UCLA
UCLA Electronic Theses and Dissertations

Title
Top-Down Mass Spectrometry Characterization of Protein-Ligand Complexes Important to Neurodegenerative Diseases

Permalink
https://escholarship.org/uc/item/1fh0d4kx

Author
Wongkongkathep, Piriya

Publication Date
2015

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Top-Down Mass Spectrometry Characterization of Protein-Ligand Complexes Important to Neurodegenerative Diseases

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biochemistry and Molecular Biology

by

Piriya Wongkongkathep

2015

ABSTRACT OF THE DISSERTATION

Top-Down Mass Spectrometry Characterization of Proteins and Protein-Ligand Complexes Important to Neurodegenerative Diseases

By

Piriya Wongkongakthep

Doctor of Philosophy in Biochemistry and Molecular Biology

University of California, Los Angeles, 2015

Professor Joseph A. Loo, Chair

Mass spectrometry (MS) has made significant contributions to protein and proteomics analysis during the past decades from its advantages of speed, sensitivity, specificity, and low sample consumption. While the proteomics field grows rapidly to identify thousands of proteins in a single analysis, “native” mass spectrometry, exploiting the unique features of electrospray ionization (ESI) for delivering large macromolecules to the mass spectrometer, has provided many potential exciting capabilities and applications to structural biology and biochemistry. It can analyze proteins in their native states, i.e., structures present in their native configurations from physiological pH solutions, with minimal sample preparation.

In this thesis, I describe the application of native ESI combined with top-down MS using electron capture dissociation (ECD) and ion mobility (IM) to characterize the molecular features of protein-ligand complexes. Binding and structural information can be comprehensively obtained

from this experimental platform. Native ESI-MS alone provides molecular mass, stoichiometry, and binding affinity, all from a single analysis. We demonstrate that top-down MS, the fragmentation of intact proteins and protein complexes using MS, offers a powerful capability to elucidate the location of ligand binding on a protein's structure and for probing the surface topology of proteins. Ion mobility mass spectrometry, a recently developed technique that yields information on the structural conformation of molecules, was used to reveal structural changes of proteins upon ligand binding.

My thesis focuses on several proteins, including α-synuclein (AS), which is a small protein related to Parkinson's disease. AS is natively unfolded at physiological pH, which makes it difficult to study by standard methods such as X-ray crystallography or NMR. Using our mass spectrometry techniques, transition metal binding (copper, cobalt, and manganese) to AS that is associated with accelerating fibril formation was monitored. The binding of a small molecule amyloid inhibitor called molecular tweezer (MT or CLR01) on two model proteins important in neurodegenerative diseases, AS and superoxide dismutase (SOD1), was studied. Tandem mass spectrometry (MS/MS) techniques such as collisionally activated dissociation (CAD) along with ECD were used to characterize the sites of binding of small molecule ligands to proteins. Ion mobility mass spectrometry was implemented to reveal the conformational changes of AS upon metal binding. It was demonstrated that copper can induce the AS protein to collapse into a more compact state, which may provide a hint of the mechanisms behind amyloid fibrillation.

Additionally, two new methods to extend the application of top-down MS for protein structure characterization were developed. First, the same molecular tweezer molecule, which has specificity to bind lysine residues, was used to probe surface residues of proteins. The lysines found to bind to the molecular tweezers identified by top-down MS correlate well with solvent

accessibility values, suggesting that the MT compound can be applied as a molecular probe to pinpoint surface active lysine residues. Lastly, supplemental activation methods by ultraviolet and infrared laser irradiation prior to ECD were applied to assist disulfide bond cleavage of complex multiple intermolecular and intramolecular disulfide bond-containing proteins. Backbone bond cleavage from top-down MS was significantly increased when the disulfide bonds were cleaved, allowing more sequence information to be obtained. The new methods described in this thesis extend the applicability of mass spectrometry to provide a more complete picture of a protein's structure.

This dissertation of Piriya Wongkongkathep is approved.

Jorge Torres

Gal Bitan

Joseph A. Loo, Committee Chair

University of California, Los Angeles

2015

I would like to dedicate this thesis to

My grandmother, Seew Sae-kwan

My father, Suriya Wongkongkathep, and my mother, Sunee Wongkongkathep

For all their lifetime support.

I also would like to dedicate my work to the scholarship from Development and Promotion of Science and Technology (DPST), Royal Thai Government for giving me an opportunity to study abroad.

Table of Contents

Abstract of Dissertation ii

List of Tables viii

List of Figures ix

Acknowledgments xi

Vita xiii

Publications xiv

Chapter 1: Characterization of Protein Non-covalently Bound Complexes by Mass Spectrometry 1

References 17

Chapter 2: Top-Down Mass Spectrometry Reveals the Sites of Metal Binding Important in Fibrillation of α-Synuclein 24

References 56

Chapter 3: Structural Elucidation of Molecular Tweezer Binding to Neurodegenerative Disease Proteins 61

References 89

Chapter 4: Mapping Protein Surface Residues by Monitoring Molecular Tweezer Binding 92

References 110

Chapter 5: Enhancing Protein Disulfide Bond Cleavage by UV Excitation and Electron Capture Dissociation for Top-Down Mass Spectrometry 114

References 145

Chapter 6: Conclusion and Perspective 149

References 154

List of Tables

Table 2-1	Mass accuracies of AS-metal complexes	36
Table 2-2	AS-metal binding affinity results from MS and previous reports	55
Table 4-1	Lysine solvent accessibility of ubiquitin and myoglobin	106

List of Figures

Figure 1-1	Techniques for protein structural characterization	3
Figure 1-2	Ion desolvation under native ESI conditions	7
Figure 1-3	Product ions from electron capture dissociation (ECD)	11
Figure 1-4	Comparison of ECD versus traditional MS/MS	12
Figure 1-5	Separation of oligomers by ion mobility	16
Figure 2-1	Sequence of α-synuclein	28
Figure 2-2	MS spectra of α-synuclein and metal complexes	37
Figure 2-3	Fragmentation plot of α-synuclein	38
Figure 2-4	Fragmentation plot of α-synuclein/copper complexes	41
Figure 2-5	Fragmentation plot of α-synuclein/cobalt complexes	47
Figure 2-6	Fragmentation plot of α-synuclein/cobalt complexes	48
Figure 2-7	MS^3 of α-synuclein/copper and manganese complexes	49
Figure 2-8	Ion mobility of α-synuclein/copper complexes	52
Figure 2-9	Ion mobility of α-synuclein/cobalt and managese complexes	53
Figure 3-1	Structure of SOD1	67
Figure 3-2	Molecular structure of MT	68
Figure 3-3	Native ESI-MS spectra of α-synuclein-MT complexes	72
Figure 3-4	Determination of dissociation constants of α-synuclein-MT complexes	74
Figure 3-5	Native ESI mass spectra of noCys-SOD1 and MT complexes	77
Figure 3-6	Determination of dissociation constants of SOD1-MT complexes	78
Figure 3-7	ECD of α-synuclein-MT complexes	81
Figure 3-8	Fragmentation map of α-synuclein-MT complexes by top-down MS	82
Figure 3-9	Representative ECD fragments of α-synuclein-MT complexes	83
Figure 3-10	Native ESI mass spectra of MT-SOD1	86

Figure 3-11 ECD fragmentation maps of 1:1 MT-SOD1 complex 87

Figure 4-1 MS spectrum and ECD-MS of native MT-bound ubiquitin complex 102

Figure 4-2 MS spectrum and ECD-MS of native MT-bound apo-myoglobin complex 103

Figure 4-3 Solvent surface accessibility calculation of ubiquitin 107

Figure 4-4 Solvent surface accessibility calculation of myoglobin 108

Figure 5-1 Instrument diagram of the modified 15-Tesla FT-ICR MS 121

Figure 5-2 Amino acid sequence of insulin and ribonuclease A 124

Figure 5-3 ESI mass spectra of insulin and ribonuclease A 125

Figure 5-4 UVPD (266 nm) mass spectra of insulin and ribonuclease A 126

Figure 5-5 Cleavage pattern of insulin by ECD and UV-ECD 128

Figure 5-6 Fragmentation mass spectra of insulin using ECD and UV-ECD 129

Figure 5-7 Fragmentation plot of ribonuclease A using different ion activations 132

Figure 5-8 ECD mass spectra of ribonuclease A with and without UV/IR-activation 135

Figure 5-9 Representative UV/IR-activated ECD fragments from ribonuclease A 136

Figure 5-10 %SS abundance for ECD of ribonuclease A with different activations 139

Figure 5-11 Fragmentation plot of lysozyme and β-lactoglobulin 142

Figure 5-12 Comparing %SS abundance of lysozyme and β-lactoglobulin 143

Acknowledgments

I would foremost like to thank my mentor, Professor Joseph Loo, and Dr. Rachel Loo for their warm welcome and the opportunity to join the lab, for their great support during my graduate career, and most importantly for giving me guidance, motivation, research inspiration, encouragement, and freedom to follow my interests. I couldn't have done it without their support. I would like to thank my committee members, Professor Steve Clarke, Professor Jorge Torres, Professor Gal Bitan, and Professor Feng Guo, for their guidance, suggestions, and comments given to me during my meetings with them.

Additionally, I would like to thank my family for their lifetime support. Although it is a trend in Thailand that most parents want their children to study in medical school or engineering school, they continuously encourage me to study what I like. They directly and indirectly taught me the importance of learning English, and supported me to go aboard. I would like to specially thank the scholarship from the Development and Promotion of Science and Technology Talents Project (DPST) for financial support for my graduate studies at UNC-Chapel Hill and UCLA. This program has given me long-time support since my undergraduate study.

Since I was originally trained as an analytical chemist, I would like to thank Nae Chantaravisoot for helping me understand molecular biology. She later on became a good collaborator at work and my closest friend in real life. Outside the lab, we have been traveling together from the southernmost point of South America to the north above the Arctic Circle.

Finally, I am grateful to have Jiang Zhang and Huilin Li, who are the FTMS experts, for showing and training me the state-of-the-art 15-Tesla FT-ICR MS mass spectrometer, or the so-called giant monster that we have in our lab. My skills couldn't have been improved and my work couldn't have been completed without their help and great discussions. I would like to thank all of Loo lab members for a great time during the past four years, especially Dyna

Shirasaki, Nalaka Rannulu, Raji Lakshmanan, Jonathan Erde, Carly Ferguson, Hong Nguyen, Sean Shen, and Reid O'Brien Johnson. I also want to thank Professor Gary Glish from UNC-Chapel Hill and the Glish lab members for teaching me instrumentation skills. Especially I want to thank Alessandra Ferzoco, Mark Ridgway, Atim Enyenihi, Alfred Zhong, and Sandra Spencer for providing suggestions to help me troubleshoot my many problems.

Chapter two of this dissertation is from a manuscript in preparation entitled "Top-down mass spectrometry reveals the sites of metal binding important in fibrillation of α-synuclein". I would like to acknowledge Dr. Sheng Yin who pioneered this work. I would like to thank Professor Hugh Kim, Jong Yoon Han, and Tae Su Choi for their contribution to the ion mobility study.

Some of the work presented in Chapter three of this dissertation was published in the Journal of Biological Chemistry, entitled "Molecular basis for preventing α-synuclein aggregation by a molecular tweezer", in collaboration with Professor Gal Bitan, and Professor Lisa Lapidus. Also, I would like to acknowledge Professor Joan Valentine, Madhuri Chattopadhyay, and Becky Chan for their contribution to the SOD1 work.

Chapter five of this dissertation is published in the International Journal of Mass Spectrometry: Piriya Wongkongkathep, Huilin Li, Xing Zhang, Rachel R. Ogorzalek Loo, Ryan R. Julian, Joseph A. Loo, "Enhancing protein disulfide bond cleavage by UV excitation and electron capture dissociation for top-down mass spectrometry," 2015, volume 390, pp. 137-145 (reprinted with permission from Elsevier). I would like to acknowledge Xing (Vic) Zhang, Ben Moore, and Professor Ryan Julian for their help with the set up of the UV laser system to our FT-ICR mass spectrometer.

Vita

March 2008	Bachelor of Science in Chemistry Mahidol University, Bangkok, Thailand
June 2011	Master of Science in Analytical Chemistry University of North Carolina, Chapel Hill
2010	Teaching Assistant Department of Chemistry University of North Carolina, Chapel Hill
2012-2015	Teaching Assistant Department of Chemistry and Biochemistry University of California, Los Angeles
2013	ASMS Sanibel Conference Student Travel Stipend American Society for Mass Spectrometry
2015	ASMS Asilomar Conference Student Travel Stipend American Society for Mass Spectrometry

Publications

1. **Wongkongkathep P**, Li H, Zhang X, Ogorzalek Loo RR, Julian RR, Loo JA, Enhancing Protein Disulfide Bond Cleavage by UV Excitation and Electron Capture Dissociation for Top-Down Mass Spectrometry, *International Journal of Mass Spectrometry* (2015), 390, 137-145
2. Chantaravisoot N, **Wongkongkathep P**, Loo JA, Mischel PS, Tamanoi F, Significance of Filamin A in mTORC2 Function in Glioblastoma, *Molecular Cancer* (2015), 14, 127 (14 pp.)
3. Zhang X, Li H, Moore B, **Wongkongkathep P**, Ogorzalek Loo RR, Loo JA, Julian RR, Radical-Directed Dissociation of Peptides and Proteins by Infrared Multiphoton Dissociation and Sustained Off-Resonance Irradiation Collision-Induced Dissociation with Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, *Rapid Communications in Mass Spectrometry* (2014), 28, 24, 2729–2734
4. Li H, **Wongkongkathep P**, Van Orden SL, Ogorzalek Loo RR, Loo JA, Revealing Ligand Binding Sites and Quantifying Subunit Variants of Noncovalent Protein Complexes in a Single Native Top-Down FTICR MS Experiment, *Journal of the American Society for Mass Spectrometry* (2014), 25, 2060-2068
5. Acharya S, Safaie BM, **Wongkongkathep P**, Ivanova MI, Attar A, Klärner F-G, Schrader T, Loo JA, Bitan G, Lapidus LJ, Molecular Basis for Preventing α-Synuclein Aggregation by a Molecular Tweezer, *Journal of Biological Chemistry* (2014), 289, 10727-10737
6. Dabir DV, Hasson SA, Setoguchi K, Johnson ME, **Wongkongkathep P**, Douglas CJ, Zimmerman J, Damoiseaux R, Teitell MA, Koehler CM, A Small Molecule Inhibitor of Redox-Regulated Protein Translocation into Mitochondria. *Developmental Cell* (2013), 25, 81–92

CHAPTER ONE

Characterization of Protein Non-covalently Bound Complexes by Mass Spectrometry

Characterization of protein interactions by conventional biophysical approaches

Most proteins almost never work alone. They bind to substrates, cofactors, small molecules, or other proteins. This binding step forms larger molecular complexes to perform their specific functions, which is very important in nearly every cellular process including metabolism. A protein may be able to bind different molecules at different sites to activate different tasks, creating a variety of multiple functions on a single molecule. Alternatively a protein can bind to several proteins to form large molecular machines for more advanced and complex functions[1]. A complete view of the structural biology of protein interactions can lead to an understanding of biological pathways and protein recognition[2]. It is also important for designing new drugs to target a specific interaction involved in human diseases[3-5]. In the past decades, many tools have been developed to characterize protein binding. For example, X-ray crystallography[6] and NMR[7] have been used as the "gold standard" to examine protein structures. They provide a comprehensive view of protein structures with high resolution. But relatively large amounts of sample and homogenously pure material are required[8, 9]. In X-ray crystallography, a large single crystal must be obtained, which can be difficult in some cases, particularly amyloid fibrils[10]. Also many biophysical methods such as analytical ultracentrifugation[11], isothermal titration calorimetry (ITC)[12], circular dichroism (CD)[13], surface plasmon resonance (SPR)[14], and several fluorescent spectroscopy based techniques[15-17] provide stoichiometry, kinetics, thermodynamic characteristics, and temporal resolution of the binding event. However, these approaches reveal some aspects of protein binding but none of them can provide everything in a single technique[18].

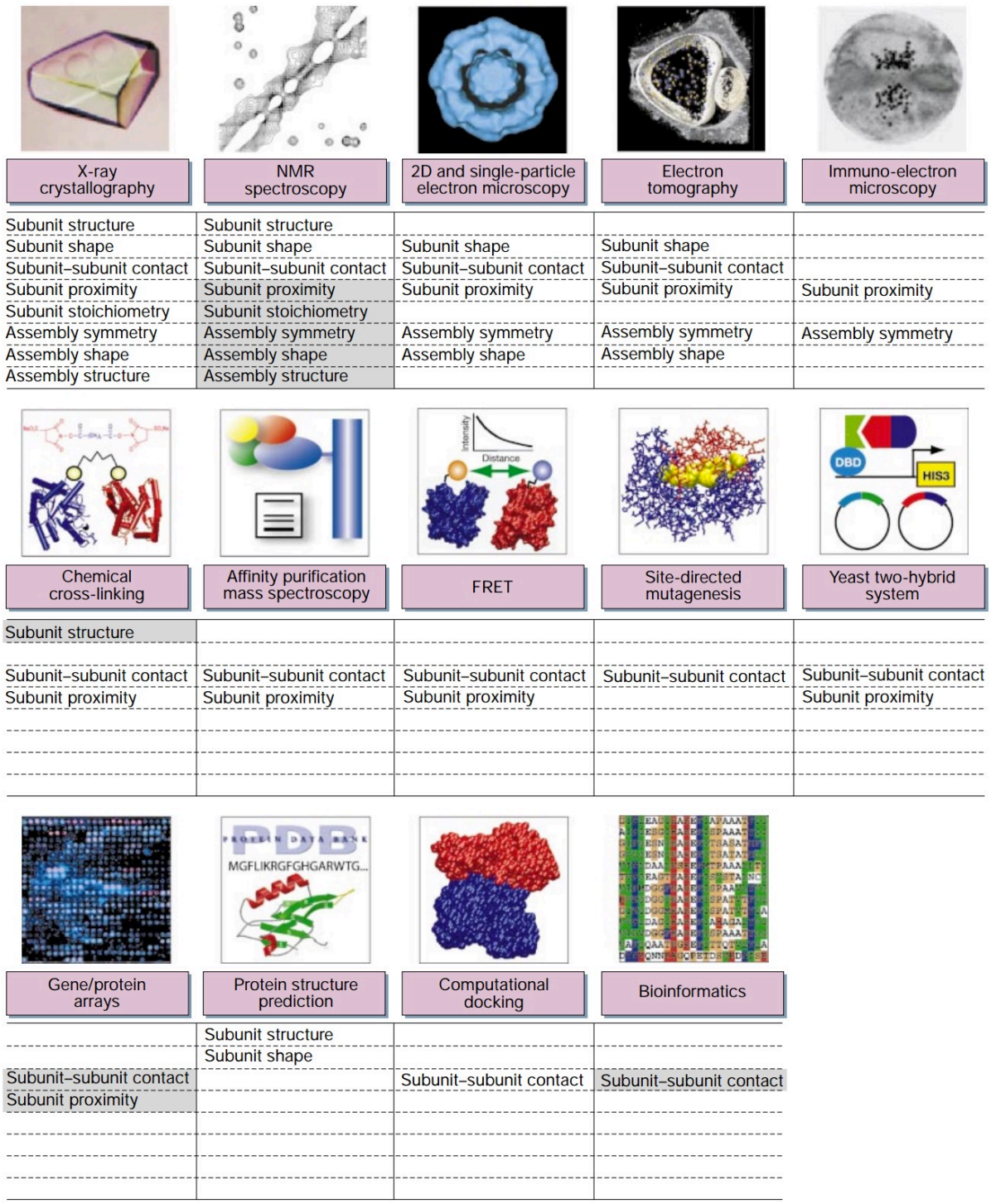

Figure 1-1. Experimental and computational approaches for characterization of protein complexes, and information that is provided by each techniques is listed below. (Reprinted from Sali et al (2003)[18] with permission from Nature Publishing Group.).

Studying proteins and native structures by mass spectrometry

Mass spectrometry (MS) was developed more than a century ago but it has been used to characterize proteins only in the past few recent decades. MS provides measurement simplicity, speed, specificity, selectivity and sensitivity, while using only small amounts of sample. It measures charged analytes in their gaseous state. In the early days, it has been used to analyze small organic molecules to reveal their molecular masses and structures. Analytes were usually vaporized by a thermal processes and they were ionized by bombarding with electrons. This technique is called electron ionization (EI). Nowadays, EI is used with GC-MS systems or with magnetic-electrostatic sector instrument, which are popular for the analysis of volatile small molecules. Before electrospray ionization (ESI)[19] and matrix-assisted laser desorption ionization (MALDI)[20] were developed, it was very difficult to find a soft desorption/ionization methods to introduce charges onto the protein molecules and also vaporize them into the gas phase without decomposing the proteins. Because of the high energy often used during the desorption and ionization steps, proteins were usually fragmented before entering the analyzer of the mass spectrometer. Thus, obtaining an intact mass of a protein was nearly impossible. Fast atom bombardment[21] and plasma desorption[22, 23] were among the “soft” ionization techniques that were used for proteins during that time. However the ionization efficiency was not very good, required relatively large amounts of sample, and sample preparation methods required for these methods were not easy to implement.

ESI was a groundbreaking advancement for introducing a wide range of biological molecules, from small metabolites, DNA, to large proteins into the mass spectrometer. The achievement was awarded the Nobel Prize in 2002, along with the development of MALDI. A decade after ESI was pioneered, nanoelectrospray (nESI) was developed[24] as another variation of ESI that has the benefit of lowering sample consumption by using a very low flow rate (around low nL

per minute)[25]. Generally, nESI-MS has a limit of detection at the femtomole level. Another advantage of using a low analyte flow rate is that it produces smaller droplets for better desolvation, yields high sensitivity and improves salt tolerance (Figure 1-2). Nowadays ESI has been widely used for studies of peptides and proteins. Depending on the solvent composition, protein structural states can be modulated between the denatured and “native” states. Normal ESI solvent conditions generally use compositions in which an organic solvent is added to aqueous solutions to achieve a ratio 50:50:1 of acetonitrile (or methanol), water, and formic acid, respectively. Sometimes 0.1% TFA (trifluoroacetic acid) or 1% acetic acid is used instead of formic acid. With this solvent mixture containing organic solvent and an acid pH (ca. 2-3), polypeptides are mostly unfolded or denatured. This condition yields maximum sensitivity for polypeptide detection and sequence analysis, but it is not suitable for studying protein folding and higher order structures that are present in physiological solution conditions.

When the analyte solution contains a neutral pH buffer and a gentle ESI spray condition is used, the “native” state of the protein can be achieved and its original solution structure is preserved. Ammonium acetate has become a popular buffer used for such purposes because of its high volatility and its pH can be maintained within the near-physiological range (pH 6.8 – 7.0). This technique is named native ESI, or native nESI if a low flow rated is used. Native ESI allows MS to observe native-like structures and also weakly non-covalent interactions. To avoid non-specific interactions (i.e., self oligomerization), which can be observed at higher analyte concentration, a typical operating sample concentration for native ESI is between 1-10 μM. When combining native ESI with MS, the same advantages mentioned previously, including speed, simplicity, sensitivity, specificity, selectivity, and low sample consumption can be realized. It provides speed because the signal appears within seconds and it does not take hours to acquire MS spectra. MS has simplicity because no chemical modification is needed for

detection. Sensitivity is very high, while sample consumption is low. It gives specificity and selectivity from an ability to measure multiple analytes in mixtures simultaneously; specific precursor ions can be isolated for tandem MS experiments. More selectivity can be achieved by coupling ESI-MS with ion mobility, which is a technique that can separate molecules based on different structural conformations.

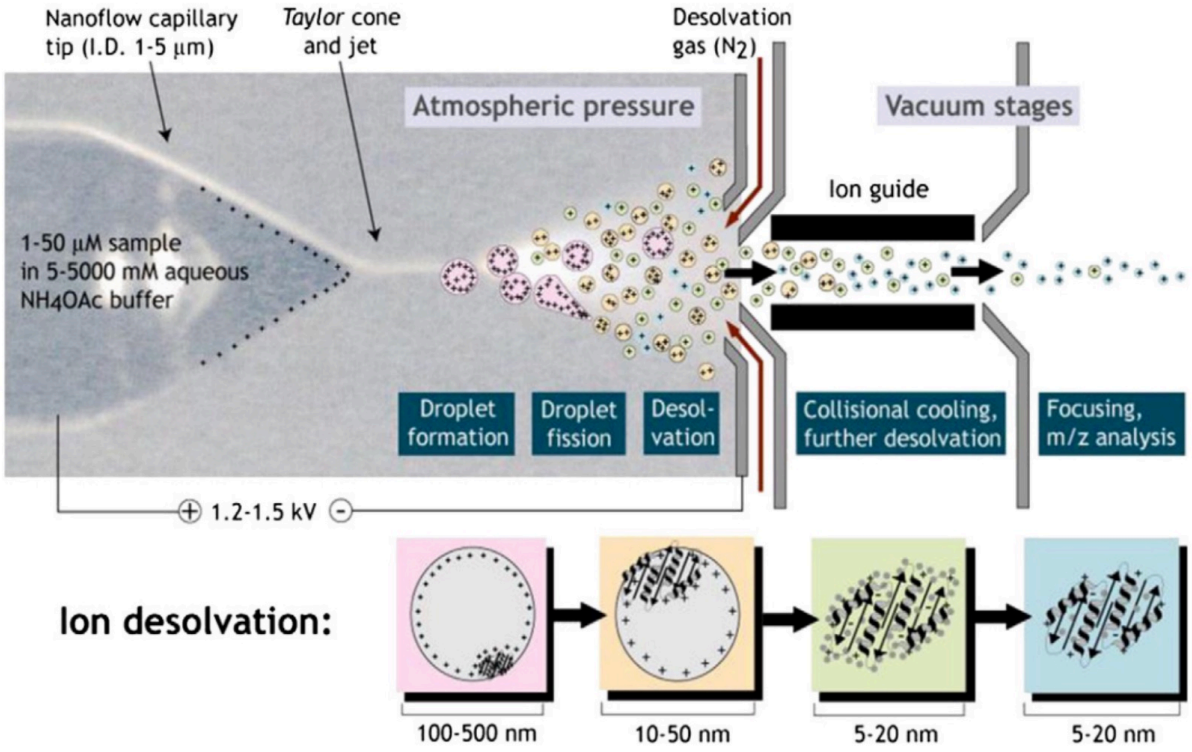

Figure 1-2. An ion desolvation process showing how protein molecules are delivered into the gas phase under native ESI conditions. (Reprinted with permission from Konijnenberg et. al. (2013)[26], Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics.)

Inside the mass spectrometer

Mass spectrometry is composed of two main parts: ionization and mass detection. ESI is basically an ionization method in which the ionization process occurs at atmospheric pressure outside the mass spectrometer analyzer. ESI can be coupled with various types of mass analyzers for detection. In the early days it was used with “EB” double focusing instruments that has an electric sector (E) in a combination with a magnetic sector (B), or with a single quadrupole mass analyzer. Nowadays, the linear ion trap (LIT)[27, 28], time-of-flight (TOF)[29], Orbitrap[30], and Fourier transform ion cyclotron resonance (FT-ICR)[31, 32] are among the popular mass analyzers for native ESI on many commercial instruments. Different mass analyzers have their own unique advantages. TOF is a beam type mass analyzer and it is known for its measurement speed, but it does not have tandem MS (MS/MS) capabilities on its own. Tandem MS is an advancement that includes additional steps before mass detection. The critical step is fragmentation, which can provide more structural and compositional information about the molecule. Some instruments may have an additional mass isolation section for selecting a precursor molecule to undergo fragmentation. Most ion trapping mass analyzers such as the ion trap and FT-ICR are capable of performing ion isolation and multi-stage tandem MS (MS^n). FT-ICR has the highest resolution, greater than one million but its detection process is slow. The Orbitrap, which is a trapping instrument, is currently a very popular instrument because it is fast and it has high resolution. Like TOF, however, the Orbitrap does not have MS/MS capability. Generally TOFs and Orbitraps need to be coupled with an additional linear ion trap or quadrupole for precursor selection and fragmentation.

Several fragmentation methods have been developed that has advanced the field of mass spectrometry. The most widely used conventional approach is collisionally activated dissociation (CAD), which was developed more than three decades ago. CAD is performed by introducing

precursor molecules into a chamber (a collision cell) containing a neutral bath gas and applying additional DC voltage to increase the kinetic energy of the molecules. Activated precursor ions collide with the neutral gas molecules (He, Ar or nitrogen). The kinetic energy of the precursor is then converted into an internal energy upon collisions, resulting in backbone bond dissociation. The more number of collisions, the more the internal energy is increased. Sometimes CAD is referred to as a slow-heating method. CAD is usually performed inside an ion trap or a quadrupole-based collision cell. But it can be done in the source region (often referring to in-source fragmentation), or inside the FT-ICR cell as well, by using a technique called Sustained Off-Resonance Irradiation CAD (SORI-CAD). Normal CAD is not preferred to be performed in the ICR cell because ultra low pressure is required for high sensitivity and resolution ICR detection. There is another variation of CAD using high energy called higher-energy collisional dissociation (HCD), which is mostly implemented in Orbitrap mass spectrometers. For proteins, CAD cleavage sites are dependant on sequence. As the molecule's internal energy from collisions is increased, the weakest bonds or interactions are dissociated first. Thus, most non-covalent interactions will fall apart during this step.

In late 1990s, McLafferty's group discovered a new fragmentation mechanism that originates from low energy electrons (0.2-1 eV). This method was coined electron capture dissociation (ECD). ECD has been heavily studied in the past decades[33-35]. The mechanism underlying the fragmentation is from a capture of electrons by a multiple charged molecule, $[M+nH]^{n+}$, at a carbonyl group of on the protein backbone. A charge-reduced odd-electron ion, $[M+nH]^{(n-1)+\bullet}$, is formed; then subsequent cleavage at N-Cα bonds occurs, resulting in production of c (even electron) and $z^{\bullet}$ (odd-electrion) fragments (Figure 1-3). Hydrogen radical migration may occur yielding $c^{\bullet}$ (odd electron) and z' (z+1, even electron) products[36, 37]. Electron capture cross sections are proportional to the square root of the charge state of a positively charged

precursor. Greater ECD efficiency will be achieved with higher charge. In the past decades, many variations have been developed, such as hot ECD[38], activated ion ECD (ai-ECD)[39-41], electron ionization dissociation (EID)[42], electron detachment dissociation (EDD)[43], negative ion ECD (ni-ECD)[44] and ETD[45]. Importantly, ETD expanded the limitation of ECD's availability to only FT-ICR instruments to become more applicable to many other instruments, such as QTOF, linear ion trap, and Orbitrap[46-48]. ECD (or ETD) has many advantages over conventional CAD. For example, ECD generates extensive fragmentation throughout the sequence because ECD cleavage is a fast process that occurs before the energy has time to randomize and be localized to a specific "weak" site. ECD generates less internal fragmentation than CAD[49], but secondary electron capture is possible. ECD facilitates disulfide bond cleavage[50-52]. Lastly and most importantly, ECD preserves post-translational modifications and weak non-covalent interactions while the protein backbone is cleaved[53, 54] (Figure 1-4). For this reason, ECD and ETD are groundbreaking discoveries that are extremely useful for structural biology, especially for the characterization of protein-ligand, protein-peptide interactions[55, 56] and protein complex assemblies[57-59].

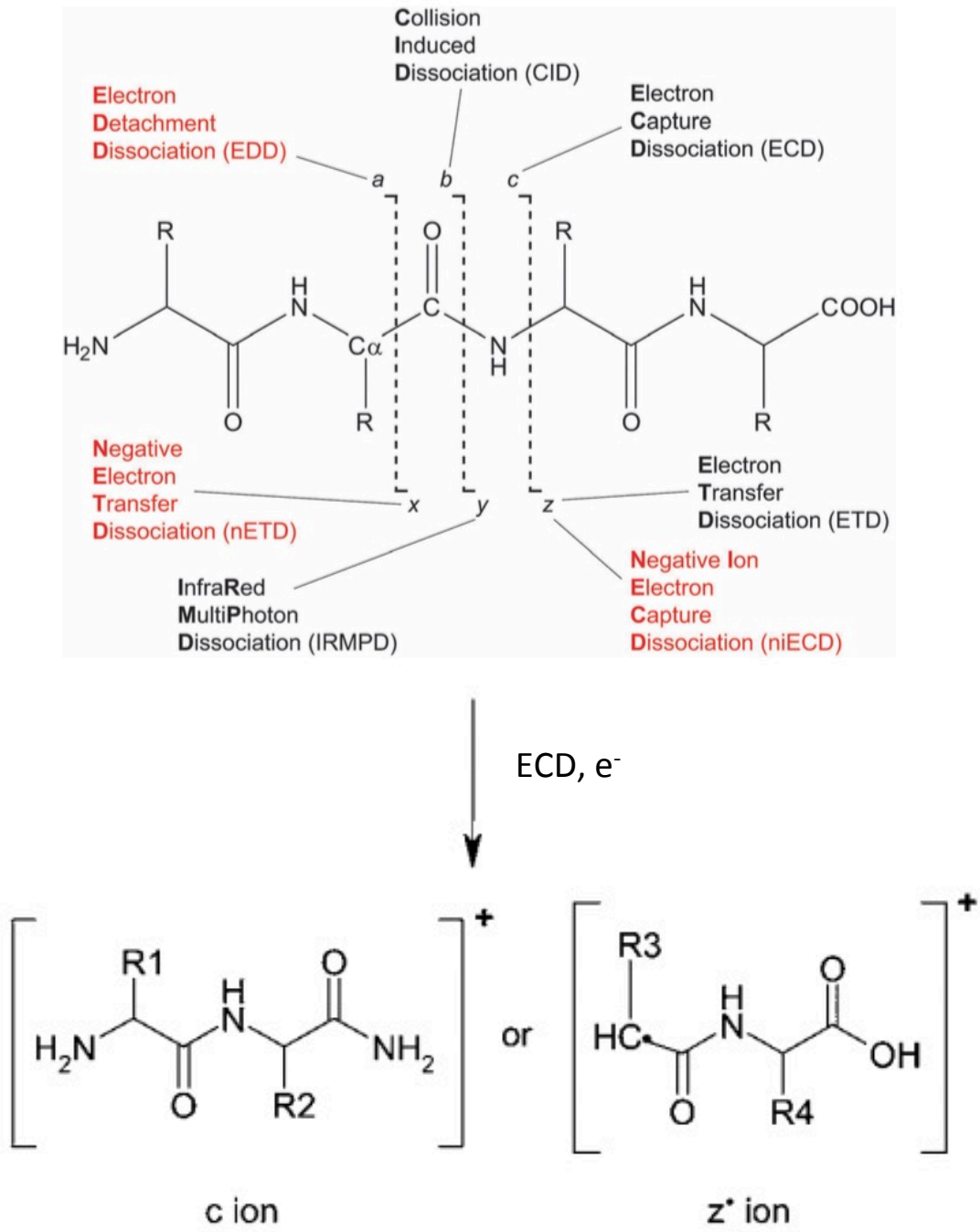

Figure 1-3. Electron capture dissociation (ECD) breaks the N-Cα bond and generates unique c- and z•-type ions (Zhurov et al (2013)[33], Jones and Cooper (2011)[60]). Reproduced from Ref. 33 and 60 with permission from The Royal Society of Chemistry.

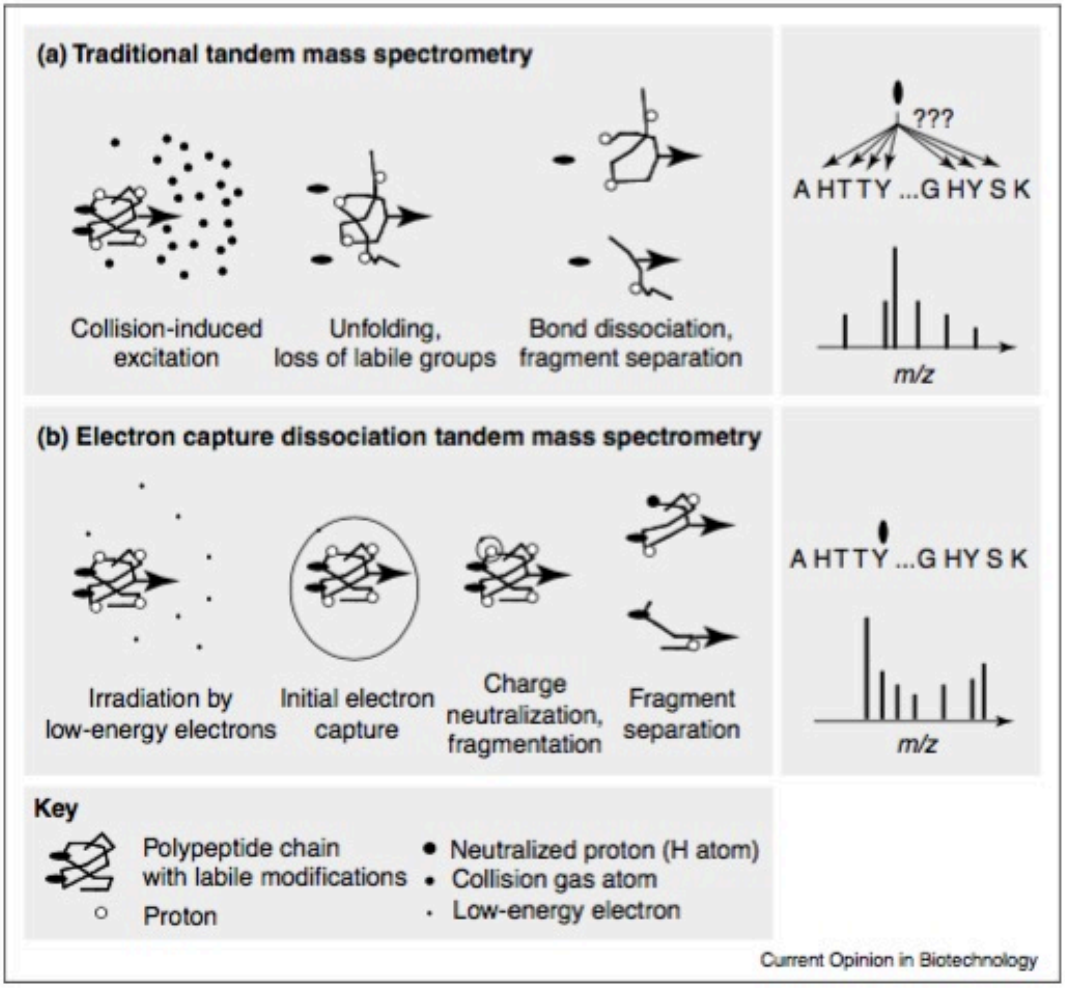

Figure 1-4. Compared to collision-based fragmentation, ECD does not disrupt labile post-translational modifications or weakly bound ligands. Thus, identifying ligand binding sites by top-down MS is feasible (from Zubarev (2004)[54]). Reproduced with permission from Current Opinion in Biotechnology (2004).

Native ESI of protein and protein-ligand complexes

Native ESI-MS studies of complexes have been reported since 1991 when the interaction between a cytoplasmic receptor, FKBP, and rapamycin was detected by ESI-MS. Since then there have been a number of studies that have been reported for systems such as metal-peptides[61, 62], nucleotide-proteins[63, 64], heme binding proteins[65], lipid-protein assemblies[66-68], and drug binding[69]. Recently native ESI has grown in popularity to study larger protein complexes and assemblies[70]. Native ESI-MS of the proteasome[71], monoclonal antibodies, membrane protein complexes[72, 73], intact ribosomes[74], and up to large megadalton-sized viruses[75] have been illustrated. Top-down MS approaches have been heavily used to reveal several structural aspects of large protein complexes. In this case, surface induced dissociation (SID)[76] has been developed as a superior technique over CAD to break down protein complexes to give more information about subunit assembly topology[77, 78]. Ion mobility mass spectrometry (IM) is a technique that is physically connected to a mass spectrometry system to provide information about the shape of molecules. IM-MS is a major tool to differentiate between different conformations and folding states[26, 79-81]. ECD has been used to identify protein surface residues and flexible regions, but fragmentation efficiency may be limited depending on sequence and subunit orientation[57, 58, 82, 83]. Recently, 193 nm ultraviolet photodissociation (UVPD) has proved that it can generate extensive fragmentation for native proteins[84, 85]. For protein-ligand interactions, ECD (or ETD) is still the most available tool for top-down MS to gain sequence information, protein structure and possible ligand binding regions[65, 86, 87].

Studies of neurodegenerative disease proteins by ESI-MS

In the past ten years, there have been several research groups interested in characterizing amyloid proteins by mass spectrometry[88-92]. Although most fibrils are insoluble, some intermediate oligomeric states are water-soluble[93] and available for studies by native ESI-MS.

Additionally, with IMS, conformational states and assembly information of protein oligomers can be easily obtained[94, 95] (Figure 1-5). Intrinsically disordered proteins (IDPs) is especially beneficial from IM-MS studies because the unstructured regions are highly flexible and not well defined. IDPs are difficult to study by traditional X-ray crystallography or NMR approaches[26, 96, 97]. There are many examples of using native ESI coupled to IM-MS to study amyloidogenic proteins, such as amyloid-β[95], α-synuclein[98], IAPP[99], and β-2-microglobulin[89, 100]. By performing screening of small molecules, several cofactors and potential inhibitors for amyloid fibrillation were identified[92]. Native ESI-MS has illustrated its capability to directly measure many substrate molecules binding to target amyloidogenic proteins. The interactions between metal ions and amyloidogenic proteins have been reported, such as α-synuclein-copper[101], amylin-copper[102], and IAPP-Zn[103]. Also, binding between small molecules and amyloid protein, or amyloid protein and other proteins have been investigated, including α-synucelin-dopamine[104]. IAPP-insulin[105], tau with amyloid-β[106], and amyloid-β and EGEG[105, 107]. Our group has utilized native ESI and top-down MS to determine both stoichiometry and binding location. We were able to characterize spermine binding to α-synuclein[55]. Molecular tweezer (MT), an inhibitor of aggregate assembly of amyloidogenic proteins, has been investigated for its binding sites on amyloid-β[108] and α-synuclein[109, 110] and the molecular mechanisms behind inhibition function of MT.

In this dissertation, heavy metals (Cu, Co, Mn) binding to α-synuclein are characterized by native ESI-MS, IM-MS and top-down MS/MS to reveal the major binding sites and conformational changes upon metal binding. CAD and ECD provided complementary information regarding binding site identification. This work is demonstrated in Chapter 2, and a manuscript is in preparation. Chapter 3 describes the use of ECD top-down MS to determine the important binding sites of molecular tweezer to amyloidogenic protein α-synuclein and

superoxide dismutase-1 (SOD1); the tweezer/α-synuclein studies was published in the *Journal of Biological Chemistry*[109]. Chapter 4 demonstrates an application of molecular tweezer to probe the surface residues of small proteins. ECD-MS/MS is, again, implemented to determine the binding residues and the results are validated using computational models. Lastly, irradiation from a 266 nm UV laser has been used to facilitate disulfide bond cleavage in conjuction with ECD-MS/MS. Because disulfide bonds have relatively high bond energies, it survives under most CAD and ECD conditions. It prevents chain separation, so that a mass analyzer is not able to detect fragments from the protein. We showed that applying UV raditation significantly improved disulfide bond cleavage of insulin (containing 3 disulfide bonds), and ribonuclease A, which contains 4 disulfide bonds. The method offers a viable means to better characterize proteins containing disulfide-bonds by mass spectrometry. Chapter 5 has been published in the *International Journal of Mass Spectrometry*[111].

