UNIVERSITY OF CALIFORNIA
RIVERSIDE

Mass Spectrometry-Based Elucidation of Protein Structure Using Noncovalent Molecular Recognition and Photo-Induced Radical Chemistry

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Tony Ly

June 2010

Dissertation Committee:

Dr. Ryan R. Julian, Chairperson
Dr. Cynthia K. Larive
Dr. Quan J. Cheng

The Dissertation of Tony Ly is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

The completion of this dissertation would have been impossible without the support of numerous people and funding agencies that I would like to mention here. Ryan has had the most significant influence in the direction of this dissertation and in my growth as a scientist. He is an ambitious mentor. Complaints that a task is too difficult are deflected with the wry repartee, "well… that might be true for a *lesser* graduate student." Delivered in a jocular tone that is uniquely Ryan's, it is nevertheless an open challenge to seek innovative strategies to overcome difficult obstacles and an accusation of greatness. How can a student refuse to comply? Entering graduate school with relatively modest expectations, I found that this challenge was exactly what I needed.

My interest in chemistry began in Ms. Galatis' IB Chemistry classroom, with Lewis dot diagrams, sprawling discussions of the Heisenberg Uncertainty Principle, and most vividly, reactions of potassium perchlorate with gummy bears after school. I thank Ms. Galatis for inspiring me to enroll in chemistry courses as a college freshman. I thank Dr. Kristine Lowe and Dr. John Callahan, for introducing me to microbiology and mass spectrometry, respectively, at the Naval Research Laboratory. I thank Dr. Gary Rice, my undergraduate research

advisor, who convincingly showed me that chemistry is beautiful, fun, *and* intellectually rigorous.

Richard Ernst argues that the "good" life requires support by multiple passions, like the legs of a chair. I could not agree more. In the personal realm, I thank my closest friends in Riverside for their companionship. I especially thank Nathan, who continues to be an enduring positive influence in my personal growth and professional career. I thank my family, whose invaluable support has allowed me to indulge in my interest in chemistry.

I thank my fellow labmates, current and former, for helpful discussions and collaborations (Aholibama, Emily, Jolene, Helen, Zhenjiu, Ben, Jackie, Claudia, Don, Eric, Geoff and Yuanqi). I thank Kim for being my partner in leading the Chemistry GSA and for her remarkable ability to turn the most burdensome chore into a fun adventure.

I thank the following institutions for funding: UC Riverside (RRJ), the National Science Foundation CAREER Award (RRJ, CHE-0747481), the National Institutes of Health (RRJ, 1R01GM084106-01A1), the American Chemical Society Analytical Division, Eastman Chemicals, and the UCR Graduate Division. I thank Helen for providing peptide fragmentation data for Figure 3.6 I thank Professor Joseph A. Loo and Dr. Sheng Yin, co-authors of the research described

in Chapter 4, who kindly provided access to an LTQ-ICR mass spectrometer. This collaboration was funded by the National Institutes of Health (JAL, RR 20004) and a National Institutes of Health/National Center for Research Resources High-End Instrumentation Award (JAL, S10 RR023045).

The text of this dissertation, in part or in full, is a reprint of the material as it appears in the following publications:

Ch. 2: Ly, T.; Julian, R. R. *J. Am. Chem. Soc.* **2008**, *130*, 351-358.

Ch. 3: Ly, T.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1148-1158 and Ly, T.; Julian, R. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 7130-7137.

Ch. 4: Ly, T.; Julian, R. R. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 2099-2101.

Ch. 5: Ly, T.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1209-1215.

Ch. 6: Ly, T.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1663-1672.

Ch. 7: Ly, T.; Liu, Z.; Pujanauski, B. G.; Sarpong, R.; Julian, R. R. *Anal. Chem.* **2008**, *80*, 5059-5064.

The co-author, Ryan R. Julian, listed in these publications directed and supervised the research that forms the basis for this dissertation. Dr. Liu prepared and performed SNAPP-MS experiments on single residue mutants of the protein ubiquitin. Dr. Sarpong and Mr. Pujanauski developed the chemistry for and synthesized the phenyl-linked bis(18-crown-6) ether molecule.

Dedication

To my parents,

who took the enormous risk of escaping Vietnam and seeking refuge in the United States in search for a better future for my siblings and me.

ABSTRACT OF THE DISSERTATION

Mass Spectrometry-Based Elucidation of Protein Structure Using Noncovalent Molecular Recognition and Photo-Induced Radical Chemistry

by

Tony Ly

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, June 2010
Dr. Ryan R. Julian, Chairperson

Advances in mass spectrometry (MS) have enabled quick and accurate protein identification, an important analytical goal. However, knowledge of the identity of a protein only skims the surface of the structural and functional complexities that are characteristic of proteins. Described in this dissertation is the development of MS-based analytical techniques to investigate the structure of proteins using noncovalent interactions and photo-induced radical chemistry.

The first half of the dissertation reports the discovery of a novel gas phase dissociation method using photo-induced radical chemistry to 'direct' backbone fragmentation to specific amino acid residues, which we have called 'radical directed dissociation' (RDD). Residue-specific dissociation is a significant

advance towards enzyme-like protein disassembly in the gas phase. Investigation of numerous peptides and proteins yields two key features of RDD. First, fragmentation frequently occurs at residues have the lowest predicted C_β-H bond dissociation energies (BDEs) of the twenty canonical amino acids. Second, removal of the C_β carbon of tyrosine abrogates backbone fragmentation at that residue. These results indicate that the C_β is the most important site for radical-induced backbone fragmentation and that radical migration in peptides appears to favor sites with the lowest C-H BDEs. We have called this latter phenomenon "radical funneling". Although useful for peptides, the "radical funneling" model becomes inappropriate for larger peptides and proteins, where structural effects cannot be ignored. Indeed, we show that radical migration is a sensitive reporter of tertiary structure of proteins *in vacuo*.

The remaining half of the dissertation reports the development of selective noncovalent adduct protein probing mass spectrometry (SNAPP-MS) to investigate the three-dimensional structure of proteins. SNAPP-MS is a fast technique that uses 18-crown-6 ether (18C6) as a molecular probe of lysine availability in proteins. The number of 18C6s that bind to protein is easily measured by MS. Interestingly, the binding of 18C6 to protein occurs in solution, which allows investigation of solution phase structure.

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

List of Acronyms

m/z – mass to charge

MS – mass spectrometry

SNAPP – selective noncovalent adduct protein probing

18C6 – 18-crown-6 ether

PBC – phenyl-linked bis(18-crown-6 ether)

RDD – radical directed dissociation

BDE – bond dissociation energy

cyt c – cytochrome c

ubi – ubiquitin

AS – α synuclein

CaM – calmodulin

Chapter 1

MASS SPECTROMETRY OF PROTEINS

1.1 Introduction

The measurement of the mass-to-charge ratio (m/z) has been the defining principle of mass spectrometry since its inception in the laboratories of Thompson and Aston in 1919. Advances in mass spectrometry instrumentation have opened the doors to new classes of analytes and have shifted the focus of the field from atoms to small organic molecules, and recently, to biological entities of ever-increasing size and complexity. Beyond the analysis of single biomolecules, it is now possible to sample intact multimeric protein complexes and large biomolecular clusters with mass spectrometry.[1,2]

Comprising one of the pillars of the central dogma of biology, proteins play a significant role in every cellular process and are the most abundant type of biomolecule in the cell. Proteins have a variety of functions, ranging from maintaining cellular structure to chemical catalysis. Proteins are polymers of amino acids covalently linked by amide bonds (also called peptide bonds). Nearly all proteins are constructed in linear fashion into a polypeptide chain.

Given the diversity of functions that proteins play, it is perhaps surprising to find that there are only twenty common amino acids (monomeric units). Compared to the analysis of organic molecules, which span a wider variety of functionalities and chemistries, the intrinsic chemistry of proteins is limited to these twenty amino acids and a number of known amino acid modifications. However, proteins and their function are not simply defined by their amino acid content. The sequence of amino acids, the three-dimensional arrangement, and the cellular environment all play a significant role in determining the ultimate chemistry of a protein.

Protein structure is organized into multiple layers of complexity. The fundamental layer is the primary structure, or the linear sequence of amino acids. The next layers are the secondary and tertiary structures, which consist of the three-dimensional arrangement of the amino acids. Finally, the quaternary structure consists of the arrangement of proteins into protein complexes. The objective of this dissertation is to develop mass spectrometry-based techniques to investigate protein structure at each level of complexity.

1.2 Primary Structure

Mass spectrometry has been increasingly used to analyze the primary structure of proteins in proteomic analyses. In a typical proteomics experiment, it

is not necessary to determine the complete amino acid sequence of a protein experimentally. A library of expected protein sequences can be generated *in silico* if the genome of the organism is known. The principal goal of mass spectrometry is to obtain enough sequence information to correctly and uniquely identify the protein against all other expected protein sequences.

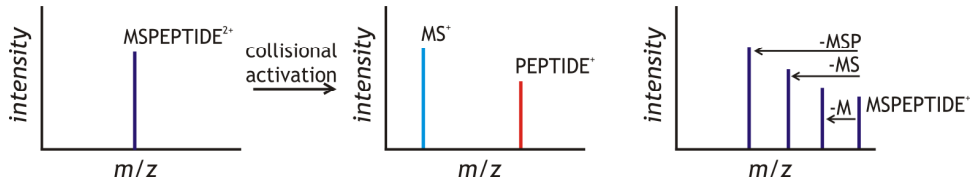


Figure 1.1 Mass spectrometry of a doubly charged peptide, MSPEPTIDE (left). Collisional activation yields fragmentation of a peptide bond (middle). Any peptide bond may be broken. The sum of numerous fragmentation events, in which a random single amide bond is broken, leads to a distribution of fragments (right). The mass differences between fragments can reveal sequence information. Intensities shown are arbitrary.

Although whole proteins can be analyzed directly by mass spectrometry, proteins are typically enzymatically digested to smaller peptides prior to mass analysis. The key peptide identifier data determined by mass spectrometry are the molecular weights (determined from measuring the m/z and charge state) of the peptide and of its fragments. Gas phase heating in the form of collisions or infrared photons ruptures the peptide bond, producing two fragments (Figure 1.1, left and middle). The location of the peptide bond that is broken is typically unpredictable, yielding a series of fragment ions. If peptide bond fragmentation

is distributed across all peptide bonds, then it is possible to extract the sequence from consecutive mass differences between adjacent fragments (Figure 1.1, right). The unpredictability of the location of the scissile amide bond can be an advantage in peptide sequencing, where high sequence coverage is optimal. However, exceptional sequence selectivity has its own unique advantages for proteomic studies.

For example, predictable gas phase fragmentation may be used as an alternative to enzymatic digestion and would eliminate a sample preparation step. In contrast to enzymatic digestions, however, selective gas phase fragmentation would provide a direct connection between the peptides and the protein from which they are derived. Analysis of whole proteins is called the "top-down" approach and is an alternative to the "bottom-up" approach described in the preceding paragraph.[3] The link between fragment peptide and source protein is especially important in studying protein post-translational modifications (PTMs). Typically, it is a unique combination of PTMs decorating the whole protein that has biochemical relevance.[4] This information is difficult to obtain using the bottom-up approach.

Selective fragmentation can be achieved by accessing unique excited states that direct fragmentation at specific residues.[5] Another approach is to lower the

threshold for backbone fragmentation at specific residues.[6] Chapter 2 of the dissertation describes the use of both approaches to direct backbone fragmentation to tyrosine residues of whole proteins. Radicals are produced at specific tyrosine residues using selective chemical modification in solution, followed by photodissociation in the gas phase. Collisional activation of the radical protein yields selective, residue-specific backbone fragmentation. This novel gas phase fragmentation method is called radical directed dissociation (RDD).

Chapter 3 of the dissertation explores the gas phase chemistry of radical peptides and proteins in detail. Specifically, it is shown that RDD proceeds through intermediates in which a radical is located at the C_β of an amino acid. In peptides, backbone fragmentation preferentially occurs at amino acids where C_β radicals are preferentially stabilized, such as residues containing alcohol and aromatic groups. Additionally, the interplay between thermodynamics and kinetics as controlling factors in radical migration and fragmentation are described.

RDD requires that the peptide analyte contains a tyrosine that is reactive towards iodination and a UV photodissociation-capable instrument. To avoid these potential limitations, alternative strategies were explored to generate RDD-

like chemistry. Chapter 4 demonstrates that similar radical chemistry is afforded using electron-induced dissociation (EID) in a commercially available FT-ICR instrument.[7,8] Results with several peptides suggest that bombardment of singly protonated peptide ions with moderate energy electrons yields sidechain losses and backbone fragmentation products that are similar in type to RDD. It is proposed that these fragmentation products are due to dissociation of a metastable doubly-charged peptide radical intermediate formed from ejection of an electron. Interestingly, the radical chemistry is not limited to peptides that contain aromatic residues. However, due to the low efficiency and regiospecificity of EID, RDD is preferred for tyrosine-containing peptides.

1.3 Higher Order Structure

The three-dimensional arrangement of the polypeptide chain is equally important as the amino acid sequence in defining the ultimate chemistry of the protein. The most advanced techniques to examine protein structure currently are nuclear magnetic resonance (NMR) spectroscopy and x-ray diffraction. High resolution data from these techniques have been used to successfully reconstruct high fidelity protein structures. Both techniques have high sample requirements and are labor intensive. In contrast, mass spectrometry experiments are typically quicker, do not require laborious sample preparation, and have low sample

requirements. Unlike NMR and x-ray diffraction however, mass spectrometry is fundamentally limited to measuring the m/z of ions, which does not intrinsically provide any direct information on protein structure. However, mass spectrometry can be coupled to a secondary structurally sensitive technique that alters the mass of the protein. For example, covalent labeling has been used to probe the chemical accessibility of specific amino acids.[9] Differences in the number and location of amino acid labels are presumed to be due to changes in structure.

Noncovalent chemistry can also be used to detect conformational changes in proteins, as described in Chapter 5 of the dissertation. 18-crown-6 ether was identified as the noncovalent probe candidate, due to its selective affinity for lysine residues.[10] Unlike nearly all covalent labeling chemistries, the binding of 18C6 to lysine residues is weak and reversible, which allows 18C6 to sample the conformational distribution at equilibrium. The acronym SNAPP (selective noncovalent adduct protein probing) has been adopted to describe this technique. Results in Chapter 5 demonstrate that SNAPP is sensitive towards the changes in protein structure in native-like vs. denaturing solvent conditions. Generally, more 18C6 molecules will attach to a given protein under denaturing conditions than under native-like conditions. Experiments using model peptides

demonstrate that the observed differences are not due to changes in the binding constant of 18C6 to lysine. Instead, these results indicate that changes in the distribution of 18C6 molecules attached to a protein, i.e. SNAPP distributions, reflect changes in three-dimensional structure in solution.

The mechanism of SNAPP has not been studied in detail. However, we can make several conjectures based on what is known about the electrospray process and the binding constants of 18C6. In water, the dissociation constant (K_D) for 18C6 and methylamine is 74.1 mM in water.[11] K_D decreases significantly with decreasing solvent polarity. For example, in methanol/water solutions and $CDCl_3$, the K_D decreases to 479. μM and 30.9 nM, respectively.[11] At the typical concentrations of a SNAPP experiment (μM of each), there should be minimal binding between 18C6 and protein. In contrast, 18C6/protein complexes are observed in significant intensity in the ESI-MS spectra of 18C6/protein in water. These results suggest that 18C6/protein binding occurs during the electrospray process. Specifically, the concentration of 18C6 and protein is expected to significantly increase as droplets evaporate in the electrospray source, which would facilitate formation of 18C6/protein complexes. In the gas phase, the dissociation enthalpy for 18C6 and lysine sidechain has been calculated to be 43.5 kcal/mol.[10] Thus, upon complete desolvation, the binding affinity between 18C6

and lysine sidechains is strong, with a *noncovalent* bond energy approaching half a typical C-H *covalent* bond!

Protein-metal interactions are prevalent in biochemistry and can modulate biomolecular function. The protein calmodulin is a classic example of how protein-metal interactions can confer biological function. Binding of calcium ions by calmodulin induces a structural rearrangement in the protein that reveals a key hydrophobic α helix, which can recognize and bind other proteins, ultimately signaling enzymatic phosphorylation of the substrate.[12] Thus, calmodulin acts as an intracellular calcium ion sensor, marking proteins for phosphorylation only when calcium ions are present. Chapter 6 of the dissertation describes the application of SNAPP-MS to examining conformational changes of the protein calmodulin. The results indicate that SNAPP-MS is sensitive towards these conformational changes. Comparison of the known structures for the Ca^{2+}-bound and Ca^{2+}-free structures of calmodulin are consistent with the idea that SNAPP-MS probes lysine availability.

Another area where SNAPP-MS holds much promise is the analysis of natively unfolded proteins. Techniques developed to determine a single, ordered structure, such as x-ray crystallography, are inappropriate for examining natively unfolded proteins. Natively unfolded proteins adopt numerous

conformations under physiological conditions. Shifts in conformational dynamics may be important in modulating the biological function of the protein.[13] SNAPP-MS is capable of monitoring multiple populations of conformers in the same experiment, due to differences in protonation state and the number of 18C6s that bind. Indeed, in Chapter 6 of the dissertation, it is demonstrated that SNAPP-MS is capable of monitoring subtle changes in the conformational dynamics of α synuclein, a protein that is implicated in neurodegenerative disease. It is known that α synuclein aggregation is accelerated *in vitro* by introducing metal ions, such as copper and aluminum.[14] Our results indicate that SNAPP-MS is sensitive towards changes in the conformational dynamics upon addition of metal. Interestingly, the interaction of α synuclein and aluminum changed the SNAPP distributions significantly, whereas copper did not. This suggests that the mechanism for accelerating aggregation is different between copper and aluminum.

