waters Inc. This technique is referred to travelling wave IMS (TWIMS).⁷⁰ In this technique, ions pass through a series of stacked-ring electrodes, where radio frequency voltage is applied to radially confine the ions.⁶³ In addition, transient DC voltage is applied to each electrode in succession, creating a travelling wave that propels the ions through a gas-filled mobility cell.⁷⁰ Simply speaking, the ions “surf” on these waves. IMS separation is achieved when higher mobility ions are carried with the wave while lower mobility ions interact more strongly with the gas, resulting in the occasional roll-over events (Figure 1.11).⁷⁰ Unfortunately, the nature of the periodic electric field distorts the relationship between Ω , and t_d as shown in equation 1.18.

$$\Omega = z \times F \times t_d^B \quad (1.18)$$

The constants F and B cannot be calculated from first principles.⁷¹ Hence, TWIMS data have to be calibrated using drift tube reference values Ω_{ref} . This requires calibrants that exhibit similar

chemical properties as the analytes of interest. For instance, a calibrant mix of ubiquitin, cytochrome c, and myoglobin is commonly used.⁷¹ TWIMS calibration is usually performed by exposing the calibrant mix to high collisional activation. Using gentle conditions may lead to systematic errors as it is unclear whether the calibrant ions in TWIMS and DTIMS share similar gas-phase conformations.⁷¹ The harsh conditions employed in calibration ensure that ions in both methods share similar expanded conformations, allowing the accurate determination of Ω .⁷¹

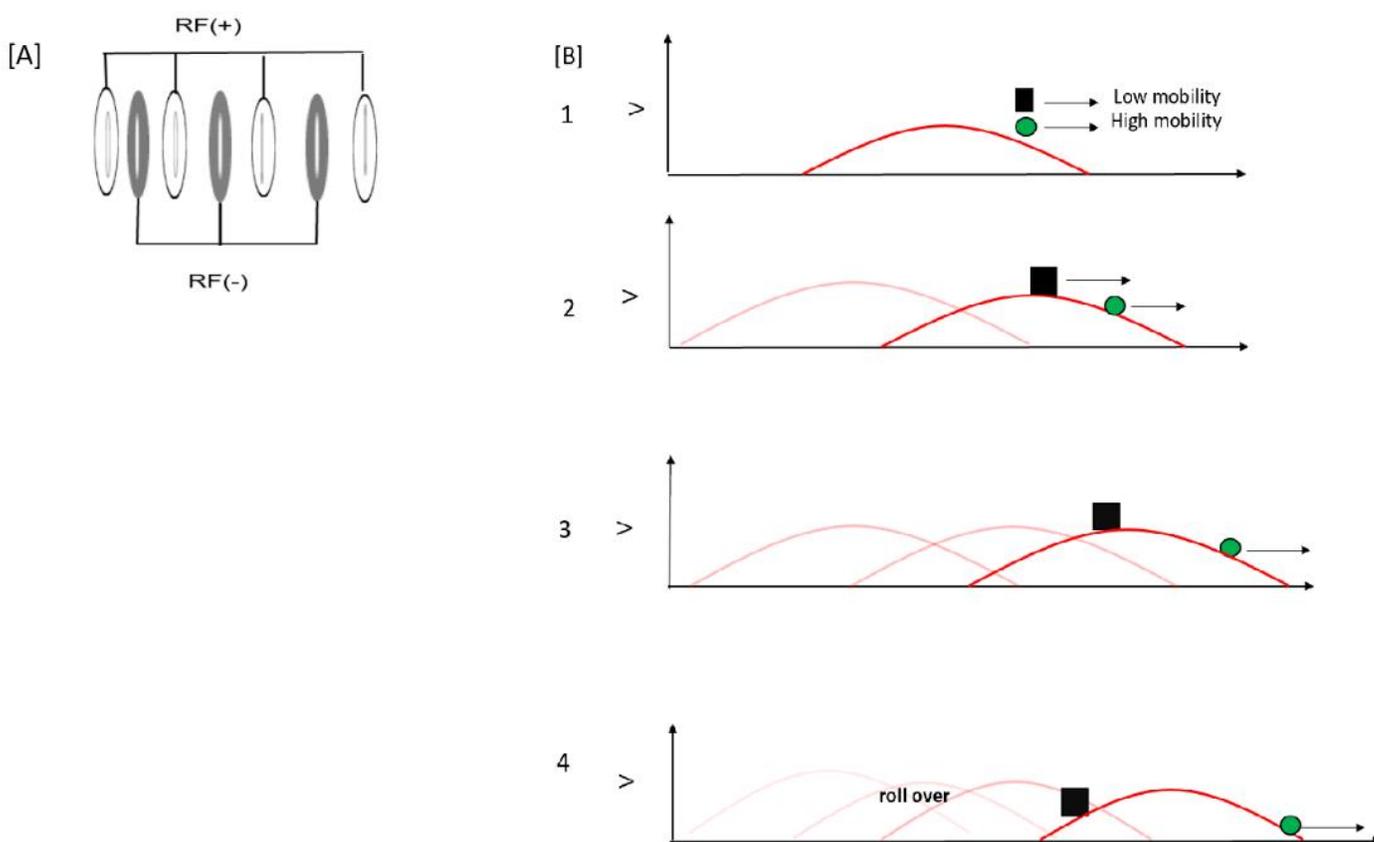


Figure 1.12. (A) A stacked ring TWIMS device, where radio frequency of opposite phases is applied. (B) A cartoon plot where ions surf in front of a travelling wave as they move along the ion guide. Smaller, more compact ions are carried with the wave while larger ions experience more drag from the inert gas and occasionally roll over the wave.

1.11 Collision Induced Unfolding

Protein stability in solution can be studied using thermal or chemical unfolding experiments as discussed in section 1.2. Analogously, the stability of electrosprayed protein ions in the gas phase can then be monitored using collision induced unfolding (CIU). In this method, protein ions enter a collision cell with increased pressure and an acceleration voltage.¹⁰ The gaseous ions will accumulate internal energy due to collisions with background gas (usually Argon). If the collision voltage is high enough, protein ions can ultimately unfold.¹⁰ For the Waters Synapt IMS/MS system used in this current work, CIU experiments can be performed by varying the trap collision energy, and the resulting gas phase unfolding events can be detected by IMS.¹⁰ It is an interesting question whether gas-phase stability correlates with the protein stability in solution. A number of studies revealed similar trends when comparing data obtained by the two approaches. For instance, ligand-induced stabilization in solution was mirrored in the gas phase for several small model proteins.⁷² However, the absence of solvent weakens the hydrophobic effects while strengthening electrostatic effects.⁶⁶ This can lead to different stability trends in solution and in the gas phase.

1.12 Scope of this Thesis

Investigating protein structure is a fundamental aspect of molecular biology because protein function depends on its structure. There are many techniques that have made it possible to gain useful insights into a protein native fold and its thermodynamic stability. These include NMR, X-ray crystallography, optical spectroscopy, and microscopy. Unfortunately, these techniques provide low resolution data or are limited to small, ordered proteins. In contrast, Native ESI-MS and protein labeling have been used to study conformational dynamics and elucidate protein structure with resounding success.

Chapter 2 describes some important aspects of native ESI. This chapter addresses how ESI can preserve the structure of proteins in the gas phase using very gentle conditions. In addition, we sought to provide a framework that considers how several salt additives can affect native ESI.

While ESI provide low resolution structural data, it is often heavily utilized in conjunction with protein labeling to gain atomistic insights into protein structure and dynamics. However, there is a general concern that labeling a protein might perturb a protein structure, introducing some unwanted bias. This is especially true when using covalent labeling methods. However, HDX labeling is often thought to be benign.

Chapter 3 scrutinizes the assumption that HDX labeling is benign and investigates the purported stabilization of deuterated proteins. We hypothesized that this stability is caused by either solvent effects or an increase in the strength of intermolecular hydrogen bonds. Hence, the stability of deuterated protein was probed in the gas and solution phase. Our findings indicate that deuterated proteins are stabilized in solution due to solvent effects as no stabilization was found in the gas phase. This implies that intermolecular hydrogen bonds remain unaffected by deuteration.

1.13 References

- [1] Alberts, A. Johnson, A. Lewis, J. Raff, M. Roberts, K. Walter, P. *Molecular Biology of The Cell*, 4th edition, Garland Science, **2002**.
- [2] FISER, A. S. *Acc. Chem. Res.* **2002**, *35*, 413-421.
- [3] Sun, P. D.; Foster, C. E.; Boyington, J. C. *Curr Protoc Protein Sci* **2004**, Chapter 17, Unit 17 11.
- [4] Anfinsen, C. B. *Science* **1973**, *181*, 223-230.
- [5] Hubbard, E.R.; Haider, K.M. . *ELS* **2010**.
- [6] Konermann, L. In *Encyclopedia of Life Sciences (www.els.net)*, Nature Publishing Group, **2003**.
- [7] Rhoades, E.; Cohen, M.; Schuler, B.; Haran . *J. Am. Chem. Soc.* **2004**, *126*, 14686-14687.
- [8] Southall, N. T.; Dill, K. A.; Haymett, A. D. J. A. *J. Phys. Chem. B* **2002**, *106*, 521-533.
- [9] Privalov, P. L.; Khechinashvili, N. N. . *J. Mol. Biol.* **1974**, *86*, 665-684.
- [10] Yin, V.; Konermann, L. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (1), 73-83.
- [11] Ng, Y. K.; Tajoddin, N. N.; Scrosati, P. M.; Konermann, L. *J. Phys. Chem. B* **2021**, *125* (48), 13099-13110.
- [12] Swint, L.; Robertson, A. D. *Protein Sci.* **1993**, *2* (12), 2037-2049.
- [13] Schmid ,F.X. In *Encyclopedia of Life Sciences (www.els.net)*, Nature publishing Group, **2001**.
- [14] Leong, Y. S.; Ker, J.P.; Jamaludin, M.Z.; Nomanbhay ,S.M.; Ismail, M.; Abdullah,F.; Looe ,M.H.; Lo, K.C. *J.Sens.* **2018**, *18* (7), 2175.
- [15] W. Zheng, N. Shan, L. Yu and X. Wang, *Dye. Pigment.*, 2008, *77*, 153–157.
- [16] F. C. Strong, *Anal. Chem.* **1952**, *24*, 338–342.
- [17]] Greenfield, N. *J.Nat. Protoc.* **2006**, *1* (6), 2876-2890.
- [18] Beychok, S. *Science* **1966**, *154*, 1288-129 .
- [19] Pelton, J. T.; McLean, L. R. *Anal. Biochem.* **2000**, *277*, 167-176.
- [20] Circular Dichroism (CD). <https://cmi.hms.harvard.edu/circular-dichroism> (accessed 2023-08-24) .
- [21] Gunther, H. 2nd ed; *Wiley-VCH Verlag GmbH & Co.* **2013**.
- [22] Marion D. *Mol Cell Proteomics.* **2013**. *12*(11): 3006-3025
- [23] Y. Hu, K. C., L. He, X. Zhang, B. Jiang, L. Jiang, C. Li, G. Wang, Y. Yang, M. Liu. *Anal.Chem.* **2021**, *93*, 1866-1879.
- [25] Maveyraud, L.; Mourey., L. *Molecules.* **2020**, *25*, 1030-1048.

- [26] Smyth, M. S.; Martin, J.H.J. *J Clin Pathol: Mol Pathol* .**2000**, 53 (8), 8-14.
- [27] Herzik, M.A. *Nature*. **2020**, 587, 39-40.
- [28] Milne, J. L.; Borgnia, M. J.; Bartesaghi, A.; Tran, E. E.; Earl, L. A.; Schauder, D. M.; Lengyel, J.; Pierson, J.; Patwardhan, A.; Subramaniam, S. *FEBS J*. **2012**, 280 (1), 28–45.
- [29] Yip, K. M.; Fischer, N.; Paknia, E.; Chari, A; Stark, H. *Nature*. **2020**, 587, 157–161
- [30] Konijnenberg, A.; Butterer, A.; Sobott, F. *Biochim. Biophys. Acta*. **2013**, 1834 (6), 1239-1256.
- [31] Percy, A. J.; Rey, M.; Burns, K. M.; Schriemer, D. C. *Anal. Chim. Acta* **2012**, 721, 7-21.
- [32] Mendoza, V. L.; Vachet, R. W. *Mass Spectrom. Rev*. **2009**, 28, 785-815.
- [33] Xu G, Chance MR. 2007. *Rev* 107:3514–3543.
- [34] Calvete, J.J.; Campanero-Rhodes, M.A; Raida, M.; Sanz, L. *C FEBS Lett*. **1999**, 444:260–264.
- [35] Goetz, M.; Iacobucci, C.; Ihling, C.H.; Sinz, A. *Anal. Chem*. **2019**, 91, 10236-10244
- [36] Vahidi, S.; Konermann, L. *J. Am. Soc. Mass Spectrom*. **2016**, 27, 1156-1164
- [37] Makhatadze, G. I.; Clore, G. M.; Gronenborn, A. M. *Nat. Struct. Biol*. **1995**, 2 (10), 852-855.
- [38] Konermann, L.; Pan, J.; Liu, Y. *Chem. Soc. Rev*. **2011**, 40, 1224-1234.
- [39] Zheng, J.; Strutzenberg, T.; Pascal, B. D.; Griffin, P. R. *Curr. Opin. Struct. Biol*. **2019**, 58, 305-313.
- [40] Chu, I.T.; Pielak, G. J. *M.R.L*. **2023** in press
- [41] James, E. I.; Murphree, T. A.; Vorauer, C.; Engen, J. R.; Guttman, M. *Chem. Rev*. **2022**, 122 (8), 7562-7623,
- [42] Connelly, G. P.; Bai, Y.; Jeng, M. F.; Englander, S. W. *Proteins: Struct. Funct. Genet*. **1993**, 17, 87-92.
- [43] Efimova, Y. M.; Haemers, S.; Wierczinski, B.; Norde, W.; van Well, A. A. *Biopolymers* **2007**, 85 (3), 264-273.
- [44] Stadtmiller, S. S.; Pielak, G. J. *Protein Sci*. **2018**, 27 (9), 1710-1716.
- [45] Kresheck, G. C.; Schneider, H.; Scheraga, H. A. *J. Phys. Chem*. **1965**, 69 (9), 3132-3144.
- [46] Scheiner, S.; Cuma. *J. Am. Chem. Soc*. **1996**, 118, 1511-1521.
- [47] Buckingham, A. D.; Fan-Chen, L. *Int. Rev. Phys. Chem*. **1981**, 1, 253–269
- [48] Kjaersgaard, A.; Vogt, E.; Christensen, F. N.; Kjaersgaard, H. G. *J.Phys.Chem.A* **2020**, 124, 1763-1774

- [49] Cioni, P.; Strambini, B. G. *Biophys. J.* **2002**, *82* (6), 3246-3253.
- [50] Pica, A.; Graziano, G. *Biopolymers* **2018**, *109*, 6.
- [51] El-Aneed, A.; Cohen, A.; Banoub, J. *Appl. Spectrosc. Rev.* **2009**, *44* (3), 210-230.
- [52] Kebarle, P.; Verkerk, U. H. *Rev.* **2009**, *28*, 898-917
- [53] Brown, R.S.; Lennon, J.J. *Anal. Chem.* **1995**, *67*: 1998–2003
- [54] Verentchikov, A. N.; Ens, W.; Standing, K. G. *Anal. Chem.* **1994**, *66* (1), 126-133.
- [55] Koppenaal, D.W.; Baringa, C. J.; Denton, M. B.; Sperline, R.P.; Hieftje, G. M.; Schilling, G. D.; Andrade, F. J.; Barnes, J.H. *Anal. Chem.* **2005**, *77* (21), 418-427.
- [56] Konermann, L.; Metwally, H.; Duez, Q.; Peters, I. *Analyst.* **2019**, *144*, 6157-6171.
- [57] Konermann, L. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 1827-1835.
- [58] Van Berkel, G. J.; Asano, K. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 853-862
- [59] McDonald, L. W.; Campbell, J. A.; Clark, S. B. *Anal. Chem.* **2014**, *86* (2), 1023-1029.
- [60] Wang, H.; Agnes, G. R. *Anal. Chem.* **1999**, *71*, 4166-4172.
- [61] Bich, C.; Baer, S.; Jecklin, M. C.; Zenobi, R. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 286-289.
- [62] Benesch, J. L. P.; Sobott, F.; Robinson, C. V. *Anal. Chem.* **2003**, *75*, 2208-2214.
- [63] Wytttenbach, T.; Bowers, M. T. *J. Phys. Chem. B* **2011**, *115*, 12266-12275.
- [64] Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Int. J. Mass Spectrom.* **1999**, *187*, 565–575.
- [65] Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. *Science.* **2005**, *310*, 1658 -1661.
- [66] Bakhtiari, M.; Konermann, L. *J. Phys. Chem. B* **2019**, *123* (8), 1784-1796.
- [67] Konermann, L.; Aliyari, E.; Lee, J. H. *Ions. J. Phys. Chem B.* **2021**, *125*, 3803-3814.
- [68] Steinberg, M. Z.; Elber, R.; McLafferty, F. W.; Gerber, R. B.; Breuker, K. *ChemBioChem* **2008**, *9*, 2417-2423.
- [69] Susa, A. C.; Xia, Z. J.; Williams, E. R. *Anal. Chem.* **2017**, *89* (5), 3116-3122.
- [70] Lanucara, F.; Holman, S. W.; Gray, C. J.; Evers, C. E. *Nat. Chem.* **2014**, *6* (4), 281-294.
- [71] Sun, Y.; Vahidi, S.; Sowole, M. A.; Konermann, L. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 31-40
- [72] Hopper, J. T. S.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1851-1858.