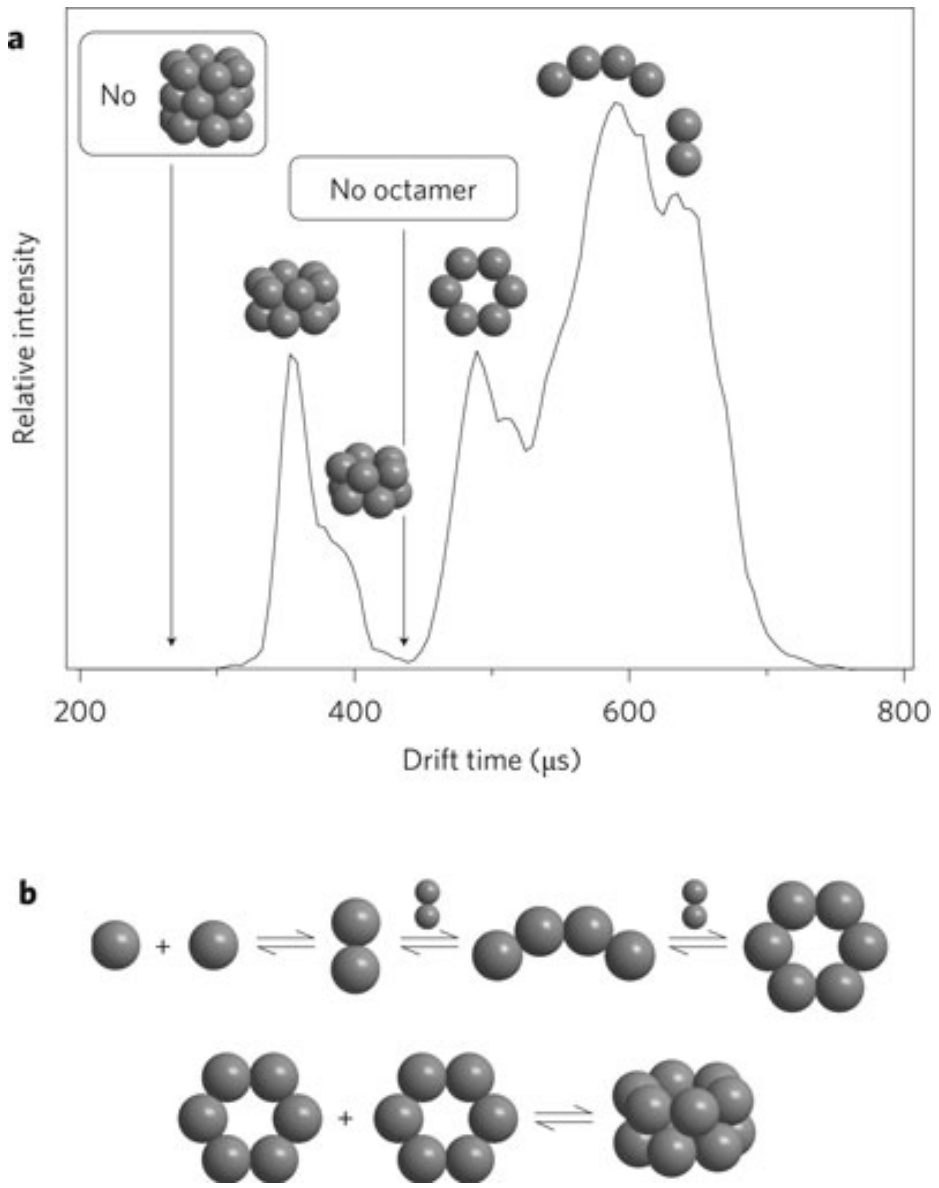

Figure 1-5. Ion mobility is capable of separating different oligomeric states of different molecular conformations of Aβ42 peptides (from Bernstein et al (2009)[95]). (Reproduced with permission from Nature Chemistry (2009).)

References

[1] Keskin, O., Gursoy, A., Ma, B., and Nussinov, R. (2008) Principles of Protein−Protein Interactions: What are the Preferred Ways For Proteins To Interact?, *Chemical Reviews 108*, 1225-1244.

[2] Heck, A. J. R. (2008) Native mass spectrometry: a bridge between interactomics and structural biology, *Nature Methods 5*, 927-933.

[3] Ideker, T., and Sharan, R. (2008) Protein networks in disease, *Genome Research 18*, 644-652.

[4] Barabasi, A.-L., Gulbahce, N., and Loscalzo, J. (2011) Network medicine: a network-based approach to human disease, *Nature Reviews Genetics 12*, 56-68.

[5] Hopkins, A. L. (2008) Network pharmacology: the next paradigm in drug discovery, *Nature Chemical Biology 4*, 682-690.

[6] Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. (1965) Structure of Hen Egg-White Lysozyme: A Three-dimensional Fourier Synthesis at 2 Å Resolution, *Nature 206*, 757-761.

[7] Wuthrich, K. (2001) The way to NMR structures of proteins, *Nature Structural and Molecular Biology 8*, 923-925.

[8] Miao, J., Ishikawa, T., Shen, Q., and Earnest, T. (2008) Extending X-Ray Crystallography to Allow the Imaging of Noncrystalline Materials, Cells, and Single Protein Complexes, *Annual Review of Physical Chemistry 59*, 387-410.

[9] Kleckner, I. R., and Foster, M. P. (2011) An introduction to NMR-based approaches for measuring protein dynamics, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1814*, 942-968.

[10] Eisenberg, D., and Jucker, M. (2012) The Amyloid State of Proteins in Human Diseases, *Cell 148*, 1188-1203.

[11] Laue, T. M. (2001) Analytical Ultracentrifugation, In *Current Protocols in Protein Science*, John Wiley & Sons, Inc.

[12] Leavitt, S., and Freire, E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry, *Current Opinion in Structural Biology 11*, 560-566.

[13] Johnson, W. C. (1990) Protein secondary structure and circular dichroism: A practical guide, *Proteins: Structure, Function, and Bioinformatics 7*, 205-214.

[14] Karlsson, R., and Fält, A. (1997) Experimental design for kinetic analysis of protein-protein interactions with surface plasmon resonance biosensors, *Journal of Immunological Methods 200*, 121-133.

[15] Piston, D. W., and Kremers, G.-J. (2007) Fluorescent protein FRET: the good, the bad and the ugly, *Trends in Biochemical Sciences 32*, 407-414.

[16] Giepmans, B. N. G., Adams, S. R., Ellisman, M. H., and Tsien, R. Y. (2006) The Fluorescent Toolbox for Assessing Protein Location and Function, *Science 312*, 217-224.

[17] Hu, C.-D., and Kerppola, T. K. (2003) Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis, *Nature Biotechnology 21*, 539-545.

[18] Sali, A., Glaeser, R., Earnest, T., and Baumeister, W. (2003) From words to literature in structural proteomics, *Nature 422*, 216-225.

[19] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules, *Science 246*, 64-71.

[20] Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons, *Analytical Chemistry 60*, 2299-2301.
[21] Morris, H. R., Panico, M., Barber, M., Bordoli, R. S., Sedgwick, R. D., and Tyler, A. (1981) Fast atom bombardment: A new mass spectrometric method for peptide sequence analysis, *Biochemical and Biophysical Research Communications 101*, 623-631.
[22] Loo, J. A., Williams, E. R., Amster, I. J., Furlong, J. J. P., Wang, B. H., McLafferty, F. W., Chait, B. J., and Field, F. H. (1987) Californium-252 plasma desorption with Fourier transform mass spectrometry, *Analytical Chemistry 59*, 1880-1882.
[23] Sundqvist, B., Kamensky, I., Håkansson, P., Kjellberg, J., Salehpour, M., Widdiyasekera, S., Fohlman, J., Peterson, P. A., and Roepstorff, P. (1984) Californium-252 plasma desorption time of flight mass spectroscopy of proteins, *Biological Mass Spectrometry 11*, 242-257.
[24] Wilm, M., and Mann, M. (1996) Analytical Properties of the Nanoelectrospray Ion Source, *Analytical Chemistry 68*, 1-8.
[25] Juraschek, R., Dülcks, T., and Karas, M. (1999) Nanoelectrospray—more than just a minimized-flow electrospray ionization source, *Journal of the American Society for Mass Spectrometry 10*, 300-308.
[26] Konijnenberg, A., Butterer, A., and Sobott, F. (2013) Native ion mobility-mass spectrometry and related methods in structural biology, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1834*, 1239-1256.
[27] March, R. E. (2009) Quadrupole ion traps, *Mass Spectrometry Reviews 28*, 961-989.
[28] Hager, J. W. (2002) A new linear ion trap mass spectrometer, *Rapid Communications in Mass Spectrometry 16*, 512-526.
[29] Guilhaus, M. (1995) Special feature: Tutorial. Principles and instrumentation in time-of-flight mass spectrometry. Physical and instrumental concepts, *Journal of Mass Spectrometry 30*, 1519-1532.
[30] Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M., and Graham Cooks, R. (2005) The Orbitrap: a new mass spectrometer, *Journal of Mass Spectrometry 40*, 430-443.
[31] Marshall, A. G., Hendrickson, C. L., and Jackson, G. S. (1998) Fourier transform ion cyclotron resonance mass spectrometry: A primer, *Mass Spectrometry Reviews 17*, 1-35.
[32] Marshall, A. G., and Hendrickson, C. L. (2002) Fourier transform ion cyclotron resonance detection: principles and experimental configurations, *International Journal of Mass Spectrometry 215*, 59-75.
[33] Zhurov, K. O., Fornelli, L., Wodrich, M. D., Laskay, U. A., and Tsybin, Y. O. (2013) Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis, *Chemical Society Reviews 42*, 5014-5030.
[34] Cooper, H. J., Håkansson, K., and Marshall, A. G. (2005) The role of electron capture dissociation in biomolecular analysis, *Mass Spectrometry Reviews 24*, 201-222.
[35] Qi, Y., and Volmer, D. A. (2015) Electron-based fragmentation methods in mass spectrometry: An overview, *Mass Spectrometry Reviews*, in press.
[36] O'Connor, P. B., Lin, C., Cournoyer, J. J., Pittman, J. L., Belyayev, M., and Budnik, B. A. (2006) Long-Lived Electron Capture Dissociation Product Ions Experience Radical Migration via Hydrogen Abstraction, *Journal of the American Society for Mass Spectrometry 17*, 576-585.
[37] Savitski, M. M., Kjeldsen, F., Nielsen, M. L., and Zubarev, R. A. (2007) Hydrogen Rearrangement to and from Radical z Fragments in Electron Capture Dissociation of Peptides, *Journal of the American Society for Mass Spectrometry 18*, 113-120.

[38] Kjeldsen, F., Haselmann, K. F., Sørensen, E. S., and Zubarev, R. A. (2003) Distinguishing of Ile/Leu Amino Acid Residues in the PP3 Protein by (Hot) Electron Capture Dissociation in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, *Analytical Chemistry 75*, 1267-1274.
[39] Horn, D. M., Ge, Y., and McLafferty, F. W. (2000) Activated Ion Electron Capture Dissociation for Mass Spectral Sequencing of Larger (42 kDa) Proteins, *Analytical Chemistry 72*, 4778-4784.
[40] Mikhailov, V. A., and Cooper, H. J. (2009) Activated Ion Electron Capture Dissociation (AI ECD) of Proteins: Synchronization of Infrared and Electron Irradiation with Ion Magnetron Motion, *Journal of the American Society for Mass Spectrometry 20*, 763-771.
[41] Tsybin, Y. O., He, H., Emmett, M. R., Hendrickson, C. L., and Marshall, A. G. (2007) Ion Activation in Electron Capture Dissociation To Distinguish between N-Terminal and C-Terminal Product Ions, *Anal. Chem. 79*, 7596-7602.
[42] Fung, Y. M. E., Adams, C. M., and Zubarev, R. A. (2009) Electron Ionization Dissociation of Singly and Multiply Charged Peptides, *Journal of the American Chemical Society 131*, 9977-9985.
[43] Budnik, B. A., Haselmann, K. F., and Zubarev, R. A. (2001) Electron detachment dissociation of peptide di-anions: an electron–hole recombination phenomenon, *Chemical Physics Letters 342*, 299-302.
[44] Yoo, H. J., Wang, N., Zhuang, S., Song, H., and Håkansson, K. (2011) Negative-Ion Electron Capture Dissociation: Radical-Driven Fragmentation of Charge-Increased Gaseous Peptide Anions, *Journal of the American Chemical Society 133*, 16790-16793.
[45] Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry, *Proceeding of the National Academy of Science USA 101*, 9528-9533.
[46] McAlister, G. C., Phanstiel, D., Good, D. M., Berggren, W. T., and Coon, J. J. (2007) Implementation of Electron-Transfer Dissociation on a Hybrid Linear Ion Trap−Orbitrap Mass Spectrometer, *Analytical Chemistry 79*, 3525-3534.
[47] McAlister, G. C., Berggren, W. T., Griep-Raming, J., Horning, S., Makarov, A., Phanstiel, D., Stafford, G., Swaney, D. L., Syka, J. E. P., Zabrouskov, V., and Coon, J. J. (2008) A Proteomics Grade Electron Transfer Dissociation-Enabled Hybrid Linear Ion Trap-Orbitrap Mass Spectrometer, *Journal of Proteome Research 7*, 3127-3136.
[48] Hartmer, R. G., Kaplan, D. A., Stoermer, C., Lubeck, M., and Park, M. A. (2009) Data-dependent electron transfer dissociation of large peptides and medium size proteins in a QTOF instrument on a liquid chromatography timescale, *Rapid Communications in Mass Spectrometry 23*, 2273-2282.
[49] Horn, D. M., Zubarev, R. A., and McLafferty, F. W. (2000) Automated de novo sequencing of proteins by tandem high-resolution mass spectrometry, *Proceedings of the National Academy of Sciences USA 97*, 10313-10317.
[50] Zubarev, R. A., Kruger, N. A., Fridriksson, E. K., Lewis, M. A., Horn, D. M., Carpenter, B. K., and McLafferty, F. W. (1999) Electron Capture Dissociation of Gaseous Multiply-Charged Proteins Is Favored at Disulfide Bonds and Other Sites of High Hydrogen Atom Affinity, *Journal of American Chemical Society 121*, 2857-2862.
[51] Ganisl, B., and Breuker, K. (2012) Does Electron Capture Dissociation Cleave Protein Disulfide Bonds?, *ChemistryOpen 1*, 260-268.
[52] Zhang, J., Loo, R. R. O., and Loo, J. A. (2015) Increasing fragmentation of disulfide-bonded proteins for top–down mass spectrometry by supercharging, *International Journal of Mass Spectrometry 377*, 546-556.

[53] Zubarev, R. A., Horn, D. M., Fridriksson, E. K., Kelleher, N. L., Kruger, N. A., Lewis, M. A., Carpenter, B. K., and McLafferty, F. W. (2000) Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations, *Analytical Chemistry 72*, 563-573.
[54] Zubarev, R. A. (2004) Electron-capture dissociation tandem mass spectrometry, *Current Opinion in Biotechnology 15*, 12-16.
[55] Xie, Y., Zhang, J., Yin, S., and Loo, J. A. (2006) Top-Down ESI-ECD-FT-ICR Mass Spectrometry Localizes Noncovalent Protein-Ligand Binding Sites, *Journal of the American Chemical Society 128*, 14432-14433.
[56] Clarke, D., Murray, E., Hupp, T., Mackay, C. L., and Langridge-Smith, P. R. (2011) Mapping a Noncovalent Protein–Peptide Interface by Top-Down FTICR Mass Spectrometry Using Electron Capture Dissociation, *Journal of the American Society for Mass Spectrometry 22*, 1432-1440.
[57] Zhang, H., Cui, W., Wen, J., Blankenship, R. E., and Gross, M. L. (2011) Native Electrospray and Electron-Capture Dissociation FTICR Mass Spectrometry for Top-Down Studies of Protein Assemblies, *Analytical Chemistry 83*, 5598-5606.
[58] Zhang, H., Cui, W., Wen, J., Blankenship, R. E., and Gross, M. L. (2010) Native Electrospray and Electron-Capture Dissociation in FTICR Mass Spectrometry Provide Top-Down Sequencing of a Protein Component in an Intact Protein Assembly, *Journal of the American Society for Mass Spectrometry 21*, 1966-1968.
[59] Geels, R. B. J., van der Vies, S. M., Heck, A. J. R., and Heeren, R. M. A. (2006) Electron Capture Dissociation as Structural Probe for Noncovalent Gas-Phase Protein Assemblies, *Analytical Chemistry 78*, 7191-7196.
[60] Jones, A. W., and Cooper, H. J. (2011) Dissociation techniques in mass spectrometry-based proteomics, *Analyst 136*, 3419-3429.
[61] Hu, P., and Gross, M. L. (1993) Gas-phase interactions of transition-metal ions and di- and tripeptides: a comparison with alkaline-earth-metal-ion interactions, *Journal of the American Chemical Society 115*, 8821-8828.
[62] Loo, J., Hu, P., and Smith, R. (1994) Interaction of angiotensin peptides and zinc metal ions probed by electrospray ionization mass spectrometry, *Journal of the American Society for Mass Spectrometry 5*, 959-965.
[63] Yin, S., Xie, Y., and Loo, J. A. (2008) Mass Spectrometry of Protein–Ligand Complexes: Enhanced Gas-Phase Stability of Ribonuclease–Nucleotide Complexes, *Journal of the American Society for Mass Spectrometry 19*, 1199-1208.
[64] Kapur, A., Beck, J. L., Brown, S. E., Dixon, N. E., and Sheil, M. M. (2002) Use of electrospray ionization mass spectrometry to study binding interactions between a replication terminator protein and DNA, *Protein Science 11*, 147-157.
[65] Enyenihi, A., Yang, H., Ytterberg, A. J., Lyutvinskiy, Y., and Zubarev, R. (2011) Heme Binding in Gas-Phase Holo-Myoglobin Cations: Distal Becomes Proximal?, *Journal of the American Society for Mass Spectrometry 22*, 1763-1770.
[66] Marty, M. T., Zhang, H., Cui, W., Blankenship, R. E., Gross, M. L., and Sligar, S. G. (2012) Native Mass Spectrometry Characterization of Intact Nanodisc Lipoprotein Complexes, *Analytical Chemistry 84*, 8957-8960.
[67] Laganowsky, A., Reading, E., Allison, T. M., Ulmschneider, M. B., Degiacomi, M. T., Baldwin, A. J., and Robinson, C. V. (2014) Membrane proteins bind lipids selectively to modulate their structure and function, *Nature 510*, 172-175.
[68] Zhou, M., Morgner, N., Barrera, N. P., Politis, A., Isaacson, S. C., Matak-Vinković, D., Murata, T., Bernal, R. A., Stock, D., and Robinson, C. V. (2011) Mass Spectrometry of

Intact V-Type ATPases Reveals Bound Lipids and the Effects of Nucleotide Binding, *Science 334*, 380-385.
[69] Ilag, L. L., Ubarretxena-Belandia, I., Tate, C. G., and Robinson, C. V. (2004) Drug Binding Revealed by Tandem Mass Spectrometry of a Protein−Micelle Complex, *Journal of the American Chemical Society 126*, 14362-14363.
[70] Sharon, M., and Robinson, C. V. (2007) The Role of Mass Spectrometry in Structure Elucidation of Dynamic Protein Complexes, *Annual Review of Biochemistry 76*, 167-193.
[71] Loo, J. A., Berhane, B., Kaddis, C. S., Wooding, K. M., Xie, Y., Kaufman, S. L., and Chernushevich, I. V. (2005) Electrospray Ionization Mass Spectrometry and Ion Mobility Analysis of the 20S Proteasome Complex, *Journal of the American Society for Mass Spectrometry 16*, 998-1008.
[72] Barrera, N. P., Di Bartolo, N., Booth, P. J., and Robinson, C. V. (2008) Micelles Protect Membrane Complexes from Solution to Vacuum, *Science 321*, 243-246.
[73] Wang, S. C., Politis, A., Di Bartolo, N., Bavro, V. N., Tucker, S. J., Booth, P. J., Barrera, N. P., and Robinson, C. V. (2010) Ion Mobility Mass Spectrometry of Two Tetrameric Membrane Protein Complexes Reveals Compact Structures and Differences in Stability and Packing, *Journal of the American Chemical Society 132*, 15468-15470.
[74] Rostom, A. A., Fucini, P., Benjamin, D. R., Juenemann, R., Nierhaus, K. H., Hartl, F. U., Dobson, C. M., and Robinson, C. V. (2000) Detection and selective dissociation of intact ribosomes in a mass spectrometer, *Proceedings of the National Academy of Sciences USA 97*, 5185-5190.
[75] Snijder, J., Rose, R. J., Veesler, D., Johnson, J. E., and Heck, A. J. R. (2013) Studying 18 MDa Virus Assemblies with Native Mass Spectrometry, *Angewandte Chemie International Edition 52*, 4020-4023.
[76] Wysocki, V., Ding, J.-M., Jones, J., Callahan, J., and King, F. (1992) Surface-induced dissociation in tandem quadrupole mass spectrometers: A comparison of three designs, *Journal of the American Society for Mass Spectrometry 3*, 27-32.
[77] Zhou, M., Dagan, S., and Wysocki, V. H. (2012) Protein Subunits Released by Surface Collisions of Noncovalent Complexes: Nativelike Compact Structures Revealed by Ion Mobility Mass Spectrometry, *Angewandte Chemie International Edition 51*, 4336-4339.
[78] Zhou, M., and Wysocki, V. H. (2014) Surface Induced Dissociation: Dissecting Noncovalent Protein Complexes in the Gas phase, *Accounts of Chemical Research 47*, 1010-1018.
[79] Freeke, J., Bush, M. F., Robinson, C. V., and Ruotolo, B. T. (2012) Gas-phase protein assemblies: Unfolding landscapes and preserving native-like structures using noncovalent adducts, *Chemical Physics Letters 524*, 1-9.
[80] Schmidt, C., and Robinson, C. V. (2014) Dynamic protein ligand interactions – insights from MS, *FEBS Journal 281*, 1950-1964.
[81] Ruotolo, B. T., Giles, K., Campuzano, I., Sandercock, A. M., Bateman, R. H., and Robinson, C. V. (2005) Evidence for Macromolecular Protein Rings in the Absence of Bulk Water, *Science 310*, 1658-1661.
[82] Li, H., Wongkongkathep, P., Van Orden, S., Ogorzalek Loo, R., and Loo, J. (2014) Revealing Ligand Binding Sites and Quantifying Subunit Variants of Noncovalent Protein Complexes in a Single Native Top-Down FTICR MS Experiment, *Journal of the American Society for Mass Spectrometry 25*, 2060-2068.
[83] Li, H., Wolff, J. J., Van Orden, S. L., and Loo, J. A. (2014) Native Top-Down Electrospray Ionization-Mass Spectrometry of 158 kDa Protein Complex by High-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, *Analytical Chemistry 86*, 317-320.

[84] O'Brien, J. P., Li, W., Zhang, Y., and Brodbelt, J. S. (2014) Characterization of Native Protein Complexes Using Ultraviolet Photodissociation Mass Spectrometry, *Journal of the American Chemical Society 136*, 12920-12928.
[85] Cammarata, M. B., and Brodbelt, J. S. (2015) Structural characterization of holo- and apo-myoglobin in the gas phase by ultraviolet photodissociation mass spectrometry, *Chemical Science 6*, 1324-1333.
[86] Yin, S., and Loo, J. A. (2010) Elucidating the Site of Protein-ATP Binding by Top-Down Mass Spectrometry, *Journal of the American Society for Mass Spectrometry 21*, 899-907.
[87] Yin, S., and Loo, J. A. (2011) Top-down mass spectrometry of supercharged native protein–ligand complexes, *International Journal of Mass Spectrometry 300*, 118-122.
[88] Mori, H., Takio, K., Ogawara, M., and Selkoe, D. J. (1992) Mass spectrometry of purified amyloid beta protein in Alzheimer's disease, *Journal of Biological Chemistry 267*, 17082-17086.
[89] Smith, D. P., Radford, S. E., and Ashcroft, A. E. (2010) Elongated oligomers in β2-microglobulin amyloid assembly revealed by ion mobility spectrometry-mass spectrometry, *Proceedings of the National Academy of Sciences USA 107*, 6794-6798.
[90] Smith, A. M., Jahn, T. R., Ashcroft, A. E., and Radford, S. E. (2006) Direct Observation of Oligomeric Species formed in the Early Stages of Amyloid Fibril Formation using Electrospray Ionisation Mass Spectrometry, *Journal of Molecular Biology 364*, 9-19.
[91] Nettleton, E. J., Tito, P., Sunde, M., Bouchard, M., Dobson, C. M., and Robinson, C. V. (2000) Characterization of the Oligomeric States of Insulin in Self-Assembly and Amyloid Fibril Formation by Mass Spectrometry, *Biophysical Journal 79*, 1053-1065.
[92] Young, L. M., Saunders, J. C., Mahood, R. A., Revill, C. H., Foster, R. J., Tu, L.-H., Raleigh, D. P., Radford, S. E., and Ashcroft, A. E. (2015) Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry–mass spectrometry, *Nature Chemistry 7*, 73-81.
[93] Haass, C., and Selkoe, D. J. (2007) Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide, *Nature Reviews Molecular Cell Biology 8*, 101-112.
[94] Woods, L. A., Radford, S. E., and Ashcroft, A. E. (2013) Advances in ion mobility spectrometry–mass spectrometry reveal key insights into amyloid assembly, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1834*, 1257-1268.
[95] Bernstein, S. L., Dupuis, N. F., Lazo, N. D., Wyttenbach, T., Condron, M. M., Bitan, G., Teplow, D. B., Shea, J.-E., Ruotolo, B. T., Robinson, C. V., and Bowers, M. T. (2009) Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease, *Nature Chemistry 1*, 326-331.
[96] Knapman, T. W., Valette, N. M., Warriner, S. L., and Ashcroft, A. E. (2013) Ion Mobility Spectrometry-Mass Spectrometry of Intrinsically Unfolded Proteins: Trying to Put Order into Disorder, *Current Analytical Chemistry 9*, 181-191.
[97] Beveridge, R., Chappuis, Q., Macphee, C., and Barran, P. (2013) Mass spectrometry methods for intrinsically disordered proteins, *Analyst 138*, 32-42.
[98] Bernstein, S. L., Liu, D., Wyttenbach, T., Bowers, M. T., Lee, J. C., Gray, H. B., and Winkler, J. R. (2004) α-Synuclein: Stable compact and extended monomeric structures and pH dependence of dimer formation, *Journal of the American Society for Mass Spectrometry 15*, 1435-1443.
[99] Dupuis, N. F., Wu, C., Shea, J.-E., and Bowers, M. T. (2009) Human Islet Amyloid Polypeptide Monomers Form Ordered β-hairpins: A Possible Direct Amyloidogenic Precursor, *Journal of the American Chemical Society 131*, 18283-18292.

[100] Hodkinson, J. P., Radford, S. E., and Ashcroft, A. E. (2012) The role of conformational flexibility in β2-microglobulin amyloid fibril formation at neutral pH, *Rapid Communications in Mass Spectrometry 26*, 1783-1792.
[101] Natalello, A., Benetti, F., Doglia, S. M., Legname, G., and Grandori, R. (2011) Compact conformations of α-synuclein induced by alcohols and copper, *Proteins: Structure, Function, and Bioinformatics 79*, 611-621.
[102] Li, H., Ha, E., Donaldson, R. P., Jeremic, A. M., and Vertes, A. (2015) Rapid Assessment of Human Amylin Aggregation and Its Inhibition by Copper(II) Ions by Laser Ablation Electrospray Ionization Mass Spectrometry with Ion Mobility Separation, *Analytical Chemistry 87*, 9829-9837.
[103] Salamekh, S., Brender, J. R., Hyung, S.-J., Nanga, R. P. R., Vivekanandan, S., Ruotolo, B. T., and Ramamoorthy, A. (2011) A Two-Site Mechanism for the Inhibition of IAPP Amyloidogenesis by Zinc, *Journal of Molecular Biology 410*, 294-306.
[104] Illes-Toth, E., Dalton, C., and Smith, D. (2013) Binding of Dopamine to Alpha-Synuclein is Mediated by Specific Conformational States, *Journal of the American Society for Mass Spectrometry 24*, 1346-1354.
[105] Susa, A. C., Wu, C., Bernstein, S. L., Dupuis, N. F., Wang, H., Raleigh, D. P., Shea, J.-E., and Bowers, M. T. (2014) Defining the Molecular Basis of Amyloid Inhibitors: Human Islet Amyloid Polypeptide–Insulin Interactions, *Journal of the American Chemical Society 136*, 12912-12919.
[106] Do, T. D., Economou, N. J., Chamas, A., Buratto, S. K., Shea, J.-E., and Bowers, M. T. (2014) Interactions between Amyloid-β and Tau Fragments Promote Aberrant Aggregates: Implications for Amyloid Toxicity, *The Journal of Physical Chemistry B 118*, 11220-11230.
[107] Hyung, S.-J., DeToma, A. S., Brender, J. R., Lee, S., Vivekanandan, S., Kochi, A., Choi, J.-S., Ramamoorthy, A., Ruotolo, B. T., and Lim, M. H. (2013) Insights into antiamyloidogenic properties of the green tea extract (−)-epigallocatechin-3-gallate toward metal-associated amyloid-β species, *Proceedings of the National Academy of Sciences USA 110*, 3743-3748.
[108] Sinha, S., Lopes, D. H. J., Du, Z., Pang, E. S., Shanmugam, A., Lomakin, A., Talbiersky, P., Tennstaedt, A., McDaniel, K., Bakshi, R., Kuo, P.-Y., Ehrmann, M., Benedek, G. B., Loo, J. A., Klärner, F.-G., Schrader, T., Wang, C., and Bitan, G. (2011) Lysine-Specific Molecular Tweezers Are Broad-Spectrum Inhibitors of Assembly and Toxicity of Amyloid Proteins, *Journal of the American Chemical Society 133*, 16958-16969.
[109] Acharya, S., Safaie, B., Wongkongkathep, P., Ivanova, M. I., Attar, A., Klarner, F. G., Schrader, T., Loo, J. A., Bitan, G., and Lapidus, L. J. (2014) Molecular basis for preventing alpha-synuclein aggregation by a molecular tweezer, *Journal Biological Chemistry* 289, 10727-10737
[110] Prabhudesai, S., Sinha, S., Attar, A., Kotagiri, A., Fitzmaurice, A. G., Lakshmanan, R., Ivanova, M. I., Loo, J. A., Klarner, F. G., Schrader, T., Bitan, G., and Bronstein, J. (2012) A novel "molecular tweezer" inhibitor of alpha-synuclein neurotoxicity in vitro and in vivo, *Neurotherapeutics 9*, 464-476.
[111] Wongkongkathep, P., Li, H., Zhang, X., Ogorzalek Loo, R. R., Julian, R. R., and Loo, J. A. (2015) Enhancing protein disulfide bond cleavage by UV excitation and electron capture dissociation for top-down mass spectrometry, *International Journal of Mass Spectrometry 390*, 137-145.

CHAPTER TWO

Top-Down Mass Spectrometry Reveals the Sites of Metal Binding Important in Fibrillation of α-Synuclein

Abstract

The 140-residue amyloid-like protein, α-synuclein (AS), involved in the pathogenesis of Parkinson's disease is known to be natively unstructured because of electrostatic repulsion from charged acidic residues at physiological pH. Upon exposure of AS to divalent and trivalent metals, the AS protein starts to aggregate. The top-down mass spectrometry (MS) approach of electrospray ionization (ESI), once combined with electron capture dissociation (ECD), can be used to characterize the non-covalent complexes, as both ESI and ECD do not interrupt weak protein-ligand and protein-protein interactions. In this work, we report the use of top-down MS with high resolution Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometry to determine the metal binding sites on α-synuclein. The binding to AS protein by three metals (copper, cobalt, and magnesium) was investigated. Product ion sequence maps generated from various fragmentation methods (e.g., nozzle skimmer dissociation, NSD; collisional activated dissociation, CAD; ECD; multistage tandem MS, MS^3) were compared. CAD and ECD product ions that retain metal-binding showed that there are two binding sites locating at the N- and C-termini of the protein. Copper has a primary binding site with high affinity at the N-terminus between sequence ^{1}MDVFMKGLSK^{11}. Both cobalt and magnesium share a lower affinity motif located near the C-terminus around residue ^{119}DPDNEAYE^{126} and ^{132}GYQDYE^{137}. By our MS methods, two primary binding sites elucidated from MS data agree with the previous studies.

Introduction

The protein, α-synuclein (AS), is significantly present in dopaminergic neurons located in the substantia nigra region of the brain. Aggregation of AS is thought to be directly linked to the neurodegenerative disorder, Parkinson's disease[1-3]. Pathological studies have characterized synucleinopathies as a major constituent of intraneuronal inclusions called Lewy bodies, which are mostly found in Parkinson's disease patients[4, 5]. However, the biological functions of AS are not fully understood[6]. AS has been suggested to be involved with dopamine biosynthesis because it is abundant in presynaptic termini. It can also bind to negatively charge lipid micelles *in vitro*[7-9], suggesting that it may be important in SNARE complex formation and synaptic vesicle trafficking[10-13]. Familial forms of Parkinson's disease and protein aggregation were shown to correlate with the presence of missense mutations, such as A30P[14], E46K[15], A53T[16], and recently discovered H50Q[17].

The native form of AS is a stable monomeric structure that can form oligomers. But many reports have suggested that it might have different structural conformers as well[18-21]. There are three major regions in the primary structure of AS (Figure 2-1): residues 1-60, containing four imperfect highly conservative hexamer motifs of sequence KTKEGV; residues 61-95 is the highly amyloidogenic sequence known as NAC (Non-Aβ Component of Alzheimer's disease amyloid) and was shown to contribute to about 10% of the amyloid plaques in Alzheimer's-related plaques; and lastly the C-terminal region of AS, residues 96-140, is enriched in proline and acidic residues that suggests a disordered conformation with an unstructured tail in this region. The first two regions exhibit a helical structure. It was shown *in vitro* that these helices are involved with vesicle tethering, implicating a role of AS in vesicle fusion at the synapse[22, 23]. Currently structural studies of AS using EPR, NMR, circular dichroism, fluorescent-base

approaches and several other biophysical techniques provide valuable information. Unfortunately, a high-resolution structure of full-length AS is currently unavailable.

Metal ions have been identified as one of the major components in neurodegeneration processes[24-29]. In case of AS, metal ions, including calcium, iron, copper, manganese, and cobalt, trigger the structural transformation, aggregation, and fibrillation of AS[29]. AS has also been shown to generate reactive oxygen species, i.e., H_2O_2, in a metal-dependent fashion, which is believed to further encourage AS aggregation[30]. It was suggested that metal ion binding neutralizes protein charges and changes the protein's conformation to a state that is more prone to aggregation.

Previous studies have given a structural picture of how metals bind to AS[31]. Many studies have been performed using truncated protein sequences, which may provide a more partial and indirect determination of metal binding sites. For example, the copper binding characteristics have been widely studied. Previous research using isothermal calorimetry (ITC) and fluorescence studies[32-35], EPR[36-38], NMR[39-41], and MALDI-MS of tryptic digested peptides revealed that there are two unrelated binding sites[42]: one is located near the N-terminus at Met1 and Asp2, while another site at His50 has been debatable[38, 41]. The solution equilibrium binding constant, K_d, was determined to be ca. 50 μM[25, 43]. Recently, N-terminal acetylation was shown to alter the copper binding characteristic[40]. On the other hand, the C-terminus was demonstrated to be a low affinity site for many other metals, including Mn, Co, Ni, Ca, and Fe[44, 45]. Available tools for characterizing the metal-binding properties of full-length protein structure are limited, but in the past ten years, mass spectrometry (MS) technologies for studying protein-ligand interactions and large protein complexes have been advanced[46-48].

1 MDVFMKGLS**KAKEGV**VAAAE**KTKQGV**AEAAG**KTKEGV**LYVGS**KTKEGV**VHGVATVAE**KTK**

61 **EQV**TNVG**GAVVTGV**TAVAQ**KTVEGA**GSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDP

121 DNEAYEMPSEEGYQDYEPEA

Figure 2-1. Sequence of AS highlighting three major structural regions. The N-terminal helix region is in green, NAC is in blue, and highly acidic C-terminal region is in yellow. KTKEGV repeats are bolded and underlined.

To study protein-ligand complexes by MS, the “native” complexes, i.e., in physiological pH aqueous solution to maintain noncovalent interactions present in the complex, can be introduced into the mass spectrometer by electrospray ionization (ESI). ESI is a soft ionization technique that is commonly used to generate multiply charged molecules while keeping peptides and proteins intact, and it can be performed to not disrupt weak covalent bonds or non-covalent interactions. Post-translational modifications and weakly bound non-covalent interactions can be observed in their native state. Protein-ligand complex binding stoichiometry can be directly measured from the intact masses of the complexes. Also their binding affinity can be quantified[46, 49, 50]. Although mass spectrometry measurements are performed by detecting gas phase molecules, there have been several studies published that confirm that stoichiometric data and binding affinity determinations from ESI-MS experiments agree with results from conventional methods measured from solution phase.

To gain sequence information for proteins, tandem mass spectrometry (MS/MS) methods can be used to break protein backbone bonds and their resulting products (fragments) can be measured. There are several means for cleaving the backbone bonds using MS. A slow heating method, collisional activated dissociation (CAD), is a popular MS/MS technique used in bottom-up methods[51] because of its robustness and ease of implementation. Briefly, bottom-up proteomics is performed by combining proteolysis, liquid chromatography separation coupled to mass spectrometry, fragmenting peptides by MS/MS and sequence database searching to identify the proteins. However, to gain sequence information directly from protein-ligand complexes and to identify sites of ligand binding, traditional bottom-up approaches cannot be used because the experimental conditions are not suitable to preserve the complexes in their native state. Bottom-up MS coupled with in-solution hydrogen-deuterium exchange (HDX) is a viable approach to gain information on sites of ligand interactions and protein-protein

interactions[52]; the HDX-MS/MS has been used by many laboratories to study protein conformation and folding[53, 54].

Top-down MS approaches, which do not require proteolysis prior to mass spectrometer, can be a powerful and direct approach to identify ligand binding sites[55]. Fragmenting the protein-ligand complex and detecting the products of dissociation can be used to map binding sites. Selecting the proper MS/MS technique for mapping protein-ligand binding sites is challenging because hydrophobic interactions and hydrogen bonding can be easily disrupted by the strong, energetic collisions in CAD. The goal would be to maintain the weak, noncovalent ligand interactions to the protein structure while cleaving covalent backbone bonds; fragments that retain binding to the ligand would be measured to yield information on the site of binding. Electron-base techniques, such as electron capture dissociation (ECD), has become a method of choice to study protein-ligand interactions[56]. ECD utilizes low energy electrons (< 10 eV) to deposit energy quickly enough to break protein backbone bonds while non-covalent binding survives.

Since the early development of ESI-MS in the mid-1980s to early 1990s, many small protein-metal complexes have been studied. These early works demonstrated the use of ESI-MS to study the interactions between short peptides (3-10 residues) and metals, such as alkali metals, calcium, and zinc using CAD fragmentation[57-60]. Copper binding to β-amyloid 16 (Aβ16), a shortened version of the Aβ40 protein that is involved in Alzheimer's disease, was studied by ESI-MS; two histidine residues were revealed to be the copper binding residues[61]. However, other previous studies were limited to small sized proteins due to limited mass resolution and insufficient mass accuracy.

ESI combined with ECD and high resolution Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry was demonstrated to be able to determine the AS binding sites of small ligands such as spermine and spermidine[48] and a compound that reduces fibril formation, molecular “tweezer” or CLR01[62]. Even for larger proteins, it was also capable of identifying the zinc binding site of the tetrameric 147 kDa yeast alcohol dehydrogenase (ADH) protein[63]. In this technique, briefly, intact protein complexes were ionized by ESI and introduced into the mass spectrometer. The ion trajectories and internal energies were manipulated by a series of ion optics and quadrupole filters. Ligand-bound molecules that were still in the intact form were isolated and subjected to ECD-MS/MS. With the special capabilities of ECD, ligands can still remain on its original binding residues even as protein backbone bonds are fragmented. The ultra high resolution of FT-ICR is advantageous for distinguishing between ligand-bound fragments (holo-fragments) and fragments without ligand (apo-fragments). Apo- and holo-fragments, which contain information on their original binding positions, can be used to map out the binding motif of ligands.

In the gas phase, electrostatic interactions is dramatically enhanced because the dielectric constant in vacuum is much lower than that in solution phase, like water. This fact leads to an increase in the *gas phase* binding affinity between protein and ligands that are held by electrostatic forces. These interactions can be retained even during collision based MS/MS techniques[64, 65]; electrostatic interactions can be as strong as covalent bonds. Complexes of nucleotide CTP binding to proteins ribonuclease A (RNaseA) and ribonuclease S (RNaseS), and ATP binding to adenylate kinase (AK) can be analyzed by ESI and CAD using the top-down MS approach; the binding sites were successfully identified[66-68].

In this paper, AS protein binding to cobalt, manganese, and copper were investigated by top-down mass spectrometry using ESI-FTICR MS. ECD and CAD were used for MS/MS experiments. Subsequently, the binding sites of the three metals were directly revealed from their intact states. Ion mobility spectrometry was also used to show structural changes of the intact conformation upon metal binding.

Experimental

Materials

Recombinant α-synuclein and three truncated variants (1-60, 61-140, and 96-140) were purchased from rPeptide (Bogart, GA). AS was dissolved in distilled water and desalted with 20 mM ammonium acetate solution by using 10 kDa MWCO centrifugal filter devices from Amicon (Billerica, MA). The final protein concentration was 10 uM in 20 mM ammonium acetate solution at pH 6.8. Aluminum acetate (basic), cobalt acetate, copper acetate, and manganese acetate were purchased from Sigma-Aldrich (St. Louis, MO) and dissolved in 10 mM ammonium acetate. Additional ultra-sonication was used to assist dissolution of some metals in solution. After protein desalting, metals were added to the AS solution at various concentration ratios (protein:metal) from 1:1 to 1:100.

Mass spectrometry

A nano-ESI source with Au/Pd-coated borosilicate emitters (Thermo Scientific, San Jose, CA) was used to deliver the analyte solution at a 0.2 uL/min flow rate using an ESI voltage of 1100-1300 V. A 15-Tesla Bruker Solarix FTMS instrument (Bruker, Billerica, MA) was used, and the estimated resolving power was set to 400,000 at *m/z* 400. The capillary temperature was set to 180 C, and 30V additional skimmer voltage was applied to activate ions to improve ECD product

ion yield while keeping the protein complexes intact. MS/MS experiments were done using ECD and CAD. For the ECD experiments, the precursor ion was isolated in the quadrupole filter and transferred to the FTICR cell to perform ECD. CAD was done inside the collision cell after quadrupole isolation. CAD fragments were detected by the FTICR cell. Additional MS^3 experiments were performed by nozzle skimmer CAD (NS-CAD) followed by CAD or ECD of the highest abundant NS-CAD fragments. MS^2 and MS^3 fragments were manually assigned against the theoretical fragments indicated by Protein Prospector (http://prospector.ucsf.edu/prospector/mshome.htm). Fragmentation maps were plotted by Matlab.

Ion mobility mass spectrometry (IM-MS) experiments were performed with a Waters Synapt G2 HDMS instrument (Waters, Milford, MA) by Professor Hugh Kim's group in Korea University (Seoul, Korea). Similar to FTMS, a nano-ESI source with Au/Pd-coated borosilicate emitters (Thermo Scientific) was used with a 0.2 μL/min flow rate and an ESI voltage of 1200-1400 V. A source temperature of 100 ˚C, a desolvation temperature of 200 ˚C, and a cone voltage of 20 V were used. For the helium cell, which is located prior to the IMS cell, a gas flow rate was set to 150 mL/min and the IMS gas flow was set to 30 mL/min. The IMS wave velocity and wave height were set to 150 m/s and 7.0 V, respectively. The pressure in the helium cell and IMS cell were 1.37 $\times 10^3$ and 5.04 $\times 10^3$ mbar, respectively.

Results and Discussion

MS and MS^2 of α-Synuclein

Alpha-Synuclein is known to be natively unfolded at physiological pH; the sequence contains many acidic residues, especially at the C-terminal end (Figure 2-1). Coulombic repulsions from negative charged acidic residues prevent a globular structure to be formed[10, 69]. Consequently, the measured charge states of AS from native ESI-MS occur are found in the lower m/z region compared to other native proteins of similar sizes, i.e., cytochrome C and myoglobin. A positive ion mass spectrum of AS in 20 mM ammonium acetate at pH 6.8 is shown in Figure 2-2a. Charge states of AS vary from 16+ to 6+ with a maximum intensity at 12+. The result suggests that AS is mainly in an unfolded state in the gas phase, even from a pH 7 protein solution. AS dimers can be observed at higher *m/z* with charging between 9+ to 6+. For top-down MS experiments, the precursor ions at *m/z* 1033 (14+), 1113 (13+), and 1206 (12+) were isolated to perform CAD and ECD. This result was similar to previous reports[70, 71].