Other SNAPP reagents have been examined, including a novel phenyl-ethylene glycol linked bis(18-crown-6-ether) (PBC). The phenyl linker provides a semi-rigid spacer between the two crown moieties, which will selectively bind to the sidechain of lysine. It is shown in Chapter 7 of the dissertation that intensity distributions of PBC-ubiquitin complexes depend on the spatial arrangement of

the lysine residues in the protein. A single PBC molecule may bind in bidentate fashion to two free lysine sidechains spaced within ~8 Å of each other, i.e. the distance of the phenyl spacer. However, if the lysine sidechains are spaced far apart, then two PBC molecules will bind and the interaction between PBC and lysine will be monodentate. Comparison of the 18C6 (which are only capable of monodentate binding) and PBC distributions reveals the number of PBC molecules that are bidentate bound to the A-state of ubiquitin.[15] The higher binding strength of bidentate interactions allows for localization of the binding sites by collision-induced dissociation. It is found that PBC binds to Lys29 and Lys33 of ubiquitin in a bidentate fashion, suggesting that these two residues are in proximity, and not participating in salt bridges in acid-unfolded ubiquitin.

Is there tertiary structure in the gas phase? Do highly ordered secondary structures such as α helices and β sheets survive upon transfer into the gas phase? Do structures sampled in the gas phase resemble those found in solution? The answers to these questions are highly sought after.[16] However, the number of techniques capable of providing the residue specific detail required to answer these questions is extremely limited. Chapter 8 of the dissertation introduces a novel technique using radical migration to investigate inter-residue contacts in the gas phase. In this context, radical migration is used much like the

phenomenon of Förster resonance energy transfer, or FRET.[17] The radical donor (tyrosine radical) is synthesized regiospecifically using radical-directed dissociation. The final radical acceptors are measured by mass spectrometry of the radical-induced fragmentation of the protein backbone. These RDD results have been integrated into a simulated annealing/molecular modeling approach to elucidate the gas phase structure of the +4 and +6 charge state of ubiquitin.

Understanding the intricacies of protein structure requires attack from multiple angles. Mass spectrometry promises to play a significant role in deciphering protein structure in combination with other techniques. It is clear that advances in mass spectrometry instrumentation have enabled high mass accuracy, resolution, sensitivity, and dynamic range.[18] In conjunction with the progress in instrumentation, there remains significant work to develop techniques that can effectively transform information contained in bioanalytical targets (e.g. primary, secondary, tertiary structure of proteins) into changes in mass, which are easily measured by mass spectrometry.

---

[1] Robinson, C. V.; Gross, M.; Eyles, S. J.; Ewbank, J. J.; Mayhew, M.; Hartl, F. U.; Dobson, C. M.; Radford, S. E. *Nature* **1994**, *372*, 646-651.
[2] Briggs, D. B.; Jones, C. M.; Mashalidis, E. H.; Nunez, M.; Hausrah, A. C.; Wysocki, V. H.; Tsao, T. S. *Biochemistry* **2009**, *48*, 12345-12357.
[3] Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom. Rev.* **2005**, *24*, 201-222.
[4] Pesavento, J. J.; Kim, Y. B.; Taylor, G. K.; Kelleher, N. L. *J. Am. Chem. Soc.* **2004**, *126*, 3386-3387.
[5] Thompson, M. S.; Cui, W. D.; Reilly, J. P. *Angew. Chem. Int. Ed.* **2004**, *43*, 4791−4794.

[6] Reid, G. E.; Roberts, K. D.; Simpson, R. J.; O'Hair, R. A. J. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1131−1150.
[7] Budnik, B. A.; Zubarev, R. A. *Chem. Phys. Lett.* **2000**, *330*, 558.
[8] Lioe, H.; O'Hair, R. A. J. *Anal. Bioanal. Chem.* **2007**, *389*, 1429.
[9] Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. *Proc. Natl. Acad. Sci*. **2000**, *97*, 5802-5806.
[10] Julian, R. R.; Beauchamp, J. L. *Int. J. Mass Spectrom.* **2001**, *210*, 613-623.
[11] Izatt, R. M.; Bradshaw, J. S.; Nielsen, S. A.; Lamb, J. D.; Christensen, J. J. *Chem. Rev.* **1985**, *85*, 271-339.
[12] Babu, Y. S.; Bugg, C. E.; Cook, W. J. *J. Mol. Bio.* **1988**, *204*, 191-204.
[13] Krishnan, S.; Chi, E. Y.; Wood, S. J.; Kendrick, B. S.; Li, C.; Garzon-Rodriguez, W.; Wypych, J.; Randolph, T. W.; Narhi, L. O.; Biere, A. L.; Citron, M.;. Carpenter, J. F. *Biochemistry* **2003**, *42*, 829-837.
[14] Uversky, V. N.; Li, J.; Fink, A. L. *J. Biol. Chem.* **2001**, *47*, 44284-44296.
[15] Brutscher, B.; Bruschweiler, R.; Ernst, R. R. *Biochemistry* **1997**, *36*, 13043-13053.
[16] Benesch, J. P.; Robinson, C. V. *Nature* **2009**, *462*, 576-577.
[17] Miyawaki, A.; Llopis, J.; Heim, R.; McCaffery, J. M.; Adams, J. A.; Ikura, M.; Tsien, R. Y. *Nature* **1997**, 388, 882-887.
[18] As reviewed in: Ly, T.; Julian, R. R. *Angew. Chem. Intl. Ed.* **2009**, *48*, 7130-7137.

Chapter 2

RESIDUE-SPECIFIC RADICAL-DIRECTED DISSOCIATION OF WHOLE PROTEINS IN THE GAS PHASE

2.1 Introduction

The chemical bonds that comprise the naturally occurring amino acids are typically stable, as are the peptide bonds that link them together in linear fashion to form a protein.[19] From a functional perspective, it is reasonable that proteins are stable under conditions that mimic the cellular environment (near neutral pH, aqueous); however, sequence identification usually requires at least partial disassembly, meaning that some of these strong bonds must be broken. There are several methods available for doing this in the condensed phase. Selective or semi-selective cleavage can be performed in solution with trypsin or a variety of other enzymes.[20] Site selective chemical digestion is also possible.[21] Importantly, the results are predictable with reasonable reliability in both cases and are used for protein identification. Random fragmentation of the protein by implementing very harsh conditions is not typically performed.

In contrast to solution-phase methods, site-directed dissociation of peptides or proteins in the gas-phase remains a formidable challenge, despite a variety of available techniques. Collision induced dissociation (CID) is the most common

method. CID is performed by depositing energy into an ion by many collisions with neutral gases.[22] Typically, this leads to bond fracture at many sites that cannot be easily determined a priori; however, preferred cleavages have been observed under certain conditions. For example, Reid and co-workers used chemical derivatization to direct very selective side-chain fragmentation at methionine residues.[23] In other experiments, dissociation of the backbone has been shown to occur preferentially at acidic residues or proline.[24,25] However, factors such as charge state and sequence are known to influence these backbone fragmentations, and there is no facile method for rationally directing dissociation at these residues, particularly with whole proteins. Electron capture dissociation (ECD)[26] and electron-transfer dissociation (ETD)[27] utilize electrons to facilitate backbone dissociation via complex and incompletely understood mechanisms.[28,29] ECD and ETD are amenable to experiments with whole proteins, but do not typically exhibit site-selective dissociation. This may relate to difficulties associated with trying to control the location where an electron will interact with a biomolecule or to conformational heterogeneity.[30] Modest preferential dissociation is observed at disulfide bonds, but this does not yield sequence information.[31] Finally, less commonly employed photodissociation experiments using various ultraviolet wavelengths generally produce a large number of

fragments in experiments with peptides.[32,33] This type of photodissociation has not been explored with whole proteins, which would probably necessitate the implementation of a multiphoton approach.

Although none of these techniques yield selective dissociation alone, it may be possible to combine the strengths of each method to achieve this goal. For example, experiments with ECD and ETD suggest that the presence of a radical facilitates backbone dissociation. It is also known that CID will follow the lowest energy dissociation pathway (if a single such pathway can be established). In fact, recent experiments employing radical precursors that can be activated by CID to generate odd electron species have yielded some fragment selectivity following re-isolation and further collisional activation, although the initial radical generation also led to many undesirable side products.[34-36] Notwithstanding, these experiments successfully combined aspects of CID and ECD or ETD to yield results not obtainable with either method alone.

There are several inherent difficulties associated with directing dissociation of a whole protein at a particular amino acid in the gas phase. First, the chemical complexity of proteins is problematic. The side chains of the twenty canonical amino acids span a variety of functionalities, such as acids, bases, hydrocarbons, alcohols, aromatic rings, amides, etc. This list does not include additional

chemical groups that can be added post-translationally.[37] Thus, rationally directed dissociation must utilize chemistry capable of functioning in the presence of any combination of these chemical groups. Second, proteins are large molecules. Even ignoring the chemical details, proteins can contain thousands of atoms with 3N − 6 vibrational degrees of freedom (N = number of atoms). Site-directed dissociation requires that sufficient energy be directed to a particular bond in order to break it without disrupting the remaining bonds. There are two ways to do this: (1) the dissociation energy for one bond must be reduced significantly below the threshold for all other bonds,[38] or (2) sufficient energy to cleave a bond must be delivered to a specific bond, which must dissociate prior to intramolecular vibrational energy redistribution (IVR).[39] IVR must be circumvented in the case of a protein because randomization of the amount of energy required to break a single bond over the number of vibrational degrees of freedom in a protein would lead to no bond dissociation.[40]

Previous experiments have demonstrated that radical mediated backbone dissociation is a low-energy process.[41] In order to utilize this chemistry to initiate site-selective protein dissociation, the radical must be generated in a specific location. To facilitate maximum selectivity, radical generation should also be achieved without heating the entire molecule. Radicals can be generated site

specifically by dissociation of a carbon-iodine (C–I) bond.[42] C–I bonds of either the alkyl or aryl variety are known to undergo direct dissociation following photoactivation in the ultraviolet.[43] Aryl iodides can also undergo rapid predissociation following conversion of the excited-state localized in the aromatic ring to a dissociative state along the C–I bond.[44] Both processes occur in less than 600 fs.[45] In proteins, there are no C–I bonds in the naturally occurring amino acids; however, tyrosine and, to a lesser extent, histidine can be iodinated under mild conditions at biological pH using the chloramine T method. Iodination under these conditions generates 3-iodotyrosine and 4(5) iodohistidine.[46] This reaction requires only minutes to complete and exhibits a high degree of selectivity, meaning that the most exposed tyrosine residues (or residue) can be selectively iodinated.[47]

Herein, it is shown that 266-nm photoactivation of an iodotyrosine-containing protein cleanly leads to the generation of a highly localized radical on the aromatic ring of the modified tyrosine side chain. This radical can then be used to direct selective a-type fragmentation of the protein backbone at the modified residue following re-isolation and further collisional activation. The technique is demonstrated successfully with a variety of proteins. Proline and histidine are shown to be susceptible sites of secondary cleavage, leading to

backbone fragments when in close proximity to a modified tyrosine. Furthermore, it is demonstrated that this type of selective dissociation can be used to identify known proteins in proteomics experiments 3 to 4 orders of magnitude faster than that with traditional techniques.

2.2 Materials and Methods

2.2.1 Materials

All reagents and proteins were used without purification unless otherwise noted. Horse heart cytochrome c, bovine ubiquitin, chicken egg lysozyme, horse heart myoglobin, human hemoglobin, and dithiothreitol were purchased from Sigma Aldrich (St. Louis, MO). Chloramine-T, sodium metabisulfite, and sodium iodide were purchased from Fisher Chemical (Fairlawn, NJ). Water was purified to 18.2 MΩ resistivity using a Millipore Direct-Q (Millipore, Billerica, MA). Dialysis membranes and clips (MWCO = 3500 Da) were purchased from Spectra Por (Rancho Dominguez, CA).

2.2.2 Chloramine-T Iodination of Proteins

Proteins were iodinated by modification of a previously published procedure.[48] I^- is oxidized by the addition of chloramine-T to I_2, which adds as I^+ to the *ortho* position of tyrosine side-chain by electrophilic aromatic addition. After a short reaction period, sodium metabisulfite, a reducing agent, is added to

quench the iodination. Stoichiometric quantities of reagent were used to limit the extent and heterogeneity of iodination (1:1:2:4 protein:sodium iodide:chloramine-T:sodium metabisulfite). Reactions were initiated by mixing 1−5 mg of protein and sodium iodide in 0.2 mL water, followed immediately by addition of chloramine-T. Sodium metabisulfite was added after 10 min to quench the reaction. We found that lysozyme, which has four disulfide bonds, required harsher reaction conditions to produce comparable iodination yield to the other proteins investigated (2× iodination reagents, 30 min reaction time). Iodinated lysozyme and hemoglobin were purified by dialysis against water. This iodination procedure typically yields 30−60% mono-iodinated protein.

2.2.3 Reduction/Alkylation of Disulfide Bonds

Disulfide bonds in iodinated lysozyme were reduced and alkylated using dithiothreitol (DTT) and iodoacetamide.[49] Iodinated lysozyme was dissolved in 6 M urea in 0.025 M triethylammonium acetic acid (TEAA) buffered at pH 8.5. DTT was added in excess (5X disulfides, 50 °C, 2 h). Reduced lysozyme was then alkylated by addition of iodoacetamide (10× DTT + free thiols, room temperature, 1 h). Reduced, alkylated lysozyme was dialyzed against 1% acetic acid, lyophilized, and reconstituted in water before mass spectrometry.

2.2.4 Electrospray Mass Spectrometry

Solutions containing 10 μM of protein in 10% methanol were infused into a standard electrospray source. Ions were transferred into an LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific, Waltham, MA) for mass analysis and/or dissociation experiments. The posterior plate of the LTQ was modified with a quartz window to transmit fourth-harmonic (266 nm) laser pulses from a flashlamp-pumped Nd:YAG laser (Continuum, Santa Clara, CA). The LTQ Diagnostics software was configured to transmit a 3.3 V differential signal to a digital delay generator (Berkeley Nucleonics, San Rafael, CA) at the activation step of each MS^2 experiment. The digital delay generator parses the signal into a TTL pulse, which triggers the laser to fire.[50]. Photodissociation (PD) of iodinated protein always resulted in loss of iodine as the most abundant product ion (>50% relative abundance). Further MS^3 experiments were performed by re-isolation and collision induced dissociation (CID) of the photodissociation (PD) product. CID experiments were performed by applying an excitation voltage on mass-selected ions using default instrument parameters. Protein fragments were assigned with the aid of UCSF Protein Prospector v.4.0.8. The nomenclature[51] adopted also includes the superscripts "*" and "°" to indicate NH_3 loss and H_2O loss, respectively (e.g., $b^*_{58} \equiv b_{58} - NH_3$). Assignments do not distinguish between closed shell vs. radical (e.g., -H·) fragments.

2.2.5 Ab Initio Calculations

All calculations were performed using hybrid density functional theory (B3LYP) at the B3LYP/6-31G* level of theory as implemented in Gaussian 03 Version 6.1 Revision D.01. Unrestricted methods (UB3LYP) were used for all open-shell systems. Spin contamination was minimal for all systems (i.e. <S> ≈ 0.75, which is characteristic of a doublet spin state). Candidate structures were built using GaussView 3.0. Transition state (TS) calculations included optimizations and frequency calculations of the reactants, products, followed by a quasi-Newton synchronous transit (QST3) calculation[52] and a frequency calculation on the resulting TS candidates. Visualization of the single imaginary frequency found (TS1: −1575.33 cm-1, TS2: −1635.26 cm-1) verified that the structures are indeed saddle points connecting the reactant and product. Zero point corrected energies were used to calculate activation barriers.