Chapter 2 Considerations for Native ESI Experiments.

2.1 Chapter 2 Introduction

Native ESI-MS is a versatile tool capable of analyzing proteins, protein-ligand complexes, and large multimeric assemblies that cannot be probed using conventional methods.¹⁻⁵ However, the question to what extent ionized proteins and complexes retain solution-like structures when transferred into the gas has been the subject of debate for decades.^{6,7} Several studies suggest that hydrophobic interactions are weakened in the gas phase, while ionic interactions are enhanced.⁸⁻¹³ In addition, the fact that electrosprayed proteins carry a net charge z suggests that their internal electrostatics are dominated by repulsive forces.¹³⁻¹⁵ Despite this, a growing body of evidence has shown that ESI is capable of retaining noncovalent complexes, as long as gentle conditions are maintained throughout the ESI process and in the vacuum of the instrument.^{3,16-20} For instance, one study highlighted the ability of ESI to generate native-like structures for the 20 S proteasome (a large protein consisting of 28 subunits).²¹ Similarly, a recent study utilizing ESI-MS in conjunction with soft landing and electron microscopy (EM) managed to obtain a three-dimensional reconstruction of the 800 kDa protein complex Gro-EL, providing direct evidence that non-covalent protein complexes are indeed preserved under native ESI conditions.⁷ The retention of solution-like structures is a remarkable phenomenon and is attributed to kinetic trapping, where large activation barriers prevent large-scale transitions on the timescale of typical ESI-MS experiments (Figure 2.1).^{13,22-24} In other words, protein ions generated by ESI are not thermodynamically stable; instead, they are metastable (kinetically stable).²²

The ESI process starts with production of electrically charged droplets that undergo multiple cycles of fission and evaporation at atmospheric pressure and ambient temperature.^{25,26} Studies have found that ESI of native globular proteins proceeds through charge residue model where proteins are released from an evaporating nanometer-sized droplet containing one protein (Figure 1.10).²⁷ When protein ions first enter the vacuum of the instrument, a drop in the pressure will result in adiabatic expansion, that is when ions gain translational energy at the expense of their internal energy.^{3,27,29} Indeed, computational experiments on cytochrome *c* have shown that in the absence of energy transfer gaseous ions undergo significant reduction in temperature.^{28,30} The rapid cooling of ions might suggest that formation of non-specific aggregates in the gas phase. In reality, however, inelastic collisions with gas molecules impart energy to gaseous ions. An additional

source of energy is blackbody infrared radiation from the instrument walls. This is achieved by employing a combination of heated desolvation gas and acceleration voltage offsets.^{3,29} Hence, the loss of internal energy is mitigated by collisions with background gas and infrared radiation.³ Depending on the extent of heating, it is also possible to induce structural perturbations such as protein unfolding or even covalent fragmentation. Hence, it is important to keep the ion source somewhat activating to remove residual solvent molecules, but gentle enough to prevent unwanted structural changes.

While the above discussion captures the ability of native MS to generate protein ions, it is important to note that inadvertent pH alterations during the ESI process can negatively affect the outcome of the experiment. Whether such pH changes affect the outcome of MS experiments is a subject of debate and depends on thermodynamic and kinetic aspects of the analyte in question. To mitigate this problem, buffers represent an obvious strategy for stabilizing the pH throughout the various ESI stages. Many biochemical experiments are conducted in neutral buffer solution to stabilize pH.³¹ Such systems typically contain added background electrolytes to provide high ionic strength necessary for protein interactions. For instance, a typical solvent for traditional biochemical assays is 50 mM aqueous phosphate buffer, with 100 mM NaCl. Under physiological conditions, salts increase the dielectric constant of water, and stabilize protein structures.^{31,32} However, such nonvolatile solvents will negatively affect the ESI process, lowering the sensitivity and causing ion suppression.³²⁻³⁴ Furthermore, salt adduction broadens the mass spectral peaks, making it harder to resolve charge state distributions.³² These non-specific adducts are also consistent with the CRM, where analytes associate with residual solutes as the droplet evaporates.³³ To circumvent these issues, experimentalists remove non-volatile salts through dialysis or ion chromatography.³⁵⁻³⁶ Instead, native ESI practitioners usually employ aqueous ammonium acetate (NH₄Ac).³⁷ The common use of this additive can be attributed to two factors: (1) NH₄Ac decomposes into NH₃ and acetic acid, which are volatile salts and do not form any adducts.³⁸ (2) NH₄Ac yields neutral pH when dissolved in water.³¹ The latter point may suggest that NH₄Ac is a buffer. However, this is not the case as buffer solutions are composed of a weak acid and conjugate base.³¹ NH₄Ac has very little buffering capacity at pH 7.00. This can be seen as the conjugate acid of Ac⁻ is HAc, while the conjugate base of NH₃ is NH₄⁺. The pK_a of acetic acid is 4.75 and that of NH₄⁺ is 9.25. Hence, the buffering capacity for NH₄Ac is best suited for pH ranges that are ±1

units away from the pKa of NH_4^+ and HAC (4.75 and 9.25). Figure 2.2 shows a titration curve of NH_4Ac . The pH of a buffer can be readily calculated using the Henderson-Hasselbalch equation.³⁹

$$pH = pKa + \log \frac{[conjugate\ base]}{[conjugate\ acid]} \quad (2.2)$$

The fact that NH_4Ac is not a buffer at pH 7.00 has been largely ignored by the ESI-MS literature.³⁴ while the popularity of this additive seems to have no chemical basis, studies have suggested that NH_4Ac converts some acetate molecules to acetic acid. This implies that pH drop is not as dramatic as in pure water.³¹

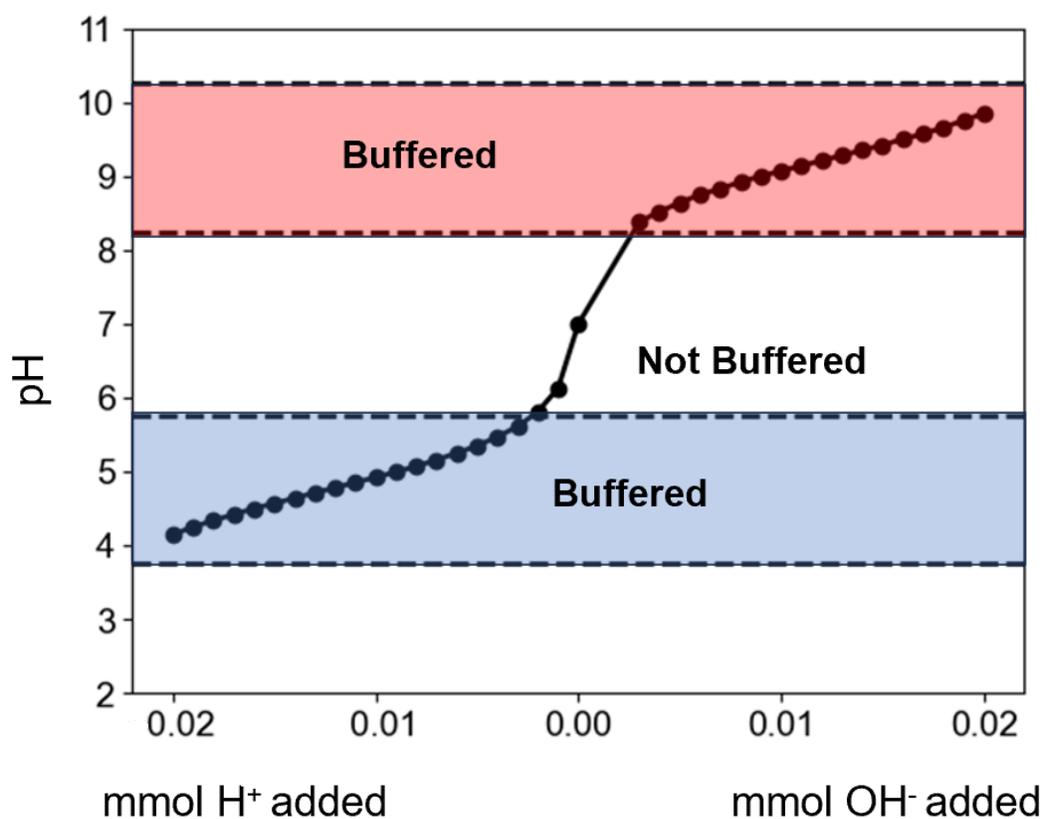


Figure 2.2. The titration curve for NH_4Ac is calculated from equation 2.2. pKa (acetic acid)=4.75 and pKa(ammonium)=9.25. Colored regions indicate the buffering range of NH_4Ac .

It is undisputed that ESI-MS remains a versatile tool for probing protein structure and function. Nonetheless, it is imperative for experimentalist to minimize structural perturbations as proteins are transferred from bulk solution into gas phase. This chapter will take a critical look at some of the parameters that can affect the outcome of native ESI-MS studies. These include collisional activation, pH drops, and buffer consideration. The aim of this work is to scrutinize common ESI workflows.

2.2 Materials and Methods

Materials and sample preparation: Equine heart cytochrome *c* (cyt *c*, 12360 Da), hen egg white lysozyme (14305 Da), bovine ubiquitin (8565 Da), and myoglobin (17,568 Da) were supplied by Millipore Sigma (St. Louis, MO). All other chemicals were purchased from Thermo Fisher Scientific (Mississauga, ON). An AB15 glass electrode pH-meter (Fisher) was used for pH measurements. Protein stock solutions (500 μ M) were initially dialyzed for 24 h against water for removal of salt contaminants using 10 kDa MWCO Millipore Sigma dialysis cassettes. Native protein was prepared at a concentration of 5 μ M (10 μ M for cyt *c*) in 10mM NH₄Ac (pH 7.00). In contrast, denatured protein was prepared at a concentration of 5 μ M in a mixture of 48% methanol:48%water: 2% acetic acid. Proteins with non-volatile salt additives were prepared in 1 mM NaCl at pH 7.00. Finally, buffered proteins at pH 5.30 were prepared in a mixture of NH₄Ac/AcH

Mass Spectrometry. ESI mass spectra were acquired on a Synapt G2 time-of-flight instrument (Waters, Milford, MA). Protein samples were infused into a Z-spray interface at 5 μ L min⁻¹ using a syringe pump. ESI was conducted in positive ion mode with the capillary set to 2.5 kV. The desolvation gas and source temperatures were 30 and 40 °C, respectively. The level of in-source activation was controlled by adjusting the sample cone voltage between 0 V (“gentle”) and 100 V (“harsh”). Unless noted otherwise, the trap collision voltage (trap CE) was set to 4 V. Cluster signals in the experimental spectra were identified using isotope models generated on the sisweb.com server

2.3 Results and Discussion:

2.3.1 Charge state Distributions and their Information Content

Typical ESI spectra display a range of charge states, often in a Gaussian-like distribution. The observed charge states arise due to a combination of factors which involve: the availability of ionizable sites, Rayleigh limit, solvent accessible surface area (SASA), and Coulombic repulsion.³ In most cases, excess charges are localized to basic amino acids (Lys, Arg, His, N-terminus).⁴⁰ Importantly, charge-state distributions can offer a glimpse into a protein's solution structure. Proteins prepared in denaturing conditions will have a broad range of charge states.³ The interplay between charge states and compactness of structure has been illustrated in many studies.^{3,41-42} For instance, it was shown that thermally heated proteins can be monitored by following the shift of the average charge states.^{3,42} Figure 2.3 highlights this phenomenon by comparing myoglobin in native and denaturing conditions.

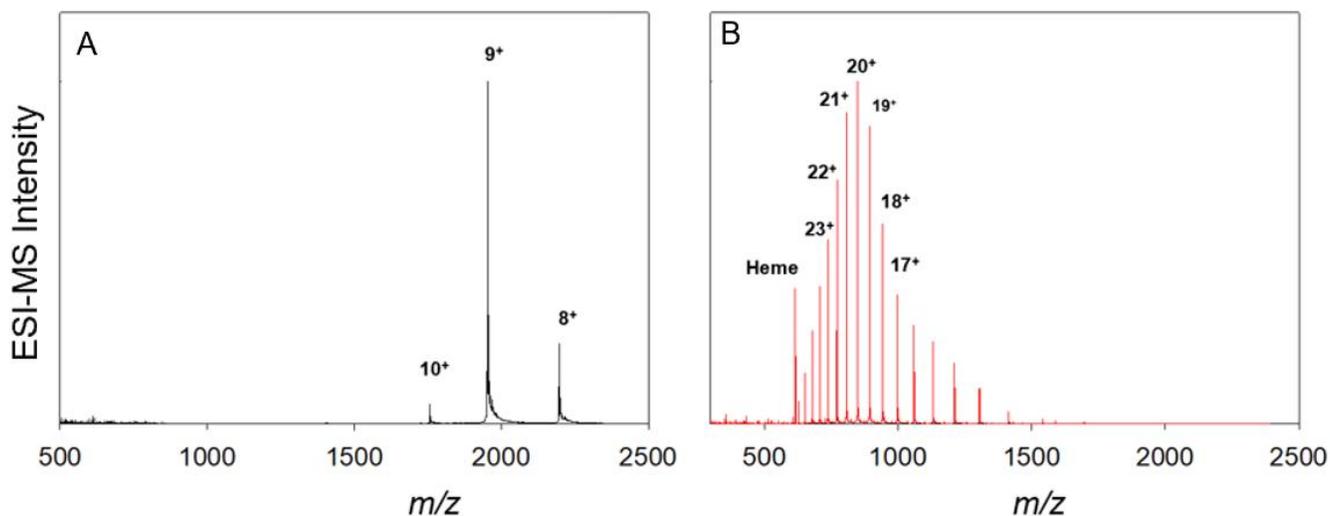


Figure 2.3. ESI mass spectra of (A) native myoglobin in 10mM aqueous NH₄Ac, (B) unfolded myoglobin in water: methanol: acetic acid in 49:49:2 ratio. Selected peaks are annotated with their charge states. Myoglobin dissolved in denaturing solvent (panel B) displays a broad distribution of highly charged ions, indicating an unfolded structure in solution.

Under native conditions, myoglobin adopts a compact structure which gives to a narrow range of charge states centered at 9^+ (Figure 2.3 A). In contrast, denatured myoglobin adopts an expanded structure. The higher charge state distributions in Figure 2.3B do not reflect the solution-phase titration state.⁴³ Instead, high charged species are formed in accordance with CEM where a Rayleigh charged droplet releases a highly protonated polymer chain. Electrostatically driven interactions result in the transfer of protons into the chain that is being ejected.⁴³ Similarly, the narrow range of m/z under native conditions is correlated with the ionization mechanism.⁴³ Globular proteins are ionized in accordance with CRM, where evaporating droplets remain close to the Rayleigh charge. In the case of a nanodroplet containing myoglobin, the Rayleigh limit is 9.5, which is close to experimentally observed values.⁴³

The low charge states produced by native ESI can be disadvantageous for some experiments.²⁵ For instance, ions with low z are not suitable for Fourier transform mass analyses, and they tend to be unreactive in top-down experiments.^{25,44} Hence, boosting a protein charge is important for such experiments. One alternative method to using denaturing conditions involves the use of supercharging agents like sulfolane ($C_4H_8SO_2$). In this approach, proteins experience native solvent environment as they enter the ESI source.^{25,45} Experiments have shown that the addition of 1% sulfolane to solution shifts the maximum charge state of myoglobin from 9^+ to 16^+ .²⁸ The mechanism by which sulfolane boosts the protein charge remains controversial.²⁵ However, computational studies have suggested a supercharging model where water evaporates more quickly than sulfolane, resulting in water-free protein/sulfolane droplets.^{25,46} This is followed by Na^+ or H^+ binding to the protein and slow sulfolane evaporation.^{25,46} Native ESI supercharging mirrors a CRM process where IEM ejections of charge carriers like Na^+ or H^+ are suppressed, boosting the protein charge.^{25,46} Other charge state enhancement techniques can be used to boost a protein signal. For instance, it was shown that trivalent metal ions like La^{3+} can form multidentate contacts involving Asp^- and Glu^- .⁴⁷ This chelation process is irreversible and precludes the ejection of La^{3+} from the droplet, resulting in the formation of charge-enhanced protein ions.⁴⁷

2.3.2. Experimental Conditions for Native ESI

The possibility of transferring proteins from bulk solution into the gas phase offers exciting opportunities for studying protein structure and dynamics.³ To preserve the native fold of a protein, it is crucial to properly optimize the ESI conditions.^{3,48} Very often it is observed that the ESI mass spectra of proteins acquired from different instruments show a degree of variability. These differences can be attributed to several instrument parameters which involve ion spray voltage, spray tip diameter, source geometry, source gas pressure, source temperature, sample cone voltage, and different ion optics voltages.^{3,48} Acceleration voltages are typically tuned to efficiently transmit ions and aid in desolvation.³ However, high acceleration voltage can lead to gas-phase unfolding. Furthermore, protein fragmentation or dissociation can also be observed at excessively high voltages.³ Under harsh ESI conditions, proteins build up sufficient internal energy from collisions with background gas, causing them to lose their native structure.³ Collision induced unfolding (CIU) and collision induced dissociation (CID) have become a common technique for studying proteins. For instance, CIU is commonly used for studying protein stability in the gas phase, while CID is heavily utilized in proteomics to produce diagnostic peptides.³ Despite this, native ESI experiments should avoid excessive collisional activation to preserve the protein's native fold.