Figure 2-3a shows a bar plot of the charge-normalized c/z fragment ion intensities resulting from ECD of AS from the combined three charge states (12+ to 14+) as a function of different cleavage residues. This is called a fragmentation map. N-terminal fragments (c ions) are plotted above the longitudinal axis, and C-terminal fragments (z ions) are plotted below the axis. The ECD data shows that fragmentation occurred equally throughout the protein backbone, but preferentially toward the N-terminal end. The observed ECD cleavages were dominant at c6, c9, c38, c46, where many of the positive charge sites are located. For CAD, a plot of b/y product ion intensities and a fragmentation map are shown in Figure 2-3b. CAD favors an extensive fragmentation around the acidic C-terminal region because of the proline effect, where cleavage of the bond N-terminal to proline residues is particularly favored[72]. The b116, b126, b127, b137

product ions are among the highest abundant products because of the five prolines in the sequence: P108, P117, P120, P128, and P138. The most frequently observed C-terminal product ion is y13 resulting from M_{126}/P_{127} cleavage. Also, cleavages of bonds C-terminal to acidic residues, including aspartic, are favorable[73] and observed as y9, y13, y14, and y21 product ions. Overall, sequence information is limited from CAD to regions mostly near the C-terminus rather than the N-terminus. Backbone cleavage efficiencies from ECD and CAD of AS were 67.14% and 35.71%, respectively. By combining the results from both techniques, the total cleavage efficiency was increased to 90%.

Metal complexes of AS

Copper, cobalt, and manganese were investigated for their binding characteristics to AS. Concentration ratios of AS to copper, cobalt, and manganese that showed significant binding for MS experiments were optimized at 2:1, 1:5, and 1:10, respectively. The concentration of AS was 10 pmol/μL and copper acetate, cobalt acetate, and manganese acetate were added to reach the desired concentration ratios. MS spectra of apo-AS and metal bound-AS are summarized in Figure 2-2. Deconvoluted mass spectra (converted from *m/z* to mass), obtained from Bruker DataAnalysis software, are displayed in the insets on the right. The measured monoisotopic and average mass of AS are 14452 and 14460 Da (Figure 2-2a), respectively. Upon metal binding, negatively charged residues are expected to be electrostatically neutralized by the metal ions. Protein structural conformations may be changed and possibly result in a change in charge state distribution. As shown in the mass spectra, however, the charge state of metal complexes does not significantly shift compared to the apo-AS. The relative abundance for the lower charge states increased for the copper complexes. This may suggest that copper binding may change protein comformations becoming more compact structures. The charge envelope distributions of cobalt and manganese are similar to apo-AS. Ion mobility spectrometry

(IMS) will be used to reveal structural changes upon metal binding. Mass shifts from metal binding can be easily viewed in the deconvoluted mass spectra. Measured mass accuracies are summarized in Table 1, demonstrating the high mass accuracies measured by high resolution FT-MS. AS-copper complexes were readily observed at a 1:0.5 protein/metal ratio. A majority of the AS species is in the 1:1 complex (AS bound with one copper; 14513 Da, monoisotopic mass), but apo-AS was barely observed indicating that copper binds to AS with high affinity and specificity (Figure 2d). On the other hand, cobalt and manganese have different affinities to AS. Even at higher AS-metal concentration ratios for AS-cobalt and AS-manganese at 1:5 and 1:10, respectively, mass spectra of both AS-cobalt and AS-manganese complexes clearly showed that intensity ratios of apo-AS and single bound complex were only 40% and 60%, respectively. This indicates that cobalt and manganese bind to AS with lower affinity than for copper. Noted that the observed masses for singly protonated 1:1 AS-Co and AS-Mn complexes were at 14510 and 14504 Da, consistent with singly bound metal as shown in Table 2-1.

Table 2-1. A summary of mass accuracies measured by FT-MS for each AS-metal complex

Compound	Theoretical mass (Da)	Observed mass (Da)	mass accuracy (ppm)
AS	14452.23	14452.34	7.61
AS+Co	14508.14	14508.24	6.89
AS+Cu	14513.14	14513.18	2.76
AS+Mn	14504.15	14504.24	6.21

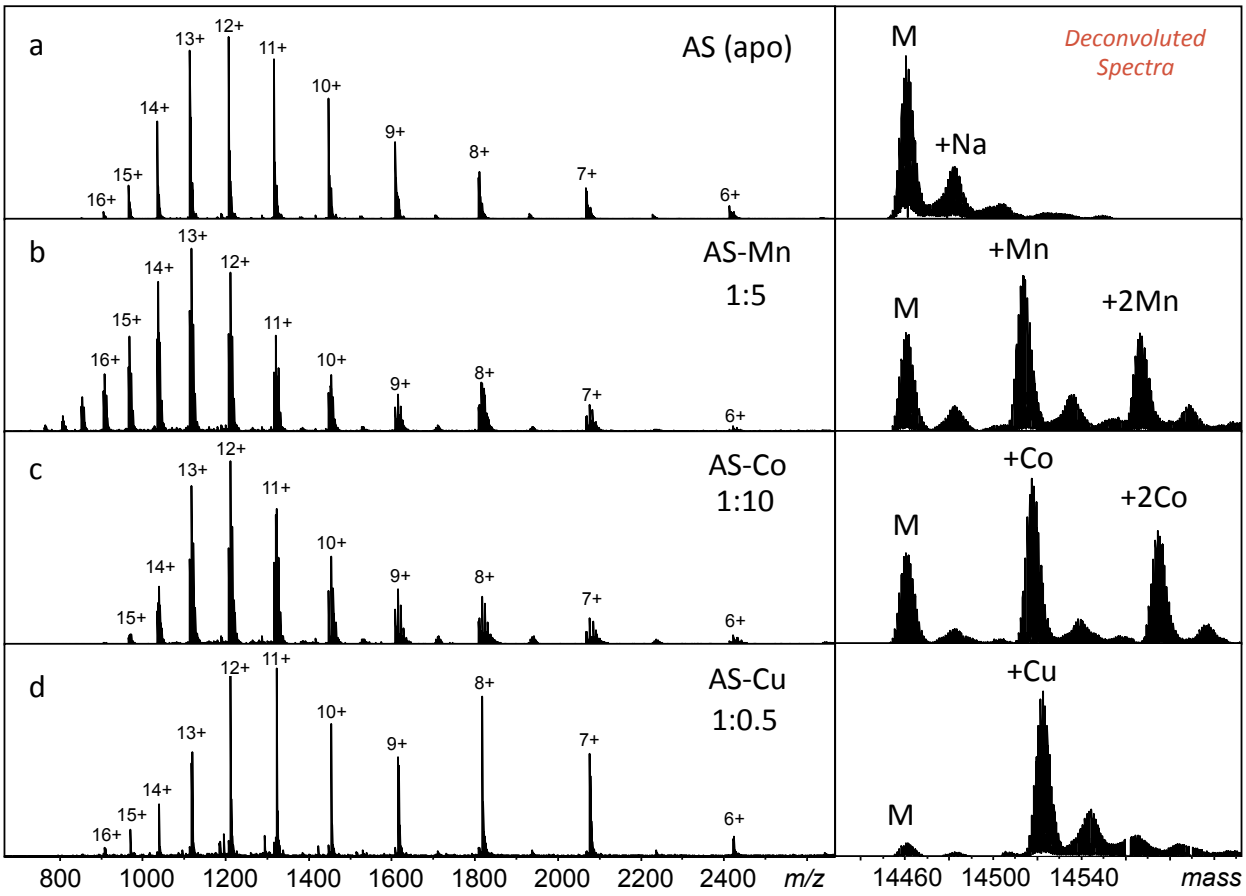

Figure 2-2. Mass spectra of apo-AS (a), AS-Manganese complex (b), AS-Cobalt complex (c), and AS-Copper complex (d). Deconvoluted spectra showing metal adducts are located on the right (e-g). Concentration ratios of protein:metal were 2:1, 1:5, and 1:10 for copper, cobalt, and manganese complexes, respectively.

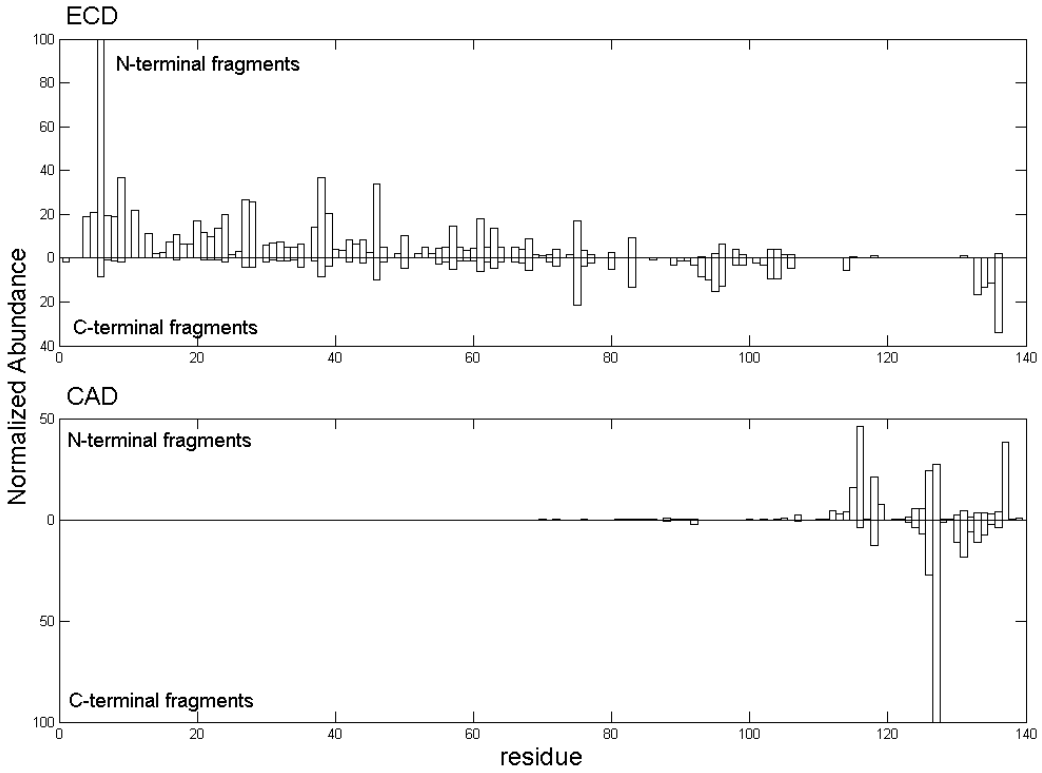

Figure 2-3. Plot of fragment ion intensities and cleavage residue number , i.e., fragmentation maps. The map of CAD shows extensive fragmentation near the C-terminus (bottom). The ECD map demonstrates equal fragmentation throughout the protein backbone (top).

MS/MS of AS-metal complexes

Three charge states, 14+, 13+, and 12+, of the 1:1 protein-metal complexes were selected for the ECD and CAD experiments. The 1:2 complexes of cobalt and manganese were not analyzed in our experiment because they are most likely due to non-specific binding (data not shown). A MS/MS scan range from *m/z* 400-4000 was selected. ECD and CAD product ion maps were generated, showing bar plots of amino acid residues of the protein against the summed charge-normalized product ion intensities from the three precursor ion charge states (Figures 2-4 to 2-6).

MS/MS reveals that copper binds to N-terminal helix

Combining c- and z-type product ions from the 14+, 13+, and 12+ charge states of the AS-Cu complexes, an ECD fragmentation map is shown in Figure 2-4 (top). N-terminal, c-product ions are displayed above the top part above the x-axis that indicates the amino acid residue number for AS. The bottom part below the axis shows C-terminal, z-type product ions. Copper-bound fragments are labeled in green, while apo-fragments were shown as hollow bars. The fragmentation map for CAD is plotted in the same manner as the ECD data (Figure 2-4, bottom). A sum of b-ion intensities is plotted above the x-axis and y-ion intensities are plotted below the axis. Ion intensities from fragments with neutral loss (-H_2O and –NH_3) were considered for the plots. From both ECD and CAD, most N-terminal ions were holo fragments, except a few apo-fragments observed toward the N-terminal end. All C-terminal ions were observed as apo-fragments, indicating that the copper-binding site is located near the N-terminal region. The CAD fragmentation pattern of AS-Cu complex looked similar to that of apo-AS. Highly abundant fragments from CAD were found near the C-terminus. However, the ECD pattern alters significantly between residues 1-20. This is again an indication that the binding of copper to the N-terminus changes the protein's structural conformation.

By examining the fragmentation pattern closely, two N-terminal apo-fragments, c_6 and c_9 and many copper-bound c ions, starting from c_{11}+Cu to c_{137}+Cu, were detected from the ECD-MS/MS experiments. This is an indication that the binding site is located somewhere around residues 1-11.

The CAD data lacks information for the region between residues 1-49. The b_{50}+Cu (copper bound b_{50}) was the first N-terminal ion detected. However, the product ion signal of b_{50}+Cu is only 0.2% relative to the highest abundant ion, b_{116}+Cu. An important point to note is that there are a number of copper-bound fragments retained from the CAD process. This result confirms that electrostatic interactions between AS and copper were enhanced in the gas phase. The enhanced non-covalent interactions could demonstrate a covalent-like affinity and they could survive high-energy gas phase collisions.

However, ECD provides greater in-depth fragmentation information near the N-terminus than CAD. We can conclude that a primary copper binding site is near the N-terminus, between residues ^{1}MDVFMKGLSK^{11}. Backbone cleavage (sequence coverage) from ECD and CAD experiments of AS-Cu were 69.3% and 48.6%, respectively. By combining data from both techniques, the total backbone cleavage was increased to 90%. CAD can be used as a complimentary technique to ECD for elucidating metal binding and other protein-ligand complexes stabilized by electrostatic interactions.

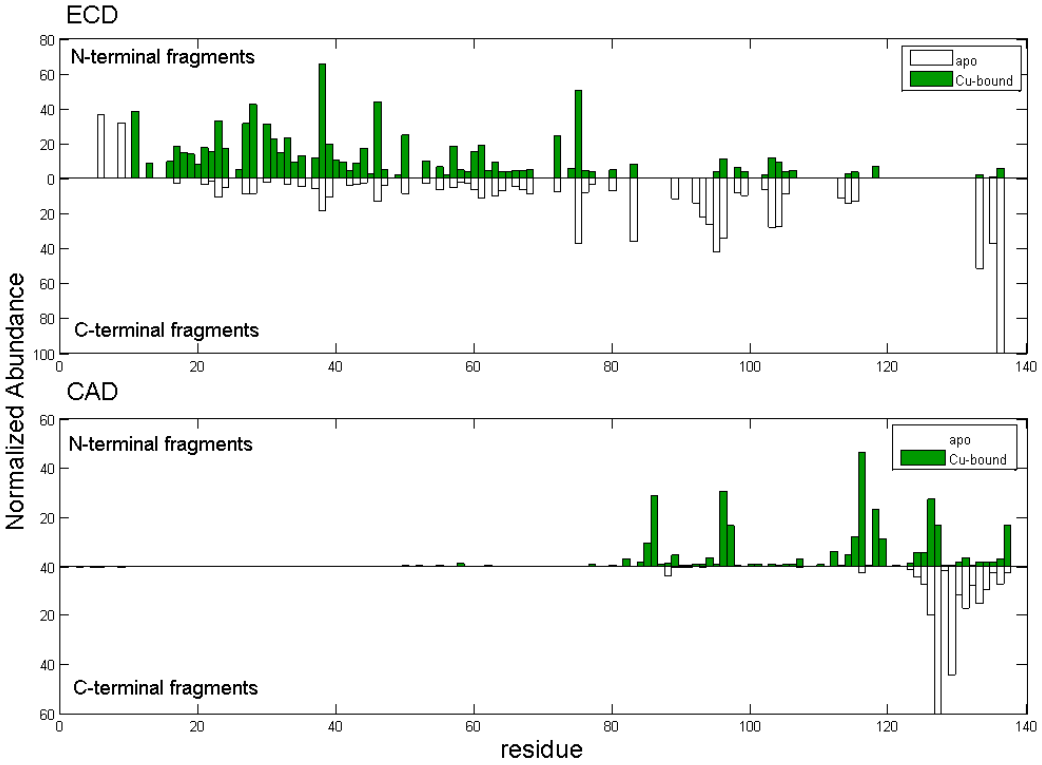

Figure 2-4. MS/MS fragmentation maps from ECD (top) and CAD (bottom) of AS-Cu complex were generated by plotting fragment ion intensities and residue number.

A possibility to perform pseudo-MS^3 experiments

Although CAD-MS/MS and ECD-MS/MS have capabilities to generate metal-binding information, there is a possibility to use a third stage of tandem MS (MS^3) to further reveal and confirm the binding sites. However, the FT-MS that was used in this study is not capable to perform "classical" MS^3 because it has a quadrupole filter followed by a collision cell coupled with the ICR cell. Multistage tandem-in-time (3D or linear ion trap) or tandem-in-space (multiple quadrupoles) MS^n experiments require additional instrument modifications. Fortunately, It is still possible to perform a "pseudo"-MS^3 experiment by ramping up the skimmer voltage in the ESI atmosphere/vacuum interface to fragment ions; this process is known as nozzle-skimmer dissociation (NSD) or in-source CAD (IS-CAD). This process yields a similar result as CAD. By utilizing the skimmer potential, the first fragmentation step was performed inside the source region. However, precursor selection is not possible for the first fragmentation step. So, all charge states generated by ESI are affected by pseudo-CAD in the source. The second fragmentation step would be inside the collision cell for CAD, or the ICR cell for ECD, similar to MS^2 experiments as mentioned above.

MS^3 of AS-copper complex

Based on the CAD data, the most abundant copper-bound fragment is b_{116}+Cu (with a 100% relative intensity). For the MS^3 experiment, after the protein-metal complex was fragmented by IS-CAD, the b_{116}+Cu product ion was isolated in quadrupole filter and was further fragmented by CAD in a collision cell, or ECD in an ICR cell. Because the b_{116}+Cu ion is relatively large, results from both MS^3 IS-CAD/CAD and IS-CAD/ECD, however, did not show any improvement in identification of the binding site (data not shown) compared to ECD-MS^2 and CAD-MS^2. However, it provided confirmation of the MS^2 data that the copper binding site remains the same as MS/MS experiment as mentioned above.

Cobalt binds to C-terminal unstructured tail

For MS/MS analysis, charge states from 14+ to 12+ of the AS-Co 1:1 complexes were examined. Fragmentation maps from ECD and CAD are shown (Figure 2-5) in a similar manner as for the AS-copper complex. Holo fragments are represented in blue and apo-fragments are shown as hollow bars. A map of ECD fragments shown in Figure 2-5 (top) revealed that most N-terminal fragments up to residue 118 do not contain the ligand, except two cobalt-bound fragments near the C-terminal end, c_{131}+Co and c_{139}+Co. The majority of the C-terminal fragments, such as z_{25}+Co, have cobalt bound. The only apo-z-fragment is z_5. Unlike copper, a pattern of ECD of cobalt complex looks similar to apo-AS, especially from the N-terminal region. The observed common fragments showing high intensity were c_6, c_9, c_{27}, c_{38}, and c_{46}. This suggests that there was no significant change in protein conformation near N-terminus upon Co-binding, and Co-binding is not found in this region. The ECD fragments indicate the C-terminal tail is the site of binding. However, ECD did not provide backbone cleavage to the same level as did CAD. Backbone cleavage by ECD was 78.6%.

If the Co-binding site is at the C-terminus, CAD of the AS-cobalt complex may yield more information for metal binding than that of the AS-copper complex due to extensive fragmentation near the C-terminus. Backbone cleavage was 60.7%, higher than for apo-AS and the AS-Cu complex. The CAD map for AS-Co shows difference compared to that for apo-AS and the AS-Cu complex (Figure 2-5, bottom). Backbone cleavage from CAD is highly dependent on protein structure. The change in the AS-Co CAD pattern compared to the apo-AS form might suggest that there is a structural change upon cobalt binding to the C-terminal residues. Overall, we observed more abundant apo- and holo- y ions for the AS-Co complex than for apo-AS. The following apo-fragments were identified: b_{115}, b_{116}, and y_{127} (low intensity). Also we observed b_{126}+Co, b_{127}+Co, b_{137}+Co, y_{21}+Co and y_{24}+Co holo products. With backbone cleavages from

CAD and ECD, we confirmed that cobalt binding region is between residues ^{119}DPDNEAYE^{126}. Backbone cleavage (sequence coverage) increased to 96.4% when combining data from both ECD and CAD. Generally, CAD provided better identification of C-terminal binding than ECD for Co-binding.

Manganese shares the same binding site as cobalt

Similarly to the experiments with the copper and cobalt complexes, three charge states from 14+ to 12+ of 1:1 AS-Mn complex were isolated for ECD-MS/MS and CAD-MS/MS. The fragmentation patterns of the manganese complex by ECD and CAD (Figure 2-6) from the three charge states are very similar to the data for the AS-cobalt complex (Figure 2-5). The maps were generated in the same manner as for copper and cobalt complexes. Manganese-bound fragments are labeled in red, and the unbound fragments are labeled with hollow bars. ECD of the 1:1 AS-Mn complex produced c_{131}+Mn and z_{25}+Mn product ions, which are key manganese-bound fragments indicating that the Mn binding site is between residues 116-131. Backbone cleavage efficiency obtained from ECD was 69.3%. Similar to the cobalt complex, CAD-MS/MS of the Mn-bound complex showed better sequence coverage (than by ECD) and more specific binding site identification near C-terminus. The CAD pattern is also similar to that from the cobalt complex. CAD provided backbone cleavage efficiency of 75.7%, and the total sequence coverage increased to 96.4% when ECD and CAD data were combined, similar to the experiments for the cobalt complex. The b_{115} and b_{116} products were among the highest abundant apo-fragment observed. Because b_{126}+Mn, y_{21}+Mn, and y_{24}+Mn ions were observed by CAD, it is reasonable to narrow down the Mn binding site to the region between ^{119}DPDNEAYE^{126}. Again, CAD generated more binding site information for manganese-binding compared to ECD because plenty of C-terminal fragments were produced. In summary, evidence from ECD and CAD suggested that manganese shares the same binding site as

cobalt. There is a possibility of a structural change near the C-terminus upon cobalt/manganese binding to AS, as indicated by the CAD fragmentation pattern or profile.

MS^3 of AS-cobalt and manganese complexes revealed another possible C-terminal binding site

Because the CAD fragmentation pattern is similar for cobalt and manganese binding, it suggested that cobalt and manganese share a similar binding motif. Similar precursors were chosen to perform MS^3 experiments to further confirm the binding site determination. As IS-CAD generated some metal bound fragments, two precursors from each metal complex: y_{21}+Co, y_{24}+Co, y_{21}+Mn, and y_{24}+Mn, were isolated for the MS^3 experiments. The four isolated precursors were transferred into the collision cell to perform the second stage CAD.

Fragmentation maps for the cobalt-bound 21-mer and 24-mer peptides were created, as shown in Figure 2-7. Interestingly, the identified cobalt-binding site by MS^3 experiments was slightly different from that by MS/MS. While the MS/MS data shows that cobalt binds to a region composed by residues 119-126, the CAD-MS^3 maps of y_{21}+Co and y_{24}+Co show another binding site between ^{132}GYQDYE^{137} (Figures 2-7a and 2-7b), suggesting that cobalt may migrate from one site to another site in the gas phase.

The data for the manganese complexes showed similar results. CAD-MS^3 fragmentation maps were generated for the y_{21}+Mn and y24+Mn products. The manganese-bound peptide MS^3 maps look different from that for cobalt binding. There is a distinct peptide backbone cleavage at E137 (relative to the AS sequence) that yields manganese-bound b-ions (b_{18}+Mn from y_{21}+Mn precursor, and b_{21}+Mn from y24+Mn precursor). The rest of the fragmentation profiles were similar that found for cobalt binding. The combined data revealed manganese binding site to be

the same as cobalt at residues ^{132}GYQDYE^{137}. The results from MS/MS and MS^3 confirmed that cobalt and manganese have a similar AS-binding behavior.

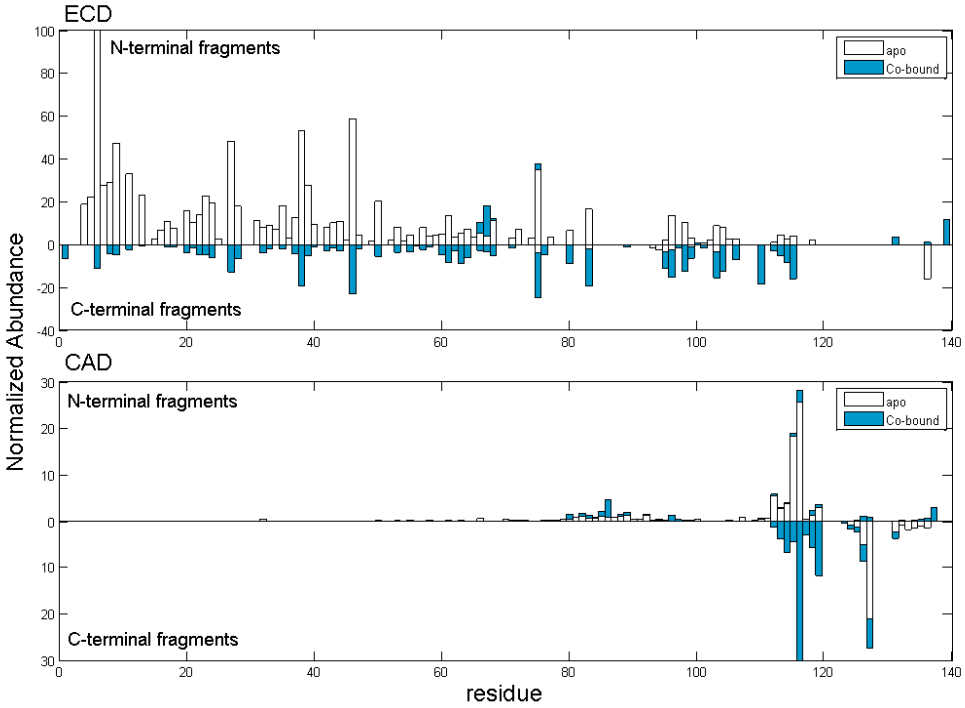

Figure 2-5. MS/MS fragmentation maps from ECD (top) and CAD (bottom) of AS-Co complex were generated by plotting fragmented ion intensities and residue number.

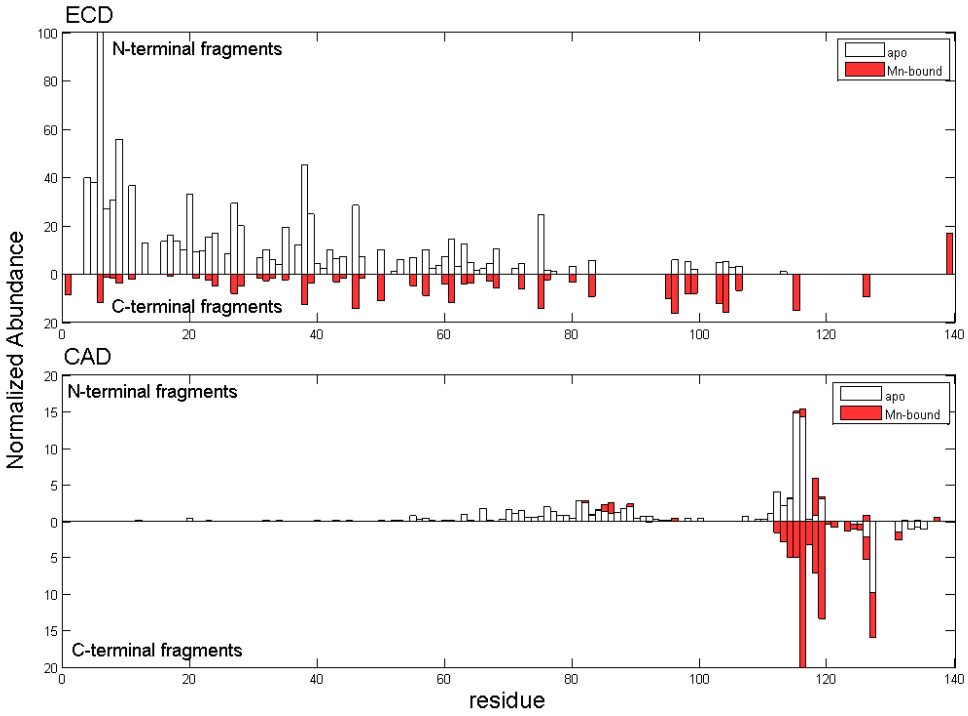

Figure 2-6. MS/MS fragmentation maps from ECD (top) and CAD (bottom) of AS-Mn complex were generated by plotting fragmented ion intensities and residue number.

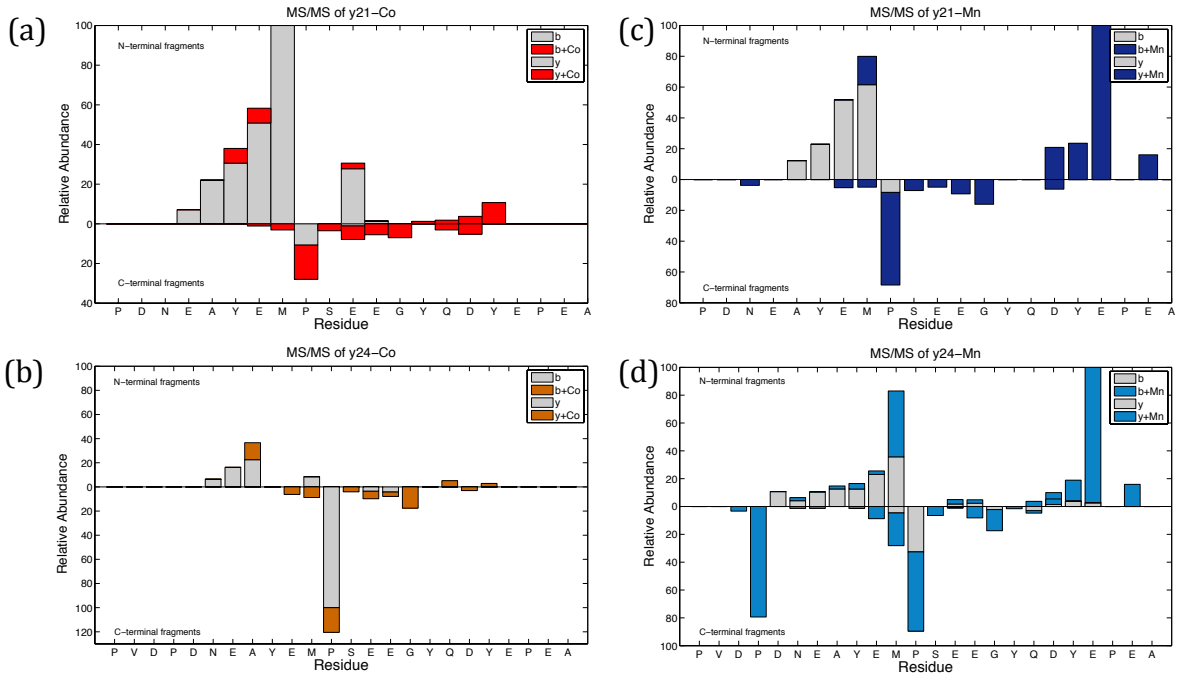

Figure 2-7. IS-CAD/CAD-MS^3 fragmentation maps of AS-Co and AS-Mn complexes. For cobalt, y21+Co (a) and y24+Co (b) were isolated and CAD spectra were acquired. AS-Mn complex was also analyzed. The y21+Mn (c) and y24+Mn (d) ions were isolated for subsequent CAD to identify the binding site.

MS/MS of truncated AS confirmed metal binding sites

Three truncated AS sequences (1-60, 61-140, and 96-140) were tested for their metal binding characteristics by MS and MS/MS. AS(1-60), which contains a helical domain of AS, showed strong binding to copper but not to cobalt and manganese (data not shown). This result agrees with Cu binding to be near the N-terminal region. However, when ECD and CAD experiments were performed to locate the binding site of Cu to AS(1-60), it showed that the site was shifted by 10 residues towards the C-terminus (data not shown). The shifted binding site suggested that the 1-60 truncated protein might have different structural conformations compared to the full length AS. As expected, both cobalt and manganese bind to AS(61-140) and AS(96-140). The binding of copper to AS(1-60) missing the acidic C-terminal region was not observed. The complex of AS(96-140) with cobalt and manganese were analyzed by CAD and ECD-MS/MS. The binding site was identified to be the same as for full-length AS.

Structural changes were observed by ion mobility spectrometry

Ion mobility was performed on qTOF (quadrupole time-of-flight) instrument with a T-wave mobility cell. Collisional cross-section (CCS) profiles of six charge states (14+ to 9+) for copper-, cobalt-, and manganese-AS complexes were examined and compared to CCSs for the apo-AS form. Without metal, AS shows one extended conformation for the high charge states (14+ and 13+). The measured CCS values decreased from 2830 $Å^2$ for the 14+ molecule to less than 2400 $Å^2$ for 9+, 10+, and 11+ molecules. Three lower charge states (11+ to 9+) have a major conformation with a similar CCS showing a more compact structure than the higher charged species. For the lower charge states, more than one conformation at a lower CCS was observed. For the 9+ molecule, three additional conformations were measured. The most abundant compact structure has a CCS of 1822 $Å^2$.

With copper (II) binding, the compact conformation proportion of the low charge states (11+, 10+, and 9+) increased, as shown with the blue arrows in Figure 2-8. The result confirmed that binding of copper at the N-terminal site induced conformational changes of AS. It is possible that there are more negatively charge sites for the lower charge state of AS. The copper ion may grab the negatively charged C-terminal tail and form a coordination that leads to more folded states. This structure may lead to an important understanding of AS aggregation with and without acceleration by heavy metals. However, the right panel of Figure 2-8 shows that there is no significant difference for high charge states (14+, 13+, and 12+) that is likely to be in extended structures.

CCS profiles of cobalt and manganese complexes are shown in Figure 2-9. The MS/MS data suggests that cobalt and manganese share the same binding motif on the unstructured C-terminal tail. Cobalt (left panel) and manganese (right panel) exhibit a similar CCS pattern. Similar to apo-AS, higher CCS values were observed for the high charge states (14+, 13+, and 12+), indicating a more extended conformation for both cobalt and manganese binding. Unlike copper, cobalt and manganese do not show conformational changes of compact structures at lower charge states (11+, 10+, and 9+). The result agrees with the fact that the C-terminus is extended and does not have a defined structure. Binding of copper or manganese to the C-terminal motif does not significantly change the big picture of protein structure, yielding similar CCS profiles between the apo- and holo-forms. There are slight shifts for the 13+ and 12+ molecules toward the lower CCS (more compact structure), which need to be further investigated.

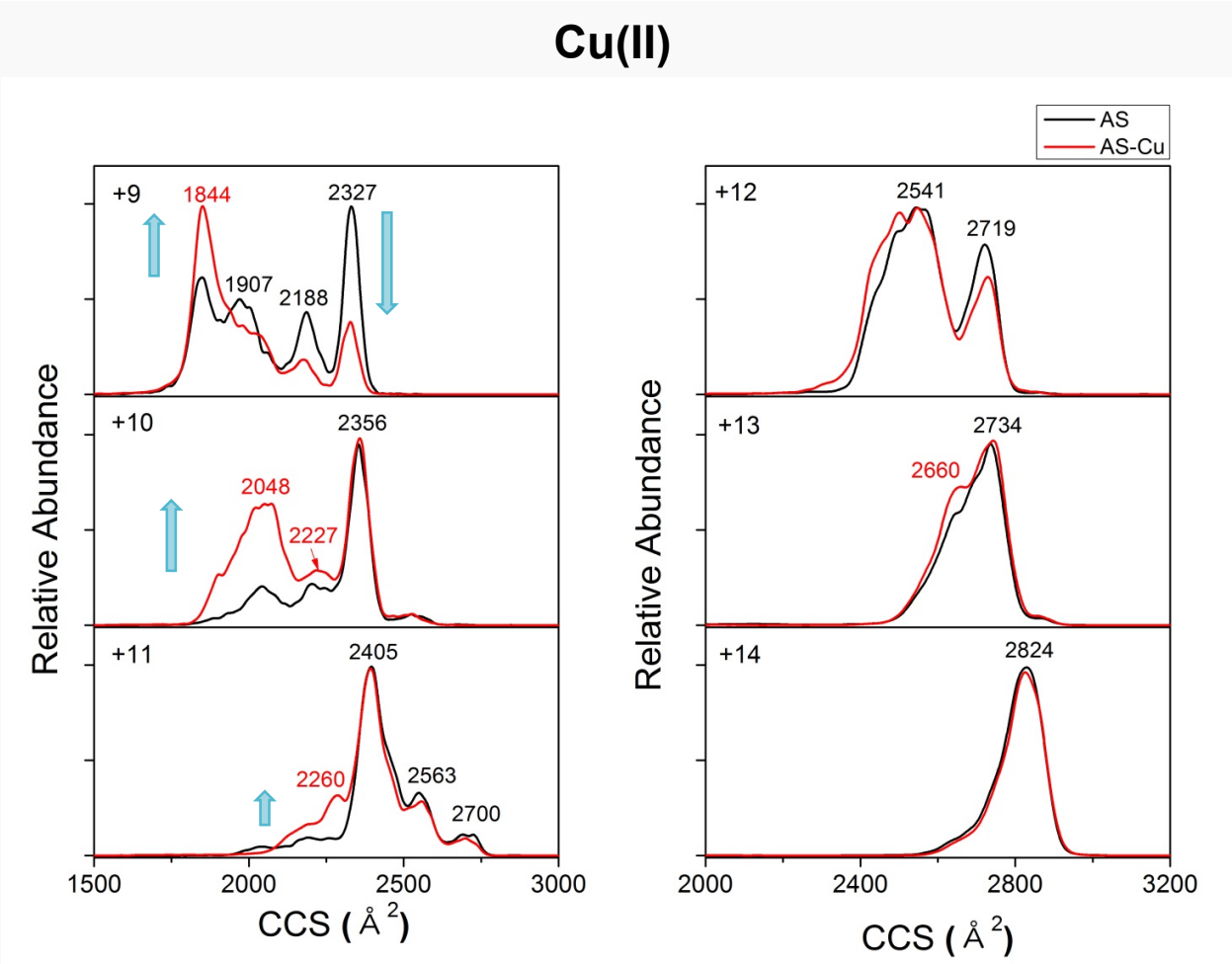

Figure 2-8. Collision cross-section (CCS) profiles from ion mobility of apo-AS (black) and copper-AS (red) complex at six different charge states ranging from 14+ to 9+. The lower charge states showed a significant change toward compact conformation upon copper binding.

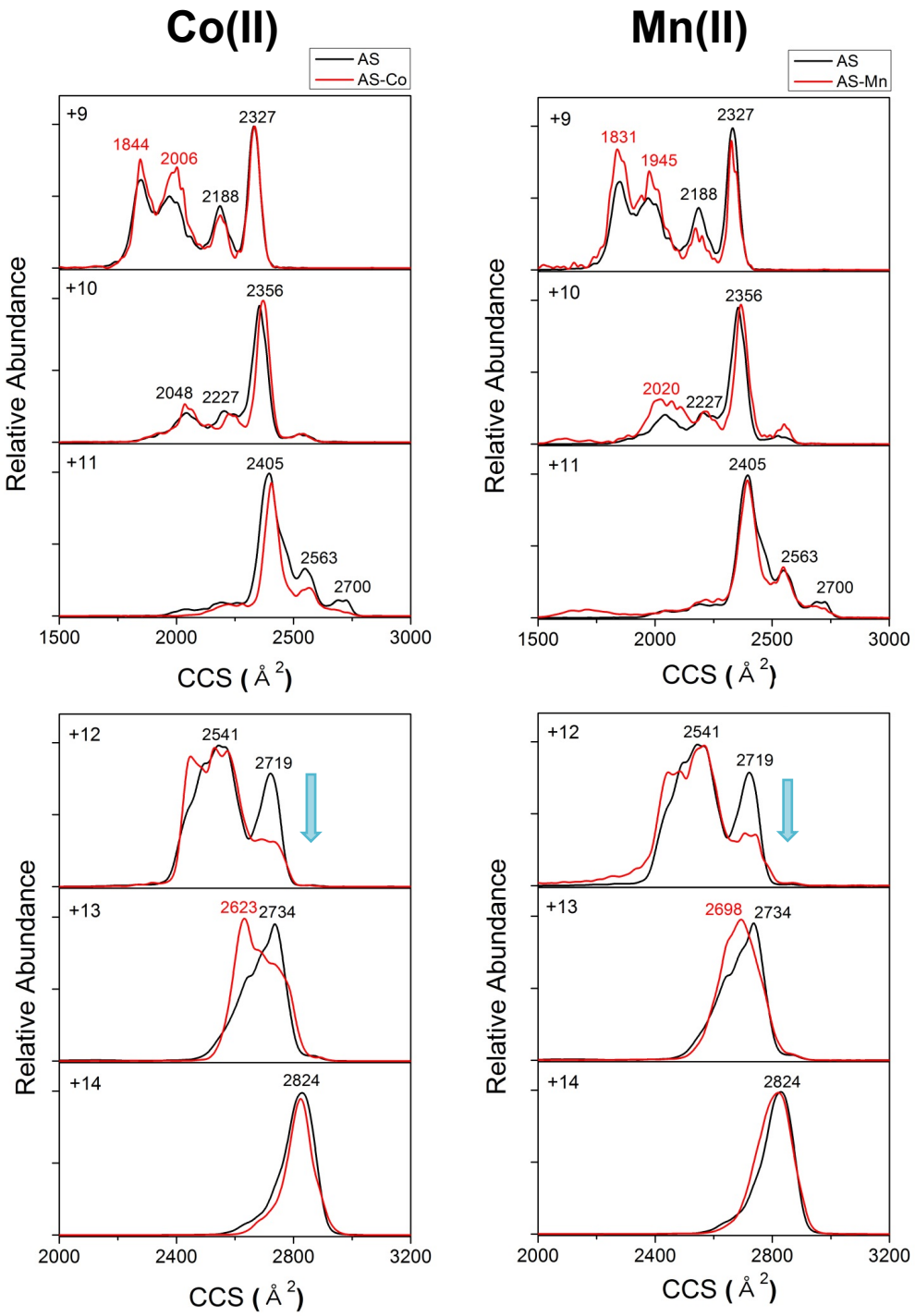

Figure 2-9. Collision cross-section of cobalt (left) and manganese (right) complexes in comparison to apo-protein (red:metal complexes, black:apo-protein). The result showed that both metals do not noticeably induce protein conformation.

Conclusion

Native top-down mass spectrometry provides important structural information for intact proteins. ESI can be used to gently ionize samples without disrupting any weak interactions. Solvent pH is kept to the physiological region by to preserve the native structure of the protein. The top-down MS approach also provides sequence information by fragmenting intact protein backbone bonds. By using soft dissociation methods like ECD, weak interactions can be retained in the protein while the backbone is cleaved. We successfully utilized these methods to study electrostatic interaction between metals and the α-synuclein protein. Our studies suggest that electrostatic interactions are enhanced in the gas phase. Table 2-2 shows the metal binding affinities to AS as determined from solution phase-based methods. Some of the previously reported data shows metal binding to be in the mid-micromolar range. CAD of protein-ligand complexes normally would dissociate the ligand from the protein, but the enhanced *gas phase* affinity allows the ligand to be retained even upon CAD. This can be advantageous for mapping metal binding sites by CAD, which uses higher energy than ECD; AS can be fragmented while metals remain bound to the protein mostly to the same region as for the native solution state. For the purposes of this paper, three metals, copper, cobalt, and manganese, were investigated. ECD- and CAD-MS/MS revealed that copper binds to the N-terminal end of AS near residues ^{1}MDVFMKGLSK^{11}. The ECD pattern of the copper complex was slightly different than for apo-AS, suggesting that a minor structural change could occur upon copper binding. The structural change was confirmed by ion mobility. Collision cross-sections of the complexes at low charge state shows a significant change toward a more compact conformation. Cobalt and manganese have similar binding sites. MS/MS unveiled the shared binding site between residues ^{119}DPDNEAYE^{126} near the C-terminus. The binding sites identified by MS agree with previous studies[74]. MS^3 experiments of cobalt and manganese complexes were performed by

IS-CAD followed by CAD to validate the binding site. However, an identified binding site from MS^3 was different from the MS/MS experiments. During MS^3, metal ions may migrate from residues ^{119}DPDNEAYE^{126} to ^{132}GYQDYE^{137}. MS studies of metal-bound truncated protein show an agreement with binding sites identified by top-down mass spectrometry.