C−H bond dissociation energies (BDEs) for tyrosine (C_β-H), asparagine (C_β-H), proline (C_δ-H), tryptophan (C_β-H), and valine (C_β-H) in the model peptide shown above were calculated using isodesmic reactions[53] with glycine as performed previously by Rauk et al.[54] The calculations were performed on the *trans* peptide, with the exception of proline, wherein both stereoisomers were considered. The C_α−H BDE for glycine was calculated to be 331.0 kJ/mol at a high

level of theory and was used here as the reference value for the isodesmic reactions.[54,55]

2.3 Results and Discussion

Experimental results obtained with cytochrome c (Cytc) are shown in Figure 2.1. Iodination was carried out under mild conditions, which produce mostly mono-iodo species. Electrospray ionization can be used to transfer the protein into the gas phase. The resulting desolvated, multiply protonated ions are collected into a linear ion trap. Photodissociation of isolated $[iodoCytc + 10H]^{10+}$ with 266 nm light yields a single peak corresponding to the loss of I· as shown in Figure 2.1a. Radical formation via direct dissociation (or fast predissociation) of the C–I bond was confirmed by comparing the results with CID experiments on the same peak. CID in an ion trap is caused by thousands of collisions with gaseous neutrals, ensuring that any observed dissociation follows a statistical process.[38,56] No loss of I· is observed with CID, suggesting that C–I cleavage is not a low-energy dissociation pathway. Therefore, photodissociation of the C–I bond must occur promptly from a dissociative excited electronic state. Photoactivation of the unmodified protein results in no appreciable dissociation, as shown in Figure 2.1b. This result indicates that after IVR, the energy from a single 266-nm photon absorbed by tryptophan or tyrosine is insufficient to

fragment the protein. In contrast, much smaller peptides containing tryptophan or tyrosine can be fragmented by 266-nm light.[33]

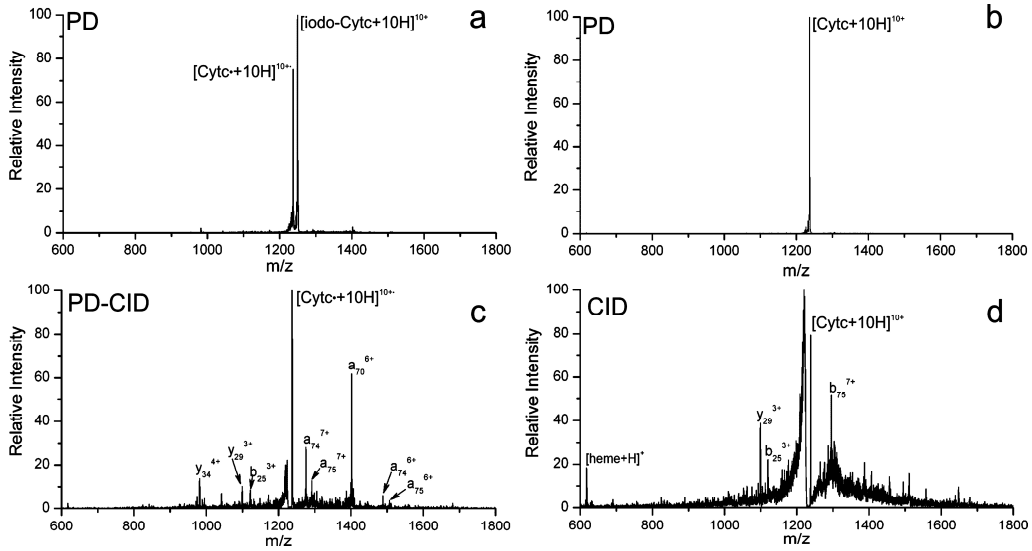


Figure 2.1 (a) Photodissociation spectrum for the +10 charge state of iodoCytc. The only significant observed loss is I· which occurs with an excellent yield. (b) In contrast, photoexcitation of the unmodified protein leads to no significant dissociation. (c) The radical protein generated in (a) is fragmented by CID. All of the labeled diagnostic fragments (except b_{25}^{3+}) represent radical-directed dissociations. Each of these radical-directed fragments is produced within four amino acids of the modified tyrosine residue. (d) CID spectrum of the unmodified protein is provided for comparison. Few fragments are resolved from the bulk of nonselective cleavages. It should be emphasized that all of the spectra shown (a–d) are not magnified in any way.

These results demonstrate that photoactivation of iodotyrosine is an effective method for generating a radical site selectively, even when attached to an entire protein. Cytc contains one tryptophan and four tyrosines, meaning that at least three tyrosine residues are unmodified. Despite the fact that these competitive chromophores are present in the protein, effective radical generation is still possible. It is likely that some absorption may occur at these residues; however,

this leads to fluorescence or internal conversion of the energy. Neither of these outcomes will dramatically affect the protein as demonstrated in Figure 2.1b, where native chromophore absorption leads to no observable change. The bond dissociation energy for an aromatic C–I bond is 280 kJ/mol.[57] The energy from a 266-nm photon is 450 kJ/mol, meaning that additional energy will be available following bond breakage. This energy can be dissipated translationally in the departing I· or by internal conversion within the side chain.

The protein radical produced by loss of I· can be re-isolated and fragmented by CID as shown in Figure 2.1c. This step of the experiment does not produce a single fragment, as observed following photodissociation; however, the number of abundant peaks is small, and examination reveals that all of them result from radical-directed fragmentation. Most importantly, the a_{74} fragment results from cleavage C-terminal to Tyr74 and is the second most abundant fragment, appearing twice in two different charge states. This fragment is not observed in a regular CID experiment performed on the unmodified protein (see Figure 2.1d for comparison). The a_{70}/y_{34} fragments are also facilitated by the tyrosine radical (as indicated by their absence in normal CID) on the n-terminal side of Pro71, which is in close proximity to Tyr74. For all other charge states studied (+9, +8),

a_{74} and a_{70} are the most intense ions observed. These and other secondary fragments are discussed further below.

In addition, the y_{29} and b_{25} ions that are generated in the standard CID experiment are observed but at much lower relative intensity, and the b_{75} ion (the most abundant product by standard CID) is not observed. Overall, these results suggest that the barriers to dissociation are lower for radical-directed cleavages. Furthermore, dissociation is heavily favored in the immediate vicinity where the radical is generated, including a diagnostic backbone cleavage at the modified tyrosine residue. Interestingly, there are four tyrosines in Cytc,[58] but the results in Figure 2.1 suggest that Tyr74 is the most reactive, allowing for selective derivatization. This is in agreement with previous observations where Tyr74 and (to a lesser extent) Tyr67 were found to be iodinated under harsher conditions.[59] Inspection of the crystal structure[58] reveals that Tyr67 is partially buried, supporting the possibility for preferential iodination of Tyr74 under mild conditions.

In Figure 2.2, the results for applying this technique to four additional proteins are shown. In these spectra, the relative contributions for fragments that appear only after CID of the protein radical are shown as a function of sequence. These plots emphasize only radical-directed fragments (which typically

dominate the spectrum). The contributions for related fragment types[51] (such as a and x, b and y, and c and z, as shown in Figure 2.3a) are summed, although the contributions from x-type fragments are very minor. The results for ubiquitin are shown in Figure 2.2a. Cleavage on the C-terminal side of Tyr 59 produces the second most abundant fragment, a_{59}. Additionally, a radical-mediated secondary fragmentation on the C-terminal side of Asn60 is observed, and several fragments occur at Arg54-Thr55. All of the abundant radical reactivity is localized in the vicinity of Tyr59 (ubiquitin has only one tyrosine). Examination of the crystal structure reveals that Arg54 is in close proximity to Tyr59.[60] This may suggest that portions of the crystal structure remain intact in the gas phase,[61] although this issue will require further examination.[62] Cleavage at His68 suggests some secondary iodination at this residue. Histidine iodination is discussed in further detail (vide infra). The identified modifications are in excellent agreement with previous iodination experiments where Tyr59 and His68 were found to be the only residues iodinated in ubiquitin.[63]

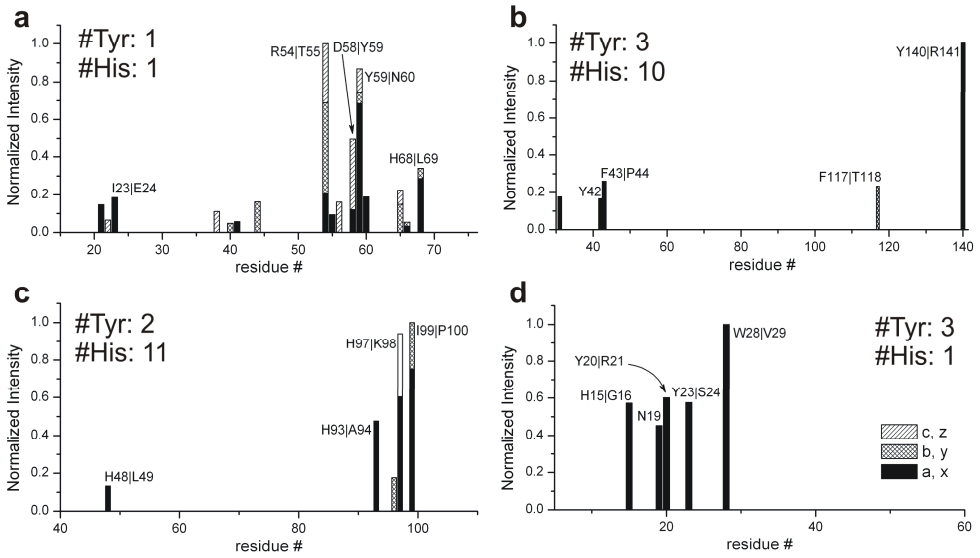


Figure 2.2 Results are shown for (a) ubiquitin (6+), (b) hemoglobin (9+), (c) myoglobin (9+), and (d) lysozyme (10+). Stack plots showing the relative contributions from different radical-directed fragments and the number of tyrosine and histidine residues for each protein. In general, fragmentation is localized around a modified tyrosine residue, producing dominant a-type fragment ions on the C-terminal side of the modified residue. X-type fragments are not common, meaning that most of the black bars represent intensity from a-type fragments. C and z and b and y type fragments are more equally distributed.

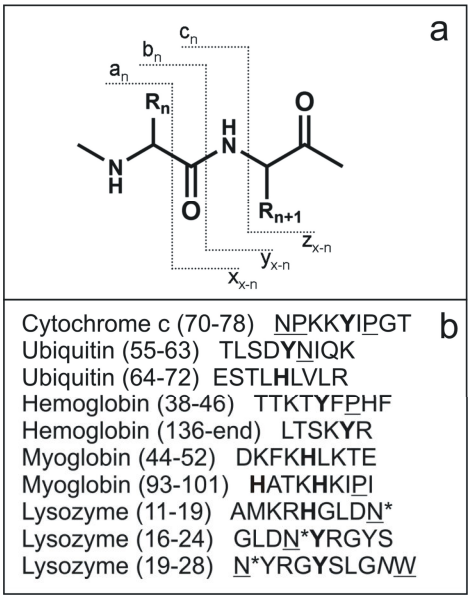


Figure 2.3 (a) Representative types of fragments typically observed in mass spectra. a/ x, b/y, c/z form complementary pairs. X = total number of amino acids in protein. (b) Local sequence around putatively iodinated residues (in bold) for each protein. Proline, asparagine, or tryptophan residues that undergo secondary fragmentation are underlined. Italicized residues represent sites where secondary fragmentations are expected but not observed. Numbers in parentheses represent amino acid numbers for the sequence shown. The "*" represents all the same asparagine residue.

In Figure 2.2b, the results for the A-chain of human hemoglobin, a substantially larger protein, are shown. The most intense peak results from a-type fragmentation on the C-terminal side of Tyr140. There are no other fragments observed in the vicinity of Tyr140, indicating that secondary fragments are sequence- or structure-dependent and will not always be observed. In fact, fewer radical-directed peaks are generated for hemoglobin than for ubiquitin, confirming that selectivity is not a function of protein size. Importantly, this should allow the technique to be extended to even larger

proteins, with the ultimate limitation being mass resolution. Although no substantial secondary backbone fragmentation is observed in the vicinity of Tyr140, there are fragments in the vicinity of Tyr42 that suggest some iodination occurs at this residue. Hemoglobin contains 3 tyrosine residues in total and 10 histidines. Tyr24 is substantially buried (precluding iodination), yet many of the histidine residues are fully exposed. Thus, observation of radical fragmentation only in the vicinity of tyrosine residues confirms the selectivity for tyrosine iodination over histidine if both residues are chemically available.[64]

Myoglobin is similar in size to hemoglobin but yields slightly different results from the other proteins, as shown in Figure 2.2c. For myoglobin, the second most abundant fragmentation is observed on the C-terminal side of His97 (a_{97}). A dominant secondary fragment (a_{99}) occurs in close proximity, N-terminal to Pro100. A combination of top-down and bottom-up MS techniques have localized the iodination site to Tyr103.[65] These results suggest that the radical may migrate from the initial tyrosine residue to other residues nearby (such as histidine).

The results for lysozyme are shown in Figure 2.2d. This protein contains disulfide bonds, which must be reduced and alkylated prior to analysis. For lysozyme, a-type fragmentation on the C-terminal side of two tyrosine residues

(Tyr20 and Tyr23) is observed.[66] Iodination is performed on the folded protein prior to alkylation, suggesting that these two residues have similar solvent accessibility. Indeed, previous reports have identified both Tyr20 and Tyr23 as sites of iodination.[67] The solvent accessibility of these two residues in the crystal structure supports these results.[68] However, recent data suggest that Tyr23 is iodinated exclusively.[65] Abundant cleavage is also observed on the C-terminal side of a proximal tryptophan residue. Tryptophan should not be iodinated by our procedure, indicating that this is a secondary fragmentation.

2.3.1 Cleavage at Iodinated Residues

It is interesting to consider why cleavage occurs specifically at modified tyrosine or at nearby histidine residues for every protein that we have examined. In both cases, the radical that is initially produced by the loss of iodine should be very reactive. This prediction can be evaluated more quantitatively in terms of relative C–H bond dissociation energies. Hydrogen atom abstraction is thermodynamically favorable if the C-H bond formed is higher in energy than the C-H bond dissociated. The C−H bond dissociation energy for either radical center (tyrosine or histidine) is ~490 kJ/ mol.[69,70] This is higher than most other nonaromatic C-H bond dissociation energies present in a protein by 80−160 kJ/mol, making nearly all possible subsequent hydrogen atom abstractions

exothermic. Given that either radical should abstract the first hydrogen that it comes in contact with, specificity would not be expected unless subsequent reactions all led to the same product. (which would seem unlikely).[71]. For tyrosine, there is a further difficulty because the initially formed radical cannot interact with the protein backbone where cleavage is observed. Previous work has shown that a-type fragments are typically produced in radical systems by abstraction of the β hydrogen, followed by cleavage of the C_α-C(=O) bond of the backbone.[36] Due to steric constraints caused by the rigid nature of the tyrosine side chain, the initially produced radical and the β hydrogen cannot interact. In fact, they are separated by >4 Å, with the radical pointed away from the hydrogen. Therefore, the selective reactivity and the direct a-type cleavages observed at tyrosine residues suggest that a fast rearrangement of the initially very reactive radical may occur.

Table 2.1 Calculated Bond Dissociation Energies (B3LYP 6-31G*)[a]

amino acid	BDE (kJ/mol)
Tyr C_β–H	381.8
Pro C_δ–H cis	393.7
Pro C_δ–H trans	391.5
Trp C_β–H	350.9
Val C_β–H	400.6
Asn C_β–H	418.6

[a] See Materials and Methods for calculation details.

2.3.2 Secondary Fragments

The local sequence for each tyrosine or histidine that exhibited an a-type fragment resulting from cleavage on the C-terminal side of the residue (i.e., residues that were likely iodinated) is shown in Figure 2.3b. In addition to the diagnostic a-fragment produced at the modified residue, the proximity of proline also appears to facilitate secondary fragments. Secondary cleavages on the N-terminal side of proline (producing y and a fragments) are observed when proline is within four amino acids of the radical site. It should also be noted that fragmentation on the C-terminal side of asparagine to produce a-type fragments is observed several times in Figure 2.3. However, one of these cleavages is in conjunction with a proline fragmentation, and one asparagine (Asn27 in lysozyme) fails to produce any fragmentation; therefore, it is unclear whether

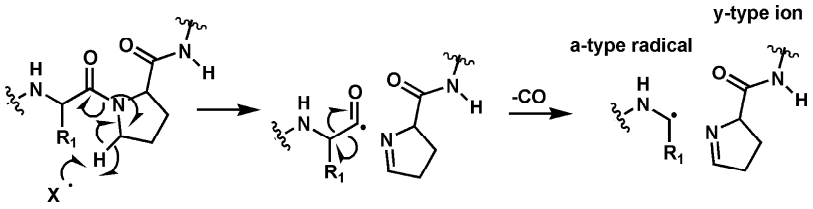
asparagine facilitates secondary dissociations. The abundant a-type fragment observed at Trp28 in lysozyme, five residues from Tyr23, is most likely a secondary fragment as well, but there are insufficient tryptophan residues in these proteins to draw any general conclusions. For proline, which represents 5% of the amino acids in proteins, there is a 44% chance that it will be within ±4 residues of a tyrosine. Therefore, it is likely that secondary cleavages will be observed frequently, potentially making them suitable for obtaining additional sequence information after initial identification of tyrosine or histidine cleavage sites.

A previously proposed mechanism invoking hydrogen abstraction by a radical can be used to rationalize the secondary fragments at asparagine and tryptophan highlighted in Figure 2.3b.[36] In this mechanism, hydrogen abstraction from the β-carbon is followed by β-elimination to produce an a-type fragment. Therefore, amino acids with relatively weakly bound β-hydrogens should be more susceptible to backbone dissociation. Previous experiments have demonstrated that abstraction of the α-hydrogen from asparagine occurs with a lower barrier than any other amino acid.[54] Furthermore, abstraction of the β-hydrogen from asparagine by OH radical is preferred over abstraction of the α-hydrogen,[72] suggesting that loss of the β-hydrogen should be facile. In contrast to

these results, calculated bond dissociation energies do not suggest that asparagine should be susceptible to this type of attack. The results in Table 2.1 suggest that abstraction of the β-hydrogen from asparagine should not be facile relative to other amino acids. Further experiments will be required to resolve this apparent contradiction.[73]

For proline, experiments in solution suggest that abstraction of a δ-hydrogen is preferred over the β-position.[74,75] A proposed mechanism which accounts for the secondary fragmentations observed in our experiments is shown in Scheme 2.1.[76] Abstraction of a δ-hydrogen leads to homolytic cleavage of the peptide bond.[77] The resulting y-type ion is stable; however, the b-type counterpart is not stable and will rapidly lose CO.[78] CO loss yields a radical a-type ion. Therefore, if this mechanism is correct, complimentary ions produced by radical-directed dissociation at proline will yield a and y ions, not b and y ions. Indeed, these are the only type of fragments detected for dissociations at the prolines underlined in Figure 2.3b. Furthermore, CID of unmodified Cytc yields complimentary b_{75} and y_{29} ions on the N-terminal side of Pro76 due to the proline effect (see Figure 2.1d).[25] CID of radical Cytc yields a_{75} and y_{29} ions; the b_{75} ion is not observed despite being the most abundant fragment for the unmodified protein. The complete absence of the b_{75} ion is best rationalized by the radical-directed

dissociation mechanism in Scheme 2.1 where b-type ions are not allowed. These observations also offer further proof that radical-directed dissociations occur with lower energy barriers than even favorable non-radical dissociation processes such as those observed at proline residues.