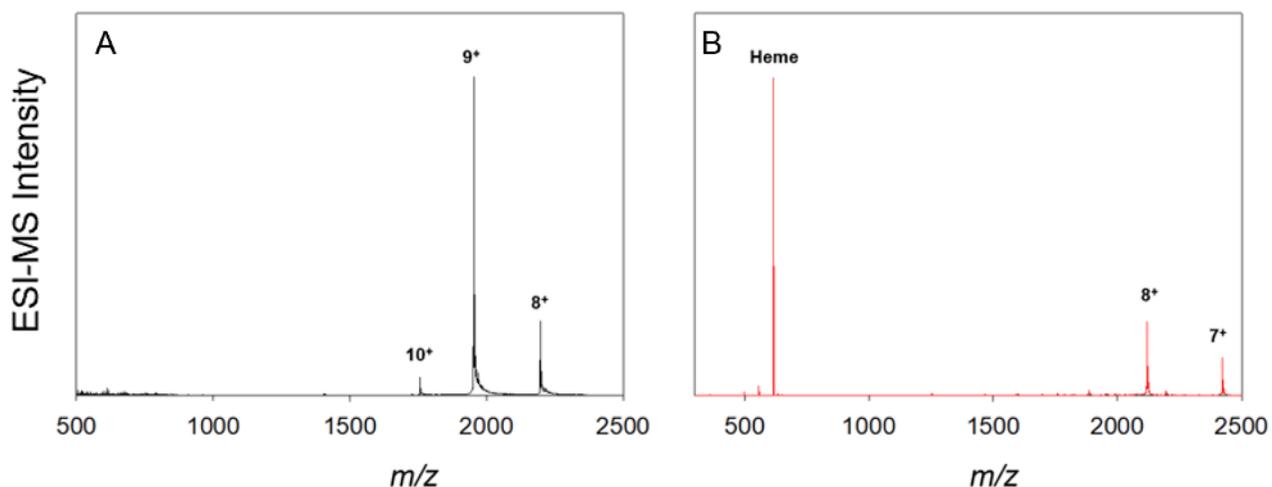


Figure 2.4. (A) mass spectra of myoglobin at low activation setting (sample cone 5V). (B) mass spectra of myoglobin at high activation setting (100 V). Selected peaks are annotated with their charge state distribution. Under harsh ESI conditions, the heme dissociates from myoglobin which appears as a singly charged species at 616 m/z.

While analyte activation can occur at several points along the ion path, we chose to focus on the sample cone because it is the component closest to the ESI source and where structural perturbations may occur. An example of undesired protein unfolding is given in Figure 2.4. In these experiments, we chose myoglobin as it is non-covalently bonded to its heme group. Hence, myoglobin can serve as a probe for tuning the sample cone. Under gentle conditions (sample cone 5 V), myoglobin retains its native fold, where heme-protein interactions remain intact (Figure 2.4 A). In contrast, harsh experimental conditions (sample cone 100 V) caused the ejection of the heme group, evident from the high intensity peak at m/z 616 (Figure 2.4 B). Also, the maximum intensity charge state of myoglobin is shifted from 9^+ to 8^+ . This can be attributed to the fact that the heme group carries a 1^+ charge. Our experiments revealed that low cone voltages are optimal for maintaining the overall native fold for globular proteins. High cone voltage can improve the sensitivity of the experiment but can cause significant structural perturbations.⁴⁸ This was shown in another study where IMS was utilized to measure the collision cross section of myoglobin at low and high cone voltages.⁴⁹ The data obtained this way demonstrated the prevalence of small Ω at low cone voltage, while high cone voltages displayed protein ions with larger Ω .⁴⁹

While tuning the cone voltage is imperative for native ESI, similar considerations have been used in other fields of chemistry. For instance, ESI of alkyl halides solutions generates salt clusters that are widely used for mass calibration.⁵⁰ Of specific relevance are the effects of sample cone voltage on the dissociation and transmission of these clusters. Mass spectra acquired after electrospraying 10 mM aqueous NaCl solution showed a wide range of salt clusters. The spectra were recorded using a moderate level of in-source activation with sample cone set to 75 V and low activation with sample cone at 0 V (Figure 2.5). The gentle conditions provided by a cone voltage of 0 V resulted in a less efficient transmission of salt clusters through the instrument as some species displayed a noticeable drop in intensity compared to the harsh conditions. However, the use of a very low activation settings enabled us to capture doubly charged clusters ($+2$). Electrostatic repulsion renders these clusters increasingly unstable as their size decrease. Consistent with this limited stability is the fact that the smallest $2+$ clusters ($\text{Na}_{23}\text{Cl}_{21}^{2+}$ to $\text{Na}_{27}\text{Cl}_{25}^{2+}$) were observable only under the most gentle conditions (Figure 2.5 B), while in-source activation caused their depletion.⁵⁰

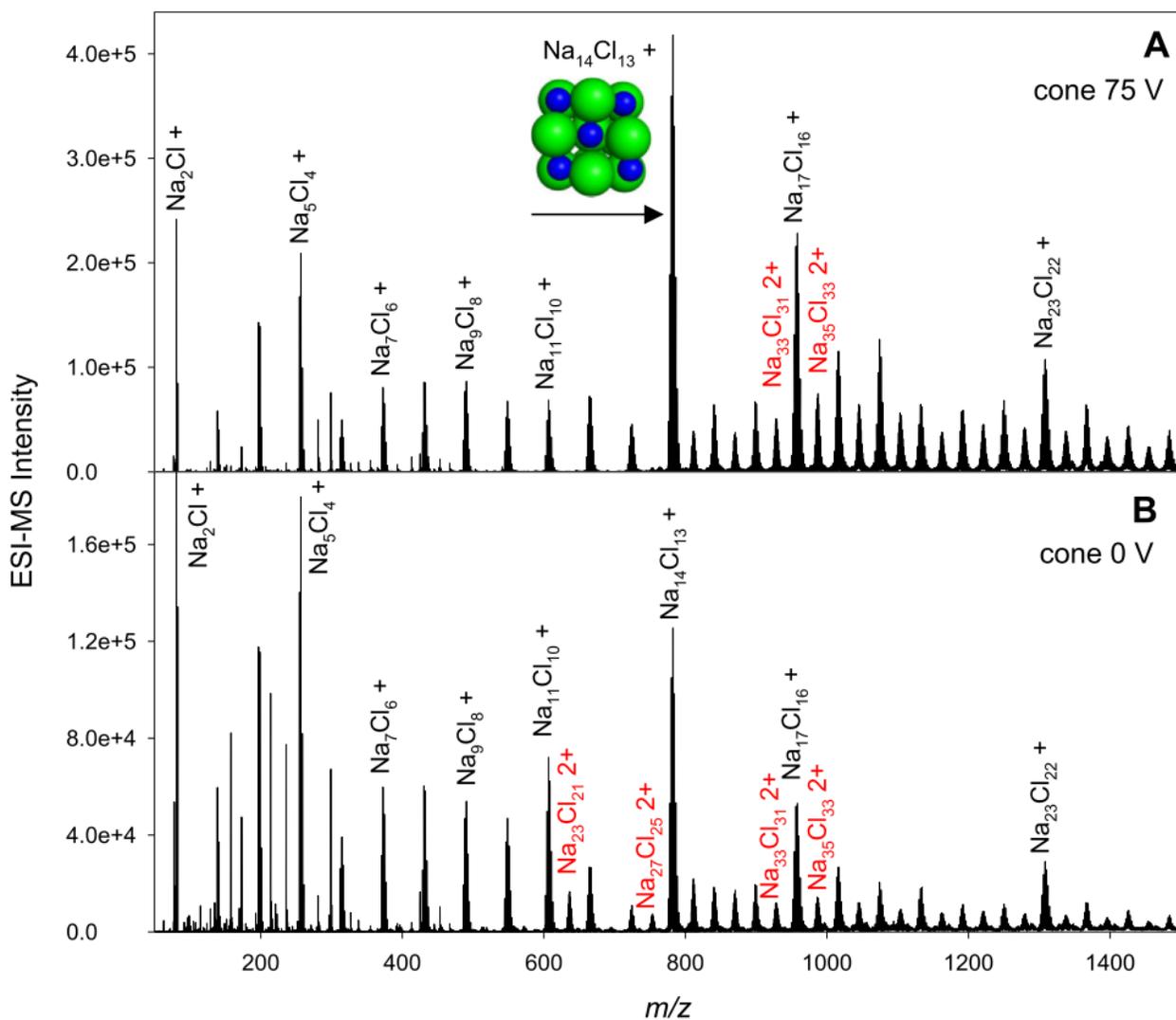


Figure 2.5. Mass spectra acquired after electrospraying 10 mM NaCl in water (A) with moderate in-source activation (sample cone 75 V) and (B) under gentle in-source conditions (sample cone 0 V). Selected peaks are annotated with their composition and charge state, that is, 1+ (black) and 2+ (red). Included in panel A is the structure of the cubic $\text{Na}_{14}\text{Cl}_{13}^+$ MNC that dominates the spectrum (Na: blue; Cl, green). For better visualization, the Na_2Cl^+ signal in panel B was truncated at 50%

2.3.3. Chemistry of Ammonium Acetate

Most proteins are stable in aqueous solution at near neutral pH. Hence, buffers can ensure a suitable pH in experiments.⁵¹ Traditional buffers are not useful for ESI, causing non-specific adducts and signal suppression. NH₄Ac is the additive of choice in native ESI experiments. NH₄Ac can undergo proton transfer to form ammonia and acetic acid. The volatility of these products ensures that electrosprayed proteins are free of undesired adducts. The volatile nature of NH₄Ac can be illustrated by direct comparison with NaCl in Native ESI-MS. Protein mass spectra generated with NaCl show extensive peak tailing due to the formation of heterogenous $[M+zH+n(\text{Na-H})+m(\text{Cl+H})]^{z+}$ adducts, illustrated in Figure 2.6 A for lysozyme. Figure 2.6 A also shows abundant chemical noise from salt clusters.⁵² In contrast, ESI of lysozyme in 100mM aqueous NH₄Ac yielded clean $[M+zH]^{z+}$ signals with greatly reduced background noise (Figure 2.6 B), highlighting the favorable properties of this volatile salt.⁵³⁻⁵⁵

Electrospraying aqueous solution without an analyte is instructive as well. NaCl solution generated a range of Na_nCl_m^{(n-m)+} clusters (Figure 2.6 C) resulting from association of Na⁺ and Cl⁻ during the final stages of droplet evaporation followed by CRM release into the gas phase. This is in contrast to spectra obtained upon electrospraying aqueous NH₄Ac, where NH₄Ac cluster ions are unobservable (Figure 2.6 D). The m/z 77 signal in Figure 2.6 D corresponds to $[M+\text{NH}_4]^+$ of acetamide, which is a contaminant in commercially supplied NH₄Ac.⁵⁶

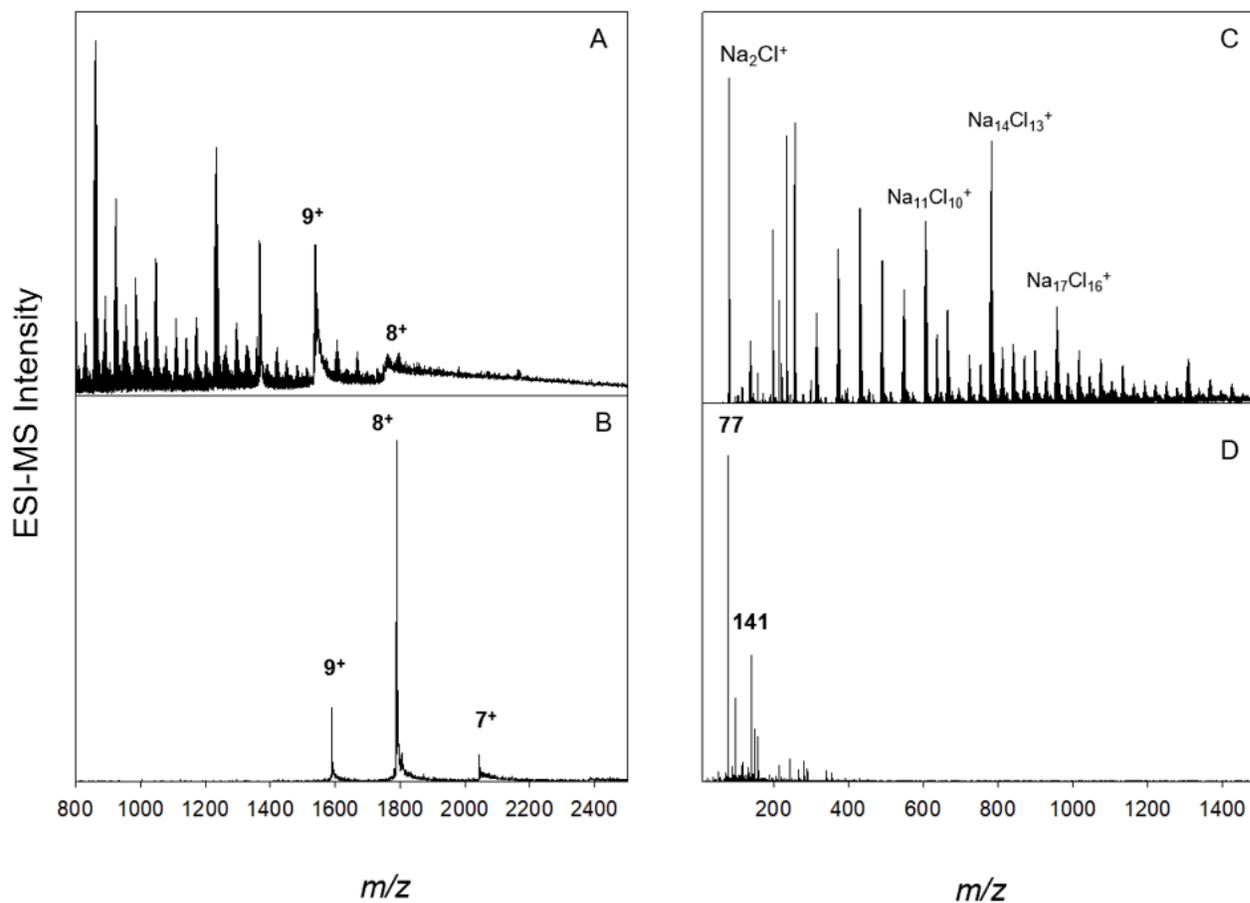


Figure 2.6. ESI mass spectra acquired for aqueous pH 7 solutions. (A) lysozyme in 5 mM NaCl. (B) lysozyme in 100 mM NH₄Ac. Charge states of protein ions are indicated as 7^+ , 8^+ , etc. (C) 10 mM NaCl. The composition of selected cluster ions is indicated. (D) 100 mM NH₄Ac. The samples used for panels C and D did not contain any analytes other than the dissolved salts.

Even though NH_4Ac is the most commonly used additive in native ESI, knowledge gaps persist for NH_4Ac containing droplets. As mentioned previously, nascent droplets released from a Taylor cone are expected to be acidic. Such pH alterations can affect the outcome of experiments. Furthermore, the buffering capacity of NH_4Ac in neutral solution is marginal, because the pKa values of both NH_4 and HAc are more than two units away from pH 7. Although the dissolution of NH_4Ac yields a neutral solution, this pH is labile.³¹ For instance, studies have shown that NH_4Ac is likely to undergo acidification to pH 5.8 or lower in the ESI plume.³¹ To investigate possible structural perturbations resulting from this pH drop, it is useful to compare NH_4Ac in neutral and acidified (pH 5.30) solutions using ESI-MS (Figure 2.7). The idea is that any pH artifacts caused by the droplet acidification in the early stages of ESI can be simulated by acquiring an ESI spectrum of proteins in a solution of NH_4Ac at pH 5.30. Since NH_4Ac can act as a buffer in this pH range, it is not expected for nascent droplets to undergo a further pH drop. Moreover, it is often expected for a droplet containing 10mM NH_4Ac to have a $\text{pH} = 4.75 \pm 1$ upon analyte release. If the charge state distributions of both samples were similar, one can conclude that ESI-induced pH changes in native ESI are not a major issue.

Protein mass spectra generated with 10 mM NH_4Ac at pH 5.30 (Figure 2.6 D-F) indicates that lysozyme, ubiquitin, and cytochrome c do not experience any major structural perturbation. This is evident as electrosprayed proteins at pH 5.30 display identical charge state distributions with proteins electrosprayed at pH 7.00 (Figure 2.6 A-C). Thus, our data demonstrates the usefulness of NH_4Ac in native ESI experiments. While NH_4Ac is not a buffer at pH 7.00, substoichiometric acidification of the droplet at the early stages of ESI converts a certain percentage of acetate into acetic acids. This stabilizes the droplet pH at 4.75 ± 1 . This acidification is significant enough to cause protonation of histidine but will leave the side chains of aspartate and glutamate negatively charged. Our data (Figure 2.7) reveals that histidine protonation is not a major issue for the native fold of globular proteins. However, we cannot discern whether the same principle applies for larger protein complexes. We recommend practitioners to use 100 mM NH_4Ac solutions where the concentration of Ac^-/AcH will stabilize the pH at 6.50. In all cases, NH_4Ac remains a useful additive in native ESI as it prevents the dramatic drop of pH in nascent droplets

that is observed in pure water (pH=1). In addition, NH₄Ac can stabilize a protein native structure by serving as background electrolyte.