Table 2-2. Comparison of results of metal binding affinity to AS protein from MS against previous literature data (modified from Binolfi, A.; Fernandez, C.O.; [63], 2012).

MS results

Metal ion	Binding region	Binding Sequence	Affinity
Cu(II)	N-term	^{1}MDVFMKGLSKA^{11}	High
Mn(II), Co(II)	C-term	^{120}PDNEAYE^{126}	Low

Previous studies by NMR and EPR

Metal ion	Binding region	Binding Sequence	Dissociation Constants
Cu(II)	N-term	^{1}MDVFMK^{6}	0.1 uM
		^{48}VVHGV^{52}	35 uM
Mn(II), Fe(II), Ni(II), Co(II)	C-term	^{119}DPDNEA^{124}	~ 1mM

References

[1] Goedert, M. (2001) Alpha-synuclein and neurodegenerative diseases, *Nature Reviews Neuroscience 2*, 492-501.

[2] Lee, V. M. Y., and Trojanowski, J. Q. (2006) Mechanisms of Parkinson's Disease Linked to Pathological α-Synuclein: New Targets for Drug Discovery, *Neuron 52*, 33-38.

[3] Masliah, E., Rockenstein, E., Veinbergs, I., Mallory, M., Hashimoto, M., Takeda, A., Sagara, Y., Sisk, A., and Mucke, L. (2000) Dopaminergic Loss and Inclusion Body Formation in α-Synuclein Mice: Implications for Neurodegenerative Disorders, *Science 287*, 1265-1269.

[4] Spillantini, M. G., Schmidt, M. L., Lee, V. M. Y., Trojanowski, J. Q., Jakes, R., and Goedert, M. (1997) [alpha]-Synuclein in Lewy bodies, *Nature 388*, 839-840.

[5] Mezey, E., Dehejia, A. M., Harta, G., Suchy, S. F., Nussbaum, R. L., Brownstein, M. J., and Polymeropoulos, M. H. (1998) Alpha synuclein is present in Lewy bodies in sporadic Parkinson's disease, *Molecular Psychiatry 3*, 493-499.

[6] Bendor, Jacob T., Logan, Todd P., and Edwards, Robert H. (2013) The Function of α-Synuclein, *Neuron 79*, 1044-1066.

[7] Bisaglia, M., Mammi, S., and Bubacco, L. (2009) Structural insights on physiological functions and pathological effects of α-synuclein, *The FASEB Journal 23*, 329-340.

[8] Ulmer, T. S., Bax, A., Cole, N. B., and Nussbaum, R. L. (2005) Structure and Dynamics of Micelle-bound Human α-Synuclein, *Journal of Biological Chemistry 280*, 9595-9603.

[9] Jo, E., McLaurin, J., Yip, C. M., St. George-Hyslop, P., and Fraser, P. E. (2000) α-Synuclein Membrane Interactions and Lipid Specificity, *Journal of Biological Chemistry 275*, 34328-34334.

[10] Sung, Y.-h., and Eliezer, D. (2007) Residual Structure, Backbone Dynamics, and Interactions within the Synuclein Family, *Journal of Molecular Biology 372*, 689-707.

[11] Liu, S., Ninan, I., Antonova, I., Battaglia, F., Trinchese, F., Narasanna, A., Kolodilov, N., Dauer, W., Hawkins, R. D., and Arancio, O. (2004) α‐Synuclein produces a long‐lasting increase in neurotransmitter release, The EMBO Journal *23*, 4506-4516.

[12] Burré, J., Sharma, M., and Südhof, T. C. (2014) α-Synuclein assembles into higher-order multimers upon membrane binding to promote SNARE complex formation, *Proceedings of the National Academy of Sciences USA 111*, E4274-E4283.

[13] Burré, J., Sharma, M., and Südhof, T. C. (2015) Definition of a Molecular Pathway Mediating α-Synuclein Neurotoxicity, *The Journal of Neuroscience 35*, 5221-5232.

[14] Polymeropoulos, M. H., Lavedan, C., Leroy, E., Ide, S. E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R., Stenroos, E. S., Chandrasekharappa, S., Athanassiadou, A., Papapetropoulos, T., Johnson, W. G., Lazzarini, A. M., Duvoisin, R. C., Di Iorio, G., Golbe, L. I., and Nussbaum, R. L. (1997) Mutation in the α-Synuclein Gene Identified in Families with Parkinson's Disease, *Science 276*, 2045-2047.

[15] Zarranz, J. J., Alegre, J., Gómez-Esteban, J. C., Lezcano, E., Ros, R., Ampuero, I., Vidal, L., Hoenicka, J., Rodriguez, O., Atarés, B., Llorens, V., Tortosa, E. G., del Ser, T., Muñoz, D. G., and de Yebenes, J. G. (2004) The new mutation, E46K, of α-synuclein causes parkinson and Lewy body dementia, *Annals of Neurology 55*, 164-173.

[16] Conway, K. A., Harper, J. D., and Lansbury, P. T. (1998) Accelerated in vitro fibril formation by a mutant α-synuclein linked to early-onset Parkinson disease, *Nature Medicine 4*, 1318-1320.

[17] Khalaf, O., Fauvet, B., Oueslati, A., Dikiy, I., Mahul-Mellier, A. L., Ruggeri, F. S., Mbefo, M. K., Vercruysse, F., Dietler, G., Lee, S. J., Eliezer, D., and Lashuel, H. A. (2014) The

H50Q mutation enhances alpha-synuclein aggregation, secretion, and toxicity, *Journal of Biological Chemistry 289*, 21856-21876.
[18] Bartels, T., Choi, J. G., and Selkoe, D. J. (2011) [agr]-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation, *Nature 477*, 107-110.
[19] Gould, N., Mor, D., Lightfoot, R., Malkus, K., Giasson, B., and Ischiropoulos, H. (2014) Evidence of Native α-Synuclein Conformers in the Human Brain, *Journal of Biological Chemistry*.
[20] Dettmer, U., Newman, A. J., Soldner, F., Luth, E. S., Kim, N. C., von Saucken, V. E., Sanderson, J. B., Jaenisch, R., Bartels, T., and Selkoe, D. (2015) Parkinson-causing [alpha]-synuclein missense mutations shift native tetramers to monomers as a mechanism for disease initiation, *Nature Commununications 6*.
[21] Pochapsky, T. C. (2015) From intrinsically disordered protein to context-dependent folding: The α-synuclein tetramer is teased out of hiding, *Proceedings of the National Academy of Sciences USA 112*, 9502-9503.
[22] Eliezer, D., Kutluay, E., Bussell Jr, R., and Browne, G. (2001) Conformational properties of α-synuclein in its free and lipid-associated states, *Journal of Molecular Biology 307*, 1061-1073.
[23] Davidson, W. S., Jonas, A., Clayton, D. F., and George, J. M. (1998) Stabilization of α-Synuclein Secondary Structure upon Binding to Synthetic Membranes, *Journal of Biological Chemistry 273*, 9443-9449.
[24] Lai, B. C. L., Marion, S. A., Teschke, K., and Tsui, J. K. C. (2002) Occupational and environmental risk factors for Parkinson's disease, *Parkinsonism & Related Disorders 8*, 297-309.
[25] Lee, E.-N., Lee, S.-Y., Lee, D., Kim, J., and Paik, S. R. (2003) Lipid interaction of α-synuclein during the metal-catalyzed oxidation in the presence of Cu^{2+} and H_2O_2, *Journal of Neurochemistry 84*, 1128-1142.
[26] Uversky, V. N., Li, J., and Fink, A. L. (2001) Evidence for a partially folded intermediate in alpha-synuclein fibril formation, *Journal of Biological Chemistry 276*, 10737-10744.
[27] Gaeta, A., and Hider, R. C. (2005) The crucial role of metal ions in neurodegeneration: the basis for a promising therapeutic strategy, *British Journal of Pharmacology 146*, 1041-1059.
[28] Yamin, G., Glaser, C. B., Uversky, V. N., and Fink, A. L. (2003) Certain metals trigger fibrillation of methionine-oxidized alpha-synuclein, *Journal of Biological Chemistry 278*, 27630-27635.
[29] Uversky, V. N., Li, J., and Fink, A. L. (2001) Metal-triggered structural transformations, aggregation, and fibrillation of human alpha-synuclein. A possible molecular NK between Parkinson's disease and heavy metal exposure, *Journal of Biological Chemistry 276*, 44284-44296.
[30] Allsop, D., Mayes, J., Moore, S., Masad, A., and Tabner, B. J. (2008) Metal-dependent generation of reactive oxygen species from amyloid proteins implicated in neurodegenerative disease, *Biochemical Society Transactions 36*, 1293-1298.
[31] Santner, A., and Uversky, V. N. (2010) Metalloproteomics and metal toxicology of alpha-synuclein, *Metallomics: Integrated Biometal Science 2*, 378-392.
[32] Bharathi, and Rao, K. S. (2007) Thermodynamics imprinting reveals differential binding of metals to alpha-synuclein: relevance to Parkinson's disease, *Biochemical and Biophysical Research Communications 359*, 115-120.
[33] Bharathi, and Rao, K. S. (2008) Molecular understanding of copper and iron interaction with alpha-synuclein by fluorescence analysis, *Journal of Molecular Neuroscience 35*, 273-281.

[34] Lee, J. C., Gray, H. B., and Winkler, J. R. (2008) Copper(II) Binding to α-Synuclein, the Parkinson's Protein, *Journal of the American Chemical Society 130*, 6898-6899.
[35] Jackson, M. S., and Lee, J. C. (2009) Identification of the minimal copper(II)-binding alpha-synuclein sequence, *Inorganic Chemistry 48*, 9303-9307.
[36] Dudzik, C. G., Walter, E. D., and Millhauser, G. L. (2011) Coordination features and affinity of the Cu(2)+ site in the alpha-synuclein protein of Parkinson's disease, *Biochemistry 50*, 1771-1777.
[37] Drew, S. C., Ling Leong, S., Pham, C. L. L., Tew, D. J., Masters, C. L., Miles, L. A., Cappai, R., and Barnham, K. J. (2008) Cu^{2+} Binding Modes of Recombinant α-Synuclein − Insights from EPR Spectroscopy, *Journal of the American Chemical Society 130*, 7766-7773.
[38] Ahmad, A., Burns, C. S., Fink, A. L., and Uversky, V. N. (2012) Peculiarities of copper binding to alpha-synuclein, *Journal of Biomolecular Structure & Dynamics 29*, 825-842.
[39] Rasia, R. M., Bertoncini, C. W., Marsh, D., Hoyer, W., Cherny, D., Zweckstetter, M., Griesinger, C., Jovin, T. M., and Fernandez, C. O. (2005) Structural characterization of copper(II) binding to alpha-synuclein: Insights into the bioinorganic chemistry of Parkinson's disease, *Proceeding of the National Academy of Science USA 102*, 4294-4299.
[40] Moriarty, G. M., Minetti, C. A., Remeta, D. P., and Baum, J. (2014) A revised picture of the Cu(II)-alpha-synuclein complex: the role of N-terminal acetylation, *Biochemistry 53*, 2815-2817.
[41] Sung, Y. H., Rospigliosi, C., and Eliezer, D. (2006) NMR mapping of copper binding sites in alpha-synuclein, *Biochimica et Biophysica Acta 1764*, 5-12.
[42] Binolfi, A., Lamberto, G. R., Duran, R., Quintanar, L., Bertoncini, C. W., Souza, J. M., Cerveñansky, C., Zweckstetter, M., Griesinger, C., and Fernández, C. O. (2008) Site-Specific Interactions of Cu(II) with α and β-Synuclein: Bridging the Molecular Gap between Metal Binding and Aggregation, *Journal of the American Chemical Society 130*, 11801-11812.
[43] Paik, S. R., Shin, H. J., Lee, J. H., Chang, C. S., and Kim, J. (1999) Copper(II)-induced self-oligomerization of alpha-synuclein, *Biochemical Journal 340*, 821-828.
[44] Nielsen, M. S., Vorum, H., Lindersson, E., and Jensen, P. H. (2001) Ca2+ binding to alpha-synuclein regulates ligand binding and oligomerization, *Journal of Biological Chemistry 276*, 22680-22684.
[45] Binolfi, A., Rasia, R. M., Bertoncini, C. W., Ceolin, M., Zweckstetter, M., Griesinger, C., Jovin, T. M., and Fernández, C. O. (2006) Interaction of α-Synuclein with Divalent Metal Ions Reveals Key Differences: A Link between Structure, Binding Specificity and Fibrillation Enhancement, *Journal of the American Chemical Society 128*, 9893-9901.
[46] Loo, J. A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry, *Mass Spectrometry Reviews 16*, 1-23.
[47] Li, H., Wolff, J. J., Van Orden, S. L., and Loo, J. A. (2014) Native Top-Down Electrospray Ionization-Mass Spectrometry of 158 kDa Protein Complex by High-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, *Analytical Chemistry 86*, 317-320.
[48] Xie, Y., Zhang, J., Yin, S., and Loo, J. A. (2006) Top-Down ESI-ECD-FT-ICR Mass Spectrometry Localizes Noncovalent Protein-Ligand Binding Sites, *Journal of the American Chemical Society 128*, 14432-14433.
[49] Loo, R. R. O., Goodlett, D. R., Smith, R. D., and Loo, J. A. (1993) Observation of a noncovalent ribonuclease S-protein/S-peptide complex by electrospray ionization mass spectrometry, *Journal of the American Chemical Society 115*, 4391-4392.

[50] Goodlelt, D., Ogorzalek Loo, R., Loo, J., Wahl, J., Udseth, H., and Smith, R. (1994) A study of the thermal denaturation of ribonuclease S by electrospray ionization mass spectrometry, *Journal of the American Society for Mass Spectrometry 5*, 614-622.
[51] Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics, *Nature 422*, 198-207.
[52] Chalmers, M. J., Busby, S. A., Pascal, B. D., He, Y., Hendrickson, C. L., Marshall, A. G., and Griffin, P. R. (2006) Probing Protein Ligand Interactions by Automated Hydrogen/Deuterium Exchange Mass Spectrometry, *Analytical Chemistry 78*, 1005-1014.
[53] Wales, T. E., and Engen, J. R. (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics, *Mass Spectrometry Reviews 25*, 158-170.
[54] Konermann, L., Pan, J., and Liu, Y.-H. (2011) Hydrogen exchange mass spectrometry for studying protein structure and dynamics, *Chemical Society Reviews 40*, 1224-1234.
[55] Kelleher, N. L., Lin, H. Y., Valaskovic, G. A., Aaserud, D. J., Fridriksson, E. K., and McLafferty, F. W. (1999) Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry, *Journal of the American Chemical Society 121*, 806-812.
[56] Zubarev, R. A. (2004) Electron-capture dissociation tandem mass spectrometry, *Current Opinion in Biotechnology 15*, 12-16.
[57] Loo, J., Hu, P., and Smith, R. (1994) Interaction of angiotensin peptides and zinc metal ions probed by electrospray ionization mass spectrometry, *Journal of the American Society for Mass Spectrometry 5*, 959-965.
[58] Hu, P., Ye, Q.-Z., and Loo, J. A. (1994) Calcium Stoichiometry Determination for Calcium Binding Proteins by Electrospray Ionization Mass Spectrometry, *Analytical Chemistry 66*, 4190-4194.
[59] Hu, P., Sorensen, C., and Gross, M. (1995) Influences of peptide side chains on the metal ion binding site in metal ion-cationized peptides: Participation of aromatic rings in metal chelation, *Journal of the American Society for Mass Spectrometry 6*, 1079-1085.
[60] Hu, P., and Gross, M. L. (1992) Strong interactions of anionic peptides and alkaline earth metal ions: bis(peptide) complexes in the gas phase, *Journal of the American Chemical Society 114*, 9161-9169.
[61] Lu, Y., Prudent, M., Qiao, L., Mendez, M. A., and Girault, H. H. (2010) Copper(I) and copper(II) binding to β-amyloid 16 (Aβ16) studied by electrospray ionization mass spectrometry, *Metallomics: Integrated Biometal Science 2*, 474-479.
[62] Acharya, S., Safaie, B. M., Wongkongkathep, P., Ivanova, M. I., Attar, A., Klärner, F.-G., Schrader, T., Loo, J. A., Bitan, G., and Lapidus, L. J. (2014) Molecular Basis for Preventing α-Synuclein Aggregation by a Molecular Tweezer, *Journal of Biological Chemistry 289*, 10727-10737.
[63] Li, H., Wongkongkathep, P., Van Orden, S., Ogorzalek Loo, R., and Loo, J. (2014) Revealing Ligand Binding Sites and Quantifying Subunit Variants of Noncovalent Protein Complexes in a Single Native Top-Down FTICR MS Experiment, *Journal of the American Society for Mass Spectrometry 25*, 2060-2068.
[64] Jackson, S. N., Moyer, S. C., and Woods, A. S. (2008) The role of phosphorylated residues in peptide-peptide noncovalent complexes formation, *Journal of the American Society for Mass Spectrometry 19*, 1535-1541.
[65] Woods, A. S., and Ferré, S. (2005) Amazing Stability of the Arginine−Phosphate Electrostatic Interaction, *Journal of Proteome Research 4*, 1397-1402.

[66] Yin, S., Xie, Y., and Loo, J. A. (2008) Mass Spectrometry of Protein–Ligand Complexes: Enhanced Gas-Phase Stability of Ribonuclease–Nucleotide Complexes, *Journal of the American Society for Mass Spectrometry 19*, 1199-1208.
[67] Yin, S., and Loo, J. A. (2010) Elucidating the Site of Protein-ATP Binding by Top-Down Mass Spectrometry, *Journal of the American Society for Mass Spectrometry 21*, 899-907.
[68] Yin, S., and Loo, J. A. (2011) Top-down mass spectrometry of supercharged native protein–ligand complexes, *International Journal of Mass Spectrometry 300*, 118-122.
[69] Uversky, V. N. (2003) A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders, *Journal of Biomolecular Structure & Dynamics 21*, 211-234.
[70] Bernstein, S. L., Liu, D., Wyttenbach, T., Bowers, M. T., Lee, J. C., Gray, H. B., and Winkler, J. R. (2004) Alpha-synuclein: stable compact and extended monomeric structures and pH dependence of dimer formation, *Journal of the American Society for Mass Spectrometry 15*, 1435-1443.
[71] Chanthamontri, C., Liu, J., and McLuckey, S. A. (2009) Charge state dependent fragmentation of gaseous α-synuclein cations via ion trap and beam-type collisional activation, *International Journal of Mass Spectrometry 283*, 9-16.
[72] Schwartz, B. L., and Bursey, M. M. (1992) Some proline substituent effects in the tandem mass spectrum of protonated pentaalanine, *Biological Mass Spectrometry 21*, 92-96.
[73] Yu, W., Vath, J. E., Huberty, M. C., and Martin, S. A. (1993) Identification of the facile gas-phase cleavage of the Asp-Pro and Asp-Xxx peptide bonds in matrix-assisted laser desorption time-of-flight mass spectrometry, *Analytical Chemistry 65*, 3015-3023.
[74] Andrés, B., and Claudio, O. F. (2012) Interactions of α-Synuclein with Metal Ions, In *Brain Diseases and Metalloproteins*, pp 327-366, Pan Stanford Publishing.

CHAPTER THREE

Structural Elucidation of Molecular Tweezer Binding to Neurodegenerative Disease Proteins

Abstract

Two amyloidogenic proteins were studied by electrospray ionization mass spectrometry (ESI-MS). The 140-residue protein, α-synuclein (AS), which is known to be involved in Parkinson's disease, and Cu/Zn superoxide dismutase (SOD1), which can form oligomeric species and is implicated in amyotrophic lateral sclerosis (ALS), were characterized by ESI-MS. A chemically synthesized compound called molecular tweezer (MT or CLR01) has been shown to disrupt aggregation and also disaggregate preformed amyloid fibril by binding to lysine residues. In this chapter, top-down mass spectrometry using physiological pH analyte solutions was used to characterize binding of MT to AS and SOD1. Electron capture dissociation was utilized to generate ligand-bound fragments, which is helpful to locate the MT binding sites. We revealed that MT has strong binding affinities to both proteins. MT binds to the K10 or K12 residues on AS. Two lysines on the zinc binding loop (K71 and K75) on SOD1 are possible binding residues of MT. The study demonstrates that ESI-MS has utility for characterizing the structures of amyloid proteins, which can be difficult using X-ray crystallography or NMR. The ligand binding data could lead to a better understanding of the mechanisms underlying amyloid fibrillation and offer additional insight to potential therapeutic approaches.

Introduction

Some peptides or proteins in their soluble forms can undergo an aberrant pathway under some conditions and be converted into higher order structured oligomers and eventually amyloid fibrils that can be deposited in extracellular tissues. These intermediate states may be toxic or lead to phathological disease conditions, which is usually referred to as protein misfolding diseases[1]. Some of them may cause neurodegenerative disorders, systemic amyloidosis, or intracellular inclusions. Many misfolded proteins contribute to similar biological functions and share propensities to aggregate. In some cases post-translational modifications may promote self-association into oligomers, such as tau[2, 3]. Many misfolded proteins are natively unstructured[1], which make structural characterization by X-ray crystallography or NMR difficult and limiting. As described in Chapter 2, α-synuclein (AS) protein is implicated in Parkinson's disease. It has been shown that AS fibrils deposit in neuronal tissue in the form of Lewy bodies[4]. This pathogenic condition is termed synucleinopathies. To date, a crystal structure of full-length AS is unknown due to the difficulties in protein crystallization. Recently, there was an effort to crystalize a small section of AS from the non-amyloid component (NAC) region[5]. This NAC fragment was shown to form insoluble intermediates and also form highly toxic fibrils. A crystal structural model revealed how NAC fragments can form a beta-sheet rich structure that leads to aggregation, suggesting that the NAC region may give rise to an initial conformation that can lead to toxic oligomers.

There have been numerous studies trying to seek new potential therapeutic approaches to prevent and treat neurodegenerative diseases. Some studies reported the use of molecular chaperones for neuroprotection induced by several mechanisms[6]. There are efforts to identify potent small molecules and develop them as preventative and intervention therapeutics. Several

compounds have been tested to disrupt protein aggregration[7-10]. Sinha *et al.* introduced a small molecule called molecular tweezer (MT or CLR01), which contains a series of nine six-membered and aromatic rings and two negatively charged phosphate groups. MT has a high affinity to lysine and arginine residues because of its specificity derived from a hydrophobic cleft and its ability to form a salt bridge with a positively charged amine group. The structure of MT is shown in Figure 3-2 showing it in a locking formation with lysine. Remarkably, MT can bind to lysine approximately 2.5 times stronger than arginine[11]. It has been shown that MT is effectively able to disrupt amyloid oligomerization and also disaggregate pre-formed amyloid fibrils of many different amyloid proteins, including amyloid-β (Aβ), α-synuclein, tau, IAPP, and β-microglobulin[10, 12]. The capability of MT to modulate or inhibit Aβ oligomerization is comparable to known inhibitors such as epigallocatechin gallate (EGCG)[13, 14]. Some potential drugs are strong inhibitors to amyloid assembly but they fail to the cross blood-brain barrier to target the disease progression in the brain. Interestingly, many studies showed that MT is able to cross blood-brain barrier as tested in vivo[15, 16].

Cu/Zn superoxide dismutase (CuZnSOD or SOD1) is an abundant enzyme in the cytosol, nucleus, mitochondria, and peroxisomes of many human cell types. The 153-residue SOD1 is natively found in a dimeric state when copper and zinc are bound and a disulfide bond is present. Both metal binding sites are in close proximity (Figure 3-1)[17]. Copper sits in the main catalytic site that convert superoxide into oxygen and hydrogen peroxide via a two-step process. The metal-loaded enzyme with a disulfide bone is very stable at high temperature up to 90°C. SOD1 mutations have been discovered that are linked with a familial form of amyotrophic lateral sclerosis (fALS) disease. ALS, also known as Lou Gehrig's disease, affects motor neuron dysfunction in the spinal chord and brain. Most ALS cases are in the form of sporadic ALS (sALS) with no genetic component, where its cause is currently unknown. Familial ALS only

contributes to 5-10% of the cases[18]. Insoluble aggregates of SOD1 has been found in spinal cords of transgenic mice, especially the one with the A4V mutation, which is the most common disease form[19]. The SOD1 structure and known pathologically related mutations are shown in Figure 3-1. Because a disulfide bond is believed to contribute to the high stability of SOD1, it has been shown that the reduced form of apo-wild-type SOD1 is more susceptible to self-aggregation even in physiological conditions[20].

To study protein-protein or protein-ligand interactions, aside from techniques that provide a complete structural view like x-ray crystallography, there are several conventional biophysical methods available, such as surface plasmon resonance, or isothermal colorimetry. Mass spectrometry also has capabilities to characterize protein complexes because it provides a direct measurement of mass, which can be used to derive molecular mass and stoichiometry. Soft conditions must be used during the ionization and ion transfer process to reduce the chances for analyte fragmentation. Generally, electrospray ionization (ESI) is preferred with physiological solution conditions (ca. pH 7 and room temperature)[21, 22]. Our lab has been utilizing a top-down MS approach to break apart protein backbone bonds, while keeping small molecular weight ligands intact[23-25]. Electron capture dissociation[26] is a low energy approach that our lab uses to generate peptide fragments that remain bound to the ligand. Measuring fragment masses has allowed us to confidentially identify ligand-bound and ligand-free fragments, which is ultimately used to pinpoint ligand-binding sites.

In this chapter, we characterize the binding of molecular tweezer binding to α-synuclein and human SOD1 using native electrospray ionization and top-down MS fragmentation analysis. Molecular tweezer reduces the tendency of AS and SOD1 to form aggregates. The structural details of tweezer binding to these proteins are provided by the MS data. Binding stoichiometry,

dissociation constants, and possible binding sites were deduced. We used a recombinant human α-synuclein and apo-wild-type SOD1 for the experiments. Originally wild-type human SOD1 has four cysteins, where two of them form an intramolecular disulfide bond between C57-C146[17]. In our experiments, a special mutated version of apo-SOD1, which has methionine at the N-terminus and four cysteine mutations (C7A, C58S C112S, and C147S) was used for all experiments. The specific cysteine free mutation (noCys SOD1) intrinsically does not contain a disulfide bond. It represents a reduced form of SOD1 that leads to an aggregation-prone conformation[20].

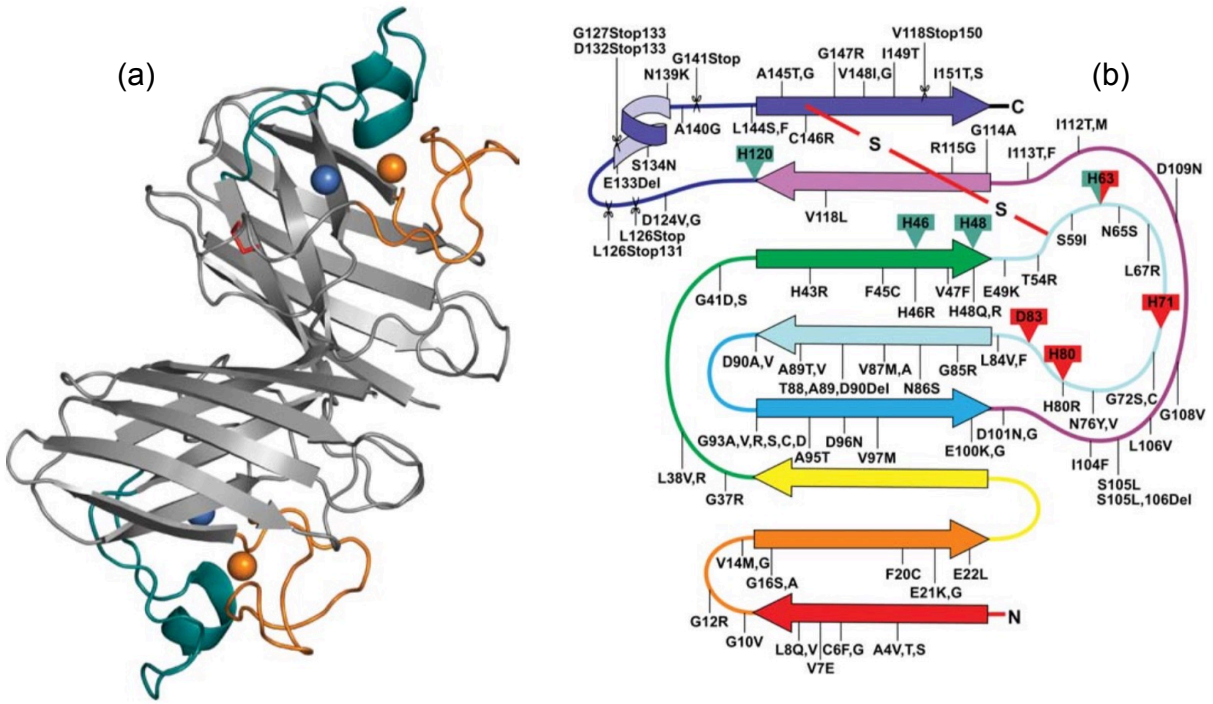

Figure 3-1. (a) A crystal structure of human Cu/Zn SOD1 showing a dimeric state when two heavy metal ions are bound. Copper ion is in blue and zinc is shown in orange. Zinc binding loops are highlighted in orange. (b) Structural representation of SOD1 showing secondary structures and a disulfide bond in the sequence. Familial ALS associated mutations are labeled. (Reprinted with permission Annual Review of Biochemistry - Valentine, J.S. *et al* (2005)[17]).

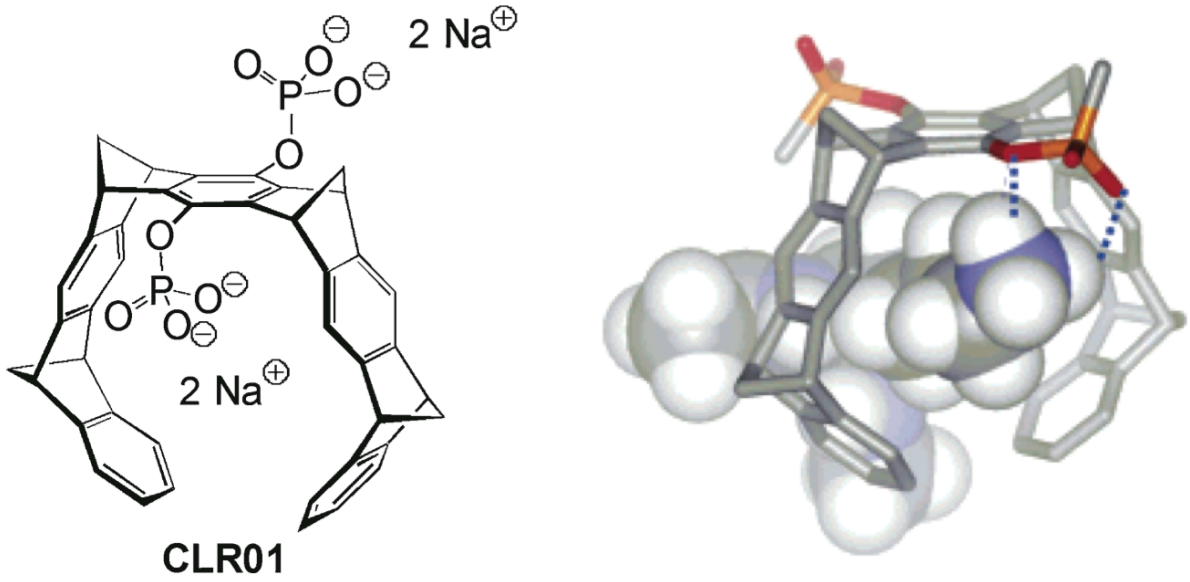

Figure 3-2. A molecular structure of MT (left) and a schematic representation showing the specificity of MT to a lysine residue.

Experimental

Alpha-synuclein was purchased from rPeptide (Athens, GA). Human SOD1 proteins were provided by Dr. Joan Valentine's lab (UCLA) in their apo-wild-type and apo-cysteine-free (NoCys) forms. Molecular tweezer (MT or CLR01) was synthesized in Germany in collaboration with Gal Bitan's lab (UCLA). Ammonium acetate (Sigma Aldrich, St. Louis, MO) buffer solution (pH 6.7) was prepared at 20 mM concentration and used for buffer exchange and desalting using 10 kDa molecular weight cut-off (MWCO) centrifugal filter devices. Protein solutions were then prepared at a concentration of 5-10 μM for analysis. Samples were loaded into a borosilicate glass capillary coated with Au/Pd (Proxeon) for nanoelectrospray. The analyte solution flow rate was in a range of 50-100 nL/min.

Electrospray ionization was performed using two different mass spectrometers. Ligand titration experiments to measure their binding affinities were performed using a Waters Synapt G1 qTOF mass spectrometer (Milford, MA) and Bruker Solarix 15-Tesla FT-ICR mass spectrometer (Billerica, MA), equipped with an Infinity ICR cell. Both instruments used the same type of nanoelectrospray capillary setups. For the Synapt system, the ESI voltage was set to 1.4 kV. Sampling cone and extraction cone voltages were tuned to 40 and 80V, respectively. In this instrument setup, nitrogen gas was used to supply a backing pressure of 0.02 bar to assist the nanoESI analyte solution flow. On the FT-ICR instrument, the ESI voltage was set to 1.1 kV. Dry gas was applied at 0.5 L/min at 160 ^{o}C in the ESI source region to aid desolvation. Transfer ion optics were tuned to the following settings: capillary 200V, deflector plate at 180V, funnel 120V, skimmer 40V.

Top-down MS experiments of protein-MT complexes were performed with the Bruker Solarix 15-Tesla FT-ICR MS. The ECD cell was tuned for top-down MS analysis using a 10ms electron pulse, a 1V bias, and 15V applied to the lens. Product ions were manually identified by the Bruker DataAnalysis software and assigned by a home-built Matlab program using a 10ppm or less mass accuracy tolerance.

Results and Discussion

Characterization of MT binding

Alpha-synuclein (AS) and SOD1 have very different structural conformations and features. AS natively unfolded at physiological pH; it has a long unstructured tail near the C-terminus, while the N-terminal region adopts mostly an alpha-helical structure. On the other hand, SOD1 has a compact structure containing two β-sheets, metal binding loops, and one disulfide bond. It forms a dimer under native solution conditions (with an oxidized disulfide bond and bound metals). Without the metal ions, apo-SOD1 shows a monomeric structure. Native ESI-MS spectra of α-synuclein were similar to the mass spectra shown in Chapter 2 (Figure 2-2), where the charge state distribution is shifted towards relatively high charge states (19+ to 9+). The highest intensity charge states are 14+ at *m/z* 1033.2 and 13+ at *m/z* 1112.6 (Figure 3-3). The observation of relatively high charge states is likely an indication of an unfolded structure of the protein in the gas phase[27].

The cysteine free mutant of SOD1 (noCys SOD1) was used for all experiments in this chapter because the disulfide bond was eliminated. NoCys SOD1 showed a relatively lower charge

state distribution, ranging from 8+ to 5+, implying that the protein remained in a more compact structure in the gas phase (Figure 3-5).

To determine the binding affinities of the MT ligand to the proteins, ligand titration experiments were performed in the protein concentrations are fixed at 10 μM, while the MT concentrations are varied from 1:0.5 to a 1:2 protein/ligand ratio for AS and a 1:3 ratio for SOD1. Figure 3-3 shows native ESI mass spectra of AS at different concentration ratios of MT. Mass spectra are displayed between *m/z* 800-1400. The 1:1 protein-ligand stoichiometry complexes were observed with 5 μM of MT in solution (1:0.5 concentration ratio). The second binding event, i.e., a 1:2 protein-ligand stoichiometry complex, appears at 7.5 μM of MT (1:0.75 molar ratio). At higher concentrations of MT, multiple binding up to 5 ligands was observed. In Figure 3-3, the apo-AS species are labeled with open circles, while the ligand-bound species are labeled with filled colored circles. The apo-protein was barely observed at a 1:2 concentration ratio due to an equilibrium shift toward the multiple ligand-bound forms.

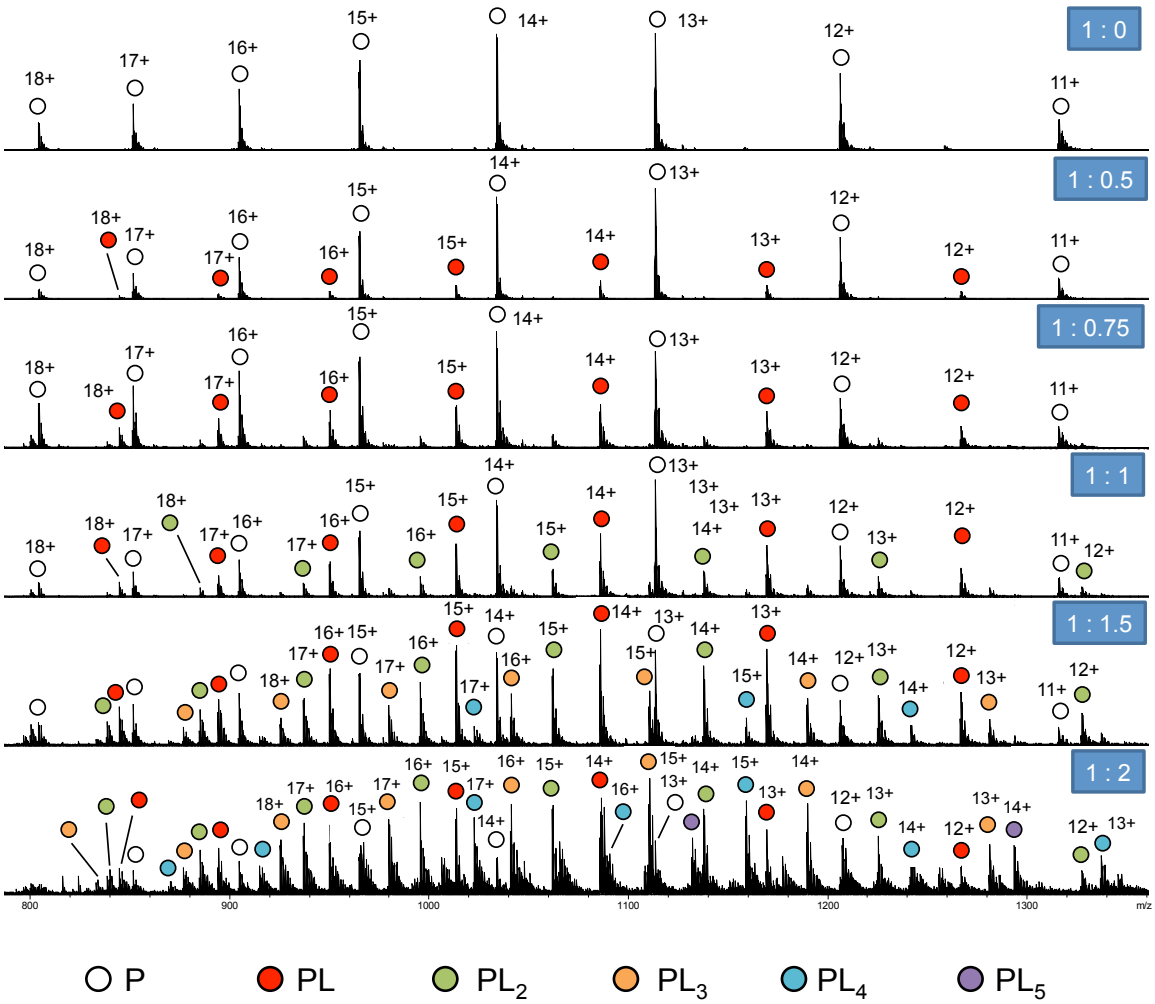

Figure 3-3. Native ESI-MS spectra of AS-MT complexes at different protein/ligand concentration ratios (apo, 1:0.5, 1:0.75, 1:1, 1:1.5, and 1:2) showing various species of protein-ligand complexes, where P is protein (AS) and L is ligand (MT).

By measuring the normalized signal intensities for each protein form against the different charge states, equilibrium dissociation constants can be calculated using a Scatchard plot. For cases in which there are multiple binding sites, the following holds:

$$P + L \rightleftharpoons PL \qquad (3a)$$

$$PL + L \rightleftharpoons PL_2 \qquad (3b)$$

$$PL_{n-1} + L \rightleftharpoons PL_n \qquad (3c)$$

Where *P* is protein and *L* is ligand. A total summed concentration of bound and free ligand must equal to the initial concentration of ligand $[L]_o$

$$[L]_0 = [L]_{free} + [PL] + [PL_2] + [PL_3] + \cdots \qquad (3d)$$

With the method described previously [28, 29], the relationship can be derived as:

$$\frac{[P]+\sum_{i=1}^{n}[PL_i]}{[P]} = \left(1 + \frac{1}{K_{D_1}} + \frac{1}{K_{D_1}\cdot K_{D_2}}\cdot + \cdots + \frac{1}{K_{D_1}\cdot K_{D_2}\cdot\cdots\cdot K_{D_n}}\right) - \frac{1}{K_{D_1}}\left([L]_0 - \sum_{i=1}^{n}[PL_i]\right)\cdot + \frac{1}{K_{D_1}\cdot K_{D_2}}\cdot$$
$$\left(-([L]_0 - \sum_{i=1}^{n}[PL_i])\right)^2 + \cdots + \frac{1}{K_{D_1}\cdot K_{D_2}\cdot\cdots\cdot K_{D_n}}\cdot\left((-1)^n([L]_0 - \sum_{i=1}^{n}[PL_i])^n\right) \qquad (3e)$$

A plot between the total free ligand or $([L]_0 - \sum_{i=1}^{n}[PL_i])$ versus the concentration ratio of total protein over free protein $\left(\frac{[P]+\sum_{i=1}^{n}[PL_i]}{[P]}\right)$ was generated as shown in Figure 3-4. From equation 3e, the first and second dissociation constants from this experiment were calculated (from 4th order polynomial fit) as $K_{D_1} = 1.8\ \mu\text{M}$ and $K_{D_2} = 4.1\ \mu\text{M}$ (Figure 3-4).

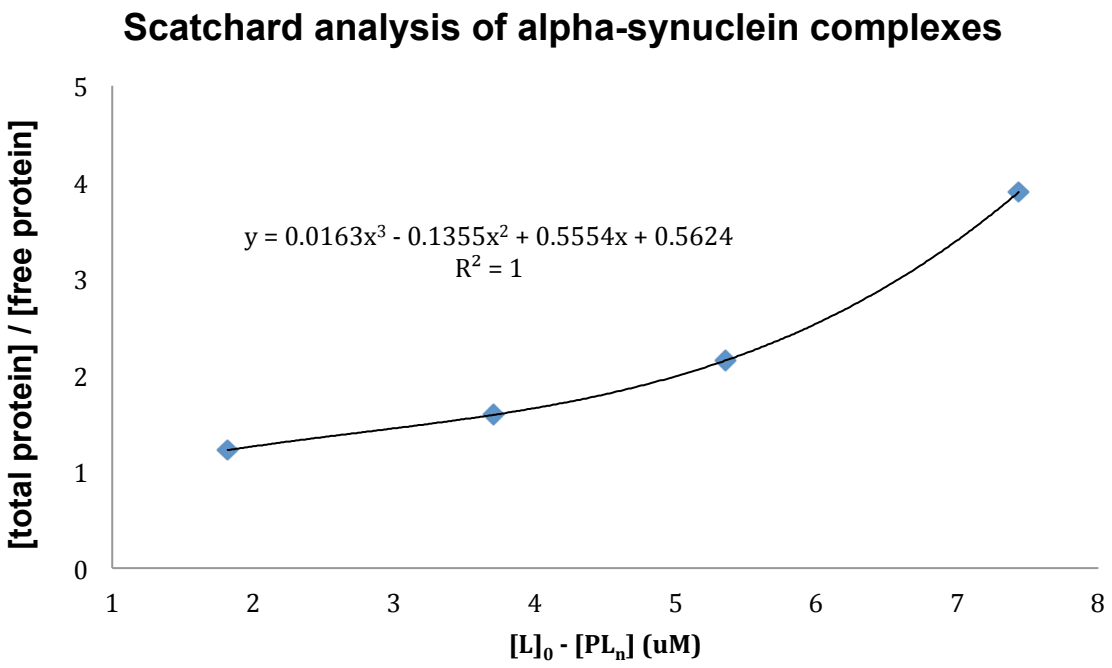

Figure 3-4. Scatchard plot of AS-MT complex was generated by varying the concentration of ligand (MT) and measuring the protein-MT complex by ESI mass spectrometry. A polynomial fit was employed and coefficients were used to calculate the first and second dissociation constants as shown in Equation 3e.