Scheme 2.1

2.3.3 Applications to Database Searching

Experimental methods currently employed in proteomics research are generally considered to be more advanced than the bioinformatics tools that have been developed to interpret the results.[79] Thus, one might conclude that the principle barrier to further progress in proteomics research is bioinformatics. However, the difficulty is also due, in part, to the type of data that is collected in the experiments. As discussed in the introduction, non-specific bond cleavage is typical among all mass spectrometric dissociation techniques used for protein identification. This leads to the production of vast amounts of raw data, but the information content of this data is frequently low.[80] Thus, the nature of the experimental results is currently creating an unduly large burden which

bioinformatics has been unable to resolve. An alternative strategy is to acquire better data with high information content that can be more easily interpreted. Radical-directed dissociation can be used to obtain specific (rather than random) information from peptide or protein fragmentation by dictating where fragmentation occurs. As a result, it is possible to reduce the amount of time needed for data analysis significantly.

Either the most intense or second most intense observed peak in each of these experiments results from dissociation on the C-terminal side of tyrosine. The one exception occurs for myoglobin where tyrosine is not modified, but in this case, the most intense cleavage is C-terminal to histidine. Given this information, four out of five proteins can be positively identified by checking two experimental peaks against all possible C-terminal a-type cleavages at tyrosine residues. There are 13 tyrosines present out of a total of 603 residues in the 5 proteins studied herein. Therefore, the numbers dictate that 10 experimental peaks would be checked against 13 possible matches for a total of 130 required calculations (assuming the protein is known). In contrast, standard techniques using dissociation results produced by random processes would require checking significantly more peaks. For example, if ECD or ETD were used and full sequence coverage was obtained, then 1196 potential peaks

(assuming c and z ions only) would need to be checked against 1196 possible matches for a total of 1.4 million calculations. The database search using the standard approach would therefore require 4 orders of magnitude more time to perform. If the sequence coverage were reduced to only 50%, then 598 peaks would be checked leading to 700,000 calculations. The potential gain would still be a factor of 5000. Inclusion of histidine residues in the radical-directed approach would require 390 calculations and lead to the identification of all five proteins. Furthermore, the secondary fragments generated at proline should be present 44% of the time, potentially allowing for independent confirmation of the identification. It should be mentioned that the presence of post-translational modifications complicates data analysis; however, this problem will scale equally for the traditional and radical-directed approaches.

2.4 Conclusions

The results presented demonstrate that proteins can be cleaved efficiently and selectively at tyrosine or histidine residues in a gas-phase experiment. This abiotic approach is achieved in milliseconds by replacing a single atom in an entire protein. This capability will enable faster bioinformatics that will greatly speed the interpretation of proteomics data. Furthermore, a side benefit of this research is rapid new method for identifying sites of iodination in a protein.

Iodination is frequently performed for protein structure analysis, and native iodotyrosines represent essential components in thyroid proteins such as thyroglobulin.[81]

---

[19] Radzicka, A.; Wolfenden, R. *J. Am. Chem. Soc.* **1996**, *118*, 6105−6109.
[20] Schnolzer, M.; Jedrzejewski, P.; Lehmann, W. D. *Electrophoresis* **1996**, *17*, 945−953.
[21] Lee, T. D.; Shively, J. E. *Methods Enzymol.* **1990**, *193*, 361−374.
[22] Wells, J. M.; McLuckey, S. A. *Methods Enzymol.* **2005**, *402*, 148−185.
[23] Reid, G. E.; Roberts, K. D.; Simpson, R. J.; O'Hair, R. A. J. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1131−1150.
[24] Tsaprailis, G.; Somogyi, A.; Nikolaev, E. N.; Wysocki, V. H. *Int. J.Mass Spectrom.* **2000**, *196*, 467−479.
[25] Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1996**, *31*, 1185−1187.
[26] Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563−573.
[27] Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528−9533.
[28] Anusiewicz, W.; Berdys-Kochanska, J.; Simons, J. *J. Phys. Chem. A* **2005**, *109*, 5801−5813.
[29] Syrstad, E. A.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 208−224.
[30] Fragmentation is reduced at low temperatures: Mihalca, R.; Kleinnijenhuis, A. J.; McDonnell, L. A.; Heck, A. J. R.; Heeren, R. M. A. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1869−1873.
[31] Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857−2862.
[32] Thompson, M. S.; Cui, W. D.; Reilly, J. P. *Angew. Chem. Int. Ed.* **2004**, *43*, 4791−4794.
[33] Oh, J. Y.; Moon, J. H.; Kim, M. S. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2706−2712.
[34] Wee, S.; Mortimer, A.; Moran, D.; Wright, A.; Barlow, C. K.; O'Hair, R. A. J.; Radom, L.; Easton, C. J. *Chem. Comm.* **2006**, *40*, 4233−4235.
[35] Masterson, D. S.; Yin, H. Y.; Chacon, A.; Hachey, D. L.; Norris, J. L.; Porter, N. A. *J. Am. Chem. Soc.* **2004**, *126*, 720−721.
[36] Hodyss, R.; Cox, H. A.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2005**, *127*, 12436−12437.
[37] Schweppe, R. E.; Haydon, C. E.; Lewis, T. S.; Resing, K. A.; Ahn, N. G. *Acc. Chem. Res.* **2003**, *36*, 453−461.
[38] Marzluff, E. M.; Beauchamp, J. L. In *Large Ions: Their Vaporization, Detection, and Structural Analysis*; Wiley: New York, 1996; pp 115−143.
[39] Stannard, P. R.; Gelbart, W. M. *J. Phys. Chem.* **1981**, *85*, 3592−3599.

[40] Andersen, L. H.; Bluhme, H.; Boye, S.; Jorgensen, T. J. D.; Krogh, H.; Nielsen, I. B.; Nielsen, S. B.; Svendsen, A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2617−2627.
[41] Haselmann, K. F.; Jorgensen, T. J. D.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 2260−2265.
[42] Thoen, K. K.; Pérez, J.; Ferra, Jr. J. J.; Kenttämaa, H. I. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1135−1140.
[43] Kavita, K.; Das, P. K. *J. Chem. Phys.* **2002**, *117*, 2038−2044.
[44] Marconi, G. *J. Photochem.* **1979**, *11*, 385−391.
[45] Cheng, P. Y.; Zhong, D.; Zewail, A. H. *Chem. Phys. Lett.* **1995**, *237*, 399−405.
[46] Miyashita, M.; Yamashita, S. *J. Chromatogr.* **1989**, *475*, 135−144.
[47] McGowan, E. B.; Stellwag, E. *Biochemistry* **1970**, *9*, 3047.
[48] Regoeczi, E. *Iodine-Labeled Plasma Proteins*; CRC Press: Boca Raton, FL, 1984.
[49] Goldberg, M. E.; Rudolph, R.; Jaenicke, R. *Biochemistry* **1991**, *30*, 2790−2797.
[50] Kim, T. Y.; Thompson, M. S.; Reilly, J. P. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 1657−1665.
[51] Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601−601.
[52] Peng, C. Y.; Schlegel, H. B. *Israel J. Chem.* **1993**, *33*, 449−454.
[53] Hehre, W. J.; Ditchfield, R.; Radom, L.; Pople, J. A. *J. Am. Chem. Soc.* **1970**, *92*, 4796−4801.
[54] Rauk, A.; Yu, D.; Taylor, J.; Shustov, G. V.; Block, D. A.; Armstrong, D. A.*Biochemistry* **1999**, *38*, 9089−9096.
[55] Armstrong, D. A.; Yu, D.; Rauk, A. *Can. J. Chem.* **1996**, *74*, 1192−1199.
[56] Laskin, J.; Futrell, J. H.; Chu, I. K. *J. Am. Chem. Soc.* **2007**, *129*, 9598−9599.
[57] Blanksby, S. J.; Ellison, G. B. *Acc. Chem. Res.* **2003**, *36*, 255−263.
[58] Bushnell, G. W.; Louie, G. V.; Brayer, G. D. *J. Mol. Biol.* **1990**, *214*, 585−595.
[59] Santrucek, J.; Strohalm, M.; Kadlcik, V.; Hynek, R.; Kodicek, M. *Biochem. Biophys. Res. Commun.* **2004**, *323*, 1151−1156.
[60] Vijaykumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531−544.
[61] Baker, E. S.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1188−1195.
[62] Clemmer, D. E.; Jarrold, M. F. *J. Mass Spectrom.* **1997**, *32*, 577−592.
[63] Pickart, C. M.; Haldeman, M. T.; Kasperek, E. M.; Chen, Z. J. *J. Biol. Chem.* **1992**, *267*, 14418−14423.
[64] Ramachandran, L. K. *Chem. Rev.* **1956**, *56*, 109−218.
[65] Sun, Q.; Yin, S.; Loo, J. A.; Julian, R. R. *Anal. Chem.* **2010**, in press
[66] Beddell, C. R.; Blake, C. C. F.; Oatley, S. J. *J. Mol. Biol.* **1975**, *97*, 643.
[67] Hayashi, K.; Shimoda, T.; Imoto, T.; Funatsu, M. *J. Biochem.* **1968**, *64*, 365.
[68] Diamond, R. *J. Mol. Biol.* **1974**, *82*, 371.
[69] da Silva, G.; Moore, E. E.; Bozzelli, J. W. *J. Phys. Chem. A* **2006**, *110*, 13979−13988.
[70] da Silva, G.; Chen, C. C.; Bozzelli, J. W. *Chem. Phys. Lett.* **2006**, *424*, 42−45
[71] Subsequent experimental investigations have suggested that radical migration can occur to a significant extent and leads to selective fragmentation. See Chapter 3 for details.
[72] Galano, A.; Alvarez-Idaboy, J. R.; Bravo-Perez, G.; Ruiz-Santoyo, M. E. *J. Mol. Struct. THEOCHEM* **2002**, *617*, 77−86.

[73] In proteins, it appears that structure is an equally important factor in determining where radical directed fragmentation occurs. See Chapter 8 for details.
[74] Easton, C. J. *Chem. Rev.* **1997**, *97*, 53−82.
[75] Nukuna, B. N.; Goshe, M. B.; Anderson, V. E. *J. Am. Chem. Soc.* **2001**, *123*, 1208−1214.
[76] A similar result is obtained by abstracting the alpha hydrogen: Fung, Y. M. E.; Chan, T. W. D. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1523−1535.
[77] Wee, S.; O'Hair, R. A. J.; McFadyen, W. D. *Int. J. Mass Spectrom.* **2004**, *234*, 101−122.
[78] Haselmann, K. F.; Budnik, B. A.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 2242−2246.
[79] Palagi, P. M.; Hernandez, P.; Walther, D.; Appel, R. D. *Proteomics* **2006**, *6*, 5435−5444.
[80] Fenyo, D.; Beavis, R. C. *Trends Biotech.* **2002**, *20*, S35−S38.
[81] Medeiros-Neto, G.; Targovnik, H. M.; Vassart, G. *Endocr. Rev.* **1993**, *14*, 165−183.

Chapter 3

RADICAL-DIRECTED DISSOCIATION IS DRIVEN BY EXOTHERMIC RADICAL MIGRATION FOLLOWED BY BETA CLEAVAGE

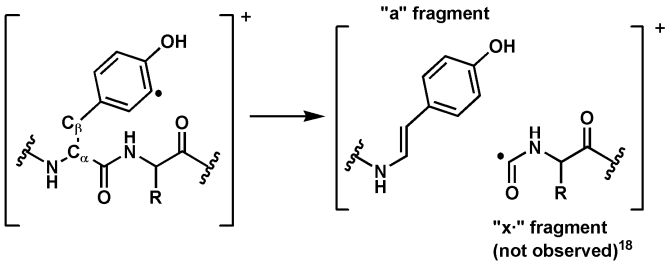
3.1 Introduction

Odd electron species play important roles in living systems, including assisting enzymatic catalysis of biologically important redox reactions.[1] Additionally, interest in biomolecular radical chemistry has been growing in the mass spectrometry field due to the development and successful application of radical-based dissociation of biomolecules. Electron capture dissociation (ECD) and electron transfer dissociation (ETD) are both radical-based techniques that have become conventional methods in proteomic analyses.[2,3] The mechanisms governing dissociation for both techniques are still under investigation, but there is significant evidence suggesting radical migration plays an important role.[4] For example, although it has been suggested that capture of a single electron leads to a single backbone cleavage,[5,6] ECD of a cyclic peptide yields separated backbone products, which can only occur if the backbone is cleaved at multiple sites. O'Connor and coworkers have proposed that hydrogen-deficient z· fragments produced by ECD further dissociate through a series of cascading radical reactions.[7] This hypothesis is corroborated by results from other experiments

where covalent attachment of radical trap moieties reduces the sequence coverage observed in ECD experiments and by evidence that extensive hydrogen atom migration occurs after electron capture.[8-10] Additional radical-based dissociation methods have been developed to interrogate peptides and proteins. In these experiments, specific bonds are modified either covalently,[11-13] or by attachment of a metal-complex,[14-16] such that the bond is predisposed towards homolytic cleavage. Activation of these precursor molecules by CID generates the radical species.

Photodissociation (PD) is an alternative method for generating radicals by electronic excitation of a suitable precursor.[17] Undesirable side-reactions are avoided with PD due to the specificity of the excitation chemistry. PD can be coupled with noncovalent or covalent delivery of the photolabile group, and a variety of precursor chemistries have been identified.[18-20] We have previously demonstrated that PD of iodine-labeled proteins with 266 nm photons results in selective homolytic dissociation of the carbon-iodine bond to generate a radical regiospecifically on the side-chain of tyrosine in high yield. Subsequent collisional activation of these radical proteins yields a-type backbone fragmentation at the modified tyrosine. Scheme 3.1 shows the reactant radical species and the backbone fragmentation products. The observed selectivity

favoring fragmentation at tyrosine is puzzling for two reasons. First, the initial radical is expected to be highly reactive, which is estimated by comparing the relative C-H bond dissociation energies (BDEs) found in peptides to the initial radical site. The C-H BDE of benzene, which should be similar to the *ortho* C-H of tyrosine, is 472 ± 2 kJ/mol.[21] Typical C-H BDEs in the remaining peptide are much lower. For example, the C_α-H BDEs have been calculated to be in the range of 316 – 369 kJ/mol.[22] Thus, abstraction of a hydrogen atom by the initial tyrosyl radical will be exothermic for all non-aromatic hydrogens, in some cases by a significant amount. Therefore, radical migration away from tyrosine should be nonselective and facile, absent any significant barriers to migration. The second reason that the observed selectivity is not expected is because the radical is separated from the scissile site by four bonds, and cannot generate the observed products directly. Furthermore, electronic rearrangement is unlikely due to the localization of the radical in a σ orbital.[23] This fact, along with the steric constraints imposed by the rigid tyrosine sidechain, suggests that radical rearrangement or migration must occur prior to dissociation.



Scheme 3.1

The present work focuses on a small group of peptides and their deuterated derivatives to examine the mechanism for selective fragmentation at tyrosine. Covalent labeling experiments reveal that the β hydrogen of tyrosine and the α hydrogen from either adjacent amino acid are involved in backbone fragmentation at tyrosine. We propose a "rebound" mechanism to account for these results. The initial tyrosyl radical abstracts a hydrogen atom from the α-carbon of an adjacent residue. This α-radical then "rebounds" and abstracts the hydrogen from the β-carbon of tyrosine due to a favorable six-membered transition state. Subsequent β-scission leads to the observed backbone fragmentation. Results from deuterated peptides also indicate that radical migration and multiple migrations can occur. Interestingly, it is observed that radicals are most likely to initiate backbone fragmentation from β-positions with the most radical stabilization (as can be quantified by the lowest C_β-H BDE). Theoretical calculations suggest that this apparent combination of kinetic and

thermodynamic control results from the facile migration of radicals, which is constrained by barriers that are substantially below dissociation thresholds.

3.2 Materials and Methods

All reagents were used as received unless specified. Diethyl ether, ethyl acetate, dichloromethane, absolute ethanol, LC-MS grade acetonitrile, methanol, chloramine-T, sodium metabisulfite, sodium chloride, sodium sulfate (anhydrous), sodium iodide, sodium bicarbonate were purchased from Thermo-Fisher Scientific (Waltham, MA). Dimethylformamide (DMF), hexanes, fluorenylmethyloxycarbonyl (Fmoc)-protected α-aminoisobutyric acid (AIB), acetyl anhydride and acetic acid were purchased from EMD Biosciences. Pyridine, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), ninhydrin, 2,6-dichlorobenzoyl chloride (DCB), (D)-4-hydroxyphenylglycine and di-isopropylethylamine (DIPEA) were purchased from Acros Organics. α-d_2-glycine was purchased from CDN Isotopes (Pointe-Claire, Quebec, Canada), while β-d_2-tyrosine was purchased from Cambridge Isotope Laboratories (Andover, MA). Fmoc-protected glycine, arginine(Pbf), leucine, and tyrosine(tBu) were purchased from LC Sciences (Houston, TX). Fmoc-protected alanine, d_4-methanol (99.8%), D_2O (99.9%) and triisopropylsilane (TIS) were purchased from Sigma-Aldrich/Fluka (St. Louis, MO). Petroleum

ether (Mallinckrodt-Baker, Phillipsburg, NJ), piperidine (Anaspec, San Jose, Ca), Fmoc-OSuccinimide (FmocOSu) (ChemPep, Miami, FL), 1-hydroxybenzotriazole monohydrate (1-HOBt) (Chem Impex-International, Wood Dale, IL) were purchased. Aluminum-backed thin-layer chromatography (TLC) plates were purchased from Sorbent Technologies (Atlanta, GA). RGYALG, RPPGYSPFR, and RPPGFSPYR peptides were purchased from American Peptide Company (Sunnyvale, CA).