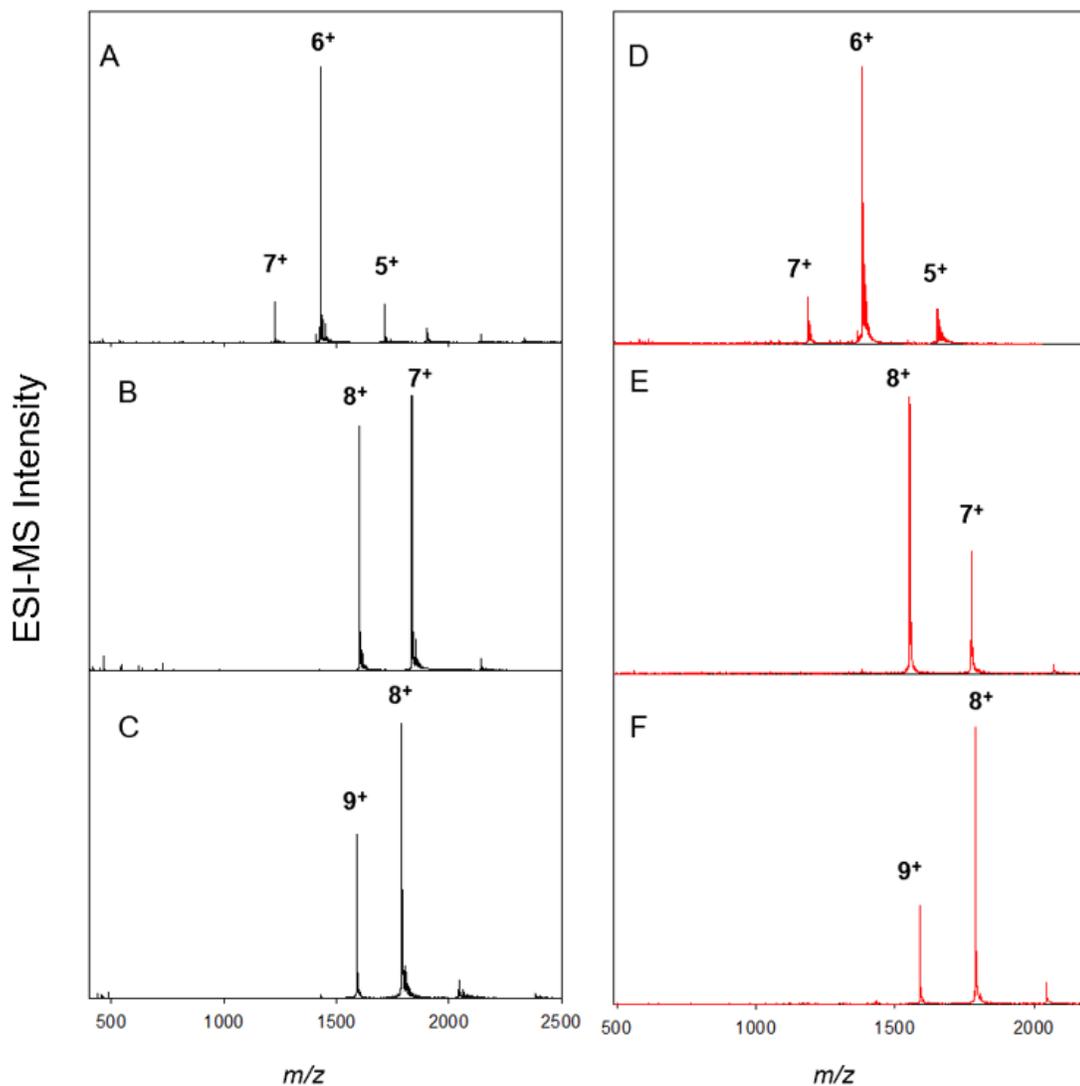


Figure 2.7. (A-C) mass spectra showing the charge state distributions of ubiquitin, lysozyme, and cytochrome c in 10mM NH₄Ac at pH 7.00. (D-F) mass spectra showing the charge state distribution of ubiquitin, lysozyme, cytochrome c and in 10mM NH₄Ac at pH 5.30.

2.4 Conclusions

Native ESI is an important tool in structural biology, that enables the transfer of proteins from bulk solution into the gas phase while preserving a native-like fold. However, practitioners should be aware of some limitations that can change the outcome of experiments. For instance, one should consider the various voltage offsets of the ion optics and account for additives suitable for ESI. In this work, we showed how charge state distributions can provide valuable information on the structure of proteins. Compact protein structures often display a narrow range of charge states. In contrast, unfolded proteins generate a broad range of charge states. This is attributed to a different ionization mechanism, i.e., the CEM as opposed to the CRM. Furthermore, our experiments reveal a simple approach to tune the sample cone voltage. Myoglobin with its weakly bound heme can serve as an indicator for measuring how activating experimental conditions are. Low sample cone voltages are less efficient in ion transmission and desolvation but ensures that the protein's native structure remain intact. Finally, we aimed to clear up some uncertainties related to NH_4Ac . This additive has a number of attractive features that make it useful for native ESI experiments. For instance, its high volatility ensures that no protein adducts are formed. Furthermore, we showed that NH_4Ac does not match the definition of a pH 7 buffer. Our data convey that ESI acidification of NH_4Ac containing droplets does not dramatically alter the outcome of ESI experiments. Future work is expected to explore compatible physiological buffers. Another interesting approach is to use sub-micrometer nano ESI tips that can impart higher salt tolerance, but this approach is not widely used.⁵⁷⁻⁵⁸

2.5. References

- [1] Leney, A. C.; Heck, A. J. R. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 5-13.
- [2] Andersen, C. B. F.; Torvund-Jensen, M.; Nielsen, M. J.; de Oliveira, C. L. P.; Hersleth, H. P.; Andersen, N. H.; Pedersen, J. S.; Andersen, G. R.; Moestrup, S. K. *Nature* **2012**, *489* (7416).
- [3] Konijnenberg, A.; Butterer, A.; Sobott, F. *Biochim. Biophys. Acta* **2013**, *1834* (6), 1239-1256.
- [4] Dyachenko, A.; Gruber, R.; Shimon, L.; Horovitz, A.; Sharon, M. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (18), 7235-7239.
- [5] Deng, L.; Broom, A.; Kitova, E. N.; Richards, M. R.; Zheng, R. B.; Shoemaker, G. K.; Meiering, E. M.; Klassen, J. S. *J. Am. Chem. Soc.* **2012**, *134* (40), 16586-16596.
- [6] Hogan, C. J.; Ruotolo, B. T.; Robinson, C. V.; de la Mora, J. F. *J. Phys. Chem. B* **2011**, *115* (13), 3614-3621.
- [7] Westphall, M. S.; Lee, K. W.; Salome, A. Z.; Lodge, J. M.; Grant, T.; Coon, J. J. *Nat. Commun.* **2022**, *13* (1), 6.
- [8] van der Spoel, D.; Marklund, E. G.; Larsson, D. S. D.; Caleman, C. *Macromol. Biosci.* **2011**, *11*, 50-59.
- [9] Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2426-2427.
- [10] Liu, L.; Michelsen, K.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2012**, *134* (6), 3054-3060.
- [11] Breuker, K.; Brüsweiler, S.; Tollinger, M. *Angew. Chem. Int. Ed.* **2011**, *50*, 873-877.
- [12] Yin, S.; Loo, J.A. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 899-907.
- [13] Konermann, L.; Aliyari, E.; Lee, J. H. *J. Phys. Chem B* **2021**, *125* (15), 3803-3814.
- [14] Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119*, 2240-2248.
- [15] Laszlo, K. J.; Munger, E. B.; Bush, M. F. *J. Am. Chem. Soc.* **2016**, *138* (30), 9581-9588.
- [16] Pagel, K.; Hyung, S.-J.; Ruotolo, B. T.; Robinson, C. V. *Anal. Chem.* **2010**, *82*, 5363-5372.
- [17] Poliakov, A.; van Duijn, E.; Lander, G.; Fu, C.-Y.; Johnson, J. E.; Prevelige, P. E. J.; Heck, A. J. R. *J. Struct. Biol.* **2007**, *157*, 371-383.
- [18] Erba, E. B.; Ruotolo, B. T.; Barsky, D.; Robinson, C. V. *Anal. Chem.* **2010**, *82* (23), 9702-9710.
- [19] Hilton, G. R.; Benesch, J. L. P. *J. R. Soc. Interface* **2012**, *9*, 801-816.

- [20] Barth, M.; Schmidt, C. *J. Mass Spectrom.* **2020**, *55* (10).
- [21] Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem.* **2007**, *76*.
- [22] Hendricks, N. G.; Julian, R. R. *Analyst* **2016**, *141* (15), 4534-4540.
- [23] Clemmer, D. E.; Russell, D. H.; Williams, E. R. *Accounts Chem. Res.* **2017**, *50* (3), 556-560.
- [24] Sever, A. I. M.; Konermann, L. *J. Phys. Chem. B* **2020**, *124* (18), 3667-3677.
- [25] Konermann, L.; Metwally, H.; Duez, Q.; Peters, I. *Analyst* **2019**, *144*, 6157-6171.
- [26] Aliyari, E.; Konermann, L. *Anal. Chem.* **2022**, *94* (21), 7713-7721.
- [27] Breuker, K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18145-18152.
- [28] Steinberg, M. Z.; Breuker, K.; Elber, R.; Gerber, R. B. *Phys. Chem.* **2007**, *9*, 4690-4697.
- [29] Gabelica, V.; De Pauw, E. *Mass Spectrom. Rev.* **2005**, *24*, 566-587.
- [30] Kebarle, P.; Tang, L. *Anal. Chem.* **1993**, *65*, 972A-986A.
- [31] Konermann, L. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 1827-1835.
- [32] Susa, A. C.; Xia, Z. J.; Williams, E. R. *Anal. Chem.* **2017**, *89* (5), 3116-3122.
- [33] Hernandez, H.; Robinson, C. V. *Nat. Protoc.* **2007**, *2* (3), 715-726.
- [34] Wang, G.; Cole, R. B. *Anal. Chem.* **1994**, *66* (21), 3702-3708.
- [35] Jiang, Y.; Hofstadler, S. A. *Anal. Biochem.* **2003**, *316* (1), 50-57
- [36] de la Mora, J. F. *Anal. Chim. Acta* **2000**, *406*, 93-104.
- [37] Tamara, S.; den Boer, M. A.; Heck, A. J. R. *Chem. Rev.* **2022**, *122* (8), 7269-7326.
- [38] Kebarle, P.; Verkerk, U. H. *Rev.* **2009**, *28*, 898-917
- [39] Nielsen, J. E. In *Meth. Enzymol.*, Johnson, M. L., Brand, L. Eds.; Methods in Enzymology, Vol. 454; Elsevier Academic Press Inc, 2009; pp 233-258.
- [40] Krusemark, C. J.; Frey, B. L.; Belshaw, P. J.; Smith, L. M. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1617-1625.
- [41] Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. *Science* **2005**, *310*, 1658-1661.
- [42] Benesch, J. L. P.; Sobott, F.; Robinson, C. V. *Anal. Chem.* **2003**, *75*, 2208-2214.
- [43] Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2013**, *85*, 2-9.
- [44] Han, X.; Jin, M.; Breuker, K.; McLafferty, F. W. *Science* **2006**, *314*, 109-112.
- [45] Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. *Anal. Chem.* **2009**, *81*, 369-377.

- [46] Metwally, H.; McAllister, R. G.; Popa, V.; Konermann, L. *Anal. Chem.* **2016**, 88 (10), 5345-5354.
- [47] Martin, L. M.; Konermann, L. *J. Am. Soc. Mass Spectrom.* **2020**, 31, 25-33.
- [48] Smith, R. D.; Loo, J. A.; Baringa, C. J.; Edmonds, C. G.; Udseth, H. R. *J. Am. Soc. Mass Spectrom.* **1990**, 1, 53-65.
- [49] Sun, Y.; Vahidi, S.; Sowole, M. A.; Konermann, L. *J. Am. Soc. Mass Spectrom.* **2016**, 27, 31-40.
- [50] Konermann, L.; Haidar, Y. *Anal. Chem.* **2022**, 94 (47), 16491–16501.
- [51] Boron, W. F. *Adv. Physiol. Educ.* **2004**, 28, 160-179
- [52] Xu, N.; Lin, Y.; Hofstadler, S. A.; Matson, D.; Call, C. J.; Smith, R. D. *Anal. Chem.* **1998**, 70, 3553-3556.
- [53] Harvey, S. R.; VanAernum, Z. L.; Wysocki, V. H. *J. Am. Chem. Soc.* **2021**, 143, 7698-7706
- [54] Walker, T. E.; Laganowsky, A.; Russell, D. H. *Anal. Chem.* **2022**, 94, 10824-10831
- [55]) Ghosh, D. K.; Rosu, F.; Gabelica, V. *Anal. Chem.* **2022**, 94, 15386–15394.
- [56] Abian, J.; Sanchezbaeza, F.; Gelpi, E.; Barcelo, D. *J. Am. Soc. Mass Spectrom.* **1994**, 5, 186-193
- [57] Jordan, J. S.; Xia, Z.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2022**, 33, 607-611
- [58] Nguyen, G. T. H.; Tran, T. N.; Podgorski, M. N.; Bell, S. G.; Supuran, C. T.; Donald, W. A. *ACS Centr. Sci.* **2019**, 5, 308–318.

Chapter 3 Effects of Hydrogen/Deuterium Exchange on Protein Stability in Solution and in the Gas Phase.

3.1 Chapter 3 Introduction:

Native proteins in solution are stabilized by numerous noncovalent contacts. Two key factors are backbone H-bonds that mediate the formation of secondary structure, and the hydrophobic effect which causes the clustering of nonpolar residues in the core.¹⁻⁴ Salt bridges and van der Waals interactions play a role as well. Some proteins also possess disulfide bonds. Together, these interactions counteract the effects of conformational entropy, such that unfolding equilibria favor the native state N over the unfolded state U ($\Delta G_U > 0$ for $N \rightleftharpoons U$). However, stabilizing, and destabilizing factors are closely balanced, such that subtle alterations can have profound effects on the protein behavior.^{5,6}

Condensed phase techniques that provide atomically resolved protein structures include X-ray crystallography, NMR spectroscopy, and cryo-em.⁷ Complementary insights come from MS-based techniques, such as covalent labeling⁸⁻¹⁰ and crosslinking.¹¹⁻¹² A concern with the latter two strategies is the possibility that covalent modifications can perturb the protein behavior, such that labeling or crosslinking patterns might not fully reflect the properties of the native state. This potential problem necessitates careful controls to ensure the absence of artifacts.¹¹⁻¹³

HDX experiments with MS or NMR detection represent another important tool for probing protein structure and dynamics.¹⁴⁻¹⁶ In these studies, the protein is incubated in D₂O-based labeling buffer, triggering the exchange of N, O, and S-linked protium (H) with deuterium (D). HDX in exposed side chains proceeds on a sub-second time scale. H-bonded backbone sites exchange more slowly, requiring seconds to weeks at physiological pH.¹⁷ This slow backbone deuteration is mediated by dynamic protein motions.¹⁴⁻¹⁶

Compared to covalent labeling or crosslinking, HDX is less intrusive. Many HDX studies implicitly assume that replacing H with D is completely benign and does not affect protein structure and dynamics.¹⁸ However, this is not necessarily true. Compared to H₂O, D₂O has a 10% higher viscosity, 10% higher maximum density, and 7 K higher temperature of maximum density.¹⁹ These differences can affect the properties of H-bonded systems.²⁰ A number of studies have reported that protein incubation in D₂O stabilizes the native state, evident from an increased T_m ²¹⁻²⁴ and from a larger (more positive) ΔG_U ²¹⁻²³ D₂O can also enhance the rigidity of the native state.²⁵

The origin^{21,22,26} and extent²⁷ of protein stabilization in D₂O remain controversial. Several studies have attributed D₂O-induced stabilization to enhancement of the hydrophobic effect in D₂O, i.e., a lower solubility of nonpolar side chains in D₂O compared to H₂O.^{23,28} This scenario may arise from stronger “H”-bonds among solvent molecules, i.e., stronger DOD··OD₂ contacts compared to HOH··OH₂.²⁸⁻³¹ Indeed, gaseous D₂O dimers are more stable than H₂O dimers, an effect that is related to shifts in the zero-point vibrational energy (ZPVE).³²⁻³⁴

Understanding ZPVE effects on H-bond stability is not straightforward. A harmonic oscillator has $ZPVE = \frac{1}{2}h\nu$ with the frequency $\nu = (2\pi)^{-1} (K/m)^{1/2}$, where h is Planck’s constant and K is the force constant.³⁵ Replacing a vibrating H-atom with D lowers ZPVE, because $m_D > m_H$. However, a lower ZPVE does not necessarily strengthen H-bonds. The enthalpy ΔH_{HB} of H-bond dissociation $XH\cdots Y \rightarrow XH + Y$ may increase, decrease, or stay the same upon deuteration. The direction and magnitude of the stability change depends on whether the bound or the unbound state experiences a larger ZPVE shift (Figure 3.1). The situation becomes even more convoluted in large systems with many vibrational modes, particularly in the presence of charges.³⁶ In such cases, the stability trend can be reversed, making D-bonds more stable than H-bonds.³⁷ It has also been noted that D- vs. H-bond stability differences are most prevalent at cryogenic temperature, while entropic factors diminish this difference under ambient conditions.³⁸ In addition, the dissociation of some H-bonds often allows the formation of other H-bonds, e.g., when backbone NH··OC contacts are replaced with water-protein bonds upon unfolding.⁶ In summary, the mechanism of protein stabilization in D₂O remains elusive, although the purported higher stability of D-bonds vs. H-bonds features prominently in most explanation attempts.