MT binding to human SOD1 was characterized using the same approach as α-synuclein. ESI mass spectra of SOD1 in its native state showed four charge states ranging from 8+ to 5+ (Figure 3-5). An apparent nonspecifically-bound dimer was observed at *m/z* 2884.3 (11+). MT was added to a protein solution to yield concentrations ranging from 5-30 μM (1:0.5 to 1:3 protein/ligand concentration ratio). Small peaks of single MT bound were readily observed when for the low concentration of MT at 5 μM. Interestingly as the concentration ratio increased, multiple ligand bound species were detected with up to 3 ligands bound at a 30 μM MT concentration. The maximum number of ligand bound was less than that observed for AS, despite the larger size for SOD1; we observed up to five ligand binding in AS. The number of ligand binding can be correlated with its structure. SOD1 would bind less MT than AS because SOD1 has a more compact structure, while AS is natively unstructured.

A dissociation constant of SOD1-MT complexes was determined in the same manner as for α-synuclein. A 4th order polynomial fit of equation 3e was applied to generate a Scatchard plot for binding of MT to SOD1 (Figure 3-6). The first and second dissociation constant was measured to be 2.93 μM and 44.9 μM, respectively. The r-squared (coefficient of multiple determination for multiple regression) value is reported to be 0.9933, which indicates a good polynomial fitting. The measured K_d of the AS-MT and SOD1-MT complexes are within the same order of magnitude, suggesting that MT binds to the amyloid proteins with a similar preference and behavior.

From these results, it clearly shows that ESI-MS can provide a comprehensive view of ligand binding characteristics that are complimentary to more conventional techniques. For example, molecular mass, stoichiometry, and binding affinity can be directly measured in a single experiment. The first K_d measured for AS-MT by ESI-MS is in good agreement with

measurements using fluorescence spectroscopy techniques[30]. MT binding to lysine residues on proteins has a higher affinity than previously reported for a small molecule containing a single lysine residue[11].

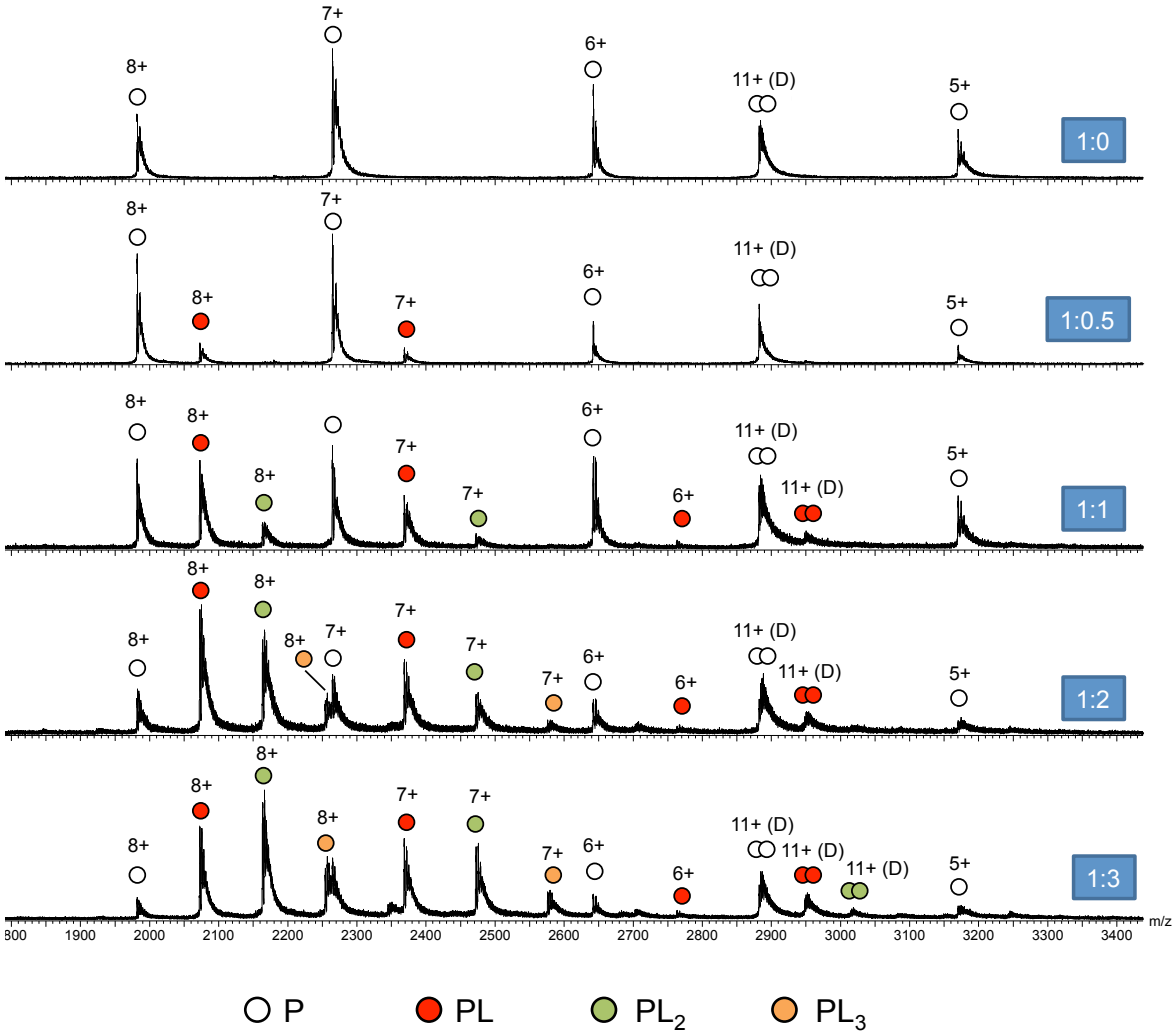

Figure 3-5. Native ESI mass spectra of protein complexes between noCys SOD1 and molecular tweezer (MT) at different protein/ligand concentration ratios, ranging from 1:0 to 1:3 (0-30 μM MT). MT-SOD1 complexes up to 3 ligand bounds were observed at the 1:3 concentration ratio.

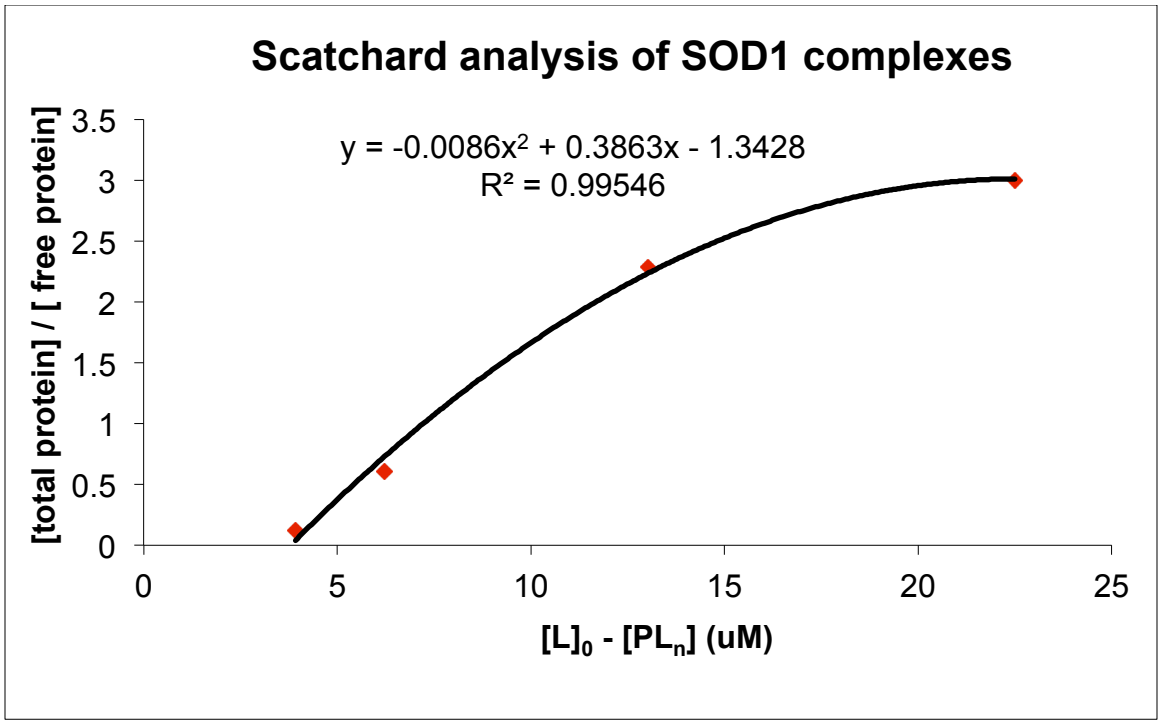

Figure 3-6. Determination of first and second dissociation constants K_{D_1} and K_{D_2} for MT binding to SOD1 by Scatchard analysis as shown in Equation 3e.

Top-down mass spectrometry for identification of MT binding sites

Previously we used electron capture dissociation (ECD) for top-down MS (Chapter 2); ECD is a relatively gentle technique to generate N-terminal (c) and C-terminal ($z^{\cdot}$) fragments. Using ECD, covalent backbone bonds of the polypeptide are cleaved while non-covalent forces holding the ligand bound to the macromolecule are maintained. Previously, this approach was used to study relatively weak solution binding of small molecule ligands to α-synuclein[23]. The sequence of these ligand-bound fragments can be mapped onto the full-length protein sequence to determine the ligand binding region(s)[23-25]. Also, this strategy was used previously to measure the binding of MT to amyloid β-protein[10] and therefore we applied ESI-MS/MS with ECD here to determine the sites of MT binding to AS and SOD1.

A 1:1 concentration ratio of AS and MT was used. The MS signal intensity decreased significantly by 3 times compared to when no MT was added to both proteins tested. The result suggested that using a high concentration of MT suppress the electrospray ionization process. For this reason, the MT was kept at an optimal concentration of 10 μM for top-down MS experiments of protein-ligand complexes of AS and SOD1.

Once protein complexes were ionized at the ESI source and gently transferred into the mass spectrometer, the 13+-charged ions for the 1:1 –α-synuclein-MT complex were isolated and subjected to ECD-MS/MS fragmentation. During the electron capture events, charge reduction occurred, yielding two more charge states (12+ and 11+). Protein fragments of the c- and $z^{\cdot}$-product-ion series were observed around the precursor *m/z* (Figure 3-7). Fragments were identified against theoretical fragments calculated from the Protein Prospector software using an in-house Matlab program and re-validated manually. Some product ions from dissociation of the polypeptide backbone corresponded to fragments showing no bound ligand (Figure 3-8, black

lines), whereas other product ions retained binding to MT (red lines). Figure 3-9 is a magnified ECD-MS/MS spectrum showing some representative MT-bound fragments, such as c_{24}+MT, c_{46}+MT, c_{57}+MT, and c_{68}+MT. At some positions along the polypeptide chain, product ions were observed in both MT-bound and unbound states. An additional voltage was applied to the skimmer region to assist ECD cleavage by increase the ions' internal energy. Observing both apo- and holo-fragments for the same residue may suggest that the MT molecule may partially fall off the protein during ion transfer or from ECD process due to the additional ion activation. Compared with our previous study with amyloid β-protein (Aβ)[10], ECD appears to induce more dissociation of MT from the protein (α-synuclein). Nonetheless, the specific observation of a MT-bound c_{20}-product ion and a MT-bound z_{131}-product ion suggests that the binding is located in the AS region spanning residues 10–20, most likely at Lys-10 and/or Lys-12.

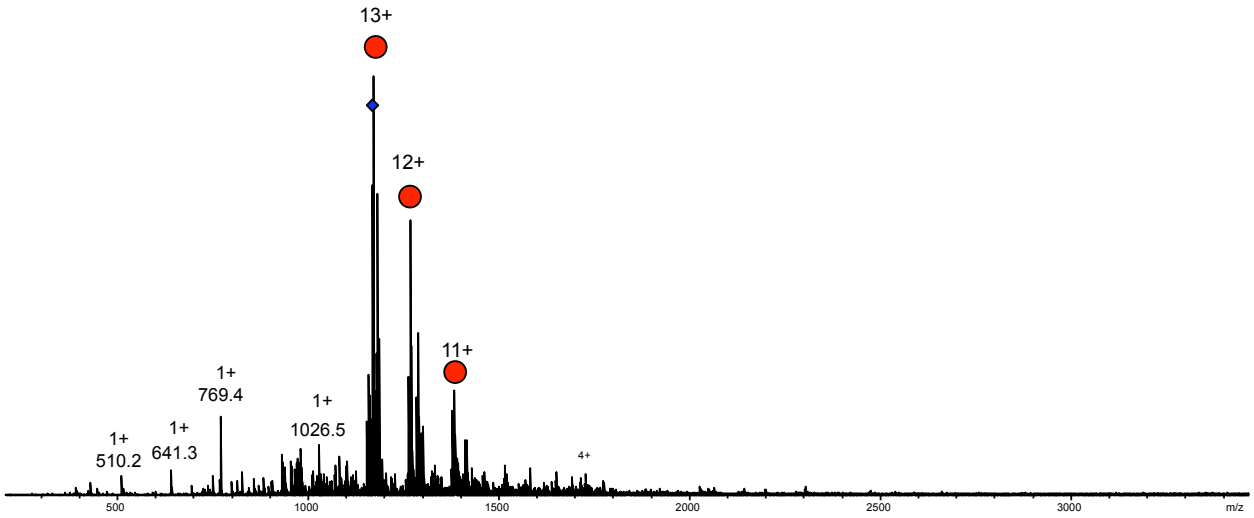

Figure 3-7. ECD-MS/MS spectrum of the 13+ ion for the 1:1 AS-MT protein-ligand complex. Charge reduced species (12+ and 11+ charged) and c/z• fragments were observed.

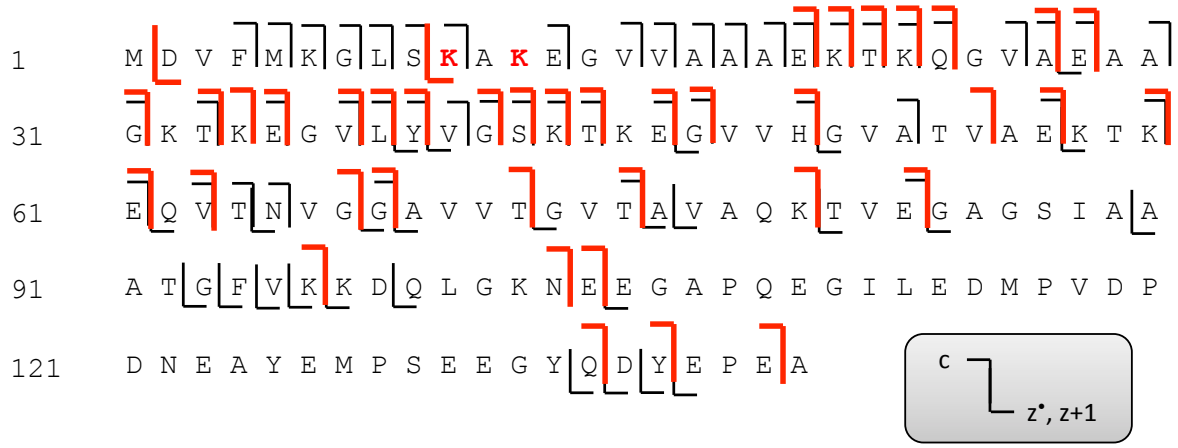

Figure 3-8. A schematic of the ECD-MS/MS fragmentation profile for the 13+-charged 1:1 -AS:MT complex on the protein sequence. Protein fragments of the c- and z•-product-ion series were observed. Some product ions from dissociation of the polypeptide backbone corresponded to unbound peptide (black line), whereas other product ions retained binding to MT (red line). At some positions along the peptide chain, product ions were observed in both MT-bound and unbound states. From the fragmentation profile, residues 10 –20 that include Lys-10 and Lys-12 are suggested to be the site(s) of MT binding.

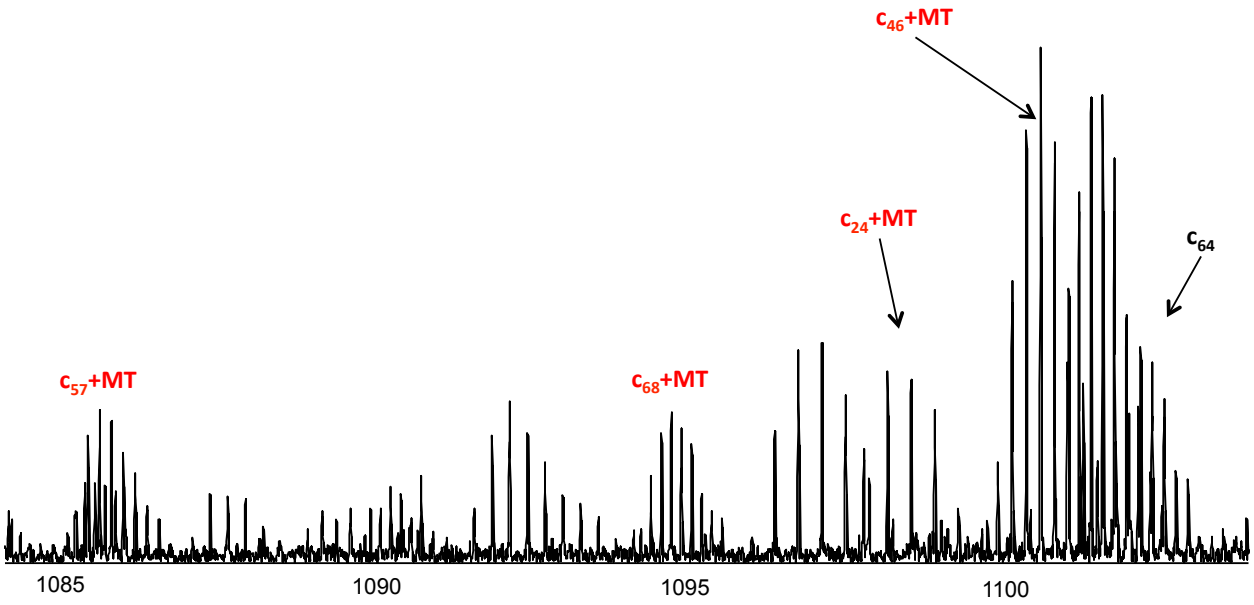

Figure 3-9. A representative MS/MS spectrum showing ECD fragments, containing four MT-bound peptide fragments (red) and an apo-fragment (black).

For MT binding to SOD1, the top-down experiments were performed using the Bruker FT-ICR mass spectrometer, using the similar approach as described for AS-MT. The ultimate goal is to locate and confirm MT binding site(s) on noCys SOD1. The experiment can be more challenging than the α-synuclein complexes because SOD1 has a more compact structure; ECD-MS/MS of compact structures may not yield as many fragments as more extended structures. Figure 3-10 shows ESI mass spectra of apo-noCys-SOD1 (top) and MT-bound noCys SOD1 (bottom). Mass spectra of noCys SOD1 acquired from the FT-ICR MS was mostly similar to what we observed using the Synapt HDMS QTOF instrument, which was used for SOD1 binding characterization. From the FT-ICR MS data, four charge states (8+, 7+, 6+, and 5+) appeared to be the most abundant species. These are similar to the mass spectra from the QTOF instrument. However, there were some higher charge molecules detected by FT-ICR MS that was not observed during the titration experiments using the QTOF MS; four higher charge states were observed at 12+ to 9+ at lower intensities than the 8+ to 5+ envelope. When MT was added to SOD1 to achieve a concentration of 10 μM or a protein/ligand 1:1 concentration ratio, MT-bound complexes were readily detected. Interestingly an additional 13+-charge state was seen. Theoretically when MT binds to a lysine residue, negatively charged phosphate groups of MT neutralize positively charge lysines, and the overall charge state should be decreased. The reason behind a charge increase is currently unknown and will need to be further investigated. Single-ligand-bound (red circles) and two-ligand-bound complexes (green circles) were observed. This is similar to the result from the QTOF mass spectrometer.

In the top-down MS experiments, six charge states of the 1:1 complexes (1 ligand bound to the protein) were examined. Each charge state was isolated and transferred into the FT-ICR cell to perform ECD top-down MS/MS separately. ECD fragments were annotated using the theoretical mass from Protein Prospector with the home-built software. By this method, MT-bound and

unbound fragments were identified. Fragments are mapped on the protein sequence, as shown in Figure 3-11. Herein, data for three out of six charge states are shown. The result from the three charge states clearly demonstrates that the ECD efficiency is highly dependent on the charge of the precursor molecules[31, 32]. The higher charge, the higher the electron capture cross-section, and more c/z$^{\bullet}$ fragments should be generated. For the 12+ precursor, there are number of MT-bound fragments (red bars), which are holo z$^{\bullet}$ ions, that were observed between residue 3 to 47, suggesting that the binding region may start from residue 47 onwards toward the C-terminus. However, when we consider also the N-terminal containing holo-fragments, they were observed between residues 31 to 153. Therefore, the binding location is inconclusive. From the 12+ data, MT may bind to one of four possible lysines: K31, K37, K71, or K76. A higher charge state precursor may display a different gas phase structure that may be more extended than a lower charged molecule, so that MT may be more accessible to some lysine residues in the third loop region or beta strand (Figure 3-1). ECD fragments from the 11+-charged precursor are more conclusive. With the observation of c_{89}+MT and z_{107}+MT, we were able to narrow down the binding region to be K71 and/or K76, which are in the zinc binding loop. This loop is located on the surface and has higher flexibility, as indicated by the Debye-Waller factor or B-factor. Lastly ECD of the 10+ charge state generate drastically less MT-bound fragments. Since only one MT-bound c-ion and 5 MT-bound z-ion were observed, the identified binding region spanned a large region of the protein. There are 7 possible lysines across this region, which results in a poorer defined region for deducing MT binding sites. The result from the 10+ charge state confirms that ECD is a process that is highly dependent on the charge state of a precursor. Two lower charge states (9+ and 8+) were tested but the coverage was weaker than 10+, which is insufficient for binding site identification (data not shown).

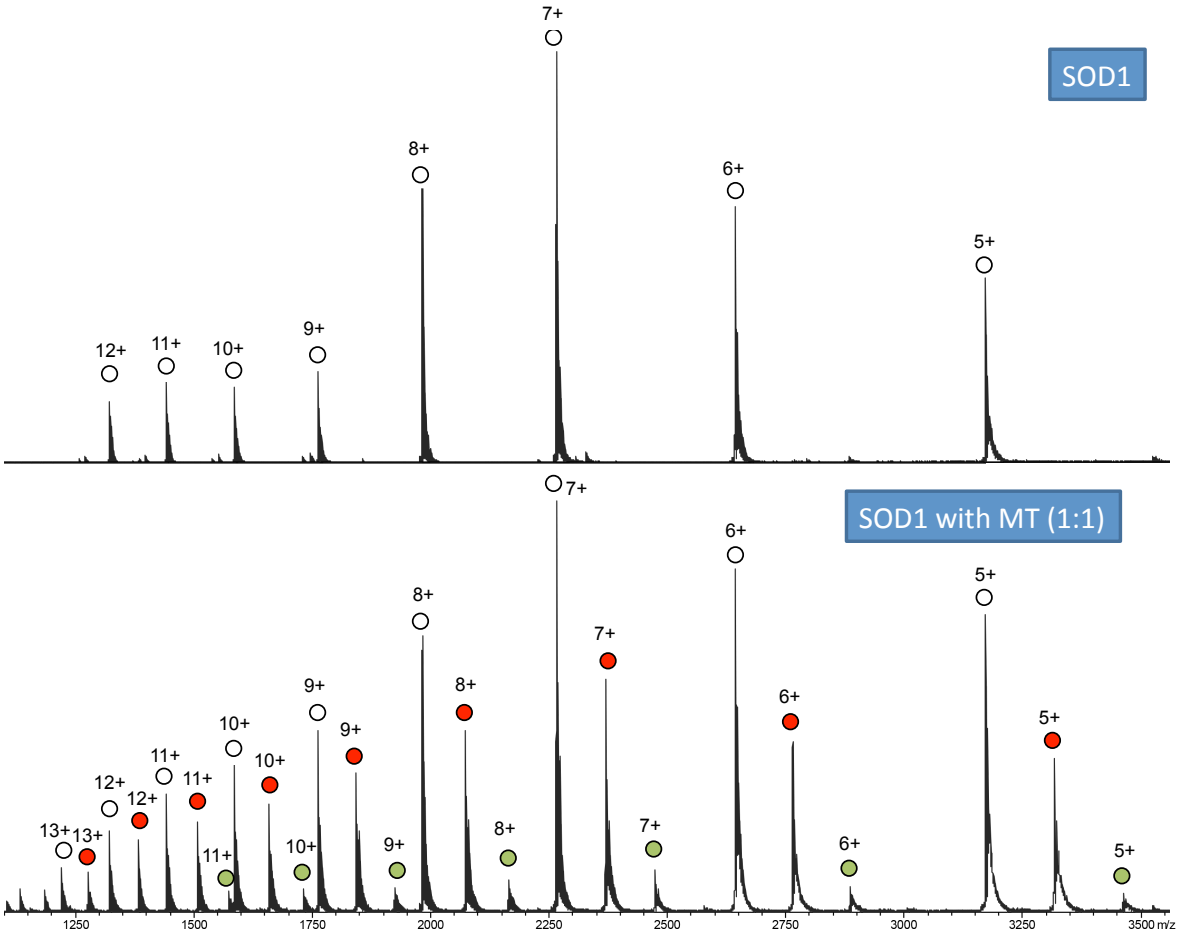

Figure 3-10. Native ESI mass spectra of SOD1 and MT-SOD1 complexes were acquired on the FT-ICR mass spectrometer. The most abundant charge states are 8+ to 5+. MT-bound forms at six highest charge states (13+ to 7+) were isolated for top-down ECD-MS/MS.

Figure 3-11. ECD fragmentation maps of 1:1 MT-SOD1 complex are shown for different precursor charge states (12+, 11+ and 10+). The higher charge state (12+) significantly generated more fragments, which is beneficial for locating the MT binding residues.

Conclusion

Native electrospray ionization mass spectrometry (ESI-MS) with a top-down approach using electron capture dissociation (ECD) is a powerful tool to provide a comprehensive view for characterizing protein-ligand interactions. In this chapter, the binding of a synthesized compound called molecular tweezer (MT or CLR01) to amyloidogenic proteins was studied. MT was previously shown that it can inhibit protein aggregation and disaggregate preformed amyloid fibrils of many proteins, such as Aβ, α-synuclein, amylin, etc. We employed top-down ESI-MS to reveal the important binding characteristics of MT to α-synuclein and SOD1, which are involved with Parkinson's and ALS diseases, respectively. Binding of MT to α-synuclein and SOD1 was readily observed by native ESI-MS and their dissociation constants (K_d) were determined to be in the low μM region. Top-down MS analysis shows that the primary MT binding site of α-synuclein occurs at K10 or K12, and MT likely binds to K71 or K76 on human SOD1. The binding information can be used to develop additional insight and a molecular mechanism behind MT binding on amyloid proteins to prevent amyloid proteins to self-aggregate and form toxic oligomer or fibrils.

References

[1] Chiti, F., and Dobson, C. M. (2006) Protein Misfolding, Functional Amyloid, and Human Disease, *Annual Review of Biochemistry 75*, 333-366.

[2] Buée, L., Bussière, T., Buée-Scherrer, V., Delacourte, A., and Hof, P. R. (2000) Tau protein isoforms, phosphorylation and role in neurodegenerative disorders, *Brain Research Reviews 33*, 95-130.

[3] Mazanetz, M. P., and Fischer, P. M. (2007) Untangling tau hyperphosphorylation in drug design for neurodegenerative diseases, *Nature Reviews Drug Discovery 6*, 464-479.

[4] Spillantini, M. G., Schmidt, M. L., Lee, V. M. Y., Trojanowski, J. Q., Jakes, R., and Goedert, M. (1997) α-Synuclein in Lewy bodies, *Nature 388*, 839-840.

[5] Rodriguez, J. A., Ivanova, M. I., Sawaya, M. R., Cascio, D., Reyes, F. E., Shi, D., Sangwan, S., Guenther, E. L., Johnson, L. M., Zhang, M., Jiang, L., Arbing, M. A., Nannenga, B. L., Hattne, J., Whitelegge, J., Brewster, A. S., Messerschmidt, M., Boutet, S., Sauter, N. K., Gonen, T., and Eisenberg, D. S. (2015) Structure of the toxic core of α-synuclein from invisible crystals, *Nature 525*, 486-490.

[6] Muchowski, P. J., and Wacker, J. L. (2005) Modulation of neurodegeneration by molecular chaperones, *Nature Reviews Neurosciences 6*, 11-22.

[7] Lashuel, H. A., Overk, C. R., Oueslati, A., and Masliah, E. (2013) The many faces of α-synuclein: from structure and toxicity to therapeutic target, *Nature Reviews Neuroscience 14*, 38-48.

[8] Zhang, X., Smith, D. L., Meriin, A. B., Engemann, S., Russel, D. E., Roark, M., Washington, S. L., Maxwell, M. M., Marsh, J. L., Thompson, L. M., Wanker, E. E., Young, A. B., Housman, D. E., Bates, G. P., Sherman, M. Y., and Kazantsev, A. G. (2005) A potent small molecule inhibits polyglutamine aggregation in Huntington's disease neurons and suppresses neurodegeneration in vivo, *Proceedings of the National Academy of Sciences of the United States of America 102*, 892-897.

[9] Fonseca-Ornelas, L., Eisbach, S. E., Paulat, M., Giller, K., Fernández, C. O., Outeiro, T. F., Becker, S., and Zweckstetter, M. (2014) Small molecule-mediated stabilization of vesicle-associated helical α-synuclein inhibits pathogenic misfolding and aggregation, *Nature Communications 5*, 5857.

[10] Sinha, S., Lopes, D. H. J., Du, Z., Pang, E. S., Shanmugam, A., Lomakin, A., Talbiersky, P., Tennstaedt, A., McDaniel, K., Bakshi, R., Kuo, P.-Y., Ehrmann, M., Benedek, G. B., Loo, J. A., Klärner, F.-G., Schrader, T., Wang, C., and Bitan, G. (2011) Lysine-Specific Molecular Tweezers Are Broad-Spectrum Inhibitors of Assembly and Toxicity of Amyloid Proteins, *Journal of the American Chemical Society 133*, 16958-16969.

[11] Fokkens, M., Schrader, T., and Klärner, F.-G. (2005) A Molecular Tweezer for Lysine and Arginine, *Journal of the American Chemical Society 127*, 14415-14421.

[12] Liu, T., and Bitan, G. (2012) Modulating Self-Assembly of Amyloidogenic Proteins as a Therapeutic Approach for Neurodegenerative Diseases: Strategies and Mechanisms, *ChemMedChem 7*, 359-374.

[13] Sinha, S., Du, Z., Maiti, P., Klärner, F.-G., Schrader, T., Wang, C., and Bitan, G. (2012) Comparison of Three Amyloid Assembly Inhibitors: The Sugar scyllo-Inositol, the Polyphenol Epigallocatechin Gallate, and the Molecular Tweezer CLR01, *ACS Chemical Neuroscience 3*, 451-458.

[14] Attar, A., Rahimi, F., and Bitan, G. (2013) Modulators of amyloid protein aggregation and toxicity: EGCG and CLR01, *Translational Neuroscience. 4*, 385-409.

[15] Attar, A., Chan, W.-T., Klarner, F.-G., Schrader, T., and Bitan, G. (2014) Safety and pharmacological characterization of the molecular tweezer CLR01 - a broad-spectrum inhibitor of amyloid proteins' toxicity, *BMC Pharmacology and Toxicology 15*, 23.
[16] Attar, A., Ripoli, C., Riccardi, E., Maiti, P., Li Puma, D. D., Liu, T., Hayes, J., Jones, M. R., Lichti-Kaiser, K., Yang, F., Gale, G. D., Tseng, C.-h., Tan, M., Xie, C.-W., Straudinger, J. L., Klärner, F.-G., Schrader, T., Frautschy, S. A., Grassi, C., and Bitan, G. (2012) Protection of primary neurons and mouse brain from Alzheimer's pathology by molecular tweezers, *Brain 135*, 3735-3748.
[17] Valentine, J. S., Doucette, P. A., and Zittin Potter, S. (2005) Copper-Zinc Superoxide Dismutase and Amyotrophic Lateral Sclerosis, *Annual Review of Biochemistry 74*, 563-593.
[18] Bruijn, L. I., Miller, T. M., and Cleveland, D. W. (2004) Unraveling the Mechanisms Involved in Motor Neuron Degeneration in ALS, *Annual Review of Neuroscience 27*, 723-749.
[19] Rosen, D. R., Bowling, A. C., Patterson, D., Usdin, T. B., Sapp, P., Mezey, E., McKenna-Yasek, D., O'Regan, J., Rahmani, Z., Ferrante, R. J., Brownstein, M. J., Kowall, N. W., Beal, M. F., Horvitz, H. R., and Brown, R. H. (1994) A frequent ala 4 to val superoxide dismutase-1 mutation is associated with a rapidly progressive familial amyotrophic lateral sclerosis, *Human Molecular Genetics 3*, 981-987.
[20] Chattopadhyay, M., Durazo, A., Sohn, S. H., Strong, C. D., Gralla, E. B., Whitelegge, J. P., and Valentine, J. S. (2008) Initiation and elongation in fibrillation of ALS-linked superoxide dismutase, *Proceedings of the National Academy of Sciences USA 105*, 18663-18668.
[21] Loo, J. A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry, *Mass Spectrometry Reviews 16*, 1-23.
[22] Ngounou Wetie, A. G., Sokolowska, I., Woods, A. G., Roy, U., Loo, J. A., and Darie, C. C. (2013) Investigation of stable and transient protein–protein interactions: Past, present, and future, *Proteomics 13*, 538-557.
[23] Xie, Y., Zhang, J., Yin, S., and Loo, J. A. (2006) Top-Down ESI-ECD-FT-ICR Mass Spectrometry Localizes Noncovalent Protein-Ligand Binding Sites, *Journal of the American Chemical Society 128*, 14432-14433.
[24] Yin, S., and Loo, J. A. (2010) Elucidating the Site of Protein-ATP Binding by Top-Down Mass Spectrometry, *Journal of the American Society for Mass Spectrometry 21*, 899-907.
[25] Yin, S., and Loo, J. A. (2011) Top-down mass spectrometry of supercharged native protein–ligand complexes, *International Journal of Mass Spectrometry 300*, 118-122.
[26] Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process, *Journal of the American Chemical Society 120*, 3265-3266.
[27] Katta, V., Chait, B. T., and Carr, S. (1991) Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry, *Rapid Communications in Mass Spectrometry 5*, 214-217.
[28] Sannes-Lowery, K. A., Griffey, R. H., and Hofstadler, S. A. (2000) Measuring dissociation constants of RNA and aminoglycoside antibiotics by electrospray ionization mass spectrometry, *Analytical Biochemistry 280*, 264-271.
[29] Zhang, S., Van Pelt, C. K., and Wilson, D. B. (2003) Quantitative determination of noncovalent binding interactions using automated nanoelectrospray mass spectrometry, *Analytical Chemistry 75*, 3010-3018.
[30] Acharya, S., Safaie, B. M., Wongkongkathep, P., Ivanova, M. I., Attar, A., Klärner, F.-G., Schrader, T., Loo, J. A., Bitan, G., and Lapidus, L. J. (2014) Molecular Basis for

Preventing α-Synuclein Aggregation by a Molecular Tweezer, *Journal of Biological Chemistry 289*, 10727-10737.
[31] Zubarev, R. A., Horn, D. M., Fridriksson, E. K., Kelleher, N. L., Kruger, N. A., Lewis, M. A., Carpenter, B. K., and McLafferty, F. W. (2000) Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations, *Analytical Chemistry 72*, 563-573.
[32] Breuker, K., Brüschweiler, S., and Tollinger, M. (2011) Electrostatic Stabilization of a Native Protein Structure in the Gas Phase, *Angewandte Chemie International Edition 50*, 873-877.

CHAPTER FOUR

Mapping Protein Surface Residues by Monitoring Molecular Tweezer Binding

Abstract

The determination of protein structures is very important for understanding their functions. The task can be difficult and labor intensive depending on the protein. Here we demonstrate the use of a molecular tweezer compound (MT or CLR01) to map the surface region of proteins in their native conformation. MT is a small, 726 Da, compound that specifically binds to lysine residues. Electron capture dissociation (ECD), which is a soft fragmentation method, with mass spectrometry was used to identify MT binding site for several globular proteins. A surface lysine of the protein ubiquitin (8.6 kDa) was identified as K11. The binding region of myoglobin (17 kDa) covers three possible lysines (K133, K145, and K147). Nonetheless, we were able to correlate the binding sites from MS/MS results with surface accessibility calculated from ASAView and GetArea programs. A good correlation was achieved. With top-down mass spectrometry and ECD of MT-protein complexes, determining protein surface lysine residues can be quickly characterized.

Introduction

Proteins undergo folding through several different pathways, but a specific pathway via a chaperone-assisted mechanism achieves the lowest energy states to be fully activated to perform their function[1]. Different regions on properly folded proteins have their own specific functions, such as cofactor binding domains and as catalytic sites. Protein surface regions are also important because they are the main regions that interact with other proteins, stimuli, or undergo some crucial modifications. Elucidating protein structures, especially surface topology is important to understand their biological functions. X-ray crystallography[2] and NMR[3] are the most common techniques to study solution phase protein structure. However, they are labor-intensive tasks and there are many challenging limitations involved depending on protein sizes and characteristics[4].

Mass spectrometry becomes an alternative choice to provide structural information because of its speed, sensitivity, and small sample amount requirements[5-7]. Since molecules in a mass spectrometer operate in vacuum, questions have been raised whether structural information obtained from a method that measures gas phase molecules relates to structures found in solution. There are a number of studies showing that if proper experimental procedures are optimized, an intact protein structure can be preserved in the gas phase that is similar to a structure in solution[8-12]. Generally electrospray ionization (ESI) is preferred because ESI is known to be a soft ionization technique[13, 14]. A normal experimental condition for ESI-MS is the use of solutions containing an organic solvent (e.g., methanol, acetonitrile) mixed with an acidic aqueous solution to maximize charging, desolvation and sensitivity. But using an ammonium acetate buffer to maintain the physiological pH is the most common approach to introduce proteins into mass spectrometer in their “native” states[15]. Native ESI-MS has limited sensitivity

compared to normal ESI-MS (using solvents that denatures the structures of proteins), but it can provide surface topology information when specific methods are applied. We will discuss more about these methods later in this chapter.

There are several alternatives besides native ESI-MS to probe surface residues of proteins, including surface labeling[16]. Labeling approaches are usually performed with additional steps in solution. A proteolytic digestion is performed after the labeling step to break apart the labeled protein into small peptides, and liquid chromatography is used to separate peptides before MS to reduce sample complexity[17]. The most common labeling techniques are hydrogen-deuterium (H/D) exchange[18, 19], oxidative footprinting[20, 21], chemical crosslinking[22], and chemical modifications[23]. These approaches are widely used because the information is truly from solution phase structures. H/D exchange is a powerful technique that takes advantage of solvent accessible amide protons. Any backbone amide protons can be exchanged with deuterium from heavy water in an equilibrium manner. So H/D exchange provides details about surface regions, or regions involved in conformational changes[24]. Ligand binding sites can be elucidated because the H/D exchange is hindered at the binding region. However, caution needs to be exercised during the exchange reaction, proteolysis, and separation steps because back-exchange can obscure the analysis. Moreover H/D scrambling is possible if collision-base fragmentation, such as CAD, is used.

Compared to H/D exchange, oxidative footprinting uses a more permanent labeling approach, using hydroxyl radicals, which can target almost any residue[25, 26]. In the past, hydrogen peroxide was used to react with ascorbate to generate radicals. It has been shown that oxidative footprinting was able to map protein surfaces and disorder regions of myoglobin[27]. Hydroxyl radical can induce unfolding of proteins in the millisecond timescale. Recently, fast

photochemical oxidative protein footprinting (FPOP) has been developed. FPOP utilizes a UV laser pulse to generate hydroxyl radicals. The labeling process is completed within a few microseconds, in which conformational changes becomes minimal[28-30]. FPOP has been used for mapping surface and conformational changes[31, 32]. Alternatively, electrochemical oxidation can be utilized[33].

Chemical crosslinking and covalent labeling are other approaches that provide similar information. These methods use chemistry to covalently label solvent accessible residues. Chemical crosslinking has also been used for mapping intermolecular protein-protein interfacial regions, as well as intramolecular structures. Data analysis is challenging due to its high complexity[22, 34]. On the other hand, conventional covalent labeling using chemical probes has less complexity because a reagent only specifically reacts to only one or a few target residues. A few examples of covalent labeling reagents (and its target residue) that have been used for surface probing include DHCH (arginine)[35], succinylation (lysine)[36], carbene[37] (any residues), DEPC[38], DMBNHS[39] (lysine), and hydrazone[40] (lysine).

On the other hand, several non-covalent binding approaches have been developed. These methods use specific compounds as molecular recognition probes that non-covalently bind to some residues. Some of the more popular compounds are 18-crown-6 (18C6) and cucurbit[n]urils (CB[n], where n = 5-8, or 10). The major benefit is minimal sample preparation and clean-up steps are required. Weakly bound ligands are unlikely to perturb the protein structure. HPLC separation is difficult to be applied because a strong organic content can open up the protein structure and disrupt the non-covalent probes. Molecular probes and proteins are mixed together in solution. But instrument settings must be carefully tuned for ionization and ion transfer so that the protein complexes are preserved in the gas phase[8]. 18C6 has been used for

a method coined “selective noncovalent adduct protein probing (SNAPP)[41]” to probe structural changes on metal binding. CB[6] has also demonstrated a similar ability to 18C6 to monitor conformational changes[42-44]. This chapter describes the use of a small molecular tweezer (MT) molecule, which has a high specificity and high affinity to bind lysine residues, to probe surface lysine residues of proteins in their native states. Previously MT has been demonstrated to be a potential inhibitor for amyloid fibrillation[45, 46].

Our lab’s current interests include the characterization of non-covalent protein-ligand interactions using mass spectrometry. Electron capture dissociation (ECD) has been used to locate ligand-binding sites. ECD is a soft dissociation method that cleaves N-C_α bonds but not labile post-translational modifications or weakly bound ligands[47, 48]. Our lab demonstrated that ECD can be used in a top-down fashion to determine the ligand binding sites of several protein complexes[49-51]. Top-down mass spectrometry also reveals the binding site of MT on two neurodegenerative disease proteins: amyloid-β and α-synuclein, which are involved in Alzheimer’s and Parkinson’s disease, respectively[52, 53]. Herein, the use of MT and top-down ECD MS/MS to probe surface lysine and provide structural information is illustrated.