Free amino acids were protected with FmocOSu according to published procedures.[24,25] Briefly, 1 eq. FmocOSu was added to 1 eq. $NaHCO_3$ and 1 eq. amino acid in water/acetone. The reaction mixture was stirred for 24 hours, after which the absence of a free amino group was checked by TLC-ninhydrin analysis. Peptides were synthesized manually using Fmoc chemistry on a Wang resin (Acros Organics).[26] The DCB method was used to load the first residue onto the resin. Subsequent chain extension reactions were achieved using PyBOP activation. Peptides were cleaved from the solid support using a solution of 95:2.5:2.5 TFA:H_2O:TIS for 2 hours, after which the solvent was evaporated. The white peptide film was carefully washed with cold diethyl ether, re-dissolved in water and lyophilized.

Peptides were iodinated using a published protocol.[18] In a typical iodination reaction, 1 eq. of chloramine-T was added to a 50-100 μL solution containing 1 eq. peptide (~1-50 nmol) and 1 eq. sodium iodide. After 30 seconds, the reaction was quenched by adding 2 eq. sodium metabisulfite. Peptide acetylation was accomplished by adding 20 μL of acetic anhydride to 5 nmol of iodinated peptide dissolved in 100 mM ammonium acetate buffer. After a 10 minute reaction time, the mixture was lyophilized. A separate aliquot of 5 nmol iodinated peptide was lyophilized and redissolved in 50/50 D_2O/CD_3OD to substitute all exchangeable hydrogens (i.e. attached to heteroatoms) with deuteriums.

Solutions containing 10 μM peptide in 50/50 water/acetonitrile were directly infused into a standard electrospray source of an LTQ linear ion trap mass spectrometer (Thermo-Fisher, Waltham, MA). The LTQ has been modified to permit photodissociation experiments as previously described.[18] Ions were collected in the linear ion trap, where mass spectrometry and dissociation experiments were performed. The electrospray needle voltage and capillary temperature were fixed at 5.0 kV and 215 °C, respectively. Voltages for the transfer optics were optimized to maximize the intensity of the singly protonated peak. We have found that the narrow isolation windows required to select the monoisotopic peak causes significant precursor ion heating. This leads to

increased hydrogen/deuterium scrambling compared to when a wide isolation width is used. Therefore, the following experimental scheme was used for all peptides to minimize experimental artifacts arising from precursor ion heating. A 1.4 m/z isolation window was used to select the monoisotopic, singly protonated peak. A second isolation with a large window (10.0 m/z) was used to allow the ions to equilibrate to approximately room temperature,[27] after which the laser was triggered. For the final step (MS^4), the radical peptide is re-isolated with a 10.0 m/z isolation window and collisionally activated. The default activation time (30.0 ms) was used for all MS^n steps.

Ab initio calculations were performed using the Gaussian 03 Version 6.1 Revision D.01 software suite using the hybrid density functional B3LYP and MPW1K methods with the 6-31++g(d,p) and 6-31+g(d,p) basis sets, respectively. Spin contamination was minimal for all systems. Candidate transition state structures were found using the quasi-Newton synchronous transit (QST3).[28] Frequency analyses at the same level of theory on all structures revealed no imaginary frequencies for minima and a single imaginary frequency for each transition state structure. Visualization of the single imaginary frequency confirmed that the structure is indeed a saddle point connecting the reactants

and products. Activation barriers are corrected for vibrational zero point energies, which are unscaled.

3.3 Results and Discussion

Figure 3.1a shows the PD spectrum for protonated iodo-RGYALG, $[RG^{I}YALG+H]^{+}$. The most abundant product is loss of iodine atom to yield a hydrogen deficient peptide radical. Hydrogen deficient is used in the context of a peptide missing a hydrogen relative to the mass of a fully protonated, even-electron ion (with no implications about the specific structure of that ion).[29] Due to the small size of this peptide, internal conversion of residual energy from the laser pulse is sufficient to fragment some ions, resulting in the additional loss of isopropene from leucine (-56 Da). PD also produces a single backbone fragment in small abundance, the a_3 ion, which corresponds to cleavage of the C_α-C(O) bond of tyrosine. The yield of all radical product ions is increased by further collisionally activating the peptide radical, which yields the spectrum shown in Figure 3.1b. In contrast to PD-CID of the iodopeptide, CID of the non-iodinated peptide produces loss of ammonia and b ions.[18] Thus, the sidechain losses and a_3 ion are due to the unique chemistry afforded by the presence of the radical. Importantly, the selective fragmentation at tyrosine for this peptide resembles the backbone fragmentation at tyrosine residues observed for whole proteins.[18]

Investigation of this smaller system should allow the details of fragmentation at tyrosine to be elucidated.

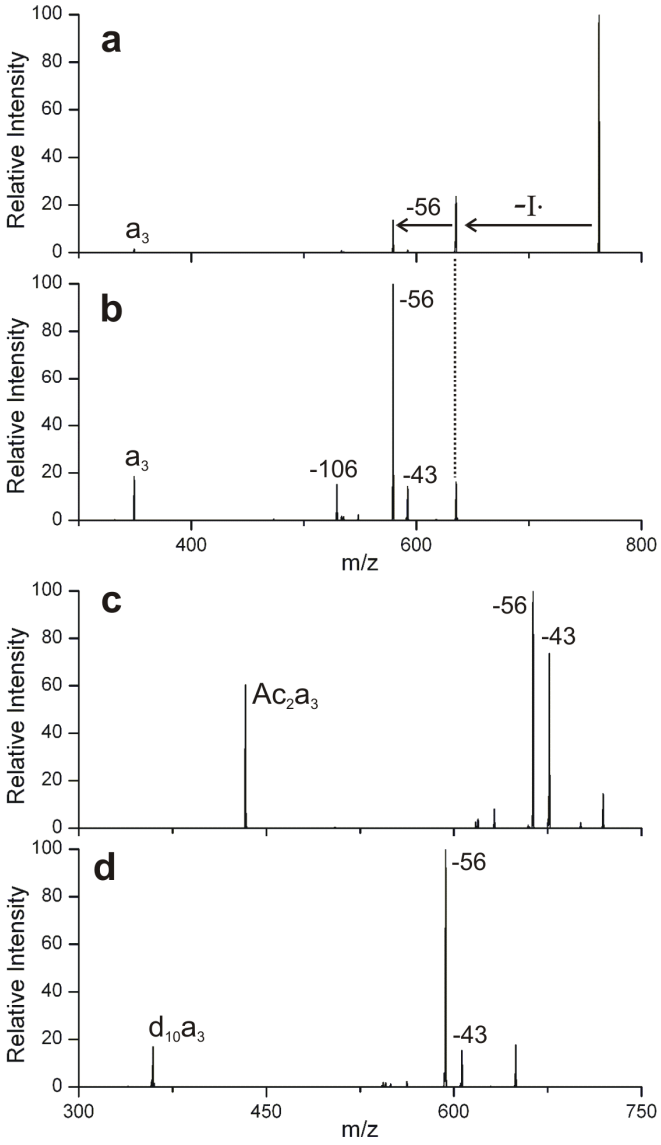


Figure 3.1 (a) Photodissociation of $[RG^{I}YALG+H]^{+}$ produces loss of I· as the most abundant product. (b) CID of $[RGYALG\cdot+H]^{+}$ yields selective backbone fragmentation at the tyrosine residue (a_3). (c) CID of $[Ac_2RGYALG\cdot+H]^{+}$. Acetylation of the tyrosine sidechain eliminates tyrosine sidechain loss. (d) CID of the fully deuterium-exchanged radical peptide, $[d_{14}RGYALG\cdot+H]^{+}$ yields a kinetic isotope effect only for the loss of

tyrosine sidechain. Thus, direct involvement of heteroatom hydrogens in backbone dissociation is unlikely.

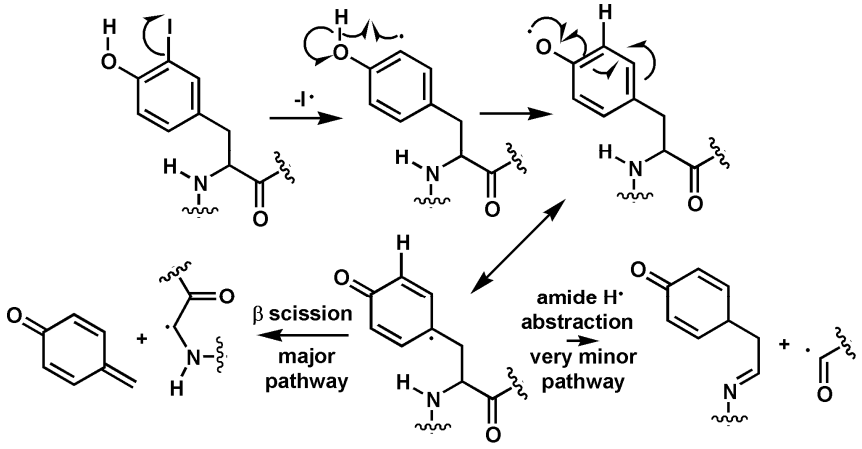
In addition to backbone fragmentation at tyrosine, sidechain fragmentation at leucine and tyrosine is observed in Figure 3.1b. Sidechain losses are commonly observed in radical-mediated peptide dissociation.[11-20] Most side chain losses are explained by hydrogen atom abstraction followed by β scission. A summary of the sites of hydrogen atom abstraction, the chemical formulae and the nominal masses of the observed sidechain fragments are given in Table 3.1.

Table 3.1 Sidechain fragmentation summary

Amino Acid	Hydrogen Abstracted	Nominal Mass (Da)	Fragment
L	α	43	$CH_3CH{\cdot}CH_3$
L	γ	56	$CH_2{=}C(CH_3)_2$
R	α	86	$CH_2CH_2NH(C{=}NH)NH_2$
R	γ	99	$CH_2{=}CHCH_2NH(C{=}NH)NH_2$
Y	ε	106	$CH_2{=}C_6H_4O$

Specific functional groups on the peptide, RGYALG, were covalently modified to determine whether these functionalities are essential to a_3 formation. The alcohol of the tyrosine sidechain was the first functionality tested. In theory, rearrangement of the radical within the tyrosine residue would yield the a_3 fragment.[18] Scheme 3.2 shows the details of this mechanism. The resonance-

stabilized radical on the γ carbon of tyrosine can abstract the amide hydrogen to yield backbone fragmentation (bottom right). Alternatively, β scission produces loss of tyrosine sidechain (bottom left). To test this hypothesis, RG^{I}YALG was acetylated at the N-terminus and the tyrosine OH. Figure 3.1c shows the PD-CID spectrum of the doubly acetylated peptide, $[Ac_2RGYALG\cdot+H]^+$. The abundance of a_3 and -43 are increased, while the loss of tyrosine sidechain vanishes completely. This suggests that the proposed rearrangement does occur and is required for tyrosine sidechain loss, but is not strictly necessary to produce a_3.



Scheme 3.2

To test whether any heteroatom hydrogens participate in this reaction, RG^{I}YALG was dissolved in 50/50 D_2O/CD_3OD to exchange all labile hydrogen atoms for deuterium. If formation of a_3 involves direct abstraction of an amide

hydrogen, replacement with deuterium should incur a kinetic isotope effect. Figure 3.1d shows the PD-CID spectrum of the fully deuterated iodopeptide, $[d_{14}RGYALG\cdot+H]^+$. Most ion abundances, including the a_3 ion, are similar to Figure 3.1b. The single major difference is the diminished intensity of tyrosine sidechain loss. This is most likely due to the replacement of the phenolic hydrogen with a deuterium, which results in a significant, attenuating kinetic isotope effect. This indicates that the phenolic hydrogen is required for tyrosine sidechain loss, which is consistent with the acetylation results mentioned above and mechanisms proposed in literature.[14,15] In summary, these results suggest that formation of a_3 does not involve the phenolic or amide hydrogens of tyrosine or any other exchangeable hydrogens. Therefore, we must conclude that the dominant pathway yielding backbone fragmentation proceeds through a different mechanism than originally proposed (Scheme 3.2).[18] This implies, counterintuitively, that the tyrosyl radical must migrate to another residue and then preferentially return to tyrosine to yield the a_3 fragment.

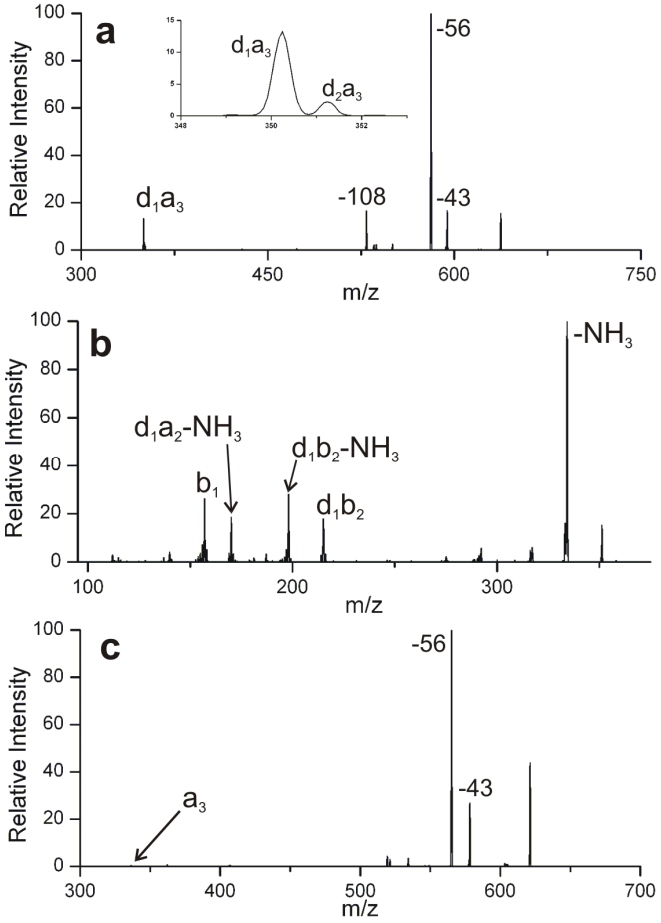


Figure 3.2 (a) CID of $[RGY_{d2}ALG\cdot+H]^+$. Examination of the a_3 fragment shows significant scrambling of the deuteriums, indicating that the β-hydrogens of tyrosine play a critical role in selective backbone dissociation. (b) Further CID of the doubly deuterated a_3 ion, $[d_2a_3+H]^+$, from (a) shows that a deuterium has migrated to Gly2 (c) Replacement of tyrosine in RGYALG with 4-hydroxyphenylglycine, which has no β hydrogens results in nearly complete elimination of selective fragmentation.

This possibility can be explored by specifically replacing C-H bonds in the peptide with C-D bonds and then monitoring migration of deuterium, which should only occur by radical initiated chemistry. It should be noted that scrambling of carbon-deuterium bonds is not expected for closed-shell peptide ions.[30,31] Substitution of hydrogen for deuterium in direct abstraction reactions

typically results in significant kinetic isotope effects. Measured kinetic isotope effects for hydrogen abstraction by ·OH range from 2.0 to 6.8 and decrease linearly with temperature.[32] Thus, monitoring deuterium transfers provides a lower limit to the extent of hydrogen transfers that would occur in an unlabeled peptide. Regardless of the magnitude, deuterium migration provides a record of where the radical has been, which can be used to elucidate radical-induced dissociation mechanisms in detail.

Abstraction of the β hydrogen from aromatic residues has been suspected to play a role in the fragmentation of radical peptides.[13,18] The peptide $RGY_{\beta d2}ALG$ was synthesized to test whether the β hydrogens of tyrosine participate in the formation of a_3 ions. Figure 3.2a shows the CID spectrum of the resulting radical ion, $[RGY_{\beta d2}ALG\cdot+H]^+$. The a_3 fragment ion intensity is split between singly and doubly deuterated a_3. The most parsimonious mechanism yielding d_1a_3 involves the migration of the initial radical to the "ALG" part of the peptide, which is followed by abstraction of one of the β deuteriums from tyrosine. The intensity of d_1a3 is much greater than d_2a_3, which unequivocally shows that radical migration occurs prior to fragmentation of this peptide. Further CID of the even-electron d_2a_3 ion (shown in Figure 3.2b) also reveals radical migration. The relative intensities of d_1b_2, d_1a_2, and d_1b_1 indicate that ~90% of d_2a_3 product contains

deuterated Gly2. To summarize, migration of one of the β-deuteriums of tyrosine nearly always precedes the formation of a_3, and can only occur following multiple radical migrations from the initial site.