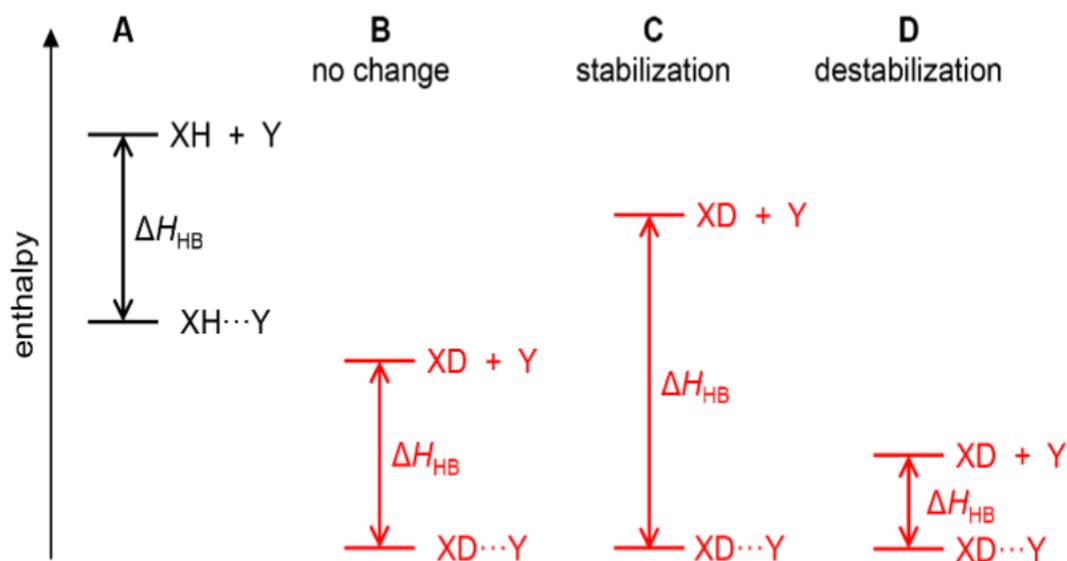


Figure 3.1. Schematic illustration of possible deuteration-induced effects on the H-bond strength ΔH_{HB} . (A) ΔH_{HB} prior to deuteration. (B) Deuteration lowers the ZPVE of both energy levels equally; ΔH_{HB} remains unchanged. (C) ZPVE of the bound state gets lowered more; ΔH_{HB} increases. (D) ZPVE of the dissociated state gets lowered more; ΔH_{HB} decreases.

We propose that it should be possible to streamline the discussion of protein stabilization in D_2O by dissecting H-bonds into three categories.^{39,40} (i) Water-water ($\text{W}\cdots\text{W}$) bonds, i.e., $\text{DOD}\cdots\text{OD}_2$ vs. $\text{HOH}\cdots\text{OH}_2$; (ii) water-protein ($\text{W}\cdots\text{P}$) bonds, i.e., $\text{D}_2\text{O}\cdots\text{protein}$ vs. $\text{H}_2\text{O}\cdots\text{protein}$; (iii) intramolecular protein-protein ($\text{P}\cdots\text{P}$) bonds, i.e., $\text{ND}\cdots\text{OC}$ vs. $\text{NH}\cdots\text{OC}$. With only a few exceptions^{39,41}, previous discussions focused on $\text{W}\cdots\text{W}$ bonds^{21-23,26,29-30} while ignoring the possible involvement of other factors. In particular, it has not been possible to uncover the relevance of $\text{P}\cdots\text{P}$ bonds, partly because many samples had either incomplete²³ or poorly controlled backbone deuteration.^{31,39,42,43}

The premise of the current work is that it should be possible to separate the role of $\text{P}\cdots\text{P}$ bonds from the solvent-linked contributions ($\text{W}\cdots\text{W}$ and $\text{W}\cdots\text{P}$) by examining solvent-free proteins. Under native ESI conditions (non-denaturing solutions, minimal collisional excitation), solution-like protein structures survive in the gas phase⁴⁴⁻⁴⁸ with retention of most backbone H-bonds.⁴⁹⁻⁵² The stability of these gaseous ions can be assessed in CIU experiments, where conformational changes are detected by IMS.⁵³⁻⁶⁰ Thus, comparative CIU experiments on deuterated and unlabeled protein ions should reveal whether the presence of $\text{ND}\cdots\text{OC}$ vs. $\text{NH}\cdots\text{OC}$ bonds stabilizes native

proteins. The observation of deuteration-induced stabilization in the gas phase would suggest that P··P bonds also cause stabilization of deuterated proteins in D₂O solution. Conversely, the absence of deuteration-induced stabilization in the gas phase would imply that solvent-linked contributions (W··W and/or W··P) are responsible for the higher stability of proteins in D₂O.

By conducting thermal unfolding experiments on fully deuterated proteins in D₂O and H₂O solution, the current work confirms the existence of D₂O-induced stabilization. However, deuterated and unlabeled gaseous protein ions generated by native ESI exhibited indistinguishable stability. We conclude that protein stabilization in D₂O is caused solely by solvent effects.

3.2 Materials and Methods

Materials and Sample Preparation. Equine heart cytochrome *c* (cyt *c*, 12360 Da), hen egg white lysozyme (14305 Da), and bovine ubiquitin (8565 Da) were supplied by Millipore Sigma (St. Louis, MO). D₂O was from Isowater (Collingwood, ON). All other chemicals were purchased from Thermo Fisher Scientific (Mississauga, ON). An AB15 glass electrode pH-meter (Fisher) was used for pH measurements; “pD” values referenced throughout this work are glass electrode readings that were corrected according to $\text{pD} = (\text{pH meter reading}) + 0.4$.⁶¹ Protein stock solutions (500 μM) were initially dialyzed for 24 h against water for removal of salt contaminants using 10 kDa MWCO Millipore Sigma dialysis cassettes. Deuteration was performed by incubating protein samples in 99% D₂O v/v at 42 °C for three weeks, at a protein concentration of 5 μM (10 μM for cyt *c*) at pD 5.3 in 10 mM D₂O-based acetate buffer. Sodium acetate was used in optical experiments to ensure consistency with earlier unfolding experiments⁶², whereas ammonium acetate was used for ESI-MS. Unlabeled samples were treated exactly the same way, except that H₂O was used in all steps instead of D₂O. Control experiments revealed that D₂O incubation periods beyond three weeks did not further enhance the deuteration percentage (data not shown).

Unfolding in Solution. Thermal unfolding experiments were performed with circular dichroism (CD) spectroscopy detection on a Jasco J-810 instrument (Easton, MD) with a 1 mm cuvette using 5 μM lysozyme and 10 μM cyt *c*. The ellipticity was measured at 222 nm (which reports on α -helicity⁶³), while heating the solutions from 21° C to 100° C at 1° C min⁻¹. Experiments in H₂O were performed at pH 4.9 and pH 5.3, and in D₂O at pD 5.3. All measurements were conducted in triplicate. Experimentally measured ellipticities θ for all temperatures T were converted to normalized ellipticity θ_{norm} according to

$$\theta_{\text{norm}} = \frac{\theta - \theta_N}{\theta_U - \theta_N} \quad (3.1)$$

where θ_N and the θ_U are the ellipticity of the native protein at 21° C and the unfolded protein at 100° C, respectively. Thermodynamic parameters were determined by fitting the experimental θ_{norm} profiles using^{6,64}

$$\theta_{\text{norm}} = \frac{(y_N + m_N T) + (y_U + m_U T) \exp\left(-\frac{\Delta G_U}{RT}\right)}{1 + \exp\left(-\frac{\Delta G_U}{RT}\right)} \quad (3.2)$$

where

$$\Delta G_U = \Delta H_U \left(1 - \frac{T}{T_m}\right) \quad (3.3)$$

is the free energy of the $N \rightleftharpoons U$ unfolding equilibrium, and ΔH_U is the corresponding enthalpy. The $(y_N + m_N T)$ and $(y_U + m_U T)$ terms in eq. 2 represent the pre-and post-transition baselines. Fitting was done using Microsoft Excel Solver. From the fitted ΔH_U and T_m parameters one can calculate the fraction of unfolded protein in solution f_{U_SOL} as

$$f_{U_SOL} = \frac{[U]}{[N] + [U]} = \frac{\exp\left(-\frac{\Delta G_U}{RT}\right)}{1 + \exp\left(-\frac{\Delta G_U}{RT}\right)} \quad (3.4)$$

Keeping in mind that $0 = \Delta H_U - T_m \Delta S_U$, the entropy of unfolding (ΔS_U) is

$$\Delta S_U = \frac{\Delta H_U}{T_m} \quad (3.5)$$

Eqs. 1-5 are widely used for analyzing thermal protein unfolding in solution^{6,64}, but the fitted parameters ΔG_U , ΔH_U , and ΔS_U are only valid in the vicinity of T_m because this strategy does not consider the temperature dependence of enthalpy and entropy.

Native Mass Spectrometry and Ion Mobility Spectrometry. ESI-IMS/MS experiments were performed on a Waters SYNAPT G2 instrument in positive ion mode (Waters, Milford, MA) with the ESI capillary held at 2.8 kV. Proteins in H₂O or D₂O solution were infused at room temperature using a syringe pump at 5 $\mu\text{L min}^{-1}$. Temperatures and voltages were adjusted to ensure minimum thermal

and collision excitation during ion sampling (source 30 °C, desolvation gas 40 °C, sampling cone 5 V, extraction cone 3 V). For a full list of instrument settings, see Table S1.

CIU was performed after quadrupole selection of the most intense charge states for each protein. Collisional excitation was implemented by varying the trap collision voltage V_{trap} between 2 V and 70 V with Ar as a collision gas. To ensure that unlabeled and deuterated protein ions experienced collisions with equivalent center-of-mass translational energies E_{COM} , we used the relationship $E_{\text{COM}} = E_{\text{LAB}} m_{\text{Ar}} / (m_{\text{prot}} + m_{\text{Ar}}) \approx E_{\text{LAB}} m_{\text{Ar}} / m_{\text{prot}}$ where m_{Ar} is the mass of Ar, m_{prot} is the mass of the protein, and $E_{\text{LAB}} = z \times e \times V_{\text{trap}}$ is the laboratory-frame translational energy.^{65,66} Accordingly, V_{trap} was increased for deuterated samples by a factor of $m_{\text{prot}}(\text{deuterated}) / m_{\text{prot}}(\text{unlabeled})$. For example, excitation of unlabeled ubiquitin (8565 Da) with $V_{\text{trap}} = 50$ V is equivalent to excitation of deuterated ubiquitin with $V_{\text{trap}} = 50 \text{ V} \times (8565 + 142) / 8565 = 50.8$ V. For simplicity, V_{trap} settings will be reported as nominal values, i.e., for the example used here, both the corrected and the uncorrected value would be given as “50” V.

Protein conformational changes triggered by collisional excitation were probed by TWIMS with N₂ as the primary buffer gas. TWIMS drift times were converted to effective He collision cross sections ($^{\text{TW}}\text{CCS}_{\text{N}_2 \rightarrow \text{He}}$, referred to as “ Ω ” throughout this work).^{67,68} Average collision cross sections $\langle \Omega \rangle$ were calculated from the measured Ω distributions. The extent of CIU was quantified by calculating the fraction of unfolding in vacuum, $f_{\text{U_VAC}}$, according to

$$f_{\text{U_VAC}} = \frac{\langle \Omega \rangle - \langle \Omega \rangle_{\text{N}}}{\langle \Omega \rangle_{\text{U}} - \langle \Omega \rangle_{\text{N}}} \quad (3.6)$$

where $\langle \Omega \rangle_{\text{N}}$ represents the average collision cross section of the folded protein ions at $V_{\text{trap}} = 2$ V, while $\langle \Omega \rangle_{\text{U}}$ represents the average collision cross section of the unfolded ions at $V_{\text{trap}} = 70$ V. All CIU experiments were performed in triplicate with independent Ω calibrations. Error bars represent standard deviations.

3.3 Results and Discussion

In some cases, protein thermal stability assays in neutral solution can be challenging as unfolding takes place close to the boiling point of water. A common strategy for mitigating this problem is to lower T_m by mild acidification.^{62,69} The current work focused on cyt *c*, lysozyme, and ubiquitin. Our experiments were conducted at pH (and pD) 5.3, which is well within the stability range of these proteins at room temperature.^{18,70} Preliminary tests (not shown) revealed that thermal unfolding in solution was straightforward for cyt *c* and lysozyme at pH 5.3. However, ubiquitin is very resilient even at low pH^{62,71}, such that we were unable to characterize thermal unfolding of this protein in solution. This high stability of ubiquitin has been attributed to its tight H-bonding network and compact hydrophobic core.⁷¹ The subsequent sections will therefore discuss solution data only for cyt *c* and lysozyme, while ESI-MS and IMS/MS results are shown for all three proteins.

3.3.1 Native ESI Mass Spectra in H₂O and D₂O

ESI mass spectra of cyt *c*, lysozyme, and ubiquitin acquired in H₂O at pH 5.3 are depicted in Figure 3.2 (black traces). All three spectra show $[M + zH]^{z+}$ ions in low charge states that are consistent with tightly folded solution conformations, as seen in earlier native ESI experiments.^{50,72-75}

Protein deuteration was performed as outlined in the Methods section, and the resulting samples were electrosprayed in D₂O solution. The charge state distributions of the deuterated $[M + zD]^{z+}$ ions (Figure 3.2, red traces) were virtually identical to those obtained with unlabeled proteins in H₂O. This high degree of similarity indicates that the release of protein ions into the gas phase and the associated charging mechanism(s) are insensitive to isotope effects. The extent of deuteration was measured from the most intense peaks in the spectra (insets of Figure 3.2). Deuteration percentages were determined using $\%D = (\Delta M_{\text{exp}} / \Delta M_{\text{max}})$, where ΔM_{max} is the maximum possible mass shift for complete deuteration of all exchangeable sites (backbone, side chains, termini, and the two heme propionates in cyt *c*; the cyt *c* N-terminus is acetylated).⁷⁶ ΔM_{max} values for cyt *c*, lysozyme, and ubiquitin are 195, 255, and 144 Da, respectively. The average $\%D$ value obtained in this way was $(96 \pm 2.5)\%$. Considering that proteins can undergo some gas phase back exchange during ESI and ion sampling^{77,78}, we conclude that the D₂O labeling strategy used here generates proteins where all their labile hydrogens are deuterated. This is in contrast to several

earlier studies on proteins in H₂O vs. D₂O, where deuteration was incomplete, poorly controlled, or unreported.^{23,31,42,43}

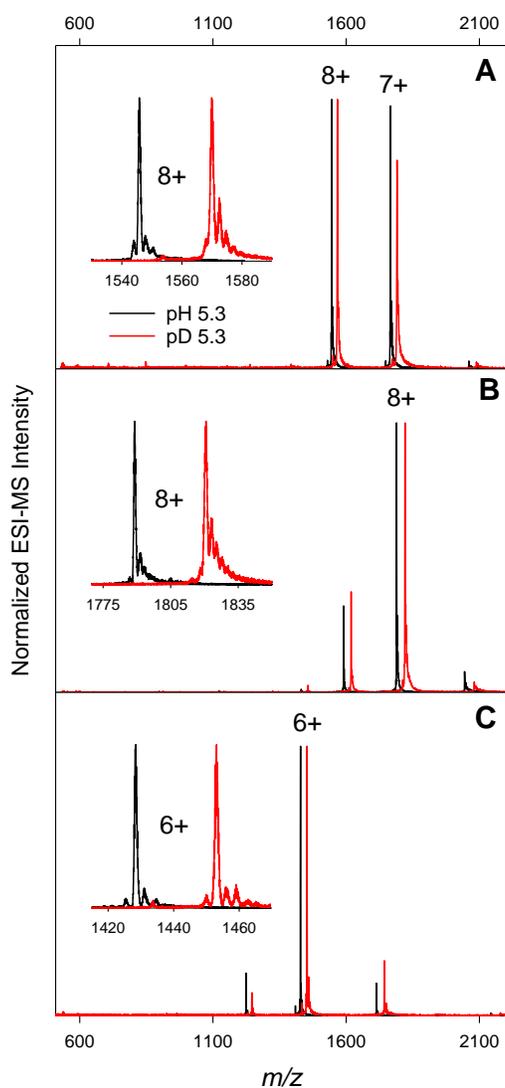


Figure 3.2. Native mass spectra of (A) cyt *c*, (B) lysozyme, (C) ubiquitin electrosprayed in H₂O solution (black, pH 5.3), and fully deuterated samples in D₂O solution (red, pD 5.3). All samples contained 10 mM ammonium acetate. Selected peaks are labeled with their charge state.

3.3.2. pH Effects on Thermal Protein Unfolding.

Prior to conducting comparative stability measurements in H₂O vs. D₂O, it is necessary to examine a potential source of artifacts. Like most earlier investigations, we used the relationship $pD = (\text{pH-meter reading}) + 0.4$ to prepare solutions with equivalent H⁺ and D⁺ activity.^{18,23-25,61} Thus, we matched samples at pH 5.3 and pD 5.3 (the latter having a pH-meter reading of 4.9). However, the appropriateness of this “+0.4 correction” has been questioned, prompting some studies to rely on uncorrected pH-meter readings in D₂O.^{22,27,79} In other words, there is a possibility that protein stability comparisons in H₂O vs. D₂O might be skewed by differences in H⁺ and D⁺ activity. To explore the severity of this issue we examined the “worst-case” scenario, where the effective acidity differs by 0.4 units. To this end, we initially performed stability measurements in H₂O solution at pH 5.3 and pH 4.9.