Experimental

Bovine ubiquitin, horse apo-myoglobin, and ammonium acetate were purchased from Sigma-Aldrich (St Louis, MO). Proteins were dissolved in 10 mM ammonium acetate solution to maintain pH at 6.8 to preserve structure in their native state. A buffer exchange was performed using a centrifugal filter with a molecular weight cut-off (MWCO) at 3 kDa (for ubiquitin) and 10 kDa (for apo-myoglobin) to remove excess salt and other matrices. Molecular tweezer (MT or CLR01) was provided by Dr. Gal Bitan's lab at UCLA. Proteins were diluted to achieve the final concentration of 10 μM for analysis by ESI-MS. Protein complex sample solutions were then inserted into a Au/Pd-coated borosilicate glass capillary (Thermo Scientific, San Jose, CA) and hooked up to a nanoelectrospray source with a 1.25 kV ESI voltage. The expected flow rate was in a range between 50-100 nL/min. There is a countercurrent nitrogen gas at a flow rate of 1 L/min and 160 ^{o}C temperature to facilitate ion desolvation. A Bruker Solarix 15-Tesla FT-ICR mass spectrometer (Billerica, MA), equipped with the infinity ICR cell, was used for this study. Source transfer optics were tuned with a voltage of capillary 200V, deflector plate 180V, funnel 120V and skimmer 50V. Top-down MS/MS experiments were performed inside the ICR cell using low energy electrons generated from the ECD cathode filament. The ECD cell was tuned to achieve 10 ms pulse, 1V bias and 15V on the ECD lens. Fragments (ligand-bound and ligand free) were manually detected by the Bruker DataAnalysis software (Billerica, MA) and annotated manually by a customized Matlab program comparing with theoretical masses obtained from Protein Prospector (http://prospector.ucsf.edu/prospector/mshome.htm), using a mass accuracy of 10 ppm or less.

Results and Discussion

Native ESI-MS of MT-bound ubiquitin and apo-myoglobin complexes

Ubiquitin and myoglobin are standard proteins that have been extensively studied by ESI-MS[39, 47, 54, 55]. Ubiquitin was chosen to be a model protein to study unfolding in the gas phase in many studies[56, 57]. Myoglobin with its heme binding has been also heavily explored by top-down MS/MS by other labs[58]. Interactions between MT and various intact proteins were observed by MS in the form of $[M+nH+726m]^{n+}$ ions, where m is a number of bound MT, n is a number of charge from protons, and the mass of a bound MT is 726 Da. Unbound MT was observed in some concentration ratios. We first tested the MT binding on 5 different proteins (ubiquitin, myoglobin, cytochrome C, ribonuclease A, and α-synuclein). Multiple binding (m = 2-4, depending on the MT concentration) was found for apo- myoglobin, holo-myoglobin (data not shown), ubiquitin and α-synuclein. Binding of α-synuclein and MT is reported in Chapter 3. We propose that the high solvent accessible lysine residues will bind with MT under native solution conditions.

Figures 4-1 and 4-2 show mass spectra of ubiquitin-MT and apo-myoglobin-MT complexes, respectively. A native ESI mass spectrum ubiquitin normally shows the 6+ and 5+ charge states[15], indicating the protein to be in the folded state because protein charging likely depends on protein surface area[59, 60]. The spectrum of ubiquitin-MT complexes was acquired from a solution that contains 10 μM of the protein and 30 μM MT (1:3 ratio). At this ratio, up to three ligand-binding events is observed. The most abundant species is the 6+-charged 1:1 complex. Because of its intensity and the highest observed charge, it will be further selected for top-down experiments by ECD. Unbound protein was barely observed at this concentration ratio. A singly charged ammonium adduct of protonated MT is detected at *m/z* 744.

The mass spectrum of apo-myoglobin-MT complexes has two charge distributions. The main distribution is at a high *m/z* region, where 10+ to 7+ charges were observed. This charge envelope represents a more compact conformation. Interestingly 16+ to 11+ charges were observed in the low *m/z* region, indicating that a fraction of the protein may be slightly unfolded. The higher charge species are expected to have more extended conformations. Although higher charge is beneficial for ECD fragmentation, it may become too unfolded to be a model conformation for surface residue probing by MT. Consequently, we selected the 9+ charged molecule for subsequent ECD experiments. The 12+ charged molecule, which is one of the lowest charge states from the high charge envelope, was also chosen for top-down MS experiments.

Top-down MS/MS to identify MT binding residues

ECD has been used to determine the amino acid residues involved in binding small ligands because it does not disrupt the non-covalent interactions. MT-bound and apo-fragments were identified. ECD of 6+-ubiquitin significantly generated less fragments than higher charge states from a denatured protein[56]. Three unbound (c_6, c_9, and c_{10}) fragments, or apo-fragments, were observed (black lines). Additionally, five MT-bound fragments, or holo fragments, (c_{21}+MT, c_{74}+MT, c_{75}+MT, $z_{66}^{\bullet}$+MT, $z_{70}^{\bullet}$+MT) were observed as depicted by bold red lines. Lysine-11 is suggested to be the tweezer-binding site because of the specific observation of c_{21}+MT and $z_{66}^{\bullet}$+MT. ECD fragmentation is shown on the ubiquitin sequence in Figure 4-1.

For apo-myoglobin, two charge states (12+ and 9+) were isolated and transferred into the ICR cell for ECD experiments. An additional skimmer voltage was needed to improve ECD fragmentation efficiency for the 9+ charge state species of myoglobin. The additional voltage

was properly tuned so that MT-protein interactions can be retained under pre-activation and ECD conditions. The ECD fragments identified from both charge states were combined and illustrated in Figure 4-2. Here, several of ligand-bound fragments were observed. Fifteen MT-bound $z^{\bullet}$ ions were found between residues 2-130, suggesting that MT binds at the C-terminal region. Two MT-bound c-type fragments (c151+MT, and c152+MT) confirmed the C-terminal binding region. The binding motif spans across residues 130-151. There are 3 lysines in this region, which are K133, K145, and K147. Although ECD efficiency was limited for the low charge states, the measured c/$z^{\bullet}$-fragments were sufficient for locating the binding sites. The identified tweezer binding sites were then compared with the known 3D protein structures to develop a model for surface mapping.

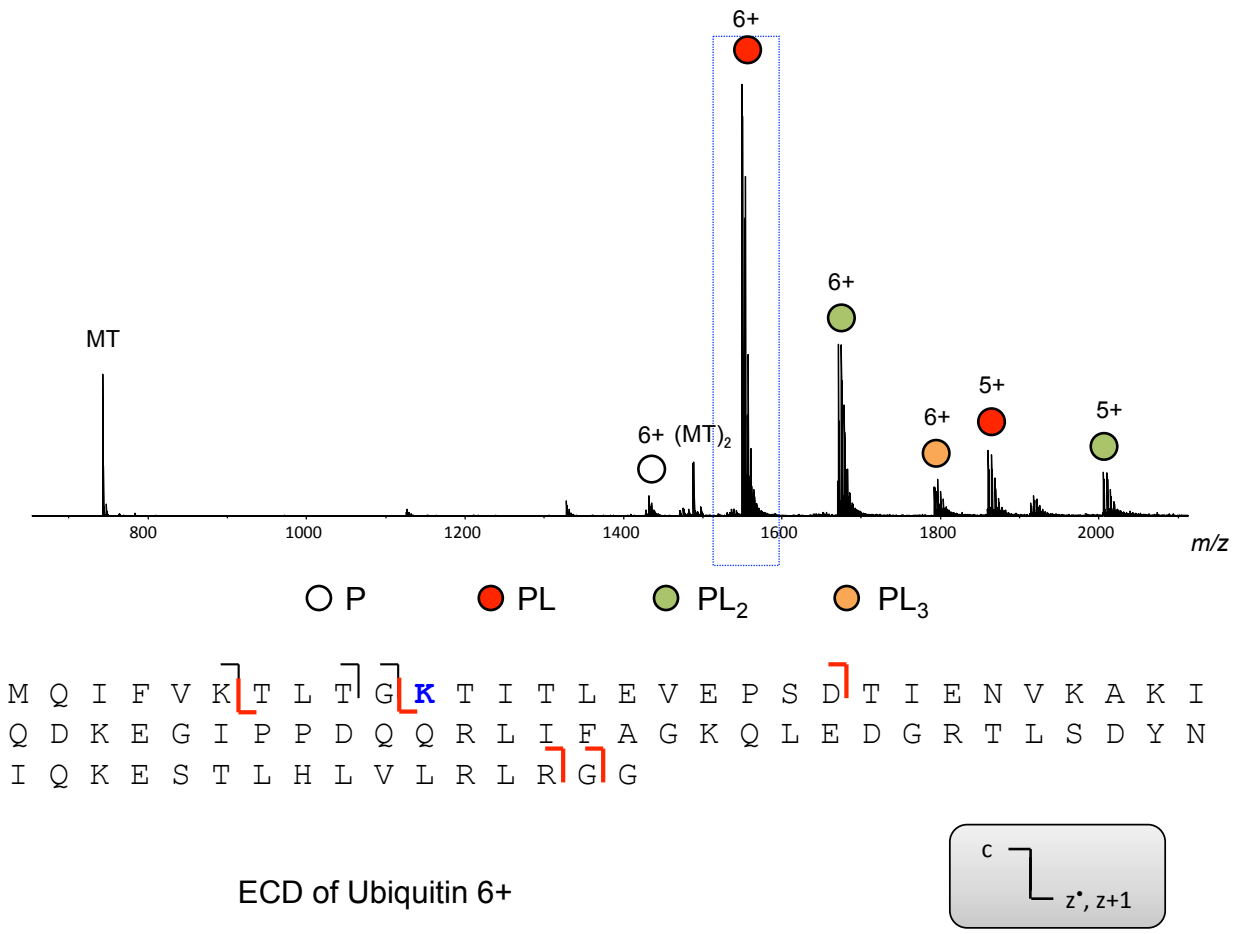

Figure 4-1. MS spectrum of native MT-bound ubiquitin complex containing five species: 1:1, 1:2, 1:3 protein-ligand complexes, ligand-free protein, and unbound ligand. The schematic below illustrates ECD fragmentation along the protein sequence. MT-bound and unbound fragments are highlighted in red and black, respectively.

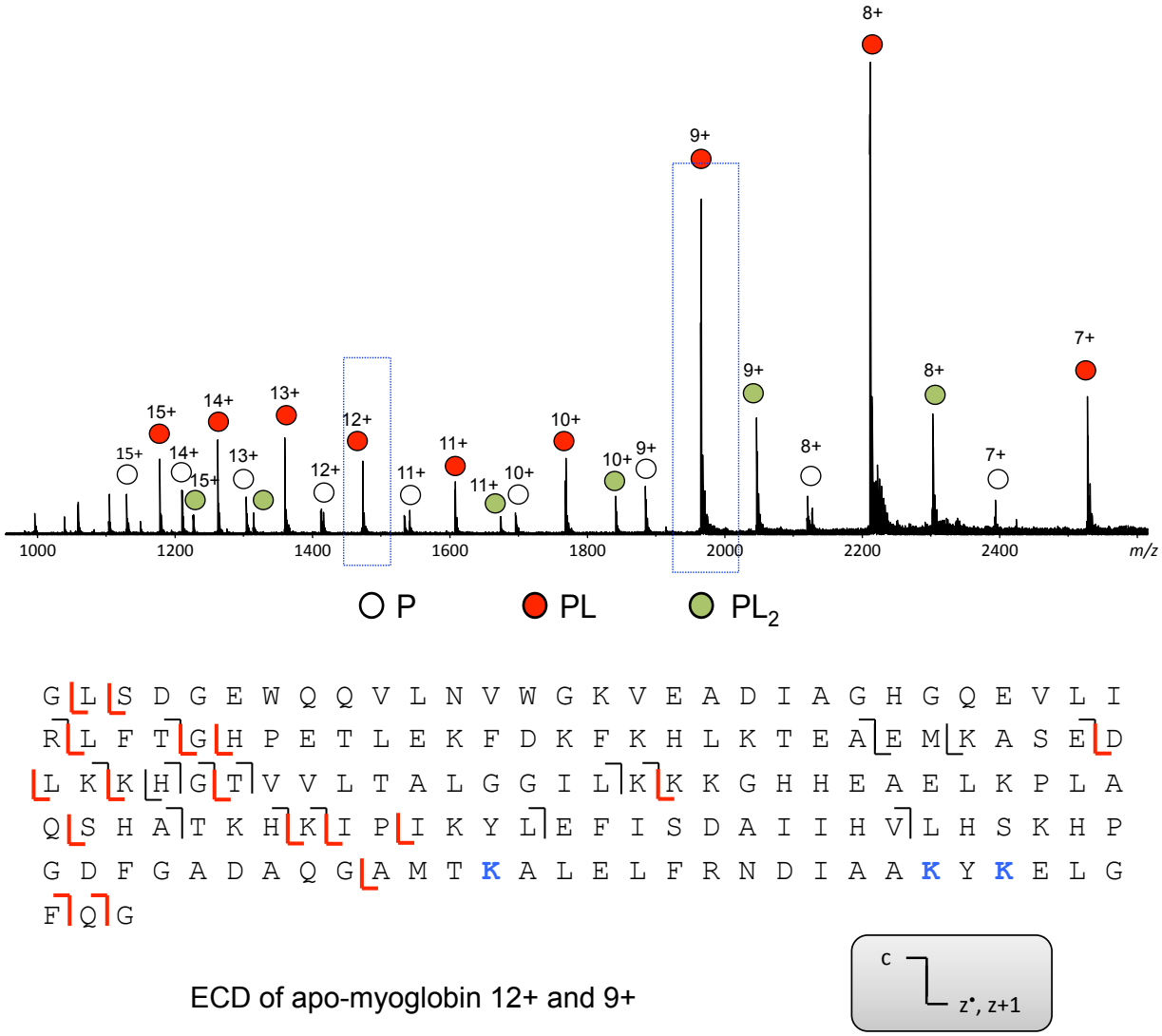

Figure 4-2. (top) MS spectrum of native apo-myoglobin-MT complex showing 1:1 and 1:2 complexes. The 12+ and 9+ charges were selected for top-down MS experiments. (bottom) A fragmentation profile showing backbone cleavage along the sequence. Red lines are ligand-bound fragment, where black lines are unbound fragments.

Surface residues calculation

Two software packages were used to calculate surface accessibility obtained from the 3D structures. 1UBQ and 1YMB are the PDB files of ubiquitin and myoglobin for the structures that were used in this experiment. First, ASA-View (http://www.abren.net/asaview/) provides accessible surface area (ASA) values and generates a graphical "spiral plot"[61]. The spiral plot sorts residues by their relative solvent accessibility. The outer ring residues are more exposed, where the inner residues are buried in the structure. The radius of each circle is proportional to the ASA value of that specific residue. Table 4-1 illustrates %ASA values, which is a relative amount of solvent accessibility. Theoretical solvent accessibilities were also confirmed by GetArea[62] (http://curie.utmb.edu/getarea.html). For each residue, GetArea reports the contribution of atoms from the backbone and sidechains. It also calculates the random coil value of X residue, which represent an average accessible surface area of a tripeptide Gly-X-Gly in random conformations. The random coil values can be found on the GetArea website. The software then tabulates a ratio of sidechain surface area over random coil value (%SA ratio). A ratio greater than 50% is considered to be solvent exposed. The residues with a %SA ratio less than 20% are buried. Percent SA ratio values are also listed in Table 4-1.

From Table 4-1, %ASA calculated from ASView of ubiquitin and myoglobin were in agreement with calculated %SA ratio from GetArea, confirming the validity of the model. Only theoretical values from lysines were displayed. Lys63 is the most solvent accessible residue of ubiquitin, but it is not the MT binding site determined by MS. ECD data showed that MT binds to Lys11, which is the second most accessible according to the %SA ratio from GetArea. A spiral plot of ubiquitin residues is shown in Figure 4-3. Basic residues are colored in blue and the Lys11 position is shown on the plot. The binding site does not correlate with the surface accessibility model very well. We then examined the 3D structure of ubiquitin (PDB: 1UBQ) and highlighted

lysines as red sticks (Figure 4-3). The crystal structure clearly showed that K63 is the most accessible. K11 does not fully point outwards. It is still possible for MT to bind K11 because K11 sits on a β-sheet surface and has room for potential MT binding. Another reason that may explain the result is from the size of ubiquitin. Compare to other mid-size proteins, ubiquitin is relatively small, so it has more surface accessible regions than larger proteins. The theoretical surface accessibility from GetArea or ASAView is only relative to the same proteins. Most lysine residues of ubiquitin have similar %AS ratio due to its size. From a molecular interaction perspective, there might not be a significant difference between accessibility to bind to K11 or K63. But K11 is surrounded by the β-sheet region that may contribute hydrophobic interactions to stabilize MT binding.

For apo-myoglobin, the ECD data suggested that MT binding occurred at three possible lysines: K133, K145, or K147. The binding site could not be narrowed down to a single residue because of a lack of backbone cleavages. The calculated surface accessibilities showed that K147 is the most accessible lysine of myoglobin. The other two are not surface exposed since their %SA ratios are below 50% (Table 4-1). The crystal structure of myoglobin (1YMB) showed that most lysines are located on the surface of proteins (Figure 4-4). But it confirmed that K147 is highest accessible residue and points out from the helical surface and its structure. Thus, the MT binding site determined from top-down ECD of MT-bound myoglobin correlates with the protein's theoretical solvent accessibility. A spiral plot of myoglobin is also shown. Basic residues are highlighted in blue and three lysines identified from top-down ECD-MS/MS are highlighted in red. In conclusion, mapping of protein surface residues can be quickly characterized by top-down ECD-MS/MS of MT-protein complexes. The most solvent accessible lysine can be assessed.

Table 4-1. Solvent accessibility of lysine residues of ubiquitin and apo-myoglobin.

Protein	Residues	%ASA	%SA ratio
Ubiquitin	K6	51	58.15
	K11	47.6	58.21
	K27	7.3	9.51
	K29	34.5	36.31
	K33	54.4	48.12
	K48	48.6	54.56
	K63	64.7	72.28
Myoglobin	K16	19.4	21.6
	K42	39.9	30.8
	K45	54.9	60.5
	K47	37.9	45.7
	K50	70	77.1
	K56	52.5	58.4
	K62	36	44.2
	K63	68.5	82.9
	K77	45.2	49.9
	K78	43.3	49.2
	K79	45.7	48.4
	K87	59.3	71.6
	K96	76.8	83
	K98	52	61.9
	K102	50.6	61.7
	K118	27.7	24
	K133	27.7	37.4
	K145	19.9	23.9
	K147	83.1	91.3

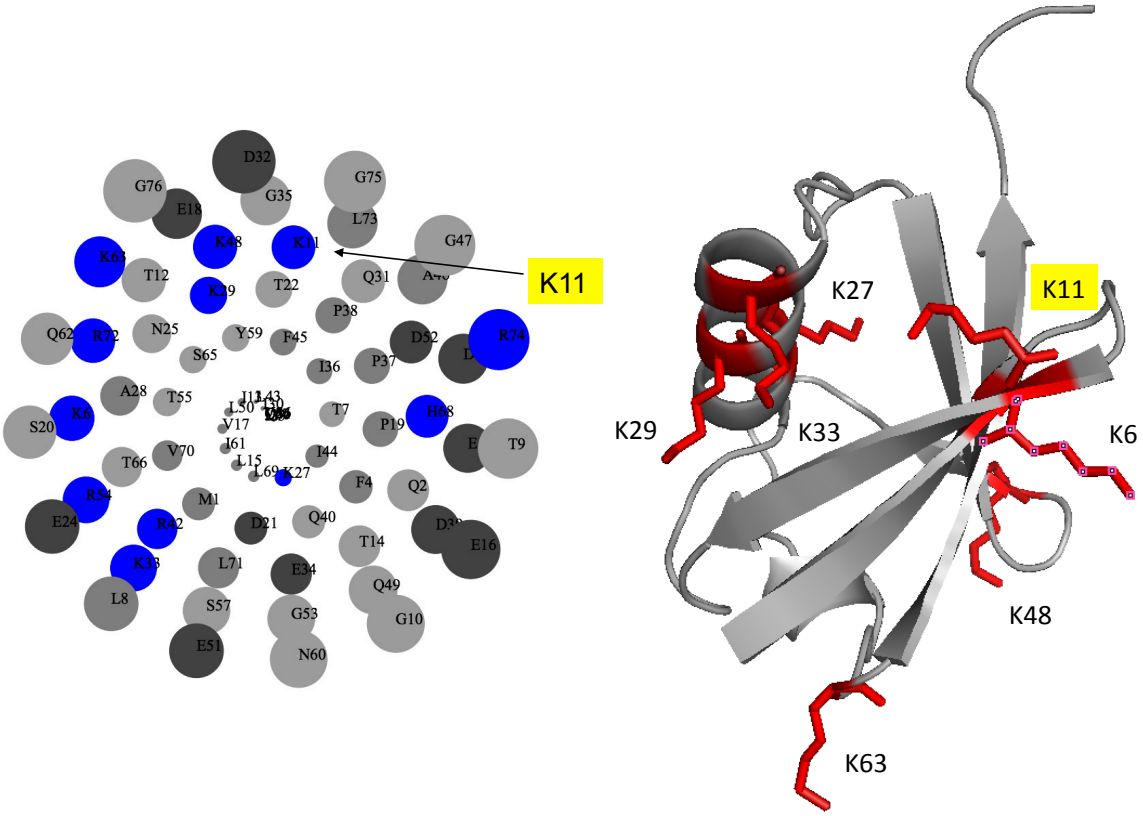

Figure 4-3. (left) A spiral plot of surface accessibility for each residue of ubiquitin. Basic residues are highlighted in blue. (right) The 3D structure of ubiquitin with lysines highlighted.

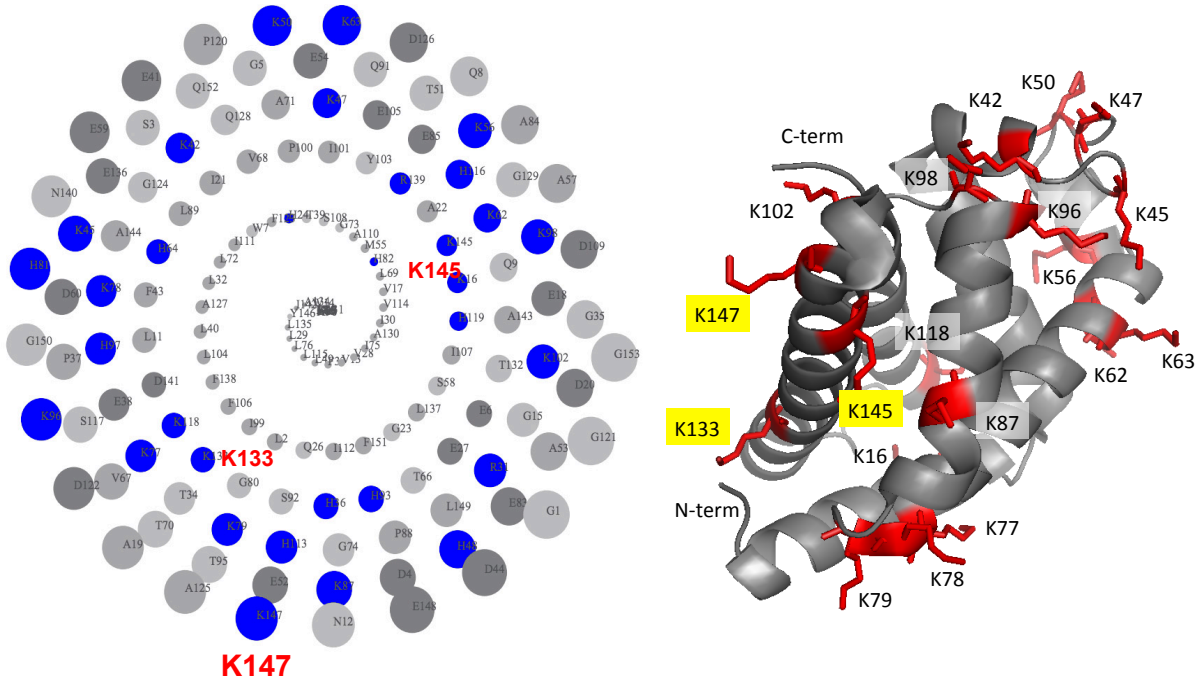

Figure 4-4. (left) A spiral plot showing solvent accessibility of myoglobin, derived from PDB file 1YMB. Blue circles are basic residues. (right) 3D structure of myoglobin. Lysine residues are highlighted in red.

Conclusion

Molecular tweezer binding sites of ubiquitin and apo-myoglobin were identified. Previously, Bitan *et al.* showed that the tweezer compound was able to inhibit α-synuclein aggregation as well as Aβ[4]. Our data suggests that α-synuclein has a potential MT binding site near the N-ternimus. For other proteins, ECD-MS data suggests the MT-binding sites to be K11 for ubiquitin and residues 130-151 for myoglobin. The known 3D structures of ubiquitin and myoglobin were used to calculate their percent relative solvent accessibility. The data indicates that the MT binding site information from ECD experiments matched to the surface residues information from ASA view. Information from FT-ICR MS with ECD of tweezer-bound protein complexes can be utilized as a method to characterize and map native protein surface residues.

References

[1] Hartl, F. U. (1996) Molecular chaperones in cellular protein folding, *Nature 381*, 571-580.

[2] Miao, J., Ishikawa, T., Shen, Q., and Earnest, T. (2008) Extending X-Ray Crystallography to Allow the Imaging of Noncrystalline Materials, Cells, and Single Protein Complexes, *Annual Review of Physical Chemistry 59*, 387-410.

[3] Wuthrich, K. (2001) The way to NMR structures of proteins, *Nature Structural and Molecular Biology 8*, 923-925.

[4] Chen, F., Gülbakan, B., Weidmann, S., Fagerer, S. R., Ibáñez, A. J., and Zenobi, R. (2015) Applying mass spectrometry to study non-covalent biomolecule complexes, *Mass Spectrometry Reviews*, in press.

[5] Loo, J. A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry, *Mass Spectrometry Reviews 16*, 1-23.

[6] Schmidt, C., and Robinson, C. V. (2014) Dynamic protein ligand interactions – insights from MS, *FEBS Journal 281*, 1950-1964.

[7] Heck, A. J. R., and van den Heuvel, R. H. H. (2004) Investigation of intact protein complexes by mass spectrometry, *Mass Spectrometry Reviews 23*, 368-389.

[8] Kitova, E., El-Hawiet, A., Schnier, P., and Klassen, J. (2012) Reliable Determinations of Protein–Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet?, *Journal of the American Society for Mass Spectrometry 23*, 431-441.

[9] Hall, Z., and Robinson, C. (2012) Do Charge State Signatures Guarantee Protein Conformations?, *Journal of the American Society for Mass Spectrometry 23*, 1161-1168.

[10] Wyttenbach, T., and Bowers, M. T. (2011) Structural Stability from Solution to the Gas Phase: Native Solution Structure of Ubiquitin Survives Analysis in a Solvent-Free Ion Mobility–Mass Spectrometry Environment, *The Journal of Physical Chemistry B 115*, 12266-12275.

[11] Kebarle, P., and Verkerk, U. H. (2009) Electrospray: From ions in solution to ions in the gas phase, what we know now, *Mass Spectrometry Reviews 28*, 898-917.

[12] Ngounou Wetie, A. G., Sokolowska, I., Woods, A. G., Roy, U., Loo, J. A., and Darie, C. C. (2013) Investigation of stable and transient protein–protein interactions: Past, present, and future, *Proteomics 13*, 538-557.

[13] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules, *Science 246*, 64-71.

[14] Loo, J. A., Edmonds, C. G., and Smith, R. D. (1990) Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry, *Science 248*, 201-204.

[15] Loo, J. A., Loo, R. R. O., Udseth, H. R., Edmonds, C. G., and Smith, R. D. (1991) Solvent-induced conformational changes of polypeptides probed by electrospray-ionization mass spectrometry, *Rapid Communications in Mass Spectrometry 5*, 101-105.

[16] Kaltashov, I. A., and Eyles, S. J. (2002) Studies of biomolecular conformations and conformational dynamics by mass spectrometry, *Mass Spectrometry Reviews 21*, 37-71.

[17] Konermann, L., Tong, X., and Pan, Y. (2008) Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches, *Journal of Mass Spectrometry 43*, 1021-1036.

[18] Wales, T. E., and Engen, J. R. (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics, *Mass Spectrometry Reviews 25*, 158-170.

[19] Konermann, L., Pan, J., and Liu, Y.-H. (2011) Hydrogen exchange mass spectrometry for studying protein structure and dynamics, *Chemical Society Reviews 40*, 1224-1234.

[20] Takamoto, K., and Chance, M. R. (2006) Radiolytic Protein Footprinting with Mass Spectrometry to Probe the Structure of Macromolecular Complexes, *Annual Review of Biophysics and Biomolecular Structure 35*, 251-276.
[21] Konermann, L., Stocks, B. B., Pan, Y., and Tong, X. (2010) Mass spectrometry combined with oxidative labeling for exploring protein structure and folding, *Mass Spectrometry Reviews 29*, 651-667.
[22] Sinz, A. (2003) Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes, *Journal of Mass Spectrometry 38*, 1225-1237.
[23] Mendoza, V. L., and Vachet, R. W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry, *Mass Spectrometry Reviews 28*, 785-815.
[24] Katta, V., Chait, B. T., and Carr, S. (1991) Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry, *Rapid Communications in Mass Spectrometry 5*, 214-217.
[25] Kiselar, J. G., and Chance, M. R. (2010) Future directions of structural mass spectrometry using hydroxyl radical footprinting, *Journal of Mass Spectrometry 45*, 1373-1382.
[26] Xu, G., and Chance, M. R. (2007) Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics, *Chemical Reviews 107*, 3514-3543.
[27] Sharp, J. S., Becker, J. M., and Hettich, R. L. (2003) Protein surface mapping by chemical oxidation: Structural analysis by mass spectrometry, *Analytical Biochemistry 313*, 216-225.
[28] Hambly, D. M., and Gross, M. L. (2005) Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale, *Journal of the American Society for Mass Spectrometry 16*, 2057-2063.
[29] Gau, B. C., Sharp, J. S., Rempel, D. L., and Gross, M. L. (2009) Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding, *Analytical Chemistry 81*, 6563-6571.
[30] Aye, T. T., Low, T. Y., and Sze, S. K. (2005) Nanosecond Laser-Induced Photochemical Oxidation Method for Protein Surface Mapping with Mass Spectrometry, *Analytical Chemistry 77*, 5814-5822.
[31] Jones, L. M., B. Sperry, J., A. Carroll, J., and Gross, M. L. (2011) Fast Photochemical Oxidation of Proteins for Epitope Mapping, *Analytical Chemistry 83*, 7657-7661.
[32] Chen, J., Rempel, D. L., Gau, B. C., and Gross, M. L. (2012) Fast Photochemical Oxidation of Proteins and Mass Spectrometry Follow Submillisecond Protein Folding at the Amino-Acid Level, *Journal of the American Chemical Society 134*, 18724-18731.
[33] McClintock, C., Kertesz, V., and Hettich, R. L. (2008) Development of an Electrochemical Oxidation Method for Probing Higher Order Protein Structure with Mass Spectrometry, *Analytical Chemistry 80*, 3304-3317.
[34] Gardner, M. W., Vasicek, L. A., Shabbir, S., Anslyn, E. V., and Brodbelt, J. S. (2008) Chromogenic Cross-Linker for the Characterization of Protein Structure by Infrared Multiphoton Dissociation Mass Spectrometry, *Analytical Chemistry 80*, 4807-4819.
[35] Suckau, D., Mak, M., and Przybylski, M. (1992) Protein surface topology-probing by selective chemical modification and mass spectrometric peptide mapping, *Proceedings of the National Academy of Sciences USA 89*, 5630-5634.
[36] Glocker, M. O., Borchers, C., Fiedler, W., Suckau, D., and Przybylski, M. (1994) Molecular Characterization of Surface Topology in Protein Tertiary Structures by Amino-Acylation and Mass Spectrometric Peptide Mapping, *Bioconjugate Chemistry 5*, 583-590.

[37] Gómez, G., Mundo, M., Craig, P., and Delfino, J. (2012) Probing Protein Surface with a Solvent Mimetic Carbene Coupled to Detection by Mass Spectrometry, *Journal of the American Society for Mass Spectrometry 23*, 30-42.
[38] Mendoza, V. L., and Vachet, R. W. (2008) Protein Surface Mapping Using Diethylpyrocarbonate with Mass Spectrometric Detection, *Analytical Chemistry 80*, 2895-2904.
[39] Reid, G. E., and McLuckey, S. A. (2002) 'Top down' protein characterization via tandem mass spectrometry, *Journal of Mass Spectrometry 37*, 663-675.
[40] Vasicek, L., O'Brien, J. P., Browning, K. S., Tao, Z., Liu, H.-W., and Brodbelt, J. S. (2012) Mapping Protein Surface Accessibility via an Electron Transfer Dissociation Selectively Cleavable Hydrazone Probe, *Molecular & Cellular Proteomics 11*, 10.1074/ mcp.O111.015826.
[41] Ly, T., and Julian, R. R. (2008) Protein–Metal Interactions of Calmodulin and α-Synuclein Monitored by Selective Noncovalent Adduct Protein Probing Mass Spectrometry, *Journal of the American Society for Mass Spectrometry 19*, 1663-1672.
[42] Heo, S. W., Choi, T. S., Park, K. M., Ko, Y. H., Kim, S. B., Kim, K., and Kim, H. I. (2011) Host–Guest Chemistry in the Gas Phase: Selected Fragmentations of CB[6]–Peptide Complexes at Lysine Residues and Its Utility to Probe the Structures of Small Proteins, *Analytical Chemistry 83*, 7916-7923.
[43] Lee, J., Heo, S., Lee, S., Ko, J., Kim, H., and Kim, H. (2013) Probing Conformational Changes of Ubiquitin by Host–Guest Chemistry Using Electrospray Ionization Mass Spectrometry, *Journal of the American Society for Mass Spectrometry 24*, 21-29.
[44] Lee, S. J. C., Lee, J. W., Lee, H. H., Seo, J., Noh, D. H., Ko, Y. H., Kim, K., and Kim, H. I. (2013) Host–Guest Chemistry from Solution to the Gas Phase: An Essential Role of Direct Interaction with Water for High-Affinity Binding of Cucurbit[n]urils, *The Journal of Physical Chemistry B 117*, 8855-8864.
[45] Prabhudesai, S., Sinha, S., Attar, A., Kotagiri, A., Fitzmaurice, A. G., Lakshmanan, R., Ivanova, M. I., Loo, J. A., Klarner, F. G., Schrader, T., Bitan, G., and Bronstein, J. (2012) A novel "molecular tweezer" inhibitor of alpha-synuclein neurotoxicity in vitro and in vivo, *Neurotherapeutics 9*, 464-476.
[46] Attar, A., Rahimi, F., and Bitan, G. (2013) Modulators of amyloid protein aggregation and toxicity: EGCG and CLR01, *Translational Neuroscience 4*, 385-409.
[47] Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process, *Jouornal of the American Chemical Society 120*, 3265-3266.
[48] Zhurov, K. O., Fornelli, L., Wodrich, M. D., Laskay, U. A., and Tsybin, Y. O. (2013) Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis, *Chemical Society Reviews 42*, 5014-5030.
[49] Xie, Y., Zhang, J., Yin, S., and Loo, J. A. (2006) Top-Down ESI-ECD-FT-ICR Mass Spectrometry Localizes Noncovalent Protein-Ligand Binding Sites, *Journal of the American Chemical Society 128*, 14432-14433.
[50] Yin, S., and Loo, J. A. (2010) Elucidating the Site of Protein-ATP Binding by Top-Down Mass Spectrometry, *Journal of the American Society for Mass Spectrometry 21*, 899-907.
[51] Yin, S., and Loo, J. A. (2011) Top-down mass spectrometry of supercharged native protein–ligand complexes, *International Journal of Mass Spectrometry 300*, 118-122.
[52] Acharya, S., Safaie, B. M., Wongkongkathep, P., Ivanova, M. I., Attar, A., Klärner, F.-G., Schrader, T., Loo, J. A., Bitan, G., and Lapidus, L. J. (2014) Molecular Basis for

Preventing α-Synuclein Aggregation by a Molecular Tweezer, *Journal of Biological Chemistry 289*, 10727-10737.
[53] Sinha, S., Lopes, D. H., Du, Z., Pang, E. S., Shanmugam, A., Lomakin, A., Talbiersky, P., Tennstaedt, A., McDaniel, K., Bakshi, R., Kuo, P. Y., Ehrmann, M., Benedek, G. B., Loo, J. A., Klarner, F. G., Schrader, T., Wang, C., and Bitan, G. (2011) Lysine-specific molecular tweezers are broad-spectrum inhibitors of assembly and toxicity of amyloid proteins, *Journal of the American Chemical Society 133*, 16958-16969.
[54] Horn, D. M., Zubarev, R. A., and McLafferty, F. W. (2000) Automated de novo sequencing of proteins by tandem high-resolution mass spectrometry, *Proceedings of the National Academy of Sciences USA 97*, 10313-10317.
[55] Shaw, J. B., Li, W., Holden, D. D., Zhang, Y., Griep-Raming, J., Fellers, R. T., Early, B. P., Thomas, P. M., Kelleher, N. L., and Brodbelt, J. S. (2013) Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation, *Journal of the American Chemical Society 135*, 12646-12651.
[56] Breuker, K., Oh, H., Horn, D. M., Cerda, B. A., and McLafferty, F. W. (2002) Detailed Unfolding and Folding of Gaseous Ubiquitin Ions Characterized by Electron Capture Dissociation, *Journal of the American Chemical Society 124*, 6407-6420.
[57] Skinner, O., McLafferty, F., and Breuker, K. (2012) How Ubiquitin Unfolds after Transfer into the Gas Phase, *Journal of the American Society for Mass Spectrometry 23*, 1011-1014.
[58] Enyenihi, A., Yang, H., Ytterberg, A. J., Lyutvinskiy, Y., and Zubarev, R. (2011) Heme Binding in Gas-Phase Holo-Myoglobin Cations: Distal Becomes Proximal?, *Journal of the American Society for Mass Spectrometry 22*, 1763-1770.
[59] Kaltashov, I. A., and Mohimen, A. (2005) Estimates of Protein Surface Areas in Solution by Electrospray Ionization Mass Spectrometry, *Analytical Chemistry 77*, 5370-5379.
[60] Konermann, L., and Douglas, D. J. (1998) Unfolding of proteins monitored by electrospray ionization mass spectrometry: a comparison of positive and negative ion modes, *Journal of the American Society for Mass Spectrometry 9*, 1248-1254.
[61] Ahmad, S., Gromiha, M., Fawareh, H., and Sarai, A. (2004) ASAView: Database and tool for solvent accessibility representation in proteins, *BMC Bioinformatics 5*, 51.
[62] Fraczkiewicz, R., and Braun, W. (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules, *Journal of Computational Chemistry 19*, 319-333.

CHAPTER FIVE

Enhancing Protein Disulfide Bond Cleavage by UV Excitation and Electron Capture Dissociation for Top-Down Mass Spectrometry

Abstract

The application of ion pre-activation with 266 nm ultraviolet (UV) laser irradiation combined with electron capture dissociation (ECD) is demonstrated to enhance top-down mass spectrometry sequence coverage of disulfide bond containing proteins. UV-based activation can homolytically cleave a disulfide bond to yield two separated thiol radicals. Activated ECD experiments of insulin and ribonuclease A containing three and four disulfide bonds, respectively, were performed. UV-activation in combination with ECD allowed the three disulfide bonds of insulin to be cleaved and the overall sequence coverage to be increased. For the larger sized ribonuclease A with four disulfide bonds, irradiation from an infrared laser (10.6 μm) to disrupt non-covalent interactions was combined with UV-activation to facilitate the cleavage of up to three disulfide bonds. UV activation poorly improved disulfide bond cleavage for lysozyme and β-lactoglobulin. We conclude that preferences for disulfide bond cleavage are dependent on protein structure and sequence. Disulfide bonds can reform if the generated radicals remain in close proximity. By varying the time delay between the UV-activation and the ECD events, it was determined that disulfide bonds reform within 10-100 msec after their UV-homolytic cleavage.

Introduction

Disulfide bonds are one of the most important post-translational modifications because they aid proteins to preserve their tertiary structure [1, 2]. The development of methods for the structural elucidation of protein disulfide bonds has been on-going for many years. As antibodies and antibody-drug conjugates become more popular in the market as therapeutic drugs, the development of efficient, rapid, and accurate methods for characterizing disulfide bond linkages become more important [3, 4].

Several mass spectrometry approaches for characterizing disulfide bonds have been reported in the literature [5, 6]. A common strategy employs chemical reduction of the S-S bond (with, for example, dithiothreitol or TCEP (*tris*(2-carboxyethyl)phosphine)), followed by alkylation of the free thiols to prevent disulfide scrambling; the resulting protein is digested with a suitable protease and the peptide products are measured by LC-MS/MS. However, information of the disulfide bond linkages (i.e., residues) can be lost [5]. In some cases, partial reduction and alkylation can be used, but this results in complex mixtures of peptides with differing amounts of alkylated residues [7]. Top-down MS analysis, i.e., direct analysis and dissociation of the intact gas phase protein without prior chemical/enzymatic fragmentation into smaller peptides, has been considered because it skips the time-consuming proteolysis and chromatographic separation steps [8, 9]. The number of disulfide bonds can be directly by determined from an accurate intact protein mass measurement and comparison to the theoretical sequence mass. However, top-down analysis of disulfide bonded proteins remains challenging due to the limited efficiency of disulfide bond fragmentation. Backbone fragments often still remain connected by S-S bridges and it becomes difficult to interpret the product ion mass spectra to obtain sequence information within disulfide-bridged regions [10-12].

More than a decade ago, McLafferty's group introduced the development and application of electron capture dissociation (ECD) with MS for protein sequencing [13, 14]. The early ECD-MS work demonstrated its capability to not only cleave backbone bonds for large proteins, but also disulfide bonds for peptides [15], while leaving labile post-translational modifications and non-covalent interactions intact. Possible mechanisms of ECD related to disulfide bonds include the migration of a hydrogen radical ($H^{\bullet}$) from a neutralized ammonium group at an electron capture site to a disulfide bond to form $–S^{\bullet}$ and –SH (Cornell mechanism), or another possibility is that an S-S bond captures an electron to generate $–S^{\bullet}$ and $–S^{-}$, which is converted to –SH later (Utah-Washington mechanism) [16-18]. Electron transfer dissociation (ETD) [19] has also shown some capabilities for disulfide bond cleavage [20, 21]. However, the efficiency for disulfide bond cleavage for *proteins* is relatively low, even with additional collisional and vibrational excitation [22].

There have been several alternative techniques proposed to cleave disulfide bonds for characterization by top-down MS. These include online electrolytic reduction prior to electrospray [23-26], activation/dissociation of metal cation-adducted species [27], and radical-driven approaches [28]. Our research group previously introduced supercharging reagents to enhance charging of native proteins and complexes [29-31]. We demonstrated that by activating higher charged proteins, ECD could access more sequence information, especially for proteins containing multiple disulfide bonds [32].

Laser-based approaches have also been tested to tackle this problem. Because the maximum absorption of a disulfide bond is in the vacuum-UV region (around 150 nm for cystine [33]), 157-nm ultraviolet photodissociation (UVPD) was used to target disulfide bond cleavage; however, some competitive dissociation channels of protein backbone were observed [34]. Julian's group

demonstrated the use of 266 nm UVPD to provide more selective disulfide bond cleavage in peptides [35]; this work was based on a possible mechanism that utilized electronic energy transfer (EET) from neighboring tyrosine or tryptophan residues [36]. However, there remains the question on whether any of these methods can be applied to larger molecules such as proteins. In this work, we demonstrate that 266 nm UVPD could be combined with vibrational excitation through infrared laser activation to give rise to backbone cleavage for small proteins, and we show evidence that suggests a strong relationship between disulfide bond cleavage and EET via adjacent tyrosine residues.