The results in Figure 3.2a strongly suggest the β-hydrogens of tyrosine play a critical role in selective backbone dissociation. This hypothesis was investigated further by substituting 4-hydroxyphenylglycine (U) for tyrosine in the RGYALG peptide sequence. The sidechain of 4-hydroxyphenylglycine is a phenol attached directly to the α carbon of the peptide backbone and does not contain β hydrogens. The lack of β hydrogens to facilitate selective fragmentation should lead to the disappearance of the a_3 fragment. Figure 3.2c shows the CID spectrum for $[RGUALG\cdot+H]^+$. Backbone fragmentation almost completely disappears and sidechain loss at 4-hydroxyphenylglycine is absent. Interestingly, the a_3 fragment is observed at 0.2% relative abundance. This indicates that an alternative pathway exists to produce a_3, which does not involve the β hydrogens of tyrosine. This channel likely involves abstraction of the amide hydrogen, as proposed previously.[18] The intensity indicates that this is a very minor channel. By comparing the a_3 fragment intensities in Figure 3.2c and Figure 3.1b, we can conclude that the β hydrogen is essential for backbone fragmentation at tyrosine.

The results above illustrate how backbone fragmentation is preceded by radical migration to the β position of tyrosine. How does the initially generated radical arrive at the β carbon? The deuterium transfer demonstrated in Figure 3.2b (β carbon of Tyr3 to the α carbon of Gly2) suggests that the adjacent amino acid may facilitate radical transfer to the β carbon. To evaluate whether the α-hydrogens of neighboring amino acids are essential to backbone fragmentation, we focused on a small tetrapeptide system. CID of $[RGYG\cdot+H]^+$ produces a_3 in good yield and loss of tyrosine sidechain, as shown in Figure 3.3a, reproducing the selective cleavage at tyrosine. In order to examine the necessity of the α hydrogens of the glycine residues, experiments were next conducted with α-aminoisobutyric acid (designated with the single letter code J). This uncommon amino acid does not contain α hydrogens and prevents radical migration to the α carbon adjacent to tyrosine. Figure 3.3b shows the CID spectrum of $[RJYJ\cdot+H]^+$. The intensity of the a_3 ion in Figure 3.3b is significantly diminished relative to Figure 3.3a, which indicates that the α hydrogens of glycine in RGYG do facilitate backbone fragmentation at tyrosine. Even-electron counterparts of all other backbone fragments in Figure 3.3b (b_2, b_3+H_2O, c_1-2) are found in the CID of the non-iodinated peptide (data not shown). The remaining a_3 ion intensity must be produced from another pathway. Radicals at other locations may also be able

to abstract the β hydrogen of tyrosine. Indeed, abstraction of a hydrogen atom from the methyl groups of α-aminoisobutyric acid is feasible (the calculated BDE for the methyl group in alanine is 428 kJ/mol).[18] However, the seven-membered transition state for hydrogen atom transfer from the methyl sidechain of J to the β-position of tyrosine is not as favorable as for glycine. Transition states involving sites separated by a greater number of bonds are also possible, e.g. the initial radical site and the sidechain of arginine. Regardless, the α-positions of the neighboring amino acid are shown to play an influential role in backbone fragmentation in this peptide model.

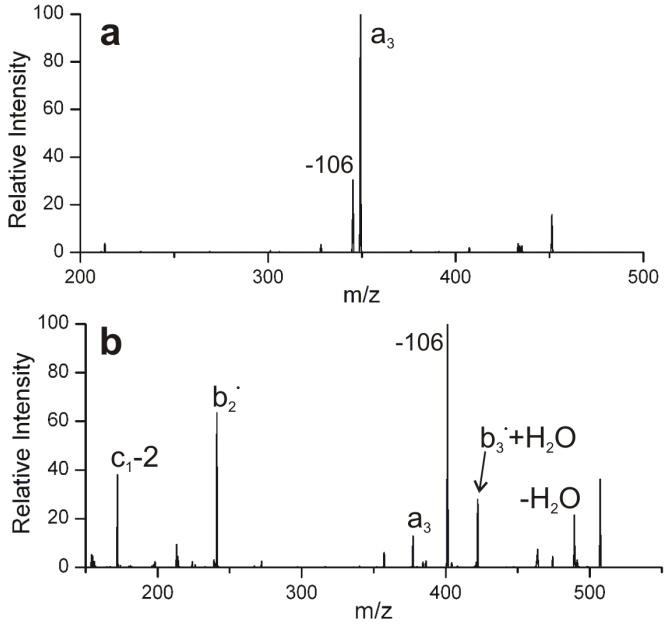
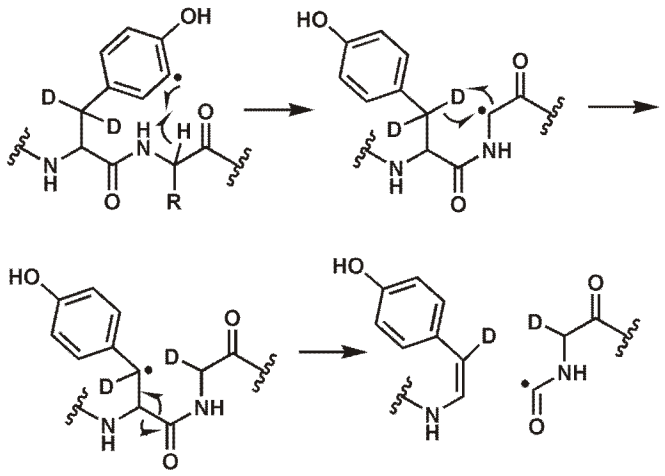


Figure 3.3 (a) CID of $[RGYG\cdot+H-I]^+$ yields backbone fragmentation at tyrosine (a_3) and loss of tyrosine sidechain. (b) The CID of $[RJYJ\cdot+H]^+$ is shown. Replacement of glycine with α-aminoisobutyric acid (J), which does not contain α-hydrogens, substantially diminishes selectivity.

We propose the following "rebound" mechanism as the primary pathway yielding backbone fragmentation at tyrosine residues where the radical originates (Scheme 3.3). The initial radical abstracts a hydrogen atom from the α carbon of the adjacent amino acid. This captodatively-stabilized radical can then abstract the β-hydrogen from tyrosine through a kinetically favorable six-membered transition state. Comparison of the calculated C-H BDEs for glycine (350 kJ/mol) and the β-position of tyrosine (367 kJ/mol) reveals that this abstraction is energetically feasible (ΔBDE = 17 kJ/mol). [19,22] This mechanism may access either the C- or N-terminal residue through a six-membered transition state. For RGYALG, it appears that scrambling to the C-terminal side is favored.



Scheme 3.3

3.3.1 Radical Migration

The CID spectrum of $[RGYG_{d2}\cdot+H]^+$ is shown in Figure 3.4a. The observation of singly deuterated a_3 ions unambiguously shows that a deuterium from the C-terminal glycine residue has migrated to the first three residues. Tandem MS can be used to pinpoint the residue where the deuterium has relocated. CID of the singly deuterated a_3 fragment (Figure 3.4b) produces a series of non- and singly deuterated b ions. Examination of d_1b_2 and b_2 reveals that the majority of the deuterium (75%) migrates to the tyrosine residue. The remaining deuterium (25%) is split between Arg1 and Gly2. Figure 3.4c shows the CID spectrum for $[RG_{d2}YG\cdot+H]^+$. All backbone fragments ($b_{2^\cdot}$, a_3) observed retain both deuteriums. In contrast to the CID of $[RGYG_{d2}\cdot+H]^+$, deuterium migration from Gly2 to Gly4 (either directly or indirectly) does not occur. The combined results suggest that radical migration not only occurs, but also does so selectively. Because deuteration is unlikely to significantly affect conformational states, radical migration must be influenced by other factors.

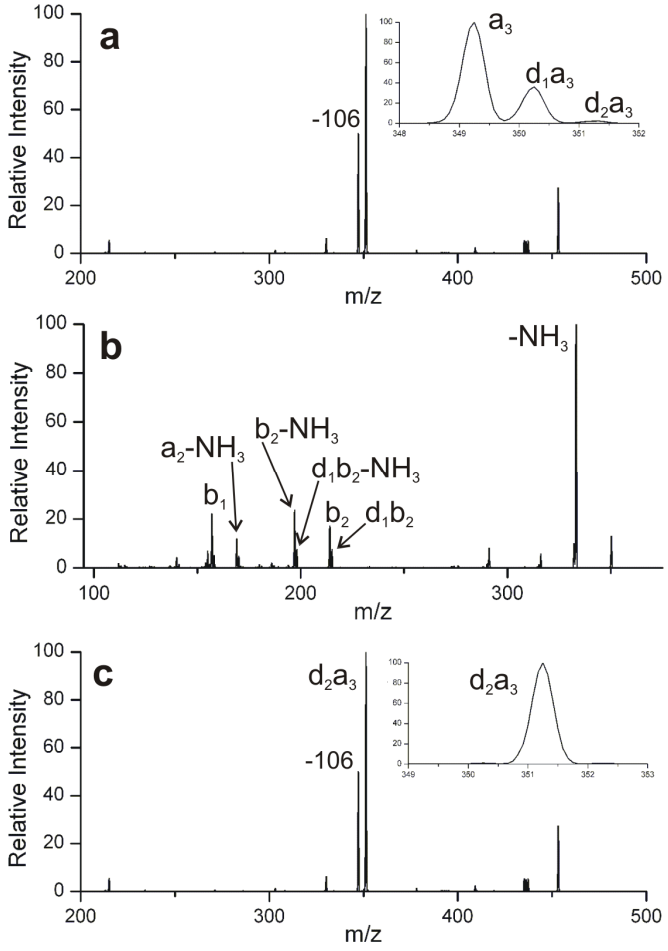
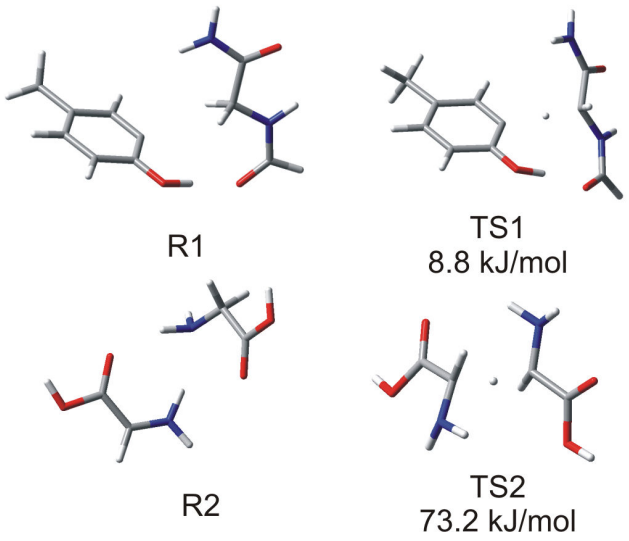


Figure 3.4 (a) CID of $[RGYG_{d2}\cdot+H]^+$ results in a deuterium transfer from the C-terminal glycine to the a_3 fragment. (b) CID of d_1a_3 from (a) indicates that 75% of the time, the scrambled deuterium resides on tyrosine (c) CID of $[RG_{d2}YG\cdot+H]^+$ yields backbone fragmentation without extensive scrambling of the deuteriums on Gly2.

There are two factors that contribute to barriers for radical migration in peptides. The first contribution results entirely from the nature of the donor and acceptor sites for the radical, or is simply a measure of the "inherent radical chemistry". The second contribution is introduced by conformational restraints which must be overcome to allow the donor and acceptor sites to interact.

Conformational restraints may also affect radical transfer by preventing optimal alignment of the donor and acceptor groups. Conformational contributions to the barriers will vary in magnitude from zero to very high energies and are not easily calculated or generally applicable (i.e. they will be numerous and different for each peptide). For example, barriers have been calculated to be as high as 94.9 kJ/mol in 1,4 H-atom transfers due to bond angle strain, or as low as 50 kJ/mol in other amino acid systems.[33-35] On the other hand, barrier contributions due to radical chemistry alone can easily be calculated by examining radical transfer between two unconstrained molecules. Barriers for abstraction by a tyrosyl radical from several relevant model molecules were calculated at the B3LYP/6-31++g(d,p) level of theory. Abstraction of the H_α of glycine by the tyrosyl radical has a calculated activation barrier of 8.8 kJ/mol. The optimized structures for the model tyrosyl radical - glycine complex (R1) and the transition state (TS1) are shown in Scheme 3.4. Similarly, abstraction of benzylic and primary hydrogens also proceeds with minimal barriers, viz. 7.1 and 15.9 kJ/mol, respectively. These results are in agreement with previous conformationally unconstrained calculations on related aromatic molecules.[36] Therefore, for a large variety of possible sites, the only substantial barriers to radical migration in peptides will originate from conformational constraints.



Scheme 3.4

In contrast, a more substantial 73.2 kJ/mol barrier is calculated for hydrogen atom migration between glycyl radical and glycine, although this value is still lower than typical proton-driven dissociation barriers (120-140 kJ/mol).[37] The optimized structures for glycyl radical - glycine complex (R2) and the transition state (TS2) are shown in Scheme 3.4. The C_α-H bond of glycyl radical must "bend" out-of-plane with respect to the adjacent amine and carboxylic acid groups in order to accommodate the incoming hydrogen atom from glycine. The energetic penalty associated with the loss of captodative stabilization in the transition state complex gives rise to the observed barrier, which is an upper limit for glycyl radicals in peptides where stabilization may be considerably less.

For example, glycyl radicals that are restricted from planarity due to secondary structure will likely have lower barriers to abstracting a hydrogen atom.

Recently, it has been reported that the B3LYP level of theory underestimates barrier heights for hydrogen atom abstraction and that MP1WK is a suitable alternative.[36-39] For the reactions shown, barrier heights calculated using the MP1WK/6-31+g(d,p)//B3LYP/6-31++(d,p) level of theory show a modest and systematic increase of ~8 kJ/mol compared to B3LYP. These barrier heights remain well below typical dissociation barriers and do not affect the main conclusions in this manuscript.

Theory therefore suggests modest to low barriers will exist for radical migration in peptides, with the dominant portion of the barrier frequently originating from conformational effects. The experimental results presented above indicate clearly that radical migration occurs prior to dissociation. Furthermore, radical directed dissociation dominates over proton driven fragmentation in most cases. Taken together, these observations suggest that understanding radical migration is particularly important for predicting the ultimate fragmentation chemistry. The preference for specific backbone and side chain fragmentations will be ultimately determined by the relevant kinetics controlling these dissociation pathways; however, because radical migration

occurs below the fragmentation threshold, dissociation will be enhanced at sites to which the radicals preferentially migrate (which correspond to thermodynamically favorable sites). Therefore, backbone fragmentation at RGYALG, for example, occurs almost exclusively at tyrosine (the thermodynamically preferred destination) despite the presence of several other side chains with β carbons, which might offer similar kinetic constraints. Once the radical migrates to tyrosine, backbone dissociation is more favorable than subsequent radical migration to leucine, alanine or arginine (which is both thermodynamically uphill and protected by kinetic barriers to fragmentation). We refer to the process of radical migration to thermodynamically favorable sites as radical funneling.

An interesting example is shown in Figure 3.5, where three related peptides with distinct radical initiation points yield essentially identical fragmentation. CID of $[RPPGYSPFR\cdot+H]^+$ yields a-type fragmentation at Tyr5 and at Phe8. The a_8 fragment is consistent with abstraction of the β hydrogen of phenylalanine, followed by β scission. Unlike fragmentation at tyrosine, dissociation at phenylalanine cannot occur by the rebound mechanism (although radical transfer to the adjacent α carbon is possible). Furthermore, secondary backbone fragmentation at phenylalanine could occur by direct abstraction of the β

hydrogen and does not strictly require an α-carbon radical intermediate. C-type fragments N-terminal to serine are also observed in Figure 3.5. This type of fragmentation has been documented previously and is produced from abstraction of the serine β hydrogen.[13,19]

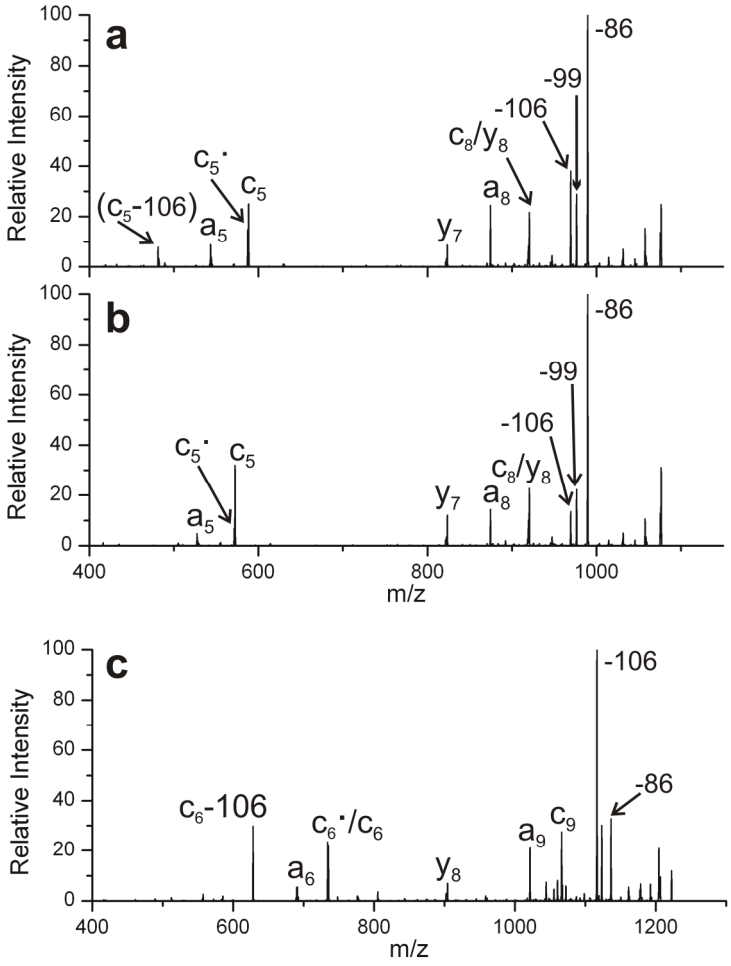


Figure 3.5 CID spectra of bradykinin analogs, (a) $[RPPGYSPFR\cdot+H]^+$, (b) $[RPPGFSPYR\cdot+H]^+$ and (c) $[YRPPGFSPFR\cdot+H]^+$ reveal similar backbone fragments despite different radical starting locations. The similarity suggests that the radical on each peptide scrambles to a similar distribution of radical isomers prior to dissociation.