CD-detected thermal unfolding curves of *cyt c* and lysozyme at pH 5.3 and pH 4.9 are depicted in Figure 3.3A, D. Visual inspection of the experimental and fitted θ_{norm} data reveals subtle differences for both proteins upon changing pH by 0.4 units (black and blue in Figure 3.3A, D). However, closer analysis reveals that the T_m values of both proteins remain unchanged, within experimental error. Both proteins exhibit a slightly higher ΔH_U at pH 5.3, but even this alteration remains close to the measurement uncertainty (Table 3.1). The f_{U_SOL} profiles (Figure 3.3B, E) as well as the corresponding $\Delta G_U(T)$ data (Figure 3.3 C, F) are nearly superimposable at pH 4.9 and pH 5.3. We conclude that the thermal unfolding behavior of *cyt c* and lysozyme is virtually identical when conducting the experiments at pH 4.9 and pH 5.3. In other words, the disputed validity^{22,27} of the “+ 0.4 correction” is not an issue under the conditions of this work. We therefore continued to rely on this correction throughout this work, consistent with most other studies in the field.^{18,23-25,61}

3.3.3 D₂O-Mediated Protein Stabilization in Solution

While numerous studies have reported that D₂O enhances the thermodynamic stability of native proteins in solution^{21-25,28}, there are also voices that have questioned existence of this effect.²⁷ Instead of relying on these partially conflicting literature data, we sought to verify the occurrence of stability difference in H₂O vs. D₂O ourselves.

CD-detected unfolding profiles of fully deuterated *cyt c* and lysozyme acquired at pD 5.3 showed a notable shift to higher temperatures, compared to profiles measured at pH 5.3 (black vs.

red data in Figure 3.3A, D). The D₂O-induced T_m increase for the two proteins was 2.0 K and 4.2 K, respectively. Thermodynamic stabilization of both proteins is evident from an upward displacement of the ΔG_U profiles in D₂O relative to H₂O (Figure 3.3 C, F), implying that the N \rightleftharpoons U equilibria were shifted toward the native state in D₂O. This D₂O-induced stabilization is consistent with the results of refs.^{21-25,28} Overall, the results of Figure 3.3 confirm that *cyt c* and lysozyme are more thermodynamically stable in D₂O than in H₂O. In contrast to some earlier studies, this result was obtained for samples that had well controlled (virtually complete) deuteration, as seen from the mass shifts in Figure 3.2. We also verified that the observed isotope effect is independent of possible differences in the H⁺ vs. D⁺ activity, as discussed in the preceding section.

The free energy of unfolding is $\Delta G_U = \Delta H_U - T\Delta S_U$, allowing us to determine the enthalpic and entropic contributions to D₂O-induced stabilization. Table 1 reveals that D₂O-exposure causes ΔH_U to increase by 40 kJ mol⁻¹ and 60 kJ mol⁻¹ for *cyt c* and lysozyme, respectively. This enthalpic stabilization of the native state in D₂O is in line with earlier reports.²¹⁻²³ Interestingly, enthalpic stabilization is counteracted by a ΔS_U increase in D₂O of ca. 100 J K⁻¹ mol⁻¹ which destabilizes the native state. The occurrence of this enthalpy-entropy compensation in H₂O vs. D₂O has been noted earlier.²⁷ However, the fact that ΔG_U is more positive in D₂O than in H₂O (Figure 3.3C, F) implies that the stabilizing effect of $\Delta(\Delta H_U) > 0$ dominates over the destabilizing effect of $\Delta(\Delta S_U) < 0$.

In the absence of additional information, it is difficult to interpret D₂O-induced ΔH_U and ΔS_U effects of Table 3.1 because proteins in solution experience numerous intra- and intermolecular contacts, all of which have enthalpic and entropic contributions.^{6,22} In particular, it is not possible to unravel whether the D₂O-induced net stabilization is related to solvent effects (W··W and W··P bonds, see Introduction), or by the strengthening of H-bonds within the proteins (P··P bonds). The gas phase experiments discussed in the following section help unravel this puzzle.

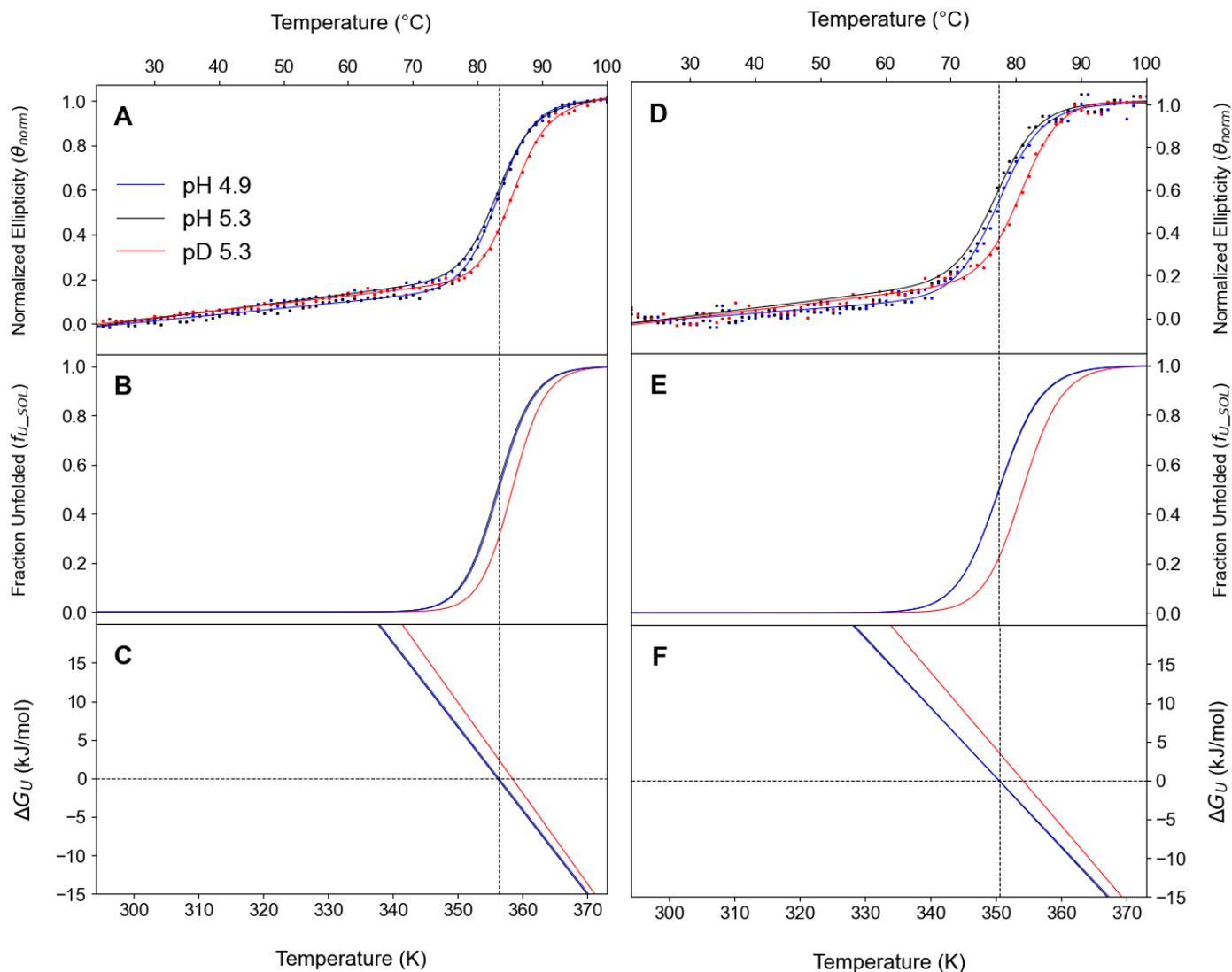


Figure 3.3. Thermodynamic analyses of *cyt c* (A-C) and lysozyme (D-F) thermal unfolding in H₂O and in D₂O solution. Colors denote pH 4.9 (blue), pH 5.3 (black), and pD 5.3 (red). Panels A, D show experimental CD unfolding profiles (dots) and the corresponding eq. 2 fits. Panels B, E depict the fraction of unfolded protein f_{U_SOL} vs. temperature (eq. 4), while panels C, F show free energy profiles. Vertical dashed lines indicate T_m values in H₂O.

Table 3.1. Thermodynamic parameters determined from thermal unfolding experiments (Figure 3.3) on *cyt c* and lysozyme in H₂O and D₂O solution.

		cyt <i>c</i>			lysozyme		
		T_m (K)	ΔH_U (kJ mol ⁻¹)	ΔS_U (J K ⁻¹ mol ⁻¹)	T_m (K)	ΔH_U (kJ mol ⁻¹)	ΔS_U (J K ⁻¹ mol ⁻¹)
H ₂ O	pH 4.9	356.4 ± 0.5	380 ± 10	1070 ± 40	350.9 ± 0.5	290 ± 10	814 ± 40
	pH 5.3	356.4 ± 0.7	390 ± 20	1080 ± 70	350.2 ± 0.7	310 ± 10	920 ± 40
D ₂ O	pD 5.3	358.4 ± 0.5	430 ± 10	1190 ± 40	354.4 ± 0.3	370 ± 20	1040 ± 30

3.3.4 Unfolding of Unlabeled and Deuterated Proteins in the Gas Phase.

We examined the CIU behavior of the most intense protein ions generated by native ESI, i.e., *cyt c* 8+, lysozyme 8+, and ubiquitin 6+. IMS profiles were acquired for V_{trap} values between 2 V and 70 V. Gas phase collisional excitation triggered large-scale unfolding, evident from shifts of the IMS distribution to higher Ω (Figure 3.4). For V_{trap} values beyond 65 V the spectral quality started to deteriorate as a result of collision-induced dissociation.

CIU of *cyt c* and ubiquitin proceeded via semi-unfolded intermediate structures, evident from features in-between the most compact and the fully unfolded species (e.g., Figures 3.4C, K). In contrast, lysozyme CIU took place without distinct intermediates. The relative increase in $\langle \Omega \rangle$ during CIU was smaller for lysozyme (21%) than for *cyt c* (49%) and ubiquitin (38%). This behavior reflects the presence of four disulfide bridges that limit the conformational freedom of unfolded lysozyme.⁸⁰ In contrast, neither *cyt c* nor ubiquitin possess disulfide bridges^{71,76}, allowing these two proteins to adopt more expanded conformations after CIU. Overall, the gas phase unfolding behavior seen in Figure 3.4 for all three proteins agrees with previous native IMS/MS data on proteins electrosprayed out of H₂O solutions^{56,59,81}

A key result of our CIU experiments is that unlabeled and deuterated protein ions exhibited IMS data that were virtually indistinguishable from one another throughout the entire range of V_{trap}

values, evident from the overlapping black and red profiles in Figure 3.4. Numerous additional V_{trap} values were tested, and these data were compiled into $f_{\text{U_VAC}}$ profiles that reflect the extent of gas phase unfolding (Figure 3.5). These $f_{\text{U_VAC}}$ data reaffirm that the CIU behavior of all three protein ions is independent of their deuteration status. The insensitivity of gas phase protein unfolding to isotope effects (Figures 3.4, 3.5) is in striking contrast to the behavior in solution, where deuteration significantly stabilizes the native state (Figure 3.3)²¹⁻²⁴. It appears that this is the first time that the CIU behavior of unlabeled vs. deuterated proteins has been compared directly.

3.3.5. Implications of H-Bonds vs. D-Bonds for Protein Stability in the Gas Phase.

The compact gas phase conformers populated in native ESI experiments with minimum collision excitation ($V_{\text{trap}} = 2$ V in Figure 3.4) retain much of their solution secondary and tertiary structure, along with preservation of most backbone $\text{NH}\cdots\text{OC}$ hydrogen bonds.^{44-47,49-52} Additionally, these gas phase proteins form side chain H-bonds as part of salt bridge networks on the protein surface.^{49,51,82,83} CIU of these compact protein ions generates significantly expanded conformers that have lost much of their secondary and tertiary structure, and where most backbone and side chain H-bonds (D-bonds) have been disrupted or rearranged.^{84,85} Our CIU data reveal that there is no stability difference in unlabeled vs. deuterated proteins, implying that the dissociation energy of $\text{P}\cdots\text{P}$ bonds in gaseous protein ions is not affected by the bridging atom (H vs. D). This finding does not support the view that D-bonds are generally more stable than H-bonds.³²⁻³⁴ Instead, our data suggest that $\text{P}\cdots\text{P}$ bonds behave in accordance with the scenario of Figure 3.1 B. D-induced stabilization has previously been found to be most prevalent in systems that are very small and neutral, such as H_2O dimers.³⁷ Thus, the absence of deuteration-induced stabilization in electrosprayed protein ions (i.e., large systems with a net charge) is not completely unexpected.

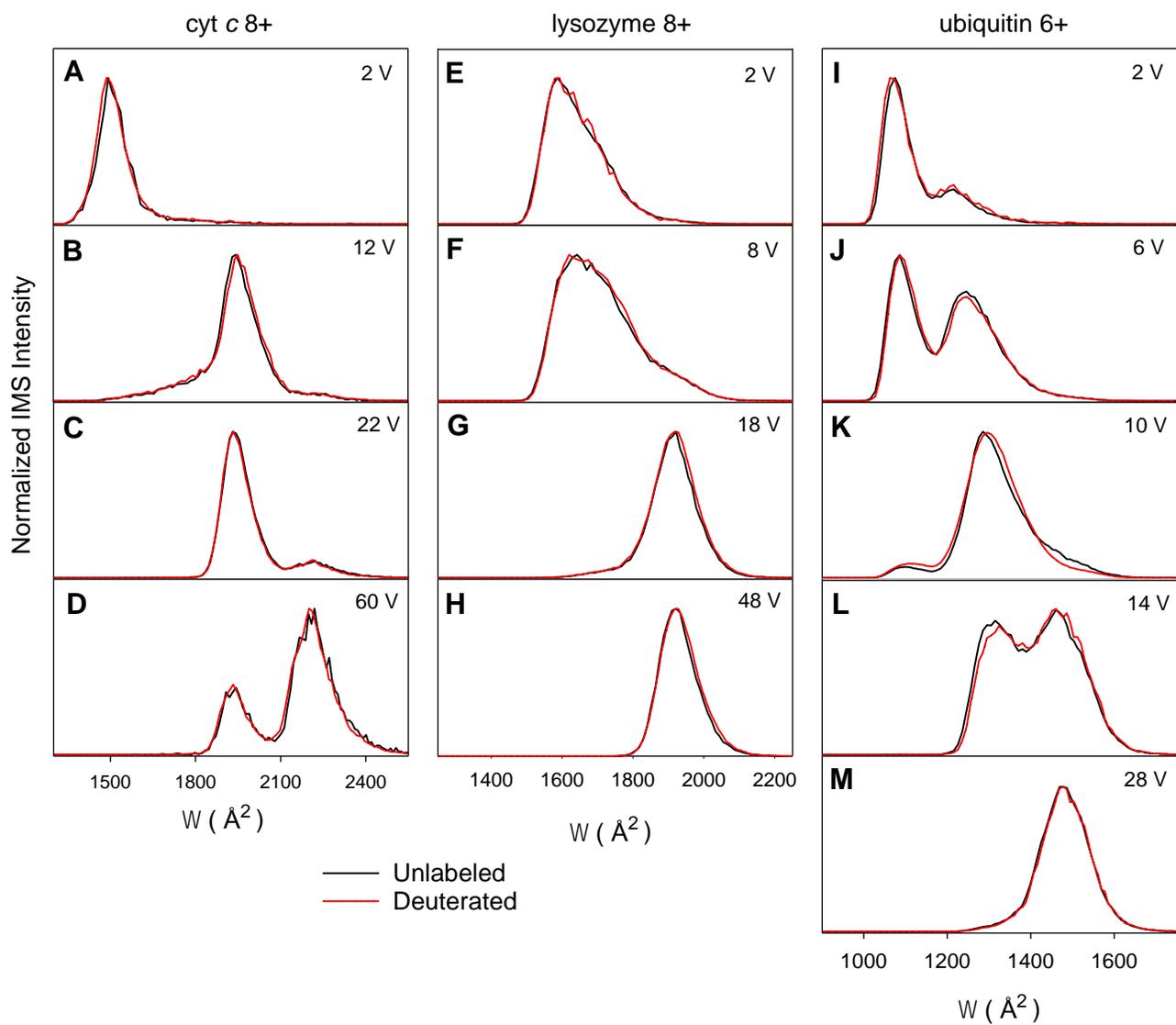


Figure 3.4. CIU data, displaying collision cross section (Ω) distributions for unlabeled (black) and deuterated (red) gaseous protein ions at different levels of collisional heating. The trap collision voltage V_{trap} is indicated in each panel. (A-D) cyt *c* 8+, (E-H) lysozyme 8+, and (I-M) ubiquitin 6+.

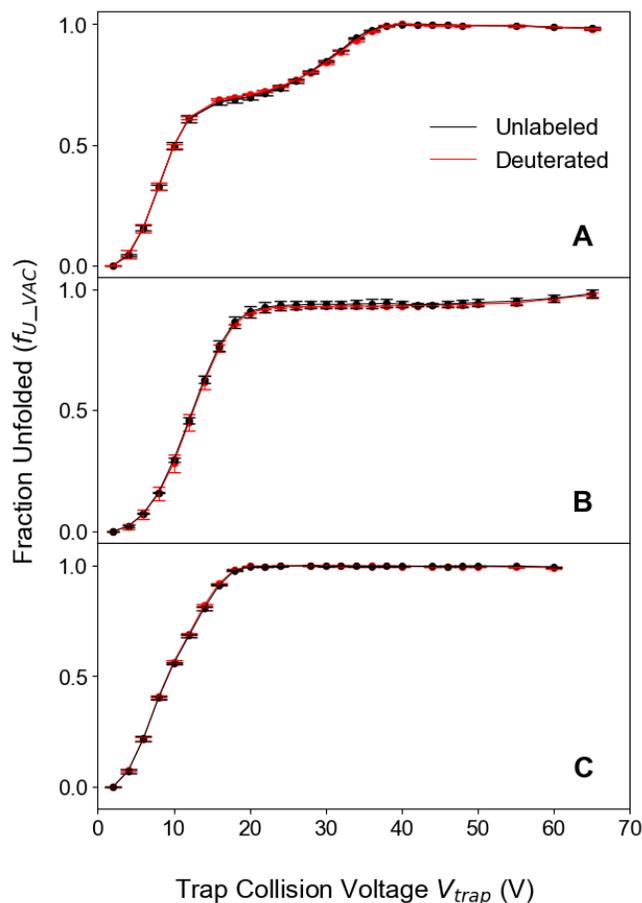


Figure 3.5. CIU profiles of deuterated vs. unlabeled (A) *cyt c* 8+, (B) lysozyme 8+, and (C) ubiquitin 6+ ions generated by native ESI. The profiles were calculated from Ω values acquired at different V_{trap} , with subsequent normalization via eq. 6.