Experimental

Materials

Two disulfide bond-containing proteins, bovine insulin (UniprotKB P01317) and bovine ribonuclease A (UniprotKB P61823) were purchased from Sigma-Aldrich (St. Louis, MO). The proteins were resuspended in water and desalted by centrifugal filtration using 3 kDa MW cutoff (MWCO) membranes for insulin and 10 kDa MWCO for ribonuclease A. The desalted proteins were dissolved in 50:50:0.1 (*v*/*v*/*v*) H_2O/CH_3CN/formic acid.

Instrumentation

A 15-Tesla SolariX FT-ICR mass spectrometer equipped with an infinity cell (Bruker Daltonics, Bremen, Germany) was used for the high-resolution top-down MS experiments. Two types of lasers were used for the experiments. A 10.6 μm (30 W) infrared beam was generated by a CO_2 laser (Synrad, Mukilteo, WA). UV radiation (266 nm) was generated from the 4th harmonic of a Nd:YAG laser (Continuum, Santa Clara, CA). A BaF_2 window was mounted on the end flange of the vacuum system to allow the laser beams to pass through to the infinity ICR cell. A fused-silica window beamsplitter (Esco Optics, Oak Ridge, NJ) was aligned 45° to both the infinity cell and the 10.6 μm IR beam, acting to reflect the IR beam partially while transmitting the 266 nm UV beam (Figure 5-1). Because the UV and IR laser beams propagated to the center of the infinity cell overlapped and were co-linearly aligned, UV and IR irradiation can be applied individually or together without realignment. Both lasers were triggered by unique TTL pulses from the FT-ICR mass spectrometer data system, controlled by a customizable pulse program. An iodinated-tyrosine containing peptide (*N-4*-iodobenzoyl-RGYALG) was used to test the alignment of the UV laser by generating the $[RGYALG]^{\bullet}$ radical from homolytic C–I bond cleavage [37].

Samples were directly infused into the FT-ICR mass spectrometer by nanoelectrospray ionization (nanoESI) using Au/Pd coated borosilicate emitters (Thermo Scientific, San Jose, CA). The ESI voltage was set to 800-1000 V. The instrument parameters were set to the following: glass transfer capillary temperature, 180 ^{o}C; applied voltages to ion funnel and skimmer were 120 and 50 V, respectively; ion source RF frequency was set to 200 V_{pp}; quadrupole and hexapole RF frequencies were 2 MHz, 1200 V_{pp}. Ions were accumulated for 0.5 sec before sending them to the ICR cell with the time-of-flight of 1.2 msec. The signal transient length was set to 2.8 sec to achieve a resolving power of 800,000 at *m/z* 400. The isolation power was to 35%. Precursor ions were isolated by the quadrupole and transferred into the ICR cell to perform ECD or activated-ion ECD (activation by UV or IR irradiation or both). The electron energy for ECD was 1 eV (10 ms pulse and 15 V on the lens). ECD fragments were detected by the SNAP centroid peak detection algorithm (Bruker Daltonics) and manually assigned against theoretical fragments obtained from Protein Prospector. Mass spectra were externally calibrated with CsI.

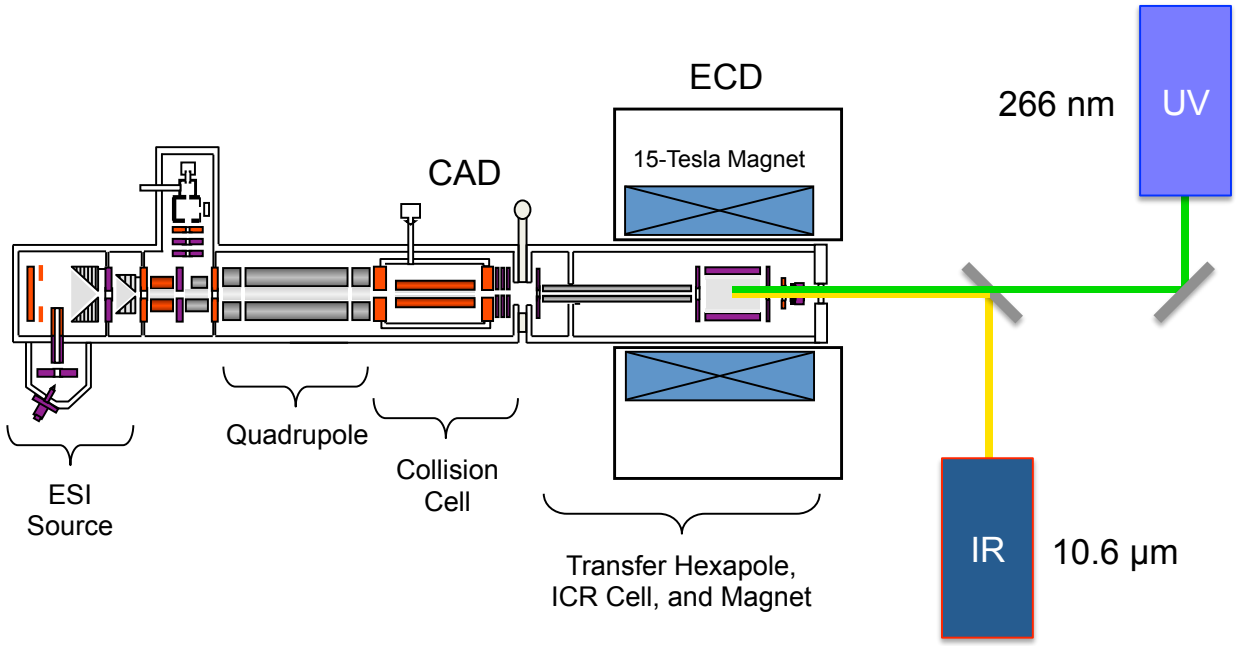

Figure 5-1. Instrument diagram of the modified 15-Tesla FT-ICR mass spectrometer interfaced to the UV and IR lasers.

Results and Discussion

Structural Overview

Insulin consists of two peptide chains (chain A: 21 residues, and chain B: 30 residues). It is first synthesized in beta cells as a single chain, proinsulin. After being processed by proteases, cleaving off the C-peptide in the middle of its sequence, it becomes mature insulin [38]. Two intermolecular disulfide bonds connect chain A and B: $C7_A$-$C7_B$, $C20_A$-$C19_B$. There is another intramolecular disulfide bond linking C6 and C11 on chain A. Ribonuclease A (RNaseA; bovine) is a small 13.7 kDa single-chain protein that has a complex intramolecular linkage, containing four disulfide bonds. The amino acid sequences of insulin and ribonuclease A are shown in Figure 5-2.

The ESI mass spectrum of denatured insulin shows a narrow charge state distribution from 6+ to 4+. The mass spectrum of denatured RNaseA showed relatively low charging from 12+ to 5+ compared to the disulfide-reduced form [39-41], with the maximum intensity measured for the 8+ charged molecule. Disulfide bonds prevent the protein to be completely unfolded and fully protonated to yield maximum charging (Figure 5-3).

UV irradiation separates insulin chains and preferentially fragments RNaseA

To study how 266 nm UV activation affects protein fragmentation, a series of short pulses (~8 nsec, 4 mJ per pulse) was fired and directed into the ICR cell where ions were trapped. Disulfide bond cleavage by photodissociation was readily observed when 10 laser shots per spectra were applied. For insulin, ions for both chains A and B were detected, confirming that the two intermolecular disulfide bonds were cleaved by PD and yielded a separation of both A and B chains (Figure 5-4a). The measured masses of both chains confirmed that homolytic

cleavage (—S• •S—) occurred. However, the PD yield was low; relative abundances for the liberated chains A and B were only 1% relative to the precursor ion abundance. This result might be expected because both intermolecular disulfide bonds need to be cleaved to separate the two chains. A number of y-product ions from the C-terminal portion of the B-chain outside of the disulfide bonds were observed, such as $y_{4,B}$, $y_{5,B}$, and $y_{6,B}$ (where the "B" denotes that it originates from chain B), in addition to b-fragments from chain B connected to intact chain A (e.g., $b_{24,B}$+A, $b_{25,B}$+A, $b_{26,B}$+A). The observed collisionally activated dissociation (CAD)-type ions (i.e., b/y) can occur following internal conversion of the photon energy into vibrational excitation and should yield fragments similar to what is observed by CAD or IRMPD.

For RNaseA, there was no direct evidence of disulfide bond cleavage from only PD because the protein mass remains the same after homolytic cleavage(s) of the disulfide bond(s). Interestingly, cleavage of a bond between Val_{116}-Pro_{117} to generate the y_8 ion ($PVHFDASV_{124}$) was found for PD of the $[M+11H]^{11+}$ molecule (Figure 5-4b). This may suggest that UV irradiation was absorbed by the protein probably from a nearby tyrosine (Tyr-115). Formation of the y_8 ion was likely from UV photoactivation resulting in an elevated internal energy to produce a favorable cleavage at the proline residue [42, 43]. The abundance of the y_8 ion was about 2% relative to the precursor. The complimentary ion, $[M-y_8+10H]^{10+}$, was observed also, in addition to charge stripped $[M+10H]^{10+}$ and $[M-y8+9H]^{9+}$. From this evidence for insulin and ribonuclease A, 266 nm irradiation can photodissociate and photoactivate small proteins, but the PD yields are relatively low.

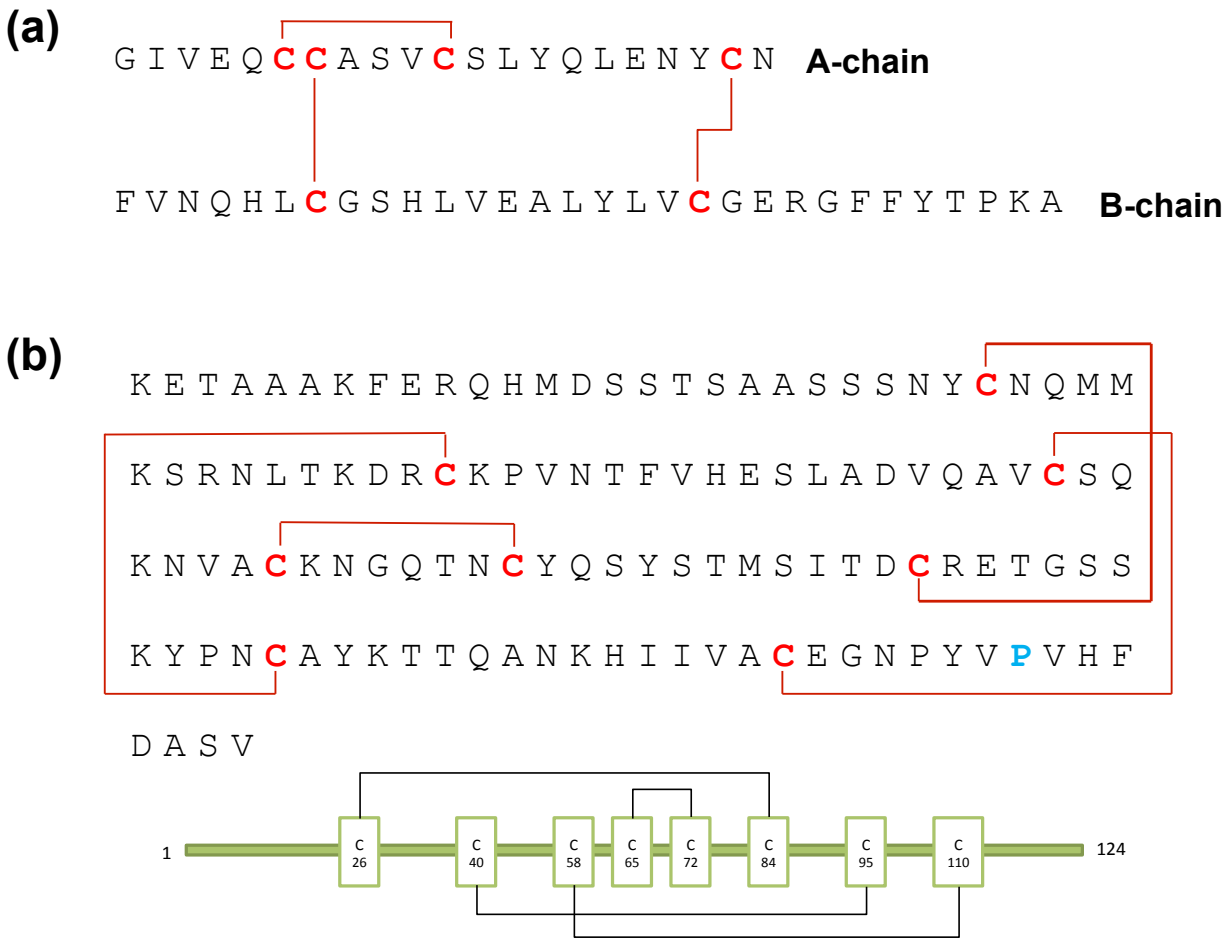

Figure 5-2. Amino acid sequence and disulfide bonds of (a) bovine insulin and (b) bovine ribonuclease A.

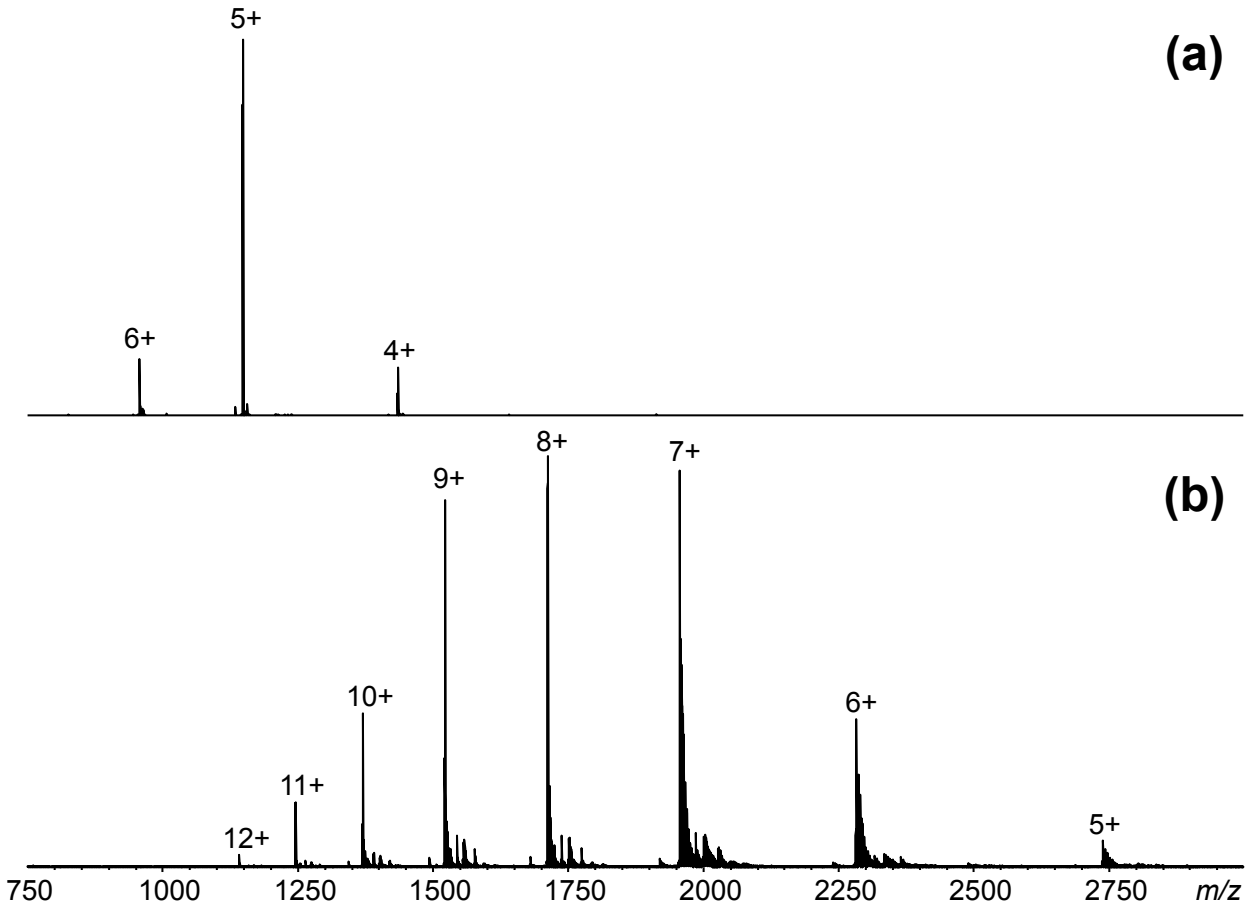

Figure 5-3. ESI mass spectra of (a) insulin and (b) ribonuclease A.

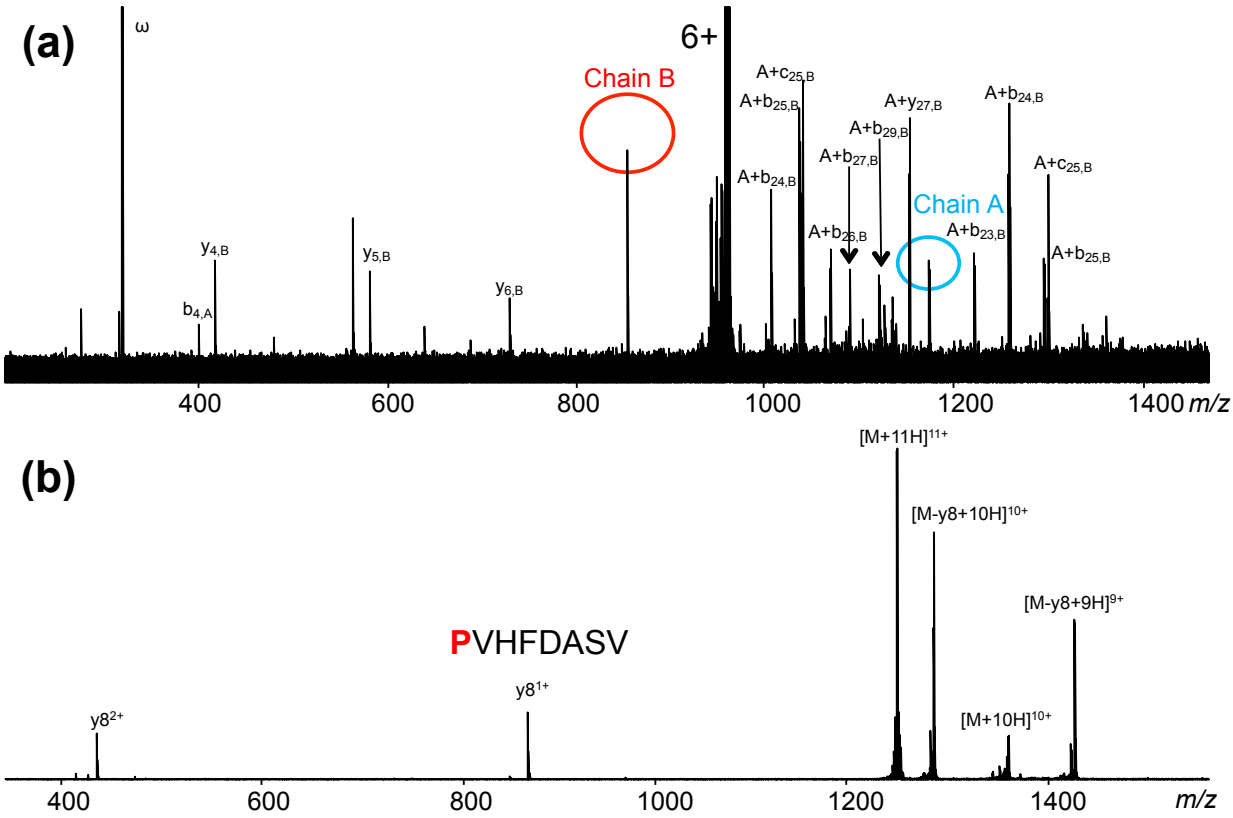

Figure 5-4. UVPD (266 nm) mass spectra of (a) 6+ charged insulin and (b) 11+ charged ribonuclease A.

Disulfide bond cleavage of insulin by UV-enhanced ECD

Although ECD alone can cleave disulfide bonds for disulfide-linked peptides, the efficiency is low for proteins and it may be sequence dependent [22]. In some cases, such as 20 kDa trypsin inhibitor, ECD does not promote disulfide bond cleavage. In our study, UV activation combined with ECD was aimed to enhance disulfide bond cleavage. Insulin and RNaseA were tested with ECD alone, and with some pre-activation by UV and IR irradiation. Backbone cleavage of regions enclosed by disulfide bonds will be indicative of the presence of disulfide links.

The 6+ charge state, which was the highest charge observed for insulin, was isolated and transferred into the ICR cell to perform ECD or UV-activated ECD. ECD mass spectra were acquired from an average of 200 microscans. Upon ECD, two charge-reduced species ($[M+6H]^{5+\bullet}$ and $[M+6H]^{4+\bullet\bullet}$) were observed. Small c/z$^{\bullet}$-type fragments were observed between *m/z* 400-1000 (e.g., $c_{5,A}$, $c_{4,B}$, $c_{6,B}$, $z_{8,B}$, $z_{9,B}$). Larger fragments linked to one of the intact chains were detected at *m/z* 1200-1800 (for instance, $c_{23,B}$+A, $c_{28,B}$+A, $z_{27,B}$+A, $z_{29,B}$+A, $c_{20,A}$+B, $z_{16,A}$+B, $z_{26,A}$+B). However, the majority of the fragments were from outside both intermolecular disulfide bonds ($C7_A$-$C7_B$, $C20_A$-$C19_B$). Also observed were b/y-type product ions ($b_{4,A}$, $y_{4,B}$, $y_{6,B}$, $y_{7,B}$, $y_{8,B}$, $y_{9,B}$, and $y_{20,A}$+B). Figure 5-5 summarizes the backbone cleavage coverage of ECD of the insulin 6+ charge state. Interestingly, six fragments showed evidence of intermolecular disulfide bond cleavage of $C20_A$-$C19_B$: $z_{13,B}$, $z_{15,B}$, $z_{16,B}$, $c_{14,B}$+A, $c_{19,B}$+A, and $c_{17,A}$+B (Figure 5-6a). The ECD results are consistent with previous reports by McLafferty, McLuckey, and Breuker [15, 22, 44]. The loss of S and •SH was observed, which is also an indication of disulfide bond cleavage. From the c/z$^{\bullet}$-fragments when only ECD was applied, only one disulfide bond ($C20_A$-$C19_B$) was cleaved, and the other two S-S bonds remained intact. Overall, 51 product ions were observed and cleavage of 51% of the total possible inter-residue linkages was achieved.

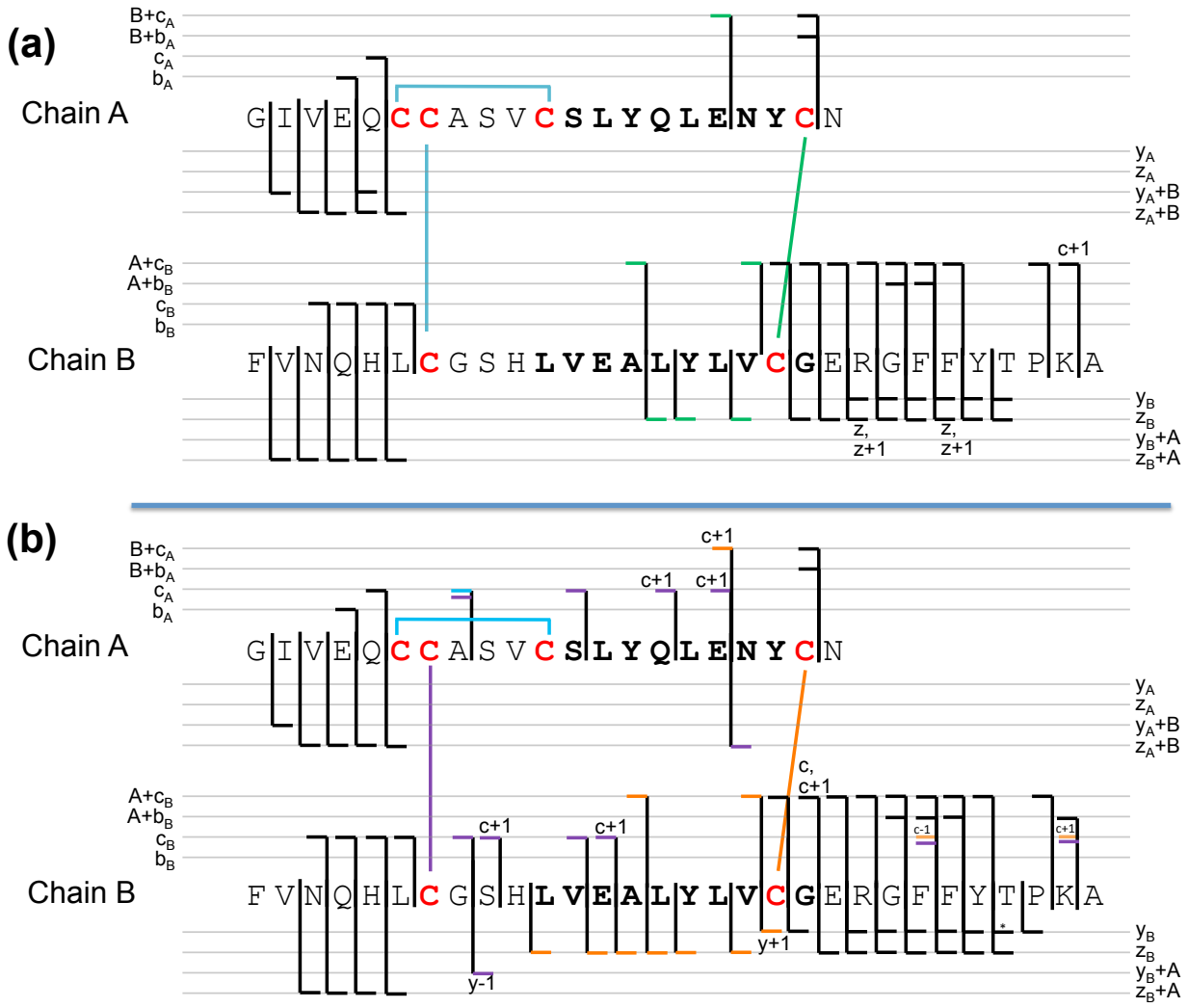

Figure 5-5. Cleavage pattern of insulin by (a) ECD and (b) UV-activation/ECD showing one and three disulfide bonds cleaved, respectively. Fragments are color coded according to the corresponding cleaved disulfide bonds

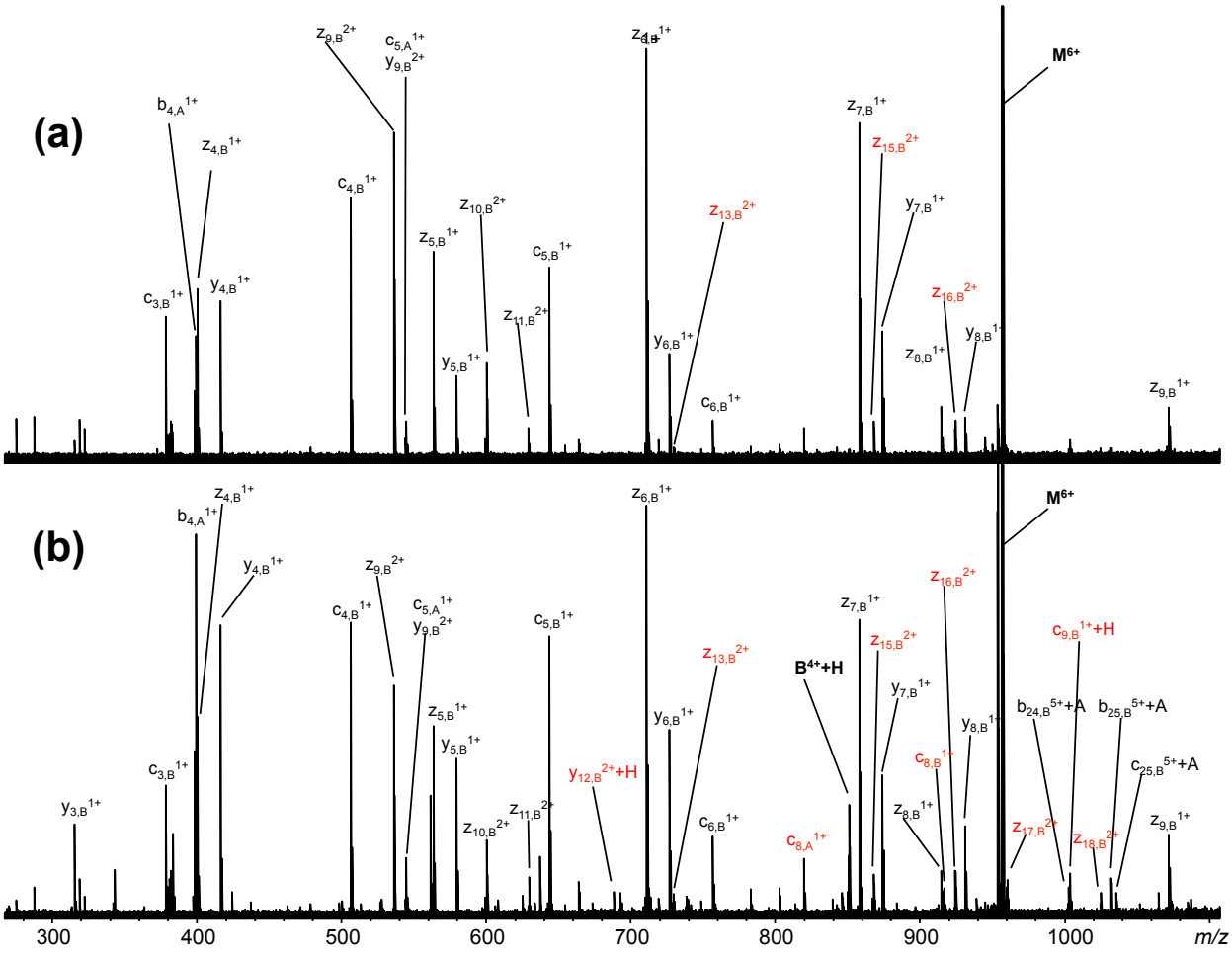

Figure 5-6. Fragmentation mass spectra of 6+ charged insulin using (a) ECD and (b) UV-activation/ECD. Product ions confirming disulfide bond cleavages are highlighted in red.

For UV-assisted ECD, 10-pulses of the UV laser were programmed to fire immediately prior to ECD, with no additional time delay between the UV laser and ECD events. From the results, it is clear that UV activation improved the ECD efficiency for disulfide bond cleavage. The total number of backbone cleavages increased slightly from 51 to 68, but most of the new products (15) were from regions within the formerly disulfide bond-enclosed area. These newly created fragments were not observed with ECD alone. The fragmentation pattern indicated that all three disulfide bonds were cleaved. Similar to ECD alone, products $z_{13,B}$, $z_{15,B}$, $z_{16,B}$, $c_{14,B}$+A, and $c_{19,B}$+A indicated dissociation of the intermolecular $C20_A$-$C19_B$ bond (Figure 5-6b). Evidence for $C7_A$-$C7_B$ bond dissociation is given by product ions $c_{8,A}$, $c_{11,A}$, $c_{14,A}$, $c_{16,A}$, $c_{8,B}$, $c_{9,B}$, $c_{12,B}$, $c_{13,B}$, $y_{22,B}$+A, and $z_{4,A}$+B. The $c_{8,A}$ ion confirmed that the $C6_A$-$C11_A$ disulfide bond was cleaved. Moreover, some fragments ($c_{25,B}$ and $c_{29,B}$) suggested (near) simultaneous cleavage of both intermolecular disulfide bonds to yield the separation of chains A and B, as found for UVPD alone (*vide supra*). Upon UV photoactivation prior to ECD, 73% of the total possible inter-residue linkages were cleaved. Out of the 68 total fragments observed, 31% (21) originated within the formerly disulfide bonded regions.

Disulfide bond cleavage efficiency was further quantified as the relative abundance of the disulfide bond derived fragments (%SS abundance) as defined by:

Equation 5-1.

$$\%SS\ abundance = \frac{\sum product\ ion\ intensity\ from\ disulfide\ bridged\ region}{total\ product\ ion\ intensity\ observed}$$

The %SS abundance increased from 1.6% for ECD alone to 8.1% for UV activation/ECD, indicating the significant improvement of using UV pre-activation for ECD of disulfide bonds for insulin.

UV with vibrational excitation can assist ECD to cleave RNaseA disulfide bonds

Ribonuclease A has four intramolecular disulfide bonds (C26-C80, C40-C95, C58-C110, and C65-C72) that predominantly cover residues 26-110, which represent 68.5% of the protein sequence (Figure 5-2). The $[M+11H]^{11+}$ and $[M+10H]^{10+}$ molecules were chosen for the ECD and UV-activated ECD experiments. The c/z$^{\bullet}$-ion abundances were normalized and displayed against each cleavage site; the eight cysteine residues that form the four intramolecular disulfide bridges are displayed as vertical lines in the fragmentation plots (Figure 5-7).

ECD of RNaseA yielded c/z$^{\bullet}$-fragments that were solely from the regions that are not bridged by disulfide bonds, toward the N- and C-terminal ends (i.e., c_{2-24}, c_{121}, c_{122}; z_6, z_9, z_{12}, z_{121}' and z_{122}'). Increasing collisional activation prior to ECD by increasing the skimmer voltage to 60 V, as suggested by Ganisl and Breuker [22], did not yield fragments in the disulfide-bridged region. We also implemented an additional vibrational excitation by IR laser irradiation prior to the ECD event, generally known as activated ion ECD (ai-ECD), to disrupt weak noncovalent interactions such as hydrogen bonds that may be holding c-/z-product pairs together [45] and can significantly improve ECD fragmentation efficiency [46-48]. Elevated internal energy from IR irradiation also enhances secondary ECD fragmentation that may play a role in disulfide bond cleavages. In this experiment, the ai-ECD of RNaseA generated four c-fragments from residues 33-36, which are in the region covered by the first disulfide bond (C26-C84), but not in the second (C40-C95) (Figure 5-7b). Three of the fragments, c_{34-36}, contained an extra hydrogen atom suggesting that the disulfide bond cleavage mechanism is from ECD and hydrogen radical migration.

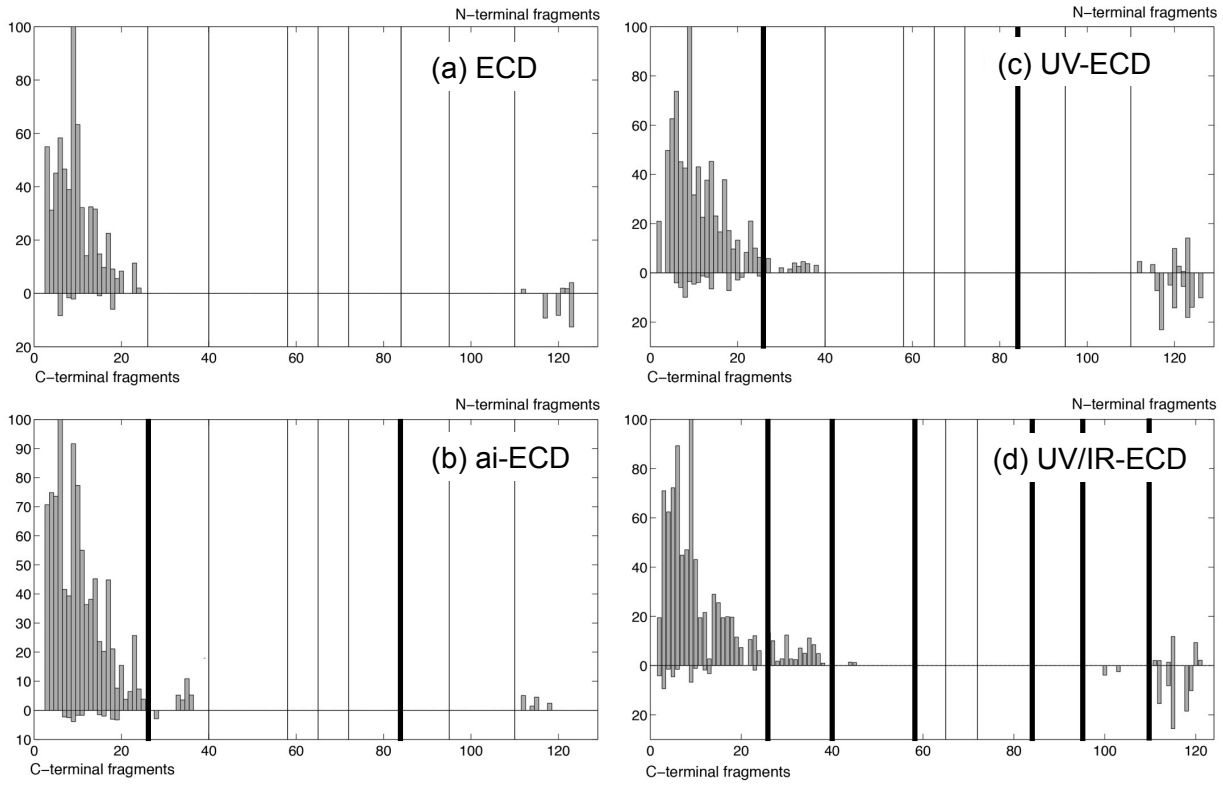

Figure 5-7. Relative abundances of backbone fragmentation for ribonuclease A using (a) ECD, (b) ai-ECD, (c) UV-ECD, and (d) UV/IR-ECD. The vertical lines represent cysteine residues that form disulfide bonds. The bolded lines indicated the cysteines where the disulfide bond cleavage occurred. Bars above the zero level represent N-terminal containing product ions (b- and c-products) and the bars below the zero level represent C-terminal containing product ions (y- and z-products).

Activated ion-ECD yielded only a single disulfide bond cleavage event in RNaseA, and so we investigated additional activation methods to facilitate disulfide bond cleavage from ECD radical cascade reactions. UV-activation prior to the ECD event was tested. Ten consecutive UV pulses were used without significantly reducing precursor signal intensity. Similar to ai-ECD, some fragments from residues 26-38 were observed (Figure 5-7c), indicating a single disulfide bond (C26-C84) was cleaved. Fragmentation from the disulfide-bridged region was slightly improved compared to ai-ECD; five newly-observed c-fragments (c_{26}+H, c_{27}+H, c_{30}+H, c_{32}+H, and c_{38}+H) were observed by UV-ECD. A total of 9 c-type ions out of 15 possible backbone cleavage sites between the first and second disulfide bonds (residues 26-39) were observed. However, the overall normalized abundance of fragments from the C26-C84 bridge region was low, around 5% relative to small fragments from the N-terminal region. The large number of small N-terminal fragments might indicate a secondary dissociation from secondary electron capture of larger fragments because most disulfide bonds were still intact. Along the C-terminal region outside the disulfide bridges, more c/z$^{\bullet}$-fragments were observed with higher abundance when UV-ECD was used.

To further improve the prospects for top-down protein disulfide bond cleavage, we utilized both UV- and IR-based activation prior to ECD by firing 10-shots of UV light followed by a 60-msec IR pulse. Because both UV and IR irradiation were used, the IR laser irradiation time was reduced from 120-150 msec (for normal ai-ECD experiments) to 60 msec to minimize CAD-type fragmentation prior to electron capture. Figure 5-8 shows the ECD-MS spectra with UV followed by IR activation (UV/IR) compared to ECD alone. Several new fragments were detected between *m/z* 1000-1500 from UV/IR-activated ECD. In the region of residues 26-39 that is covered by the first disulfide bond, 13 out of 14 available backbone cleavages (93%) were detected. Interestingly, c_{44}+2H and c_{45}+2H fragments were observed (Figure 5-7d), which

results from cleavage of *both* the first (C26-C84) and second (C40-C95) disulfide bonds. Unlike fragments from the region between residues 26-39, the cleavage at sites 44 and 45 contains additional two hydrogen atoms, which is evidence for breaking two disulfide bonds. Moreover, the detection of z_{21} and z_{24} fragments confirmed that the third disulfide bond (C58-C110) was cleaved (but only one disulfide bond cleavage is required to observe these fragments). Three representative ECD-type fragments, which indicated cleavage of the first, second, and third disulfide bonds, are shown in the mass spectra depicted in Figure 5-9. The loss of S and •SH, which is evidence of disulfide bond cleavage, was observed when UV was used and was more abundant when both UV- and IR-activation were applied. Overall, using both UV- followed by IR-activation provided the best results for RNaseA disulfide bond cleavage. The %SS abundance results are summarized in Figure 5-10a.

Further, the affects of reversing the chronological order of the two laser irradiation events prior to ECD was investigated (Figure 5-10b). With the IR excitation event immediately before UV-activation (IR/UV), the %SS abundance was slightly lower compared to UV/IR, but still higher than both single laser activation experiments (i.e., ai-ECD, UV-ECD). We considered that the radicals from UV homolytic cleavage could recombine if the two cysteine residues are in close proximity. Application of IR heating immediately after UV excitation could allow for the separation of the two thiol radicals immediately after their generation. IR activation mitigated disulfide bond reformation by reducing intramolecular noncovalent interactions. UV/IR-activation yielded more ECD fragmentation from the disulfide-bridged regions than the IR/UV approach.

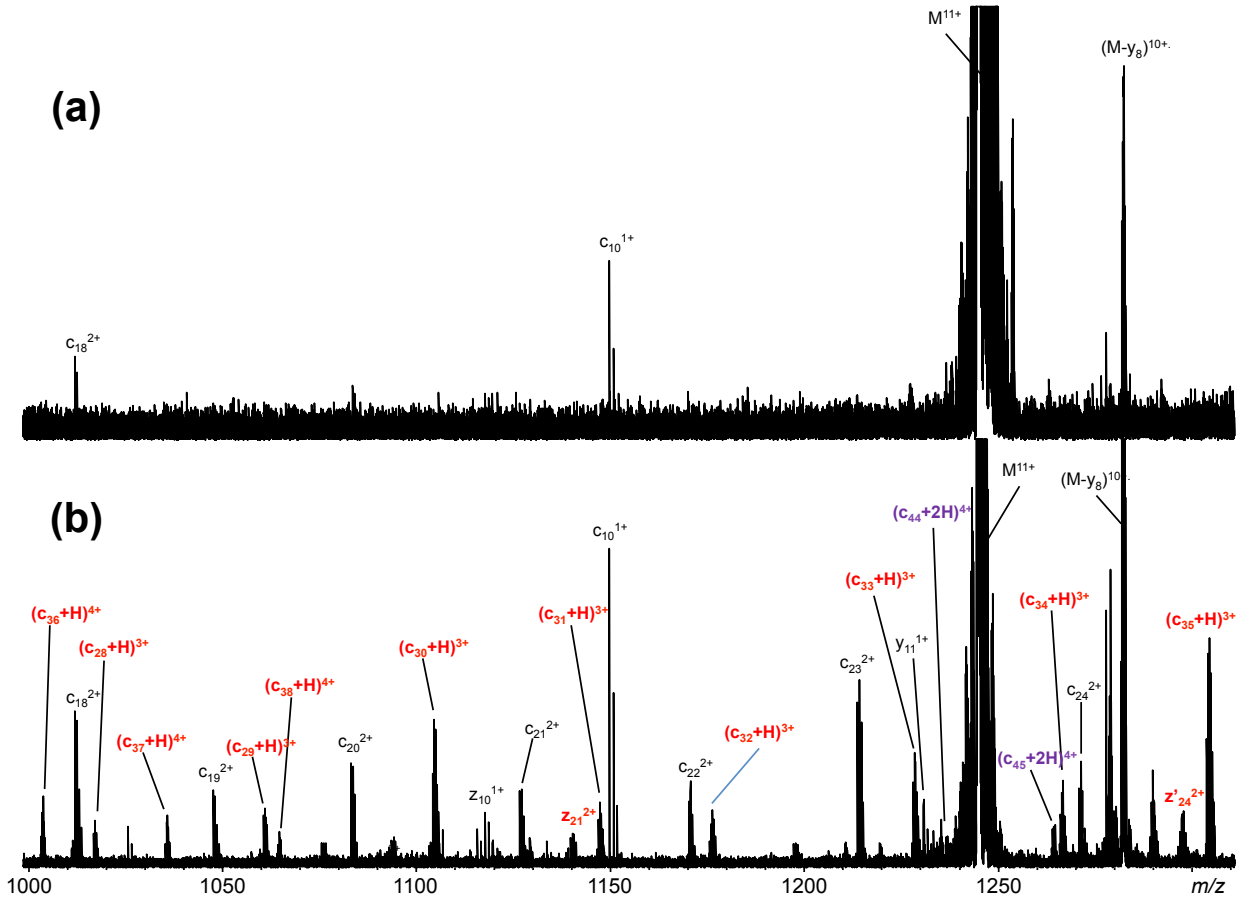

Figure 5-8. (a) ECD mass spectra of the 11+ charged ribonuclease A and (b) with UV/IR-activation prior to ECD. Fragments from disulfide bridged regions are highlighted in red and purple.