CID of an analogous peptide where the positions of tyrosine and phenylalanine have been interchanged yields strikingly similar fragmentation. CID of $[RPPGFSPYR\cdot+H]^+$ is shown in Figure 3.5b. The only significant difference is the absence of the c_5-106 fragment. This is expected given that the c_5 in this peptide does not contain tyrosine and phenylalanine sidechain loss is not typically observed. All other backbone fragments are present in similar relative abundance, despite the difference in initial radical location. Figure 3.5c shows the CID spectrum of the radical peptide, YRPPGFSPFR, where the initial radical begins at the N-terminus. Surprisingly, backbone fragmentation still occurs at in similar relative abundances despite the quite different initial radical location.

The identical backbone fragmentation observed for all three bradykinin derivatives suggests that for a larger peptide where many equivalent endpoints for radical migration may exist, the radical can migrate between these sites upon activation. Isomerization may occur through direct hydrogen atom abstraction between sidechains, or by radical migration along the peptide backbone. Regardless, the correspondence between the three spectra is consistent with rapid isomerization to similar radical intermediates, from which fragmentation occurs and yields a similar distribution of products. It is worth noting that CID of two of these radicals, $[RPPGYSPFR\cdot+H]^+$ and $[RPPGFSPYR\cdot+H]^+$, have been

shown previously and yield qualitatively the same products.[40] Dissimilarities in ion abundances between our experiment and previous work likely arise from differences in preparation of the radical and ion trapping pressures.

Given inherent peptide structure, only radicals situated at the amide nitrogens, or at α or β carbons, or the δ carbon of proline may undergo β scission to cleave the backbone directly. The hydrogen-deuterium exchange experiments above indicate that the amide nitrogens are not favorable sites for radical migration. β scission from an α-carbon radical would necessitate cleavage of the resonance-stabilized amide bond and is therefore unfavorable. The amide BDE in the model compound, $CH_3C(O)—NH_2$, is 414 kJ/mol.[21] In contrast, the C_α-C(O) bond, which is cleaved by β scission from a β-centered radical, is significantly weaker (BDE of $CH_3C(O)—CH_3$ is 354 kJ/mol). Additionally, a theoretical study of a closely related alkoxyl radical peptide predicts that β scission of the C_α-C(O) bond is essentially barrierless.[41] These results are consistent with preferential backbone fragmentation at residues with low C_β-H BDEs and the absence of backbone fragments from α-centered radicals. Thus, radical-induced fragmentation occurs only when a radical can access a thermodynamically favored site that is β to a weak bond, leading to a small number of preferred

fragmentation channels. For large peptides, migration to sites with low C_β-H BDEs appears to be facile.

3.4 Application of Radical Funnel Model to Peptides

Determining the scope of the radical funnel model requires examination of numerous peptides with diverse sequences in addition to the model peptides discussed above. One of the limitations of preparing peptide radicals by PD of carbon-iodine precursors is that the peptide must contain residues that are easily iodinated, such as tyrosine and histidine. An alternative approach was developed using the noncovalent radical precursor, 5-iodo-2-(2-hydroxymethyl-18-crown-6)-naphthoate (1), which forms adducts with peptide cations upon ESI.[19] Upon photodissociation of the peptide·**1** complexes, peptide radicals can be isolated and interrogated using collisional activation. The combined dataset from tyrosine-containing peptides and peptides examined with the noncovalent approach enables verification of the radical funnel model with a more diverse dataset.

The radical funnel model predicts that radical-directed backbone fragmentation occurs more frequently at residues that have low relative C_β-H BDEs. Figure 3.6 shows the C_β-H BDEs for all canonical amino acids (except glycine) calculated using *ab intitio* theory.[42] Interestingly, the C_β-H BDEs for

residues containing aromatic or alcohol sidechains are the lowest C_β-H BDEs, and are similar to reported C_α-H BDEs (~330 kJ/mol),[22] which are predicted to be the most stable sites for radicals in peptides. Superimposed on the same figure are the normalized intensities of backbone fragmentation by amino acid. Data from 17 peptides were used to construct the plot. Six of the peptides were derivatized by noncovalent chemistry; the remaining eleven were derivatized by iodination. The shading indicates the sampling frequency of the amino acid in the dataset. Although data points for asparagine, aspartic acid, cysteine, threonine, and glutamine may be significantly skewed due to insufficient sampling, it is clear that even without these data points, there is a general inverse relationship between backbone fragmentation and C_β-H BDE. The observed relationship is entirely consistent with the proposed radical funnel model.

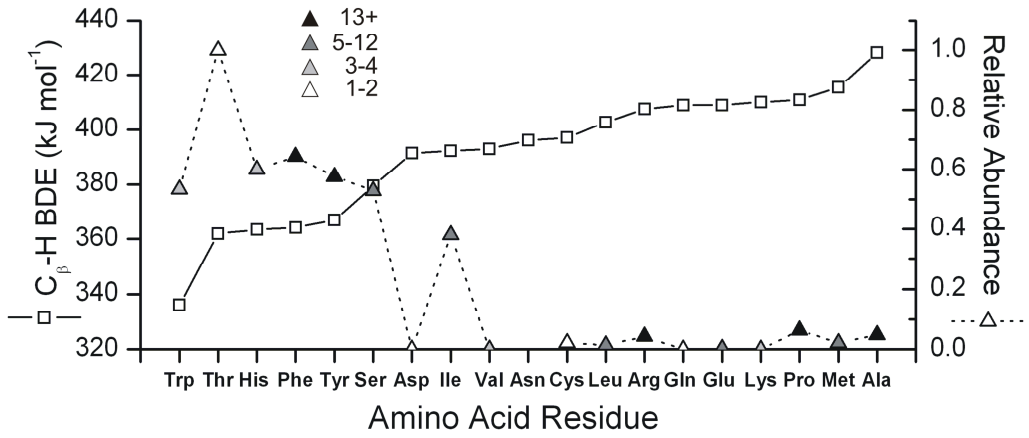


Figure 3.6 A comparison of C_β-H BDEs and observed radical-directed backbone fragmentation. The shading indicates the number of times the amino acid appears in the dataset. The interquartile ranges are shown in the legend.

3.5 Conclusions

The mechanisms underlying backbone dissociation in radical peptides were studied using chemically modified derivatives and deuterium-labeling on model peptide scaffolds. We have shown that a radical positioned on the C_β of tyrosine is essential for producing a-type fragmentation at tyrosine. Movement of the radical from the initial site to the C_β is facilitated by a rebound mechanism between tyrosine and the adjacent amino acid. Failure of the radical to rebound and return to tyrosine yields sidechain losses and backbone fragmentation at other residues with low C_β-H BDEs. The results herein also indicate that the

peptides studied are either structurally flexible or exist in highly heterogeneous ensembles, enabling the radical to sample multiple sites over multiple residues. We propose that the backbone fragmentation observed can be explained by a "radical funnel" model, where radical migration preferentially yields backbone dissociation at sites with low C_β-H BDEs. Numerous peptides were examined to evaluate the validity of the model. The observed inverse relationship between the calculated C_β-H BDEs and experimental backbone fragmentation is entirely consistent with the radical funnel model.

---

[1] Stubbe, J.; van der Donk, W. A. *Chem. Rev.* **1998**, *98*, 705-762.
[2] Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
[3] Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9528-9533.
[4] Syrstad, E. A.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 208-224.
[5] Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857-2862.
[6] Zubarev, R. A.; Haselmann, K. F.; Budnik, B.; Kjeldsen, F.; Jensen, F. *Eur. J. Mass Spectrom.* **2002**, *8*, 337-349.
[7] Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8949-8958.
[8] Jones, J. W.; Sasaki, T.; Goodlett, D. R.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 432-444.
[9] Belyayev, M. A.; Cournoyer, J. J.; Lin, C.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1428-1436.
[10] O'Connor, P. B.; Lin, C.; Cournoyer, J. J.; Pittman, J. L.; Belyayev, M.; Budnik, B. A. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 576-585.
[11] Wee, S.; Mortimer, A.; Moran, D.; Wright, A.; Barlow, C. K.; O'Hair, R. A. J.; Radom, L.; Easton, C. J. *Chem. Commun.* **2006**, *40*, 4233-4235.
[12] Masterson, D. S.; Yin, H. Y.; Chacon, A.; Hachey, D. L.; Norris, J. L.; Porter, N. A. *J. Am. Chem. Soc.* **2004**, *126*, 720-721.
[13] Hodyss, R.; Cox, H. A.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2005**, *127*, 12436-12437.
[14] Chu, I. K.; Rodriquez, C. F.; Lau, T. C.; Hopkinson, A. C.; Siu, K. W. M. *J. Phys. Chem. B.* **2000**, *104*, 3393-3397.

[15] Wee, S.; O'Hair, R. A. J.; McFadyen, W. D. *Int. J. Mass Spectrom.* **2004**, *234*, 101-122.
[16] Laskin, J.; Yang, Z.; Chu, I. K. *J. Am. Chem. Soc.* **2008**, *130*, 3218-3230.
[17] Thompson, M. S.; Cui, W. D.; Reilly, J. P. *Angew. Chem. Int. Edit.* **2004**, *43*, 4791-4794.
[18] Ly, T.; Julian, R. R. *J. Am. Chem. Soc.* **2008**, *130*, 351-358.
[19] Sun, Q.; Nelson, H.; Ly, T.; Stoltz, B. M.; Julian, R. R. *J. Proteom. Res.* **2008**, *8*, 958-966.
[20] Diedrich, J. K.; Julian, R. R. *J. Am. Chem. Soc.* **2008**, *130*, 12212-12213.
[21] Blanksby, S. J.; Ellison, G. B. *Acc. Chem. Res.* **2003**, *36*, 255-263.
[22] Rauk, A.; Yu, D.; Taylor, J.; Shustov, G. V.; Block, D. A.; Armstrong, D. A. *Biochemistry* **1999**, *38*, 9089-9096.
[23] Kavita, K.; Das, P. K. *J. Chem. Phys.* **2002**, *117*, 2038-2044.
[24] Paquet, A. *Can. J. Chem.* **1982**, *60*, 976-980.
[25] Carpino, L. A.; Han, G. Y. *J. Org. Chem.* **1972**, *37*, 3404-3408.
[26] Chan, W. C.; White, P. D., Eds. Fmoc Solid Phase Peptide Synthesis: A Practical Approach. Oxford University Press: New York, NY, 2004.
[27] Gronert, S. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 845-848.
[28] Peng, C. Y.; Schlegel, H. B. *Israel J. Chem.* **1993**, *33*, 449-454.
[29] Nielsen, M. L.; Budnik, B. A.; Haselmann, K. F.; Olsen, J. V.; Zubarev, R. A. *Chem. Phys. Lett.* **2000,** 330, 558-562.
[30] Kenny, P. T. M.; Nomoto, K.; Orlando, R. *Rapid Commun. Mass Spectrom.* **1992,** 6, 95-97.
[31] Harrison, A. G.; Yalcin, T. *Int. J. Mass Spectrom. Ion Processes* **1997**, *165*, 339-347.
[32] Dunlop, J. R.; Tully, F. P. *J. Phys. Chem.* **1993**, *97*, 11148-11150.
[33] Moran, D.; Jacob, R.; Wood, G. P. F.; Coote, M. L.; Davies, M. J.; O'Hair, R. A. J.; Easton, C. J.; Radom, L. *Helv. Chim. Acta* **2006**, *89*, 2254-2272.
[34] Fung, W. M. E.; Chan, T. W. D. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1523-1535.
[35] Panja, S.; Nielsen, S. B.; Hvelplund, P.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1726-1742.
[36] Jing, L.; Nash, J. J.; Kenttamaa, H. I. *J. Am. Chem. Soc.* **2008**, *130*, 17697-17709.
[37] Paizs, B.; Suhai, S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 103-113.
[38] Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 4811-4815.
[39] Lingwood, M.; Hammond, J. R.; Hrovat, D. A.; Mayer, J. M. Borden, W. T. J. *Chem. Theory Comput.* **2006**, *2*, 740-745.
[40] Laskin, J.; Yang, Z.; Lam, C.; Chu, I. K. *Anal. Chem.* **2007**, *79*, 6607-6614.
[41] Wood, G. P. F.; Easton, C. J.; Rauk, A.; Davies, M. J.; Radom, L. *J. Phys. Chem. A* **2006**, *110*, 10316-10323.
[42] Hehre, W. J.; Ditchfield, R.; Radom, L.; Pople, J. A. *J. Am. Chem. Soc.* **1970**, *92*, 4796−4801.

Chapter 4

ELECTRON INDUCED DISSOCIATION OF PROTONATED PEPTIDES YIELDS BACKBONE FRAGMENTATION CONSISTENT WITH A HYDROGEN DEFICIENT RADICAL

4.1 Introduction

Successful application of electron capture and transfer dissociation to peptide sequencing has inspired considerable interest in their fundamental fragmentation chemistry.[1] Typically, capture of an electron by a peptide yields radical species that fragment the backbone generating c- and z- fragment ions.[2] We have recently discovered that peptide and protein radicals can also be generated by photodissociation (PD).[3-5] Homolytic fissure of a non-native carbon-iodine bond yields a carbon-centered radical. [6] Unlike the radicals described above, the species generated by PD are "hydrogen-deficient" and yield very different fragmentation. In proteins, backbone dissociation appears to occur selectively at aromatic residues.[3] In peptides, sidechain losses are prevalent and backbone selectivity is observed at residues with low C_β-H bond dissociation energies.[4,7] Similar fragmentation behavior has been observed for photoionized peptide radicals. [8,9]

One of the drawbacks of using carbon-iodine photochemistry is that the peptide must be chemically modified prior to introduction into the gas phase. Additionally, the iodination reaction is only facile for tyrosyl-containing peptides. In an attempt to eliminate the covalent modification step, we turned to electron-induced dissociation (EID) methods.[10-12] It is known that bombardment of protonated peptides with energetic electrons will detach an electron and generate a cationic radical species.[13,14] Competing processes such as vibrational excitation can occur, which would yield fragmentation that is similar to collisional activation.[15] The purpose of the current study is to examine whether EID may be used to generate the residue-specific fragmentation that is observed in the PD of iodopeptides.

4.2 Materials and Methods

To compare the fragmentation generated by PD and EID, tyrosyl-containing peptides were iodinated using well-established oxidation chemistry in the presence of sodium iodide.[16] Briefly, stoichiometric equivalents of peptide, sodium iodide, and chloramine-T were mixed. After a reaction time of ten minutes, sodium metabisulfite was added to quench the reaction. The iodination products were analyzed by electrospray mass spectrometry without purification. Samples containing ~ 5 μM iodopeptide in 50/50 water/acetonitrile were directly

infused into the standard electrospray source of a Thermo Fisher LTQ linear ion trap mass spectrometer. The back plate of the vacuum housing was modified with a quartz window to transmit a single pulse (3-5 nsec, 6 mJ) of 266 nm photons from a fourth-harmonic, flashlamp-pumped Nd:YAG laser. Laser shots were fired after isolation of the precursor ion. The radical product was subsequently isolated, collisionally activated and mass analyzed. EID experiments were performed with a Thermo Electron LTQ-FT mass spectrometer. Solutions containing ~ 5 μM of unmodified peptide were electrosprayed and trapped in an ICR cell, where ions are irradiated with ~ 23 eV electrons for 75 ms after a 5 ms delay. Standard peptide nomenclature is used to label fragments. Loss of ammonia and water from sequence ions are denoted with the suffixes * and ^{o}, respectively.