3.3.6 Dissecting Isotope Effects on Protein Stability.

Toy models can illustrate basic protein concepts.^{5,86} Here, we use a two-dimensional lattice chain model for examining a $N \rightleftharpoons U$ equilibrium in solution (Figure 3.6). Within the model, water molecules (blue spheres) can form up to four H-bonds, hydrophilic residues (red spheres) can form two H-bonds, while hydrophobic residues (green spheres) can form only one H-bond. We assume that for any protein structure, the system will form the maximum possible number of H-bonds, i.e., $W \cdots W$, $W \cdots P$, and $P \cdots P$ contacts. The number of H-bonds in each category is n_{WW} , n_{WP} , and n_{PP} , respectively. The corresponding H-bond dissociation enthalpies are $\Delta H_{\text{HB}}(\text{WW})$, $\Delta H_{\text{HB}}(\text{WP})$, and $\Delta H_{\text{HB}}(\text{PP})$, all of which are positive (Figure 3.1).

Just like for actual proteins^{71,76,80}, the native state in our model has a hydrophobic core and a hydrophilic exterior (Figure 3.6A). The total number of H-bonds in this structure is 110. Unfolding exposes hydrophobic residues to water, thereby lowering the total number of H-bonds to 108. Thus, the model correctly captures the fact that unfolding in solution leads to a net loss of H-bonds, a factor that contributes to the hydrophobic effect.⁶ Because D₂O-induced protein stabilization is caused by enthalpy (Table 1)²¹⁻²³, our discussion only focuses on ΔH_U effects, while not examining ΔS_U -related factors. ΔH_U in our model is given by

$$\Delta H_U = -\Delta n_{WW} \Delta H_{HB}(WW) - \Delta n_{WP} \Delta H_{HB}(WP) - \Delta n_{PP} \Delta H_{HB}(PP) \quad (3.7)$$

Comparison of Figure 3.6A, B reveals that the number of P··P bonds decreases as the protein unfolds. These broken intramolecular contacts are then replaced with newly formed W··P bonds. Intrusion of the unfolded chain into the water network decreases the number of W··W bonds. These trends also apply to actual proteins,⁶ although the magnitude of the Δn terms is system dependent. For our model, $\Delta n_{WW} = -9$, $\Delta n_{WP} = 14$, and $\Delta n_{PP} = -7$, such that

$$\Delta H_U = 9 \times \Delta H_{HB}(WW) - 14 \times \Delta H_{HB}(WP) + 7 \times \Delta H_{HB}(PP) \quad (3.8)$$

The data in Table 1 demonstrate that protein stabilization in D₂O results from a shift of ΔH_U to more positive values ($\Delta(\Delta H_U) > 0$). Eq. 3.8 reveals that this stabilization may be caused by three factors, i.e., an increase of $\Delta H_{HB}(WW)$, a decrease of $\Delta H_{HB}(WP)$, or an increase of $\Delta H_{HB}(PP)$.

Which of these three possibilities is most likely? The CIU data of Figure 3.4, 3.5 show that $\Delta H_{HB}(PP)$ is insensitive to isotope effects, such that this possibility can be excluded (Figure 3.1B). Early work indicated that D₂O enhances the hydrophobic effect by lowering the solubility of nonpolar side chains, suggesting that $\Delta H_{HB}(WW)$ increases in D₂O (Figure 3.1C).²⁸ However, subsequent studies found the opposite trend, i.e., higher or identical solubilities of nonpolar molecules in D₂O vs. H₂O.²⁶ These later findings cast doubt on the traditional belief that $\Delta H_{HB}(WW)$ increases in bulk D₂O, even though this stability trend holds for isolated D₂O dimers.³²⁻³⁴

A possible resolution of this conundrum is that $\Delta H_{HB}(WP)$ decreases in D_2O solution, a scenario that has not thus far been considered in the literature. While our data do not provide conclusive proof for weakened $W \cdots P$ contacts as the cause of protein stabilization in D_2O , it appears that this scenario is consistent with all of the available data. As noted in Figure 3.1D, such a destabilization of H-bonds is well within the realm of possible outcomes after HDX.^{36,37}

3.4 Conclusions

60+ years after its discovery,²⁸ the stabilization of proteins in D_2O remains poorly understood. In agreement with earlier work, we found that this stabilization is caused by enthalpic effects, i.e., a larger (more positive) value of ΔH_U in D_2O than in H_2O .²¹⁻²³ Like those earlier studies, we attribute this stabilization to changes in the dissociation enthalpy of H-bonds. While previous studies focused almost exclusively on $W \cdots W$ bonds,^{21-23,26,28-30} we took a broader approach and also considered the role of $W \cdots P$ and $P \cdots P$ bonds, because stability changes in all three categories can affect the protein behavior in D_2O (Figure 3.6).

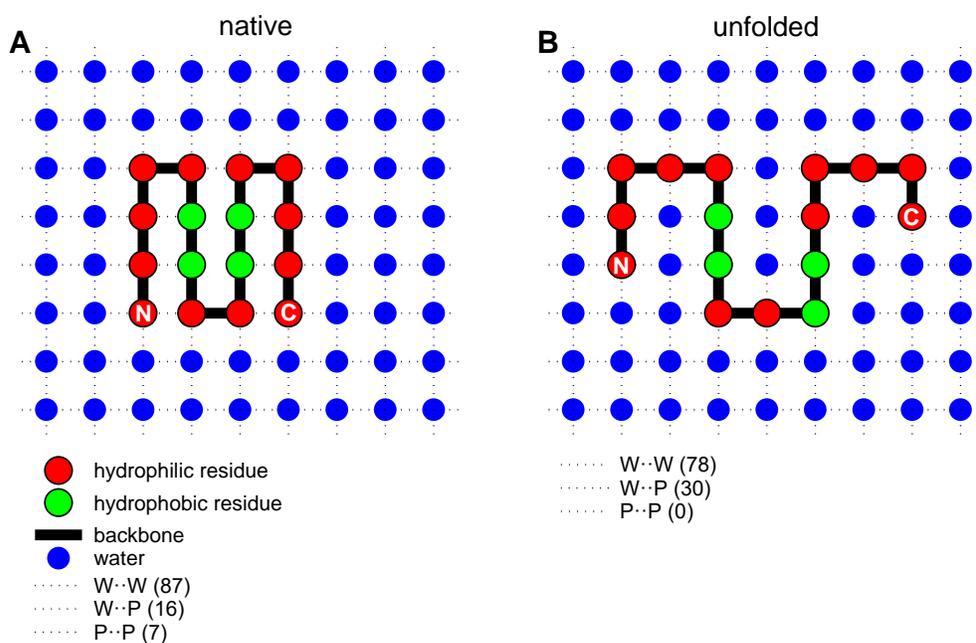


Figure 3.6. Two-dimensional lattice bead chain model of a protein in water. (A) Native protein. (B) Example of an unfolded conformation. Hydrophilic (red) and hydrophobic (green) residues are linked by backbone bonds (solid black line). Termini are marked as “N” and “C”. Blue spheres represent water. H-bonds are indicated as dotted lines of three types ($W \cdots W$, $W \cdots P$, $P \cdots P$). The

corresponding n_{HB} values are shown in brackets. Bonds around the periphery of the lattice were not included in the n_{HB} counts.

As far as we are aware, this work marks the first time that the stability of deuterated and unlabeled proteins has been examined *in vacuo* (although isotope effects on gaseous protein-ligand complexes have been explored earlier).⁸⁷ Our CIU experiments revealed that the stability of gaseous protein ions is indistinguishable before and after deuteration, demonstrating that P··P bonds are insensitive to isotope effects. It can be concluded that protein stabilization in D₂O arises either from strengthened W··W bonds, or from weakened W··P bonds (a combination of both scenarios is possible as well). Strengthening of W··W bonds has been favored in the earlier literature.^{21-23,28-30} However, weakening of W··P bonds seems just as likely, especially when considering the results of more recent solubility studies.²⁶ Thus, while we cannot conclusively determine the mechanistic basis of protein stabilization in D₂O, our results show that this stabilization is caused by the solvent, rather than H-bonds within the protein.

For HDX-MS and HDX-NMR experiments, our results imply that exposure to D₂O may alter certain aspects of protein structure and dynamics. The stability differences seen in our solution experiments were detected at relatively high temperatures, around T_m . Although typical HDX experiments use ambient temperature,¹⁴⁻¹⁶ there is great interest in using HDX/MS for high temperature measurements as well.⁸⁸⁻⁸⁹ D₂O-induced stability enhancements should be taken into account for the interpretation of such high-temperature HDX data. It is likely that D₂O-induced stabilization makes its presence felt already at room temperature, e.g., as a rigidification of the native state²⁵, but more work is required to characterize the extent of these changes. Careful comparison of protein HDX and DHX kinetics⁹⁰ are a possible way to explore this aspect in the future.

Complementary to HDX in solution, gas phase HDX can probe electrosprayed proteins in a solvent-free environment.⁹¹⁻⁹⁴ It is reassuring that our CIU data did not show any difference for the gas phase unfolding of deuterated and unlabeled protein ions, implying that gas phase HDX represents a truly “benign” labeling methods that does not perturb protein behavior *in vacuo*. This is in contrast to HDX in solution, where deuteration causes stability changes that are clearly observable.

3.5 References:

- [1] Udgaonkar, J.B. *Arch. Biochem. Biophys.* **2013**, 531, 24-33
- [2] Pace, C.N., Fu, H.L., Fryar, K.L., Landua, J., Trevino, S.R., Schell, D., Thurlkill, R.L., Imura, S., Scholtz, J.M., Gajiwala, K., Sevcik, J., Urbanikova, L., Myers, J.K., Takano, K., Hebert, E.J., Shirley, B.A., Grimsley, G.R. *Protein Sci.* **2014**, 23 (5), 652-661.
- [3] Brini, E., Fennell, C.J., Fernandez-Serra, M., Hribar-Lee, B., Luksic, M., Dill, K.A. *Chem. Rev.* **2017**, 117 (19), 12385-12414
- [4] Tanford, C. *J. Am. Chem. Soc.* **1962**, 84 (22), 4240-4247
- [5] Kazlauskas, R. *Chem. Soc. Rev.* **2018**, 47 (24), 9026-9045 .
- [6] Fersht, A.R. W. H. Freeman & Co.: New York, **1999**.
- [7] Dobson, C.M. *Annu. Rev. Biochem.* **2019**, 88, 25-33.
- [8] Ziemianowicz, D.S., MacCallum, J.L., Schriemer, D.C. *J. Am. Soc. Mass Spectrom.* **2020**, 31(2), 207-216.
- [9] Sharp, J.S., Chea, E.E., Misra, S.K., Orlando, R., Popov, M., Egan, R.W., Holman, D., Weinberger, S.R. *J. Am. Soc. Mass Spectrom.* **2021**, 32 (7), 1601-1609.
- [10] Liu, X.R., Zhang, M.M., Gross, M.L. *Chem. Rev.* **2020**, 120, 4355-4454.
- [11] Piersimoni, L., Kastritis, P.L., Arlt, C., Sinz, A. *Chem. Rev.* **2022** 122 (8), 7500-7531.
- [12] Leitner, A., Bonvin, A., Borchers, C.H., Chalkley, R.J., Chamot-Rooke, J., Combe, C.W., Cox, J., Dong, M.Q., Fischer, L., Gotze, M., Gozzo, F.C., Heck, A.J.R., Hoopmann, M.R., Huang, L., Ishihama, Y., Jones, A.R., Kalisman, N., Kohlbacher, O., Mechtler, K., Moritz, R.L., Netz, E., Novak, P., Petrotchenko, E., Sali, A., Scheltema, R.A., Schmidt, C., Schriemer, D., Sinz, A., Sobott, F., Stengel, F., Thalassinou, K., Urlaub, H., Viner, R., Vizcaino, J.A., Wilkins, M.R., Rappsilber, J. *Structure* **2020**, 28 (11), 1259-1268.
- [13] Mendoza, V.L., Vachet, R.W. *Mass Spectrom. Rev.* **2009**, 28, 785-815.
- [14] Masson, G.R., Burke, J.E., Ahn, N.G., Anand, G.S., Borchers, C.H., Brier, S., Bou-Assaf, G.M., Engen, J.R., Englander, S.W., Faber, J.H., Garlish, R.A., Griffin, P.R., Gross, M.L., Guttman, M., Hamuro, Y., Heck, A.J.R., Houde, D., Jacob, R.E., Jorgensen, T.J.D., Kaltashov, I.A., Klinman, J.P., Konermann, L., Man, P., Mayne, L., Pascal, B.D., Reichmann, D., Skehel, M., Snijder, J., Strutzenberg, T.S., Underbakke, E.S., Wagner, C., Wales, T.E., Walters, B.T., Weis, D.D., Wilson, D.J., Wintrode, P.L., Zhang, Z., Zheng, J., Schriemer, D.C., Rand, K.D. *Nat. Methods.* **2019**, 16 (7), 595-602.

- [15] James, E.I., Murphree, T.A., Vorauer, C., Engen, J.R., Guttman, M. *Chem. Rev.* **2022**, 122 (8), 7562-7623
- [16] Deng, B., Lento, C., Wilson, D.J. *Anal. Chim. Acta.* **2016**, 940, 8-20.
- [17] Skinner, J.J., Lim, W.K., Bedard, S., Black, B.E., Englander, S.W. *Protein Sci.* **2012**, 21 (7), 996-1005.
- [18] Goto, Y., Hagihara, Y., Hamada, D., Hoshino, M., Nishii, I. *Biochemistry.* **1993**, 32, 11878-11885.
- [19] Lide, D.R. *CRC Handbook of Chemistry and Physics* 82nd ed.; CRC Press: Boca Raton, London, New York, Washington, 2001.
- [20] Shi, C., Zhang, X., Yu, C.-H., Yao, Y.-F., Zhang, W. *Nat. Commun.* **2018**, 9,481.
- [21] Pica, A., Graziano, G. *Biopolymers.* **2018**, 109 (10), 6.
- [22] Stadmiller, S.S., Pielak, G.J. *Protein Sci.* **2018**, 27 (9), 1710-1716.
- [23] Efimova, Y.M., Haemers, S., Wierczinski, B., Norde, W., van Well, A.A. *Biopolymers.* **2007**, 85 (3), 264-273.
- [24] Hattori, A., Crespi, H.L., Katz, J.J. *Biochemistry.* **1965**, 4, 1213.
- [25] Cioni, P., Strambini, G.B. *Biophys.J.* **2002**, 82, 3246-3253
- [26] Graziano, G. *J. Chem. Phys.* **2004**, 121, 1878-1882
- [27] Makhatadze, G.I., Clore, G.M., Gronenborn, A.M. *Nat. Struct. Biol.* **1995**, 2 (10), 852-855
- [28] Kresheck, G.C., Schneider, H., Scheraga, H.A. *J. Phys. Chem.* **1965**, 69, 3132.
- [29] Nemethy, G., Scheraga, H.A. *J. Chem. Phys.* **1964**, 41 (3), 680.
- [30] Parker, M.J., Clarke, A.R. *Biochemistr.* **1997**, 36, 5786-5794 (1997)
- [31] Henderson, R.F., Henderson, T.R., Woodfin, B.M. *J. Biol. Chem.* **1970**, 245 (15), 3733.
- [32] Engdahl, A., Nelander, B. *J. Chem. Phys.* **1987**, 86 (4), 1819-1823.
- [33] Clark, T., Heske, J., Kuhne, T.D. *ChemPhysChem.* **2019**, 20 (19), 2461-2465.
- [34] Wilkinson, F.E., Peschke, M., Szulejko, J.E., McMahon, T.B. *Int. J. Mass Spectrom.* **1998**, 175 (2), 225-240.