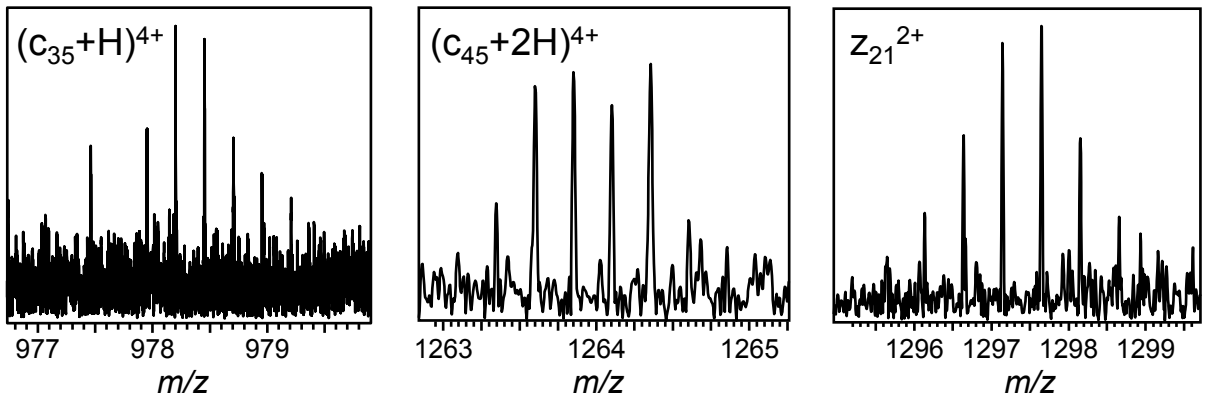

Figure 5-9. UV/IR-activated ECD products (c_{35}+H, c_{45}+2H, and z_{21}) represent cleavage of the first, second, and third disulfide bonds, respectively, from ribonuclease A.

We further investigated the kinetics of disulfide bond reformation. A short time delay was set between the UV laser (both 5 and 10 shots) and the ECD event with no additional vibrational activation applied (i.e., UV-ECD). The purpose of this experiment was to estimate the timescale of radical reformation. The delay ranged from 10 msec to 1 sec. The results (Figure 5-10b) showed that number of fragments from disulfide-bridged regions dropped significantly even with a 10 msec delay. Disulfide bond cleavage was no longer observed with a delay of longer than 100 msec, suggesting that the timescale of disulfide bond reformation was between 10-100 msec. The %SS abundance decayed slower with the greater number of UV laser shots (10) because more radicals from homolytic cleavage were generated.

Disulfide bonds can be cleaved by UV activation prior to ECD, but why are some bonds preferentially cleaved compared to others? The 266 nm wavelength of the UV laser used is close to an absorption band for aromatic residues like tyrosine and tryptophan. Disulfide bonds have some absorption in this region, but it is far less than for the aromatic residues [36, 49]. Julian's group described that energy absorbed by aromatic residues can be transferred to nearby cysteine residues by excitation energy transfer (EET) and it is strictly distance dependent [36]. Tryptophan has a working range of around 15 Å to facilitate disulfide bond cleavage, whereas tyrosine requires a shorter distance of less than 6 Å. Another report suggested that 270-290 nm UV excitation of tryptophan residues in solution is followed by transfer of electrons to the S-S bonds, resulting in their reduction, confirming the importance of tryptophan and tyrosine residues [50]. Insulin has no tryptophans, but $Y19_A$ is next to the $C20_A$ of the $C20_A$-$C19_B$ disulfide bond. Interestingly, fragments from $C20_A$-$C19_B$ cleavage exhibited higher signal intensities compared to other disulfide bond cleavages, and thus suggesting that the $C20_A$-$C19_B$ bond is the first disulfide bond that is preferentially cleaved by UV-ECD. The C26-C84 disulfide bond of RNaseA is preferentially cleaved by UV-ECD and UV/IR-ECD; C26 is adjacent to Y25. The

second disulfide bond (C40-C95) is in close proximity to Y97, where the distance is two residues apart. Unfortunately, the efficiency of EET would be much lower, resulting in no fragments observed between sites 40-57 when only UV activation was used. Fewer fragments were observed when IR was combined with UV activation. Although C65-C72 is next to Y73, evidence for disulfide bond cleavage was not observed because the C65-C72 bridge is located in the middle of the sequence that is covered by the other three disulfide bonds. The single C65-C72 disulfide bond and some N-C_α bonds might be cleaved, but the fragments likely remained bound to the protein. All four disulfide bonds must be cleaved in order to separate the ECD fragments within C65-C72 region. Overall our results are in agreement with the EET model described previously.

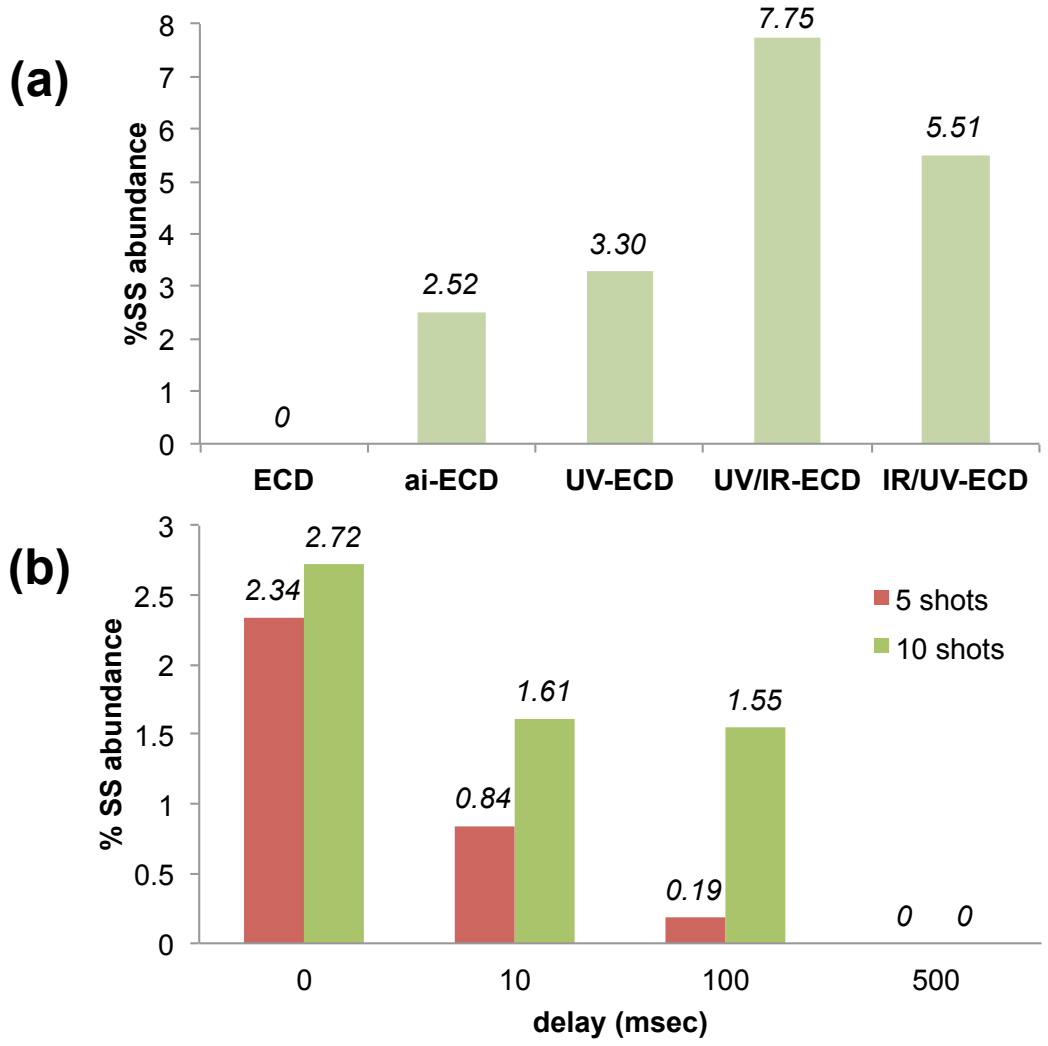

Figure 5-10. Comparison of %SS abundance for ECD of ribonuclease A (a) between different activation conditions and (b) as a function of delay time between UV-activation and the ECD event for UV-ECD.

To further investigate the UV activation, an additional two intramolecular disulfide bond proteins were tested. The first protein was lysozyme, which contains four intramolecular disulfide bonds. Lysozyme is a good system to test because it has the same number of bonds as RNaseA, but the larger disulfide-bonded region is covering 94.5% of the sequence, while RNaseA is 68.5% covered. The outermost bond links between C6 and C127, creating the tight lock in its structure starting from near N-terminus to C-terminal end. Thus it makes lysozyme difficult to study by top-down MS. Another protein in this study is β-lactoglobulin. There are two forms of β-lactoglobulin, A and B. In this experiment we focused only on the B form, which has two intramolecular disulfide bond (C66-C160 and C106-C119). ECD with additional activation approaches were used in the same manner as RNaseA, including ECD only, ai-ECD (with IR vibrational activation), UV-ECD, and IR/UV-ECD. Fragments were plotted against residue number (Figure 5-11).

For lysozyme, ECD and ai-ECD yielded similar results. The outside disulfide bond is the only bond cleaved. UV-ECD generated more fragments between C6 and C30, where the second bond is. It indicated that the C6-C127 bond is cleaved more efficient than ECD and ai-ECD, However, there is no further improvement when IR and UV were combined. The best achievement is from UV-ECD or IR/UV-ECD, which were able to to break one disulfide bond. %SS abundance from 10+ and 11+-charge state were summarized in Figure 5-12a. Overall, additional activation methods were not able to improve ECD to cleave disulfide bonds. In this case, %SS abundance is relatively high compared to other proteins because the majority of sequence is covered by disulfide bonds. Most observed fragments (>50%) are from at least one disulfide bond cleavage. The order of UV and IR activation did not significantly alter the result.

UV activation did not appear to enhance disulfide bond cleavage in β-lactoglobulin. Figure 5-11 shows c/z$^{\bullet}$ fragments of β-lactoglobulin from different activation methods plotted against a residue number. Again, one disulfide bond (C66-C160) was successfully cleaved by all techniques. Number of z$^{\bullet}$ and z+H were observed between residue 149-159, near the C-terminus. A small abundance of z_{65}+H fragment locating in the middle of the sequence was also observed in both ai-ECD and UV-ECD. This fragment is also from a single bond cleavage. However, z_{65}+H was not observed in IR/UV-ECD, suggesting that the combination between two-laser might not always give the best result. UV/IR-ECD, where the UV laser was fired before the IR event, was not tested for this protein. From %SS abundance values from three charge states in Figure 5-12b, it clearly demonstrated that UV-ECD and IR/UV-ECD did not assist disulfide bond cleavage for β-lactoglobulin.

Insulin and ribonuclease A have tyrosine in a close proximity to a cysteine residue (Y19-C20 on insulin A chain, and Y23-C24 on RNaseA). Lysozyme has three tysosines where β-lactoglobulin has four. However, none of tyrosine are located adjacent to cysteine residues. The distance from tyrosine, which is a chromophore for 266 nm, is too far for the EET process. Lysozyme has many tryptophan residues. One of them is W28, which is close to the second disulfide bond (C30-C115). W62 and W63 are also located right next to C64-C80 disulfide linkage. These three tryptophan residues may assist UV absorption and transfer energy to break disulfide bond via EET. However, to generate fragments in this region requires a cleavage of the outermost disulfide bond C6-C127 first. It may explain why there were not many fragments observed between residues 30-115.

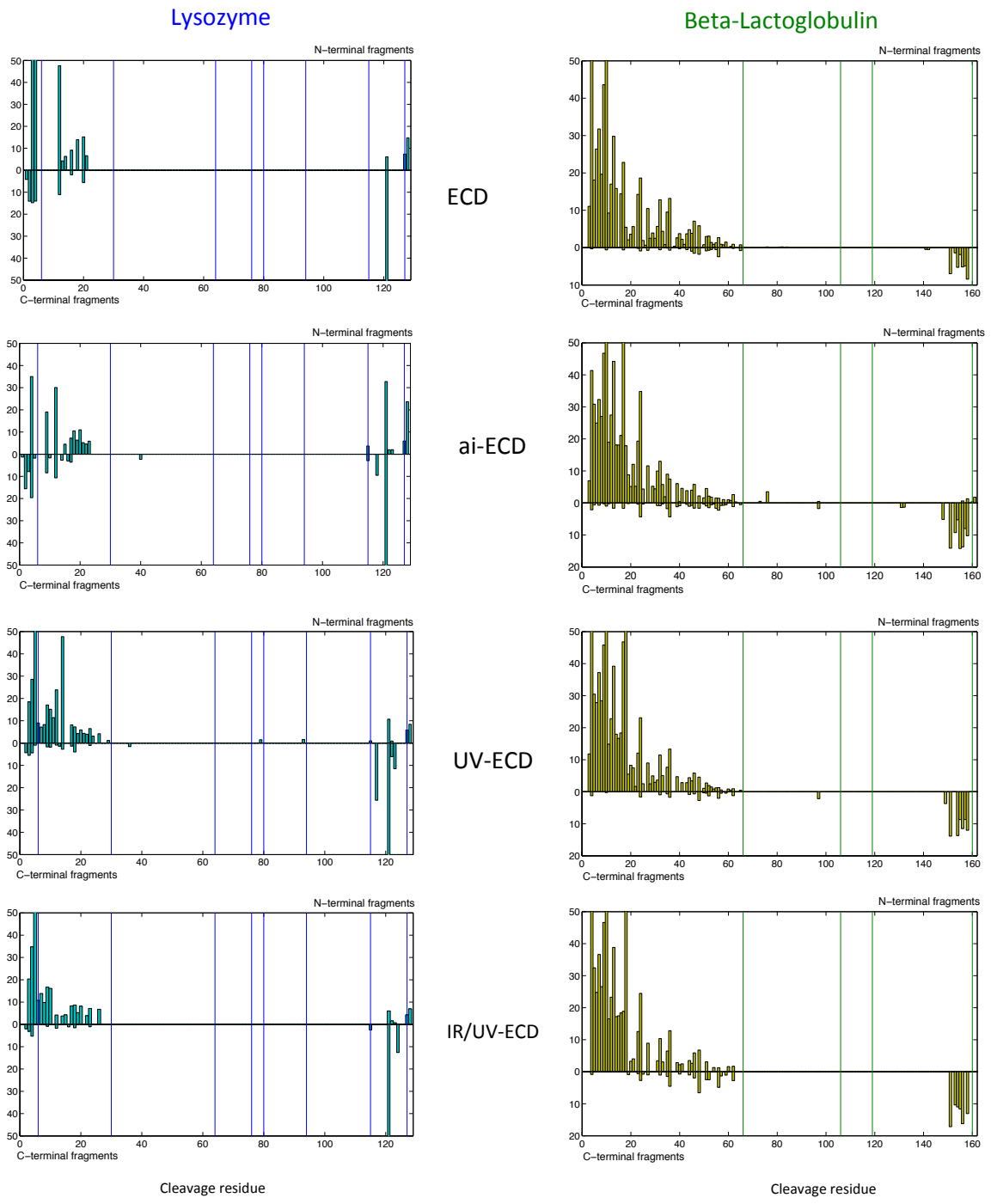

Figure 5-11. Relative abundance plot of c/z• fragments from different activation methods for lysozyme (left) and β-lactoglobulin (right). Vertical lines are the cysteine residues of the disulfide bonds.

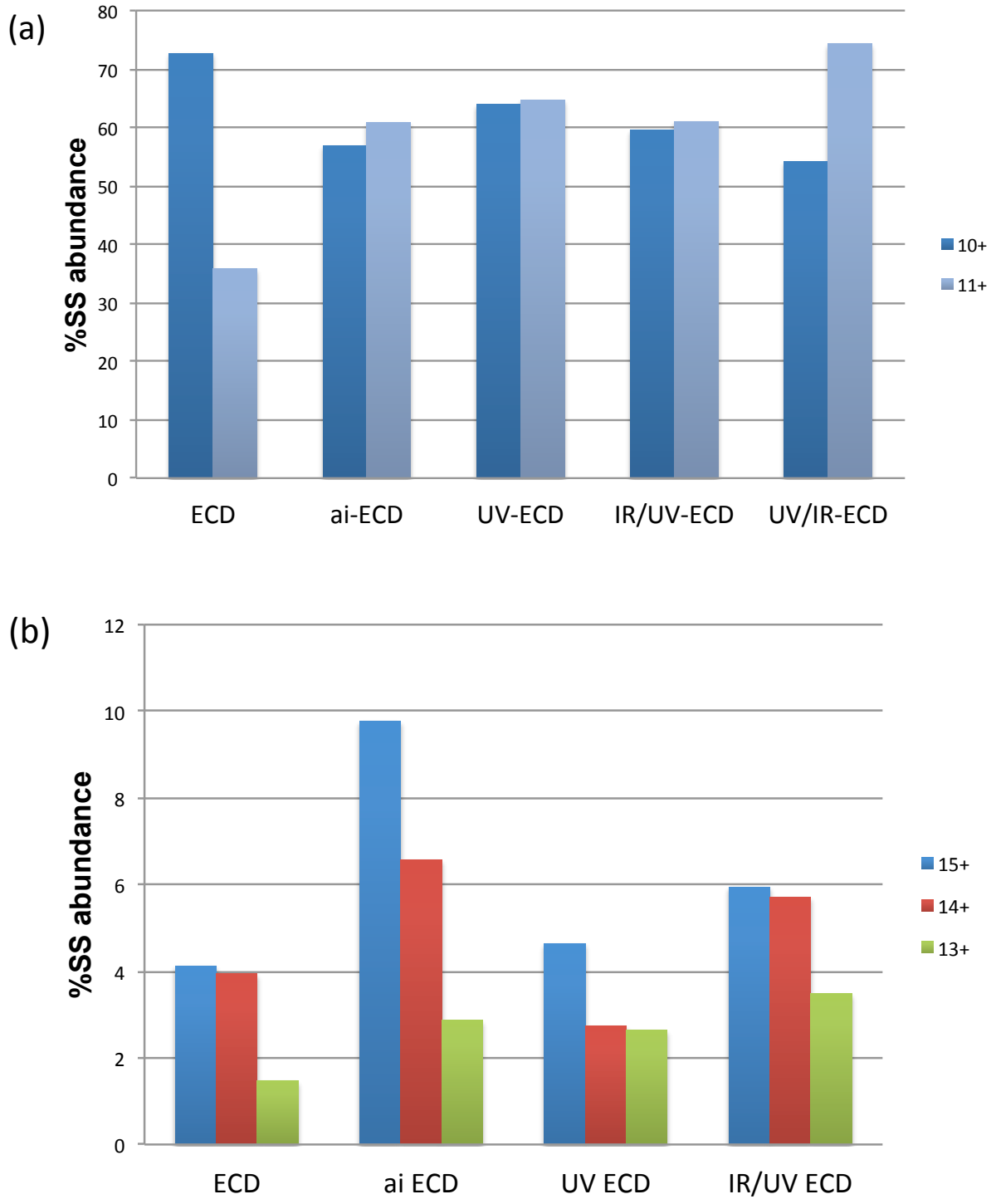

Figure 5-12. A comparison between different ECD activation methods for lysozyme (a) and β-lactoglobulin (b) does not show a significant improvement to assist disulfide bond cleavage.

Conclusion

Two separate laser systems, a UV laser operating at 266 nm in combination with an IR laser was interfaced to an FT-ICR mass spectrometer for promoting disulfide bond dissociation of gas phase proteins. UVPD was able to initiate homolytic cleavage of disulfide bonds as indicated by chain separation of insulin. With ECD only, one intermolecular disulfide bond can be cleaved. With UV pre-activation prior to ECD, all three disulfide bonds of insulin were cleaved. For ribonuclease A, none of the four intramolecular disulfide bonds could be cleaved by ECD alone. A significant improvement was shown by applying 10 shots of UV irradiation followed by vibrational activation before electron capture to cleave up to three disulfide bonds. Unfortunately UV and a combination with IR activation failed to assist disulfide bond cleavage of lysozyme and β-lactoglobulin. Thus, preferential bond cleavage is strongly affected by the sequence of the protein. For insulin and ribonuclease A, tyrosine appears to facilitate disulfide bond cleavage by EET if in a close proximity. In summary, we have demonstrated that UV excitation can enhance disulfide bond cleavage of intermolecular disulfide linked proteins and this method should be beneficial for protein characterization by top-down MS approaches.

References

[1] Wedemeyer, W. J., Welker, E., Narayan, M., and Scheraga, H. A. (2000) Disulfide Bonds and Protein Folding, *Biochemistry 39*, 4207-4216.

[2] Matsumura, M., Signor, G., and Matthews, B. W. (1989) Substantial increase of protein stability by multiple disulphide bonds, *Nature 342*, 291-293.

[3] Wypych, J., Li, M., Guo, A., Zhang, Z., Martinez, T., Allen, M. J., Fodor, S., Kelner, D. N., Flynn, G. C., Liu, Y. D., Bondarenko, P. V., Ricci, M. S., Dillon, T. M., and Balland, A. (2008) Human IgG2 Antibodies Display Disulfide-mediated Structural Isoforms, *Journal of Biological Chemistry 283*, 16194-16205.

[4] Zhang, W., Marzilli, L. A., Rouse, J. C., and Czupryn, M. J. (2002) Complete disulfide bond assignment of a recombinant immunoglobulin G4 monoclonal antibody, *Analytical Biochemistry 311*, 1-9.

[5] Wiesner, J., Resemann, A., Evans, C., Suckau, D., and Jabs, W. (2015) Advanced mass spectrometry workflows for analyzing disulfide bonds in biologics, *Expert Review of Proteomics 12*, 115-123.

[6] Gorman, J. J., Wallis, T. P., and Pitt, J. J. (2002) Protein disulfide bond determination by mass spectrometry, *Mass Spectrometry Reviews 21*, 183-216.

[7] Foley, S. F., Sun, Y., Zheng, T. S., and Wen, D. (2008) Picomole-level mapping of protein disulfides by mass spectrometry following partial reduction and alkylation, *Analytical Biochemistry 377*, 95-104.

[8] Chait, B. T. (2006) Mass Spectrometry: Bottom-Up or Top-Down?, *Science 314*, 65-66.

[9] Doerr, A. (2008) Top-down mass spectrometry, *Nature Methods 5*, 24-24.

[10] Chen, J., Shiyanov, P., Zhang, L., Schlager, J. J., and Green-Church, K. B. (2010) Top-Down Characterization of a Native Highly Intralinked Protein: Concurrent Cleavages of Disulfide and Protein Backbone Bonds, *Analytical Chemistry 82*, 6079-6089.

[11] Kleinnijenhuis, A. J., Duursma, M. C., Breukink, E., Heeren, R. M. A., and Heck, A. J. R. (2003) Localization of Intramolecular Monosulfide Bridges in Lantibiotics Determined with Electron Capture Induced Dissociation, *Analytical Chemistry 75*, 3219-3225.

[12] Peng, Y., Chen, X., Sato, T., Rankin, S. A., Tsuji, R. F., and Ge, Y. (2012) Purification and High-Resolution Top-Down Mass Spectrometric Characterization of Human Salivary α-Amylase, *Analytical Chemistry 84*, 3339-3346.

[13] Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process, *Journal of the American Chemical Society 120*, 3265-3266.

[14] Kruger, N. A., Zubarev, R. A., Carpenter, B. K., Kelleher, N. L., Horn, D. M., and McLafferty, F. W. (1999) Electron capture versus energetic dissociation of protein ions, *International Journal of Mass Spectrometry 182–183*, 1-5.

[15] Zubarev, R. A., Kruger, N. A., Fridriksson, E. K., Lewis, M. A., Horn, D. M., Carpenter, B. K., and McLafferty, F. W. (1999) Electron Capture Dissociation of Gaseous Multiply-Charged Proteins Is Favored at Disulfide Bonds and Other Sites of High Hydrogen Atom Affinity, *Journal of the American Chemical Society 121*, 2857-2862.

[16] Moore, B. N., Ly, T., and Julian, R. R. (2011) Radical conversion and migration in electron capture dissociation, *Journal of the American Chemical Society 133*, 6997-7006.

[17] Simons, J. (2010) Mechanisms for S–S and N–Cα bond cleavage in peptide ECD and ETD mass spectrometry, *Chemical Physics Letters 484*, 81-95.

[18] Zhurov, K. O., Fornelli, L., Wodrich, M. D., Laskay, U. A., and Tsybin, Y. O. (2013) Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis, *Chemical Society Reviews 42*, 5014-5030.

[19] Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry, *Proceedings of the National Academy of Science USA 101*, 9528-9533.
[20] Cole, S. R., Ma, X., Zhang, X., and Xia, Y. (2012) Electron transfer dissociation (ETD) of peptides containing intrachain disulfide bonds, *Journal of the American Society for Mass Spectrometry 23*, 310-320.
[21] Ni, W., Lin, M., Salinas, P., Savickas, P., Wu, S.-L., and Karger, B. (2013) Complete Mapping of a Cystine Knot and Nested Disulfides of Recombinant Human Arylsulfatase A by Multi-Enzyme Digestion and LC-MS Analysis Using CID and ETD, *Journal of the American Society for Mass Spectrometry 24*, 125-133.
[22] Ganisl, B., and Breuker, K. (2012) Does Electron Capture Dissociation Cleave Protein Disulfide Bonds?, *ChemistryOpen 1*, 260-268.
[23] Zhang, Y., Cui, W., Zhang, H., Dewald, H. D., and Chen, H. (2012) Electrochemistry-Assisted Top-Down Characterization of Disulfide-Containing Proteins, *Analytical Chemistry 84*, 3838-3842.
[24] Zhang, Y., Dewald, H. D., and Chen, H. (2011) Online Mass Spectrometric Analysis of Proteins/Peptides Following Electrolytic Cleavage of Disulfide Bonds, *Journal of Proteome Research 10*, 1293-1304.
[25] Nicolardi, S., Giera, M., Kooijman, P., Kraj, A., Chervet, J.-P., Deelder, A., and van der Burgt, Y. M. (2013) On-Line Electrochemical Reduction of Disulfide Bonds: Improved FTICR-CID and -ETD Coverage of Oxytocin and Hepcidin, *Journal of the American Society for Mass Spectrometry 24*, 1980-1987.
[26] Nicolardi, S., Deelder, A. M., Palmblad, M., and van der Burgt, Y. E. M. (2014) Structural Analysis of an Intact Monoclonal Antibody by Online Electrochemical Reduction of Disulfide Bonds and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, *Analytical Chemistry 86*, 5376-5382.
[27] Mentinova, M., and McLuckey, S. A. (2011) Cleavage of multiple disulfide bonds in insulin via gold cationization and collision-induced dissociation, *International Journal of Mass Spectrometry 308*, 133-136.
[28] Lee, M., Lee, Y., Kang, M., Park, H., Seong, Y., June Sung, B., Moon, B., and Bin Oh, H. (2011) Disulfide bond cleavage in TEMPO-free radical initiated peptide sequencing mass spectrometry, *Journal of Mass Spectrometry 46*, 830-839.
[29] Lomeli, S. H., Peng, I. X., Yin, S., Ogorzalek Loo, R. R., and Loo, J. A. (2010) New Reagents for Increasing ESI Multiple Charging of Proteins and Protein Complexes, *Journal of the American Society for Mass Spectrometry 21*, 127-131.
[30] Yin, S., and Loo, J. A. (2011) Top-down mass spectrometry of supercharged native protein–ligand complexes, *International Journal of Mass Spectrometry 300*, 118-122.
[31] Loo, R. R. O., Lakshmanan, R., and Loo, J. A. (2014) What Protein Charging (and Supercharging) Reveal about the Mechanism of Electrospray Ionization, *Journal of the American Society for Mass Spectrometry 25*, 1675-1693.
[32] Zhang, J., Loo, R. R. O., and Loo, J. A. (2015) Increasing fragmentation of disulfide-bonded proteins for top–down mass spectrometry by supercharging, *International Journal of Mass Spectrometry 377*, 546-556.
[33] Preiss, J. W., and Setlow, R. (1956) Spectra of Some Amino Acids, Peptides, Nucleic Acids, and Protein in the Vacuum Ultraviolet, *Journal of Chemical Physics 25*, 138-141.
[34] Fung, Y. M., Kjeldsen, F., Silivra, O. A., Chan, T. W., and Zubarev, R. A. (2005) Facile disulfide bond cleavage in gaseous peptide and protein cations by ultraviolet photodissociation at 157 nm, *Angewandte Chemie International Edition 44*, 6399-6403.

[35] Agarwal, A., Diedrich, J. K., and Julian, R. R. (2011) Direct elucidation of disulfide bond partners using ultraviolet photodissociation mass spectrometry, *Analytical Chemistry 83*, 6455-6458.
[36] Hendricks, N. G., Lareau, N. M., Stow, S. M., McLean, J. A., and Julian, R. R. (2014) Bond-specific dissociation following excitation energy transfer for distance constraint determination in the gas phase, *Journal of the American Chemical Society 136*, 13363-13370.
[37] Zhang, X., Li, H., Moore, B., Wongkongkathep, P., Loo, R. R. O., Loo, J. A., and Julian, R. R. (2014) Radical-directed dissociation of peptides and proteins by infrared multiphoton dissociation and sustained off-resonance irradiation collision-induced dissociation with Fourier transform ion cyclotron resonance mass spectrometry, *Rapid Communications in Mass Spectrometry 28*, 2729–2734.
[38] Chance, R. E., and Ellis, R. M. (1969) Proinsulin: Single-chain precursor of insulin, *Archives of Internal Medicine 123*, 229-236.
[39] Loo, J. A., Edmonds, C. G., and Smith, R. D. (1990) Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry, *Science 248*, 201-204.
[40] Henry, K. D., Williams, E. R., Wang, B. H., McLafferty, F. W., Shabanowitz, J., and Hunt, D. F. (1989) Fourier-transform mass spectrometry of large molecules by electrospray ionization, *Proceedings of the National Academy of Science USA 86*, 9075-9078.
[41] Przybylski, M., and Glocker, M. O. (1996) Electrospray Mass Spectrometry of Biomacromolecular Complexes with Noncovalent Interactions—New Analytical Perspectives for Supramolecular Chemistry and Molecular Recognition Processes, *Angewandte Chemie International Edition 35*, 806-826.
[42] Yu, W., Vath, J. E., Huberty, M. C., and Martin, S. A. (1993) Identification of the facile gas-phase cleavage of the Asp-Pro and Asp-Xxx peptide bonds in matrix-assisted laser desorption time-of-flight mass spectrometry, *Analytical Chemistry 65*, 3015-3023.
[43] Gardner, M., and Brodbelt, J. (2008) Impact of proline and aspartic acid residues on the dissociation of intermolecularly crosslinked peptides, *Journal of the American Society for Mass Spectrometry 19*, 344-357.
[44] Liu, J., Gunawardena, H. P., Huang, T.-Y., and McLuckey, S. A. (2008) Charge-dependent dissociation of insulin cations via ion/ion electron transfer, *International Journal of Mass Spectrometry 276*, 160-170.
[45] Tsybin, Y. O., He, H., Emmett, M. R., Hendrickson, C. L., and Marshall, A. G. (2007) Ion Activation in Electron Capture Dissociation To Distinguish between N-Terminal and C-Terminal Product Ions, *Analytical Chemistry 79*, 7596-7602.
[46] Breuker, K., Jin, M., Han, X., Jiang, H., and McLafferty, F. W. (2008) Top-Down Identification and Characterization of Biomolecules by Mass Spectrometry, *Journal of the American Society for Mass Spectrometry 19*, 1045-1053.
[47] Lin, C., Cournoyer, J. J., and O'Connor, P. B. (2008) Probing the Gas-Phase Folding Kinetics of Peptide Ions by IR Activated DR-ECD, *Journal of the American Society for Mass Spectrometry 19*, 780-789.
[48] Mikhailov, V. A., and Cooper, H. J. (2009) Activated Ion Electron Capture Dissociation (AI ECD) of Proteins: Synchronization of Infrared and Electron Irradiation with Ion Magnetron Motion, *Journal of the American Society for Mass Spectrometry 20*, 763-771.
[49] Ly, T., and Julian, R. R. (2009) Ultraviolet Photodissociation: Developments towards Applications for Mass-Spectrometry-Based Proteomics, *Angewandte Chemie International Edition 48*, 7130-7137.

[50] Permyakov, E. A., Permyakov, S. E., Deikus, G. Y., Morozova-Roche, L. A., Grishchenko, V. M., Kalinichenko, L. P., and Uversky, V. N. (2003) Ultraviolet illumination-induced reduction of α-lactalbumin disulfide bridges, *Proteins: Structure, Function, and Bioinformatics 51*, 498-503.

CHAPTER SIX

Conclusion and Perspectives

Conclusion and perspectives

Native electrospray ionization (ESI) mass spectrometry has delivered a potentially powerful technique for protein structural characterization that can be useful for the structural biology community. Native ESI introduces gas phase protonated proteins in their nearly native states to the mass spectrometer[1]. Not only can proteins be successfully ionized and transferred to MS, but protein-ligand complexes, protein-protein interactions, membrane proteins, and large protein complexes can be analyzed by native ESI-MS[2-5]. Membrane proteins are difficult to study by x-ray crystallography because of problems with solubility and finding proper conditions to promote crystallization, but recently it has been shown that native ESI-MS with a proper sample handling can be used to measure membrane proteins and membrane protein complexes[6-8]. Additionally, native ESI-MS can provide comprehensive data for binding events from a single analysis, such as molecular mass, stoichiometry, and binding affinity[9]. When native ESI-MS is combined with other on-line techniques, significantly more structural information can be obtained. For example, ion mobility provides structural information such as shape and conformation of protein molecules. Top-down MS reveals sequence specific information including binding sites, interfacial regions, and possibly protein subunit orientation (for multisubunit complexes). Labeling techniques can be used to illustrate surface topology. Herein, we showed several applications of native ESI combined with top-down mass spectrometry to characterize protein-ligand complexes that are involve with neurodegenerative diseases. We extended the use of protein-ligand binding to map out the surface residues of small proteins. Lastly, a photon-based activation method was presented to improve top-down MS of disulfide bond containing proteins.

The Parkinson's disease related protein α-synuclein (AS) is a small protein (14 kDa), which is natively unstructured at physiological pH[10]. Previous research has shown that several heavy

divalent metal ions can initiate self-oligomerization of the protein[11, 12]. Native ESI-MS was used to monitor AS protein complexes with copper, cobalt, and manganese. Metal binding affinities can be roughly estimated from ESI mass spectra. We showed that copper is tightly bound to AS. Top-down MS using both electron capture dissociation (ECD) and collisionally activated dissociation (CAD) identified the binding site of copper to the N-terminal helix region of AS. Copper and manganese have similar behavior. They have less affinity to AS compared to cobalt. They share the same two binding locations near the C-terminus. Ion mobility spectrometry (IMS) was also used to probe structural changes upon metal binding. The IMS data suggested that copper induces conformational changes for the low charge state molecules, but no conformational changes were observed for cobalt and manganese binding. Top-down MS with ECD was also able to locate the lysine residues of AS and SOD1 that is bound to molecular tweezer (MT) to form complexes. It is likely that these lysines are key residues that are involved with fibrillation because MT binding was able to inhibit self-oligomerization for both AS[13] and SOD1 (unpublished data). More *in vitro* biochemical experiments will be needed to further investigate the importance of these specific lysines. We had no significant difficulties to generate top-down ECD fragments from AS-MT complexes because AS is natively unstructured[14]. Some native protein complexes are difficult to fragment in top-down MS experiments because they are tightly folded.

We further extended ECD to assess and identify surface residues of small proteins. It was demonstrated that MT binds non-covalently to proteins besides those involved in neurodegeneration as well. The MT compounds were designed to bind to positively charged residues, such as lysine and arginine, with preference for lysine[15]. MT-bound fragments were generated to help locate the sites of binding, which are then compared with theoretical solvent accessibility values. We showed that the lysines found to bind to MT correlated well with solvent

accessibility, suggesting that MT binds to lysine residues found on the surface topology of the protein and that MT could be a useful protein surface probe.

However, there are significant challenges to the top-down MS analysis of native proteins and complexes. Backbone bond cleavages generated from ECD is limited due to the low charge state from native ESI. This challenge limited the amount of structural information that could be obtained for the study of SOD1-MT binding. ECD efficiency is proportional to the charge of the precursor ion. One way to improve ECD efficiency is to isolate the entire charge envelope to increase number of precursor ions for simultaneous fragmentation instead of selecting only one charge state. ECD data from multiple charge states simultaneous was useful to gather sufficient fragmentation information to confidently conclude the MT binding site to be in the zinc binding loop of SOD1. However, this approach is not practical if more than one species are present in a precursor spectrum, e.g., overlapping charge distributions from multiple species in a mixture. Another way to increase ECD efficiency is to increase precursor charge by supercharging reagents[16, 17]. Supercharging agents are molecules that are added to the analyte solution to directly increase its ESI multiple charging. The use of supercharging agents for native ESI-MS has been shown in our laboratory to increase the charge state of the native complex without significantly disrupting the binding of most proteins and ligands and other larger macromolecules. But an extra caution must be considered on whether the additives may suppress non-covalent interactions for some types of complexes. Moreover, the concentration of the supercharging reagent may need to be optimized; too much supercharging reagent may suppress analyte signal.

Another challenge to be considered that affects the efficiency of ECD is the requirement for the peptide fragments to be released upon dissociation. Some proteins may have especially strong

intramolecular interactions, such as hydrogen bonds, salt bridges, or hydrophobic interactions. Hydrogen bonds and salt bridges are based on electrostatic interactions, which are greatly enhanced in strength in vacuum because of the low dielectric constant of the media. Such enhanced intramolecular bonds prohibit ECD-generated fragments to be clearly separated from the rest of the molecule and be detected by mass spectrometry. This is possibly one of the explanations why some proteins do not appear to fragment well by ECD. The same situation applies for disulfide bond proteins. Some research groups have developed additional activation methods to increase the internal energy of the protein molecule to break down intramolecular interactions. Radiation from an infrared laser is a well-known approach to provide the supplemental activation prior to or after ECD. Backbone cleavage efficiency can be significantly improved by this method; this method is referred to as activated ion ECD (ai-ECD). However, the use of supplemental activation may need to be reconsider when protein-ligand complexes are analyzed because the additional energy may disrupt the weak interactions holding the protein-ligand complex together. Instrumental settings will need to be carefully optimized. For disultide bond proteins, we also utilized a 266 nm UV activation step to facilitate disulfide bond cleavage and improve ECD fragmentation[18]. Alternatively, electron ionization dissociation (EID)[19] and ultraviolet photodissociation (UVPD) using 193 nm wavelength radiation[20] have been shown very recently to generate extensive backbone cleavages that is greatly improved from ECD or CAD alone. EID and UVPD are promising techniques that in the future can be applied potentially to study non-covalent complexes. The application of top-down MS for characterizing native protein complexes has already shown utility for structural biology applications, but future improvements to the experimental approach will surely expand the applicability of the approach.

References

[1] Kebarle, P., and Verkerk, U. H. (2009) Electrospray: From ions in solution to ions in the gas phase, what we know now, *Mass Spectrometry Reviews 28*, 898-917.

[2] Sharon, M., and Robinson, C. V. (2007) The Role of Mass Spectrometry in Structure Elucidation of Dynamic Protein Complexes, *Annual Review of Biochemistry 76*, 167-193.

[3] Schmidt, C., and Robinson, C. V. (2014) Dynamic protein ligand interactions – insights from MS, *FEBS Journal 281*, 1950-1964.

[4] Marcoux, J., and Robinson, Carol V. (2013) Twenty Years of Gas Phase Structural Biology, *Structure 21*, 1541-1550.

[5] Chen, F., Gülbakan, B., Weidmann, S., Fagerer, S. R., Ibáñez, A. J., and Zenobi, R. (2015) Applying mass spectrometry to study non-covalent biomolecule complexes, *Mass Spectrometry Reviews*, in press.

[6] Benesch, J. L. P., and Robinson, C. V. (2006) Mass spectrometry of macromolecular assemblies: preservation and dissociation, *Current Opinion in Structural Biology 16*, 245-251.

[7] Barrera, N. P., and Robinson, C. V. (2011) Advances in the Mass Spectrometry of Membrane Proteins: From Individual Proteins to Intact Complexes, *Annual Review of Biochemistry 80*, 247-271.

[8] Laganowsky, A., Reading, E., Hopper, J. T. S., and Robinson, C. V. (2013) Mass spectrometry of intact membrane protein complexes, *Nature Protocols 8*, 639-651.

[9] Loo, J. A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry, *Mass Spectrometry Reviews 16*, 1-23.

[10] Uversky, V. N. (2003) A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders, *Journal of Biomolecular Structure & Dynamics 21*, 211-234.

[11] Uversky, V. N., Li, J., and Fink, A. L. (2001) Metal-triggered structural transformations, aggregation, and fibrillation of human alpha-synuclein. A possible molecular NK between Parkinson's disease and heavy metal exposure, *Journal of Biological Chemistry 276*, 44284-44296.

[12] Yamin, G., Glaser, C. B., Uversky, V. N., and Fink, A. L. (2003) Certain metals trigger fibrillation of methionine-oxidized alpha-synuclein, *Journal of Biological Chemistry 278*, 27630-27635.

[13] Prabhudesai, S., Sinha, S., Attar, A., Kotagiri, A., Fitzmaurice, A. G., Lakshmanan, R., Ivanova, M. I., Loo, J. A., Klarner, F. G., Schrader, T., Bitan, G., and Bronstein, J. (2012) A novel "molecular tweezer" inhibitor of alpha-synuclein neurotoxicity in vitro and in vivo, *Neurotherapeutics 9*, 464-476.

[14] Acharya, S., Safaie, B. M., Wongkongkathep, P., Ivanova, M. I., Attar, A., Klärner, F.-G., Schrader, T., Loo, J. A., Bitan, G., and Lapidus, L. J. (2014) Molecular Basis for Preventing α-Synuclein Aggregation by a Molecular Tweezer, *Journal of Biological Chemistry 289*, 10727-10737.

[15] Fokkens, M., Schrader, T., and Klärner, F.-G. (2005) A Molecular Tweezer for Lysine and Arginine, *Journal of the American Chemical Society 127*, 14415-14421.

[16] Yin, S., and Loo, J. A. (2011) Top-down mass spectrometry of supercharged native protein–ligand complexes, *International Journal of Mass Spectrometry 300*, 118-122.

[17] Zhang, J., Loo, R. R. O., and Loo, J. A. (2015) Increasing fragmentation of disulfide-bonded proteins for top–down mass spectrometry by supercharging, *International Journal of Mass Spectrometry 377*, 546-556.

[18] Wongkongkathep, P., Li, H., Zhang, X., Ogorzalek Loo, R. R., Julian, R. R., and Loo, J. A. (2015) Enhancing protein disulfide bond cleavage by UV excitation and electron capture dissociation for top-down mass spectrometry, *International Journal of Mass Spectrometry 390*, 137-145.
[19] Fung, Y. M. E., Adams, C. M., and Zubarev, R. A. (2009) Electron Ionization Dissociation of Singly and Multiply Charged Peptides, *Journal of the American Chemical Society 131*, 9977-9985.
[20] Shaw, J. B., Li, W., Holden, D. D., Zhang, Y., Griep-Raming, J., Fellers, R. T., Early, B. P., Thomas, P. M., Kelleher, N. L., and Brodbelt, J. S. (2013) Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation, *Journal of the American Chemical Society 135*, 12646-12651.