4.3 Results and Discussion

PD of singly protonated iodo-RGYALG yields loss of iodine atom. The CID of the radical product, $[RGYALG\bullet+H]^{+}$, has been published and is shown here in Figure 4.1a to facilitate comparison with EID data.[7] In addition to abundant sidechain losses from tyrosine and leucine, a single sequence ion (a_3) is observed corresponding to cleavage at the tyrosine residue. The a_3 ion is not present in the CID of the even electron counterpart, $[RGYALG+H]^{+}$. Indeed, evaluation of a vast

number of peptides has revealed that backbone fragmentation at tyrosine and other residues with low C_β-H bond dissociation energies is a hallmark of hydrogen deficient radical fragmentation. [3,4]

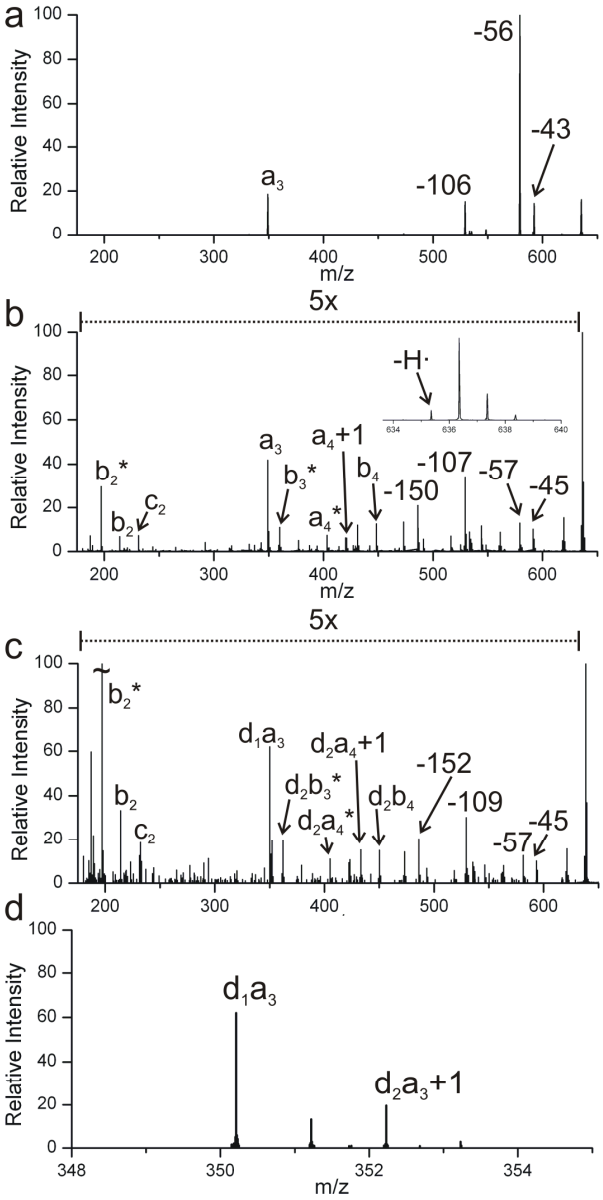


Figure 4.1 a) CID of $[RGYALG\bullet+H]^+$, which was generated by PD of the iodo-precursor, yields selective backbone dissociation at tyrosine to form a_3. b) Similarly, EID of $[RGYALG+H]^+$ yields a_3 as the most abundant backbone fragment. Loss of a hydrogen

atom from the precursor may be an intermediate (inset). c) EID of an analogous peptide where the C_β of tyrosine is deuterated, $[RGYd_2ALG+H]^+$, yields scrambling of a deuterium to generate d_1a_3. d) A closer view of the d_1a_3 fragment shows that d_2a_3+1 is formed by a competitive process that does not yield scrambling.

EID of $[RGYALG+H]^+$, shown in Figure 4.1b, yields a more complex dissociation spectrum than observed in Figure 4.1a. The low product ion intensities suggest that the most frequent ion-electron interaction leads to non-dissociative scattering of the electrons. Of the product ions observed, sidechain losses and the a_3 and b_2-NH_3 (b^*) sequence ions dominate. A series of b and radical a+1 ions are present at much lower intensities.

Only products that are derived from odd-electron species are of interest in this manuscript. Therefore, the b series of ions (including b_2^*), which are also present in low energy CID of the even-electron precursor, will not be discussed in detail here. Instead, we focus our attention on the radical-induced backbone fragmentation, namely the a-type fragment ions. The most intense backbone fragment in Figures 4.1a and b is a_3, which corresponds to cleavage of the C_α-C(=O) bond of tyrosine. This similarity led us to speculate that a hydrogen deficient radical is an important intermediate in EID. Indeed, a zoomed-in view of the molecular ion (Figure 4.1b, inset) shows that hydrogen atom detachment does occur, which may yield a similar radical as the precursor in Figure 4.1a. Determining the location of hydrogen atom detachment is not straightforward.

Hydrogen detachment may arise from many potential sites on the peptide. However, experiments with deuterated RGYALG, where the C_β of tyrosine is doubly deuterated, may provide some information on the location of the radical prior to fragmentation.

EID of [$RGY_{d2}ALG$+H]+ is shown in Figure 4.1c. The d_1a_3 fragment is observed, indicating that a deuterium has scrambled to the C-terminal portion of the peptide (ALG) prior to cleavage of the C_α-C(=O) bond of tyrosine. The peak at +1 m/z unit is consistent with the expected ^{13}C intensity. A second peak observed at +2 m/z units corresponds to d_2a_3+1. The differences in deuterium scrambling are best explained by two competitive mechanisms where the radical is situated at different locations on the peptide. Clearly, a radical at the β position of tyrosine is involved in generating the d_1a_3 ion. The absence of significant d_2a_3 indicates that the radical that abstracts the β hydrogen is typically on the C-terminal side. In contrast, the formation of a+1 ions has been suggested to originate from homolytic bond dissociation or β scission from an amide nitrogen-centered radical. Neither mechanism is expected to scramble the Cβ deuterium atoms; indeed, both deuteriums are retained in the a_3+1 ion.

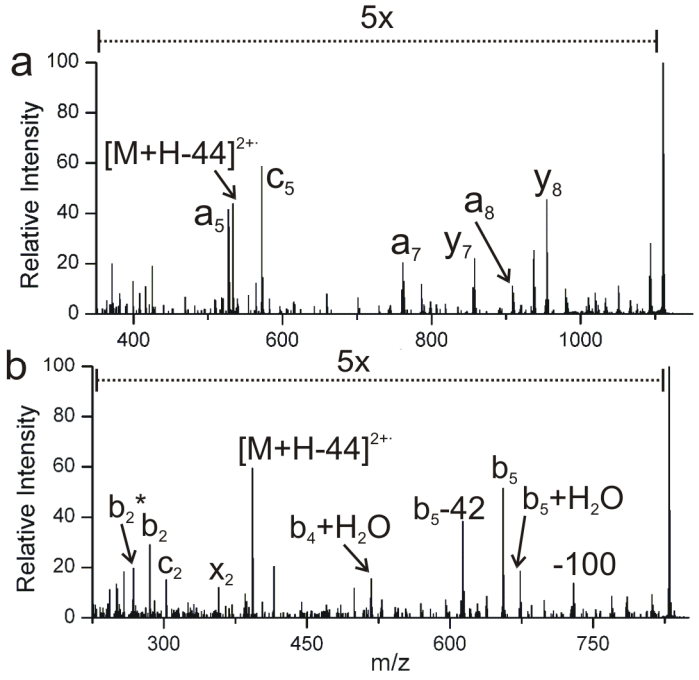


Figure 4.2 a) EID of [RPPFSFFR+H]+ yields a- and c-type ions at several residues with low C_β-H bond dissociation energies. Additionally, the $[M+H-44]^{2+\bullet}$ ion demonstrates that ionization occurs. b) EID of [KRTLRR+H]+ yields ionization of the peptide and significant radical induced fragmentation, despite having no aromatic residues.

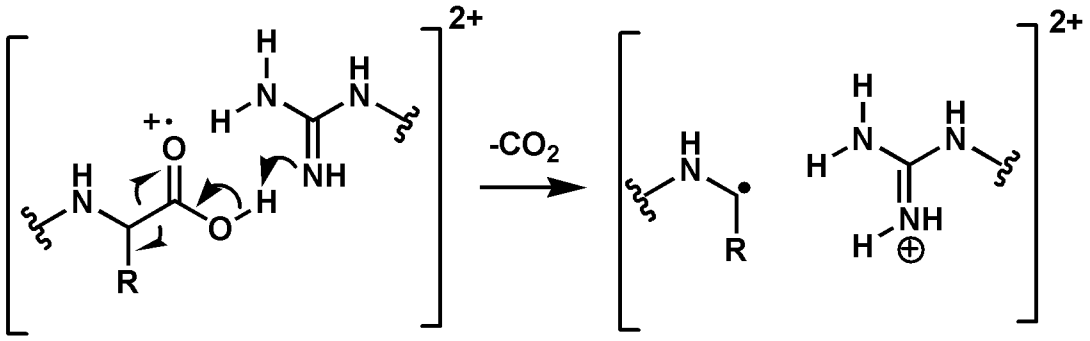
EID of $[RPPFSFFR+H]^+$, shown in Figure 4.2a, indicates that radical rearrangement to the β carbon appears to be a general feature of this dissociation method. Radical induced fragmentation occurs at residues with low C_β-H bond dissociation energies (Y, F, S) to yield a- and c- type fragments. The same fragments are present in the PD-CID of this peptide. Additionally, a doubly charged fragment ion is observed, corresponding to loss of CO_2 from the precursor ion. The observation of electron ionized fragments may be due to the presence of a second arginine residue that can stabilize the charge. This would be

consistent with the lack of stable doubly cationized fragments in RGYALG. Regardless, the RPPFSFFR results indicate that ionization of the peptide is another route for generating hydrogen deficient radicals and associated fragments.

It has been suggested previously that EID of peptides yields ionization of amino acids with the lowest ionization energies, i.e., aromatic residues. To test whether an aromatic residue is required, EID of KRTLRR was performed. KRTLRR contains a residue with a low C_β-H bond dissociation energy (T) and no aromatic residues. The EID spectrum of $[KRTLRR+H]^+$ is shown in Figure 4.2b. Radical induced fragments include c_2 and x_2, which are due to dissociation at threonine and leucine, respectively. The low ion abundances for these ions indicate that an aromatic residue may facilitate radical induced fragmentation, but is not required. Interestingly, the $[M+H-44]^{2+\bullet}$ ion is still observed and indicates that ionization does not strictly require an aromatic residue.

The data from Figures 4.2a and b suggest that ionization may occur from the C-terminus to yield a radical cation, which is expected to have different chemistry than the distonic radicals produced by PD-CID. However, the similarities in backbone fragmentation indicate that the radical cation generated by EID can isomerize to a distonic radical. Scheme 4.1 shows a postulated

mechanism for generating a distonic radical. Electron ionization of the C-terminus is followed by proton transfer to a neutral arginine residue and loss of CO_2. The remaining C_α radical is expected to have similar chemistry to the radicals generated by PD.



Scheme 4.1

EID appears to be an alternative method to generate residue-specific fragmentation for singly charged peptides. Compared to PD of iodinated peptides, EID suffers from low product yield and competitive statistical backbone fragmentation. However, if PD is not available to generate a hydrogen deficient radical, other methods of radical initiation, such as EID, must be employed.[17]

---

[1] Cooper, H.J., Hakansson, K., Marshall, A.G. *Mass Spectrom. Rev.* **2005**, *24*, 201.
[2] Zubarev, R. A. *Mass Spectrom. Rev.* **2003**, *22*, 57.
[3] Ly, T., Julian, R. R. *J. Am. Chem. Soc.* **2008**, *130*, 351. [Chapter 2]
[4] Sun, Q.; Nelson, H.; Ly, T.; Stoltz, B. M.; Julian, R. R. *J. Proteome Res.* **2009**, *8*, 958.
[5] Liu, Z.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 965-971.
[6] Dzvonik, M.; Yang, S.; Bersohn, R. *J. Chem. Phys.* **1974**, *61*, 4408.

[7] Ly, T.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2099-2101. [Chapter 3]
[8] Kalcic, C. L.; Gunaratne, T. C.; Jonest, A. D.; Dantus, M.; Reid, G. E. *J. Am. Chem. Soc.* **2009**, *131*, 940.
[9] Antoine, R.; Joly, L.; Tabarin, T.; Broyer, M.; Dugourd, P.; Lemoine, J. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 265.
[10] Lioe, H.; O'Hair, R. A. J. *Anal. Bioanal. Chem.* **2007**, *389*, 1429.
[11] Budnik, B. A.; Zubarev, R. A. *Chem. Phys. Lett.* **2000**, *330*, 558.
[12] Nielsen, M. L.; Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Int. J. Mass Spectrom.* **2003**, *226*, 181.
[13] Budnik, B. A.; Zubarev, R. A. *Chem. Phys. Lett.* **2000**, *316*, 19.
[14] Budnik, B. A.; Tsybin, Y. O.; Hakansson, P.; Zubarev, R. A. *J. Mass Spectrom.* **2002**, *37*, 1141.
[15] Fedor, D. M.; Cody, R. B.; Burinsky, D. J.; Freiser, B. S.; Cooks, R. G. *Int. J. Mass Spectrom. Ion Phys.*, **1981**, *39*, 55.
[16] Regoeczi, E. Iodine-Labeled Plasma Proteins; CRC Press: Boca Raton, FL, 1984.
[17] Barlow, C. K.; O'Hair, R. A. J. *J. Mass Spectrom.* **2008**, *43*, 1301.

Chapter 5

USING ESI-MS TO PROBE PROTEIN STRUCTURE BY SITE-SPECIFIC NONCOVALENT ATTACHMENT OF 18-CROWN 6 ETHER

5.1 Introduction

Several techniques that rely heavily on electrospray ionization mass spectrometry (ESI-MS) are used to explore the structure and folding states of proteins. Hydrogen/deuterium (H/D) exchange is one method that can be used as an elegant probe of the backbone amide groups.[1,2] In these experiments, H/D exchange of some amide hydrogens (typically on the interior of the protein) is hindered by intramolecular hydrogen bonding or other factors that are related to macromolecular structure.[3] The degree of exchange for a given set of conditions can be observed directly in a mass spectrum. Under appropriate conditions, specific sub-regions of a protein can be examined individually. This is accomplished by proteolytically digesting the protein and determining the extent of H/D exchange occurring in each individual peptide.[4] Binding sites can be ascertained by recording H/D exchange in the presence and absence of a ligand.[5] In fact, this approach has led recently to the development of MS based techniques for the determination of solution phase binding constants.[6,7] In

addition, the kinetics of protein folding can also be observed by pulsed exchange experiments.[8]

Changes in charge state distributions can also be used to monitor protein structure on a coarse level.[9,10] Numerous experiments have shown that protein charge states tend to increase for unfolded structures when compared with native or folded structures.[11,12] Under appropriate conditions, both structures for proteins that are two state folders can be observed simultaneously in a bimodal charge state envelope.[13] Nevertheless, charge state distributions yield only limited information and can be misleading. For example, the charge state envelope for cytochrome c (cyt c) reveals an unfolded state only when the protein is sampled as a cation.[14] Furthermore, it is not possible to distinguish contributions from different structures that give rise to similar charge state distributions. Another drawback is that protein structure cannot be examined on an individual basis for each observed charge state.

Despite these limitations, examining protein structure as a function of charge state remains a desirable methodology because of the elegant simplicity of the experiment itself, i.e., simply electrospraying proteins under various solvent conditions. Although H/D exchange experiments provide detailed information, they are additionally much more complicated to execute and frequently do not

directly sample equilibrium states. It is therefore desirable to devise an experiment for examining protein structure, which retains the simplicity of the charge state distribution approach, yet provides additional information.

The present work demonstrates a new approach for examining protein structure and folding based on side-chain availability. In these experiments, 18-crown-6 ether (18C6) is used to probe the availability of lysine side chains through a specific, three hydrogen bond mediated interaction.[15-17] The new approach is similar to H/D exchange because specific interactions are probed, yet also complimentary because the focus is on side chains and exposed regions rather than the peptide backbone and the protein interior. Simultaneously, the experiment retains the simplicity of a charge state distribution type experiment. Experiments with cyt c, ubiquitin, and melittin illustrate that the number of 18C6s that attach to a protein is correlated with the macromolecular structure. In these experiments, the structure for each charge state can be probed independently through observation of the unique pattern of multiple 18C6s that attach to the protein. The relative intensity of each peak in the distribution is determined by lysine side-chain availability, which is a function of the structure of the protein in solution. This approach reveals that changes in the 18C6 distribution can be observed under varying solvent conditions even though no

significant change is observed in the overall charge state distribution. The roles that protein secondary and tertiary structure play in lysine availability are explored. Control experiments with pentalysine, which lacks secondary or tertiary structure, are also presented.

5.2 Materials and Methods

Mass spectra were obtained using a Finnigan LTQ linear ion trap mass spectrometer (Thermo Electron, San Jose, CA) equipped with a standard electrospray ionization source with no modification. Voltages in the source region were optimized for observation of noncovalent complexes. This was achieved by electrospraying a solution of cyt c and 18C6 and optimizing adduct ion intensities by tuning individual voltages. The critical parameters for observing protein-adduct formation are the capillary voltage 9.0 V, capillary temperature 215 - 275 °C, and tube lens offset 145–160 V. The tube lens voltage was optimized for maximum total ion count within this range. Peptide mass spectra were taken under instrument settings reoptimized for pentalysine-18C6 adduct ion intensities. Critical parameters yielding peptide-adduct formation are the capillary voltage 49.0 V, capillary temperature 215 °C, and tube lens offset 60 V for maximal ion intensity. The spray voltage was varied from 3.6 to 4.5 kV to achieve best shot-to-shot stability of mass spectra.

Protein and peptide solutions were made in solvents ranging from pure water to mixtures of water/methanol as specified. No acetic acid was added to solutions unless indicated. Analytical reagent-grade methanol (Riedel de Häen, Seelze, Germany) was specifically purchased to limit 18C6 adducts with ammonium, sodium, and potassium cations. The concentrations for all proteins were kept in the 3 μM to 5 μM range, and 15 μM for pentalysine. For all solutions where 18C6 (Alfa Aesar, Pelham, NH) was added, the 18C6 concentration was matched to two times the number of lysine residues to ensure stoichiometric excess. All samples were prepared using NANOpure DIamond (Macalaster Bicknell Co., New Haven, CT) purified water. All chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO) or EMD (Gibbston, NJ) unless noted otherwise and used without further purification. The biogenic source for ubiquitin was bovine erythrocytes and cyt c was obtained from horse heart. Melittin was obtained synthetically from American Peptide Company (Sunnyvale, CA). Solution pHs were measured using a VWR symphony meter. The reported pHs for water/methanol mixtures were measured directly and are uncorrected for methanol content.[18]

5.3 Results and Discussion

5.3.1 Control Experiments

The experiments described in this work measure the extent to which the binding of a small molecular probe is sensitive to changes in macromolecular structure. Manipulation of the pH and addition of organic cosolvents are commonly utilized