- [35] Atkins, P. *Physical Chemistry* 10th ed.; W. H. Freeman & Co.: New York, 2010.
- [36] Buckingham, A.D., Fan-Chen, L. *Int. Rev. Phys. Chem.* **1981**, 1 (2), 253-269 .
- [37] Scheiner, S., Cuma, M. *J. Am. Chem. Soc.* **1996**, 118 (6), 1511-1521.
- [38] Kjaersgaard, A., Vogt, E., Christensen, N.F., Kjaergaard, H.G. *J. Phys. Chem. A.* **2020**, 124 (9), 1763-1774.
- [39] Itzhaki, L.S., Evans, P.A. *Protein Sci.* **1996**, 5 (1), 140-146
- [40] Frauenfelder, H., Chen, G., Berendzen, J., Fenimore, P.W., Jansson, H., McMahon, B.H., Stroer, I.R., Swenson, J., Young, R.D. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, 106, 5129-5134.
- [41] Guzzi, R., Arcangeli, C., Bizzarri, A.R. *Biophys. Chem.* **1999**, 82 (1), 9-22 .
- [42] Scheraga, H.A. *Ann. NY Acad. Sci.* **1960**, 84 (16), 608-616.
- [43] Huyghues-Despointes, B.M.P., Scholtz, J.M., Pace, C.N. *Nat. Struct. Biol.* **1999**, 6, 910-912.
- [44] Tamara, S., den Boer, M.A., Heck, A.J.R. *Chem. Rev.* **2022**, 122 (8), 7269-7326.
- [45] Mehmood, S., Allison, T.M., Robinson, C.V. *Annu. Rev. Phys. Chem.* **2015**, 66, 453-474.
- [46] Christofi, E., Barran, P. *Chem. Rev.* **2023**, 123 (6), 2902–2949.
- [47] Clemmer, D.E., Russell, D.H., Williams, E.R. *Accounts Chem. Res.* **2017**, 50 (3), 556-560.
- [48] Karch, K.R., Snyder, D.T., Harvey, S.R., Wysocki, V.H. *Ann. Rev. Biophys.* **2022**, 51, 157-179
- [49] Bakhtiari, M., Konermann, L. *J. Phys. Chem. B.* **2019**, 123 (8), 1784-1796 .
- [50] Wyttenbach, T., Bowers, M.T. *J. Phys. Chem. B* **2011**, 115 (42), 12266-12275
- [51] Breuker, K., Brüschweiler, S., Tollinger, M. *Angew. Chem. Int. Ed.* **2011**, 50, 873-877.
- [52] Bleiholder, C., Liu, F.C. *J. Phys. Chem. B.* **2019**, 123, 2756-2769 (2019)
- [53] Shelimov, K.B., Clemmer, D.E., Hudgins, R.R., Jarrold, M.F. *J. Am. Chem. Soc.* **1997**, 119 (3), 2240-2248
- [54] Dixit, S.M., Polasky, D.A., Ruotolo, B.T. *Curr. Opin. Chem. Biol.* **2018**, 42, 93-100 .
- [55] Nouchikian, L., Lento, C., Donovan, K.A., Dobson, R.C., Wilson, D.J. *J. Am. Soc. Mass Spectrom.* **2020**, 31(3), 685-692.

- [56] Hopper, J.T.S., Oldham, N.J. *J. Am. Soc. Mass Spectrom.* **2009**, 20 (10), 1851-1858.
- [57] Donor, M.T., Shepherd, S.O., Prell, J.S. *J. Am. Soc. Mass Spectrom.* **2020**, 31 (3), 602-610.
- [58] Zheng, X.Y., Kurulugama, R.T., Laganowsky, A., Russell, D.H. *Anal. Chem.* **2020**, 92 (10), 7218-7225.
- [59] Borotto, N.B., Osho, K.E., Richards, T.K., Graham, K.A. *J. Am. Soc. Mass Spectrom.* **2022**, 33 (1), 83-89.
- [60] Gadkari, V.V., Ramirez, C.R., Vallejo, D.D., Kurulugama, R.T., Fjeldsted, J.C., Ruotolo, B.T. *Anal. Chem.* **2020**, 92 (23), 15489-15496.
- [61] Glasoe, P.K., Long, F.A. *J. Phys. Chem.* **1960**, 64 (1), 188-190.
- [62] Wintrode, P.L., Makhatadze, G.I., Privalov, P.L. *Struct. Funct. Genet.* **1994**, 18 (3), 246-253.
- [63] Greenfield, N.J. *Nat. Protoc.* **2006**, 1 (6), 2876-2890
- [64] Swint, L., Robertson, A.D. *Protein Sci.* **1993**, 2 (12), 2037-2049.
- [65] Douglas, D.J. *J. Am. Soc. Mass Spectrom.* 1998, 9 (2), 101-113 .
- [66] Mayer, P.M., Poon, C. *Mass Spectrom. Rev.* **2009**, 28 (4), 608-639 .
- [67] Sun, Y., Vahidi, S., Sowole, M.A., Konermann, L. *J. Am. Soc. Mass Spectrom.* **2016**, 27, 31-40 .
- [68] Gabelica, V., Shvartsburg, A.A., Afonso, C., Barran, P., Benesch, J.L.P., Bleiholder, C., Bowers, M.T., Bilbao, A., Bush, M.F., Campbell, J.L., Campuzano, I.D.G., Causon, T., Clowers, B.H., Creaser, C.S., De Pauw, E., Far, J., Fernandez-Lima, F., Fjeldsted, J.C., Giles, K., Groessl, M., Hogan, C.J., Hann, S., Kim, H.I., Kurulugama, R.T., May, J.C., McLean, J.A., Pagel, K., Richardson, K., Ridgeway, M.E., Rosu, F., Sobott, F., Thalassinos, K., Valentine, S.J., Wyttenbach, T. *Mass Spectrom. Rev.* **2019**, 38 (3), 291-320.
- [69] Privalov, P.L., Khechinashvili, N.N. *J. Mol. Biol.* **1974**, 86 (3), 665-684.
- [70] Fink, A.L., Calciano, L.J., Goto, Y., Kurotso, T., Palleros, D.R. *Biochemistry.* **1994**, 33 (41), 12504-12511.
- [71] Vijay-Kumar, S., Bugg, C.E., Cook, W.J. *J. Mol. Biol.* **1987**, 194 (3), 531-544 .
- [72] Grandori, R. *Protein Sci.* **2002**, 11 (3), 453-458.
- [73] Konermann, L., Metwally, H., Duez, Q., Peters, I. *Analyst.* **2019**, 144 (21), 6157-6171.
- [74] Kaltashov, I.A., Abzalimov, R.R. *J. Am. Soc. Mass Spectrom.* **2008**, 19 (9), 1239-1246 .

- [75] Chowdhury, S.K., Katta, V., Chait, B.T. **1990**, 112 (24), 9012-9013.
- [76] Bushnell, G.W., Louie, G.V., Brayer, G.D. *J. Mol. Biol.* **1990**, 214 (2), 585-595.
- [77] Katta, V., Chait, B.T. *J. Am. Chem. Soc.* **1993**, 115 (14), 6317-6321.
- [78] Guttman, M., Wales, T.E., Whittington, D., Engen, J.R., Brown, J.M., Lee, K.K. *J. Am. Soc. Mass Spectrom.* **2016**, 27, 662-668.
- [79] Bundi, A., Wuthrich, K. *Biopolymers.* **1979**, 18 (2), 285-297.
- [80] Cheetham, J.C., Artymiuk, P.J., Phillips, D.C. *J. Mol. Biol.* **1992**, 224 (3), 613-628.
- [81] Shi, H., Atlasevich, N., Merenbloom, S.I., Clemmer, D.E. *J. Am. Soc. Mass Spectrom.* **2014**, 25 (12), 2000-2008.
- [82] Bonner, J.G., Lyon, Y.A., Nellessen, C., Julian, R.R. *J. Am. Chem. Soc.* **2017**, 139 (30), 10286-10293.
- [83] Zhang, Z., Browne, S.J., Vachet, R.W. *J. Am. Soc. Mass Spectrom.* **2014**, 25 (4), 604-613.
- [84] Zhou, M.W., Liu, W.J., Shaw, J.B. *Anal. Chem.* **2020**, 92 (2), 1788-1795
- [85] Bartman, C.E., Metwally, H., Konermann, L. *Anal. Chem.* **2016**, 88 (13), 6905-6913 .
- [86] Dill, K.A., Chan, H.S. *Nat. Struct. Biol.* **1997**, 4, 10-19.
- [87] Liu, L., Michelsen, K., Kitova, E.N., Schnier, P.D., Brown, A., Klassen, J.S. *J. Am. Chem. Soc.* **2012**, 134 (13), 5931-5937.
- [88] Gao, S., Klinman, J.P. *Curr. Op. Struct. Biol.* **2022**, 75, 102434.
- [89] Tajoddin, N.N., Konermann, L. *Anal. Chem.* **2022**, 94 (44), 15499-15509.
- [90] Xiao, H., Hoerner, J.K., Eyles, S.J., Dobo, A., Voigtman, E., Mel'Cuk, A.I., Kaltashov, I.A. *Protein Sci.* **2005**, 14 (2), 543-557.
- [91] Chaturvedi, R., Webb, I.K. *Anal. Chem.* 2022, 94 (25), 8975–8982.
- [92] Pan, J., Heath, B.L., Jockusch, R.A., Konermann, L. *Anal. Chem.* **2012**, 84 (1), 373-378.
- [93] Karanji, A.K., Khakinejad, M., Kondalaji, S.G., Majuta, S.N., Attanayake, K., Valentine, J.S. *J. Am. Soc. Mass Spectrom.* **2018**, 29 (12), 2402-2412.

[94] Cox, H.A., Julian, R.R., Lee, S.-W., Beauchamp, J.L. *J. Am. Chem. Soc.* **2004**, 126 (20), 6485-6490.

Chapter 4 Summary and conclusion

ESI-MS remains a popular technique for interrogating biological molecules. This thesis sheds light on some important aspects of native ESI and protein labeling. Native ESI has opened exciting avenues in biophysical and medicinal chemistry as the soft ionization afforded by ESI enables the analysis of large intact molecules and complexes.¹ However, the extent to which these gaseous ions retain native like structures has been the subject of continuous debate.² Proteins under denaturing conditions generate $[M+zH]^{z+}$ protein ions with high charge states z .³ These ions likely form according to CEM where unfolded proteins are electrostatically ejected from a shrinking droplet.³ In contrast, native ESI utilizes aqueous solutions at pH 7, generating protein ions with low z . Protein ions formed by native ESI conditions likely form according to CRM where the final nanodroplets undergo evaporation to dryness.³ While Native ESI is a valuable tool that enables the characterization of biomolecules in the gas phase, practitioners should be aware of some of its limitations.

In the second chapter, we aimed to explore the tuning of experimental parameters and address how ESI can be used to generate native-like structures. In addition, we discussed how various salt additives can affect the outcome of native ESI experiments. Our results demonstrate that Low sample cone voltages were found to be less efficient in ion transmission and desolvation but ensured that the protein's native structure remained intact. Finally, we sought to clear up some misconceptions related to NH_4Ac . Our data convey that ESI acidification of NH_4Ac containing droplets does not dramatically alter the outcome of ESI experiments.

Chemical labeling of proteins in conjunction with MS is a commonly used tool that provides insights into protein structure and dynamic.¹ This technique is especially important for proteins that are not amenable to other tools like X-ray crystallography or NMR spectroscopy. The three most common labeling methods are covalent labeling, covalent cross-linking, and H/D exchange.¹ While covalent labeling and cross-linking are thought to potentially perturb protein structures, H/D exchange is usually assumed to be a benign labeling method where proteins are incubated in D_2O , and labile hydrogens are exchanged with deuterium.⁴ HDX occurs via transient

opening/closing fluctuations that disrupt hydrogen bonds and provide access to NH sites.⁵ Several studies reported an increase in melting temperature T_m for proteins in D_2O . The origin of this effect is poorly understood.³

Chapter 3 scrutinized the assumption that HDX is a benign labeling technique and investigated the purported stability of proteins when exposed to D_2O . The results of this study reveal that deuterated proteins are stabilized in solution. This was seen as cytochrome c was stabilized by **2 K**, whereas lysozyme was stabilized by **4 K**. Conversely, these proteins exhibited different behaviour in the gas phase, where no structural stabilization was observed. The absence of the solvent for electrosprayed proteins is expected to weaken the hydrophobic effect while strengthening any electrostatic interactions. The fact that no gas-phase stabilization was observed in cytochrome-c, lysozyme, and ubiquitin implies that protein stability enhancements in D_2O arise from solvent effects brought about by the different physiochemical properties of D_2O . This can be attributed to 1. Weaker W-P contacts or 2. Stronger W-W contacts.

Future work: sub-micrometer nano ESI tips

Electrospray tips with submicron emitters have the advantage of desalting ESI droplets from non-volatile salts. Sub-micrometer emitters can be made from borosilicate capillaries with very small diameters (in the μm range).⁶ When small-tip diameters are used for solutions containing high concentrations of non-volatile salts, desalting usually occurs spontaneously and is characterized by a drop in the total ion signal, resolved protein charge state-distributions, and a reduction in salt cluster.⁶ However, one of the major problems for such emitters is that they are easily clogged, degrading ESI efficiency.⁷ Hence, future work should be focused on fabricating emitters that have a reduced risk of clogging. This enables the spray of solutions containing typical biochemical buffers like phosphate or Tris buffers without compromising the quality of the spectra.

Future work: HDX Silent Dynamics.

Having studied the stability effects of labeling proteins with deuterium, subsequent work will focus on the shortcomings of HDX. Previous studies by Konermann et al. have shown that some major cytochrome c fluctuations cannot be detected by HDX.⁸ The occurrence of such dynamics illustrates that deuteration may not always provide a comprehensive view of protein dynamics. Hence, we aim to build on previous work by studying HDX-silent fluctuations in myoglobin. The

goal is to complement classical HDX models to account for transitions that are not (or weakly) coupled to changes in H-bonding.⁸

References:

- [1] Konijnenberg, A.; Butterer, A.; Sobott, F. *Biochim. Biophys. Acta* **2013**, *1834* (6), 1239-1256.
- [2] Westphall, M. S.; Lee, K. W.; Salome, A. Z.; Lodge, J. M.; Grant, T.; Coon, J. J. *Nat. Commun.* **2022**, *13* (1), 6
- [3] Konermann, L.; Metwally, H.; Duez, Q.; Peters, I. *Analyst* **2019**, *144*, 6157-6171
- [4] Makhatadze, G. I.; Clore, G. M.; Gronenborn, A. M. *Nat. Struct. Biol.* **1995**, *2* (10), 852-855.
- [5] James, E. I.; Murphree, T. A.; Vorauer, C.; Engen, J. R.; Guttman, M. *Chem. Rev.* **2022**, *122* (8), 7562-7623,
- [6] Westphall, M. S.; Lee, K. W.; Salome, A. Z.; Lodge, J. M.; Grant, T.; Coon, J. J. *Nat. Commun.* **2022**, *13* (1), 6
- [7] Chen, S.; Zeng, J.; Zhang, Z.; Xu, B.; Zhang, B. J. *Chromatogr.open.* **2022**, *2*, 100064.
- [8] Scrosati, P. M.; Yin, V.; Konermann, L. *Anal. Chem.* **2021**, *93* (42), 14121-14129.

Appendix-Permissions

Effects of Hydrogen/Deuterium Exchange on Protein Stability in Solution and in the Gas Phase



Author: Yousef Haidar, Lars Konermann

Publication: Journal of the American Society for Mass Spectrometry

Publisher: American Chemical Society

Date: Jul 1, 2023

Copyright © 2023, American Society for Mass Spectrometry. Published by the American Chemical Society. All rights reserved.

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and 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 RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

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

Mechanism of Magic Number NaCl Cluster Formation from Electrosprayed Water Nanodroplets



Author: Lars Konermann, Yousef Haidar

Publication: Analytical Chemistry

Publisher: American Chemical Society

Date: Nov 1, 2022

Copyright © 2022, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and 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 RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

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

Yousef Haidar

Curriculum Vitae

Research Experience

Master's Candidate

London, ON

Western University, Konermann Lab

September 2021 – September 2023

- Study protein stability in the gas using Q-TOF MS.
- Perform native mass spectrometry techniques to study proteins.
- Utilize collision induced unfolding to monitor protein stability.
- Investigate the mechanism of salt cluster formation in ESI.
- Use LC-MS for peptide separation.
- Explore HDX-silent dynamics using various proteomic techniques.

Undergraduate Research Student

Mississauga, ON

University of Toronto, Prosser Lab

September 2020 – May 2021

- Investigated the role of allosteric networks in enzymatic catalysis.
- Performed protein and plasmid purifications.
- Expressed recombinant proteins in E.coli cells.
- Examined F-NMR data to analyze conformational changes of proteins.

Teaching Experience

Graduate Teaching Assistant – Chemistry

London, ON

Western University

September 2021 – April 2023

- Supervise lab activities for undergraduate students.
- Hold office hours to further explain course material to over 80 undergraduate students.
- Proctor tests and exams to maintain Western University's academic conduct policies.

Education

Western University

London, ON

Master of Sciences in Physical/Analytical Chemistry.

September 2021 – September 2023

University of Toronto

Mississauga, Ontario

Seminars and Conferences

71st ASMS conference on Mass Spectrometry and Allied Topics.

Volunteering

Event Coordinator

Mississauga, ON

Young Liberals

September 2019 – April 2020

- Coordinated events on campus for over 30 students
- Built networks with local members of the parliament to further advocate for local issues

Medical First Responder

Mississauga, ON

University of Toronto

September 2017 – March 2019

- Responded to medical emergencies on UTM campus
- Assessed patient conditions through physical and verbal exams
- Provided first aid if needed
- Attended bi-weekly training sessions to ensure readiness when responding to a call.

Certifications

- First aid (CPR/C and AED training)
- Workplace Hazardous Material Information System training (WHIMS)
- Supporting Disclosure of Gender-based Violence (Western employee training)
- Building Inclusivity through Anti-Racism (Western employee training)

Selected Publications

1. Mechanism of Magic Number NaCl Cluster Formation from Electrosprayed Water Nanodroplets" L. Konermann and Y. Haidar *Anal. Chem.* 94, 16491-16501 (2022).

2. Effects of Hydrogen/Deuterium Exchange on Protein Stability in Solution and in the Gas Phase" Y. Haidar and L. Konermann *J. Am. Soc. Mass Spectrom.* 34, 1447-1458 (2023).

3. On the Chemistry of Aqueous Ammonium Acetate Droplets During Native Electrospray Ionization Mass Spectrometry" L. Konermann, Z. Liu, Y. Haidar, M. J. Willians, and N. A. Bainbridge *Anal. Chem.* 95, 37,13957-13966 (2023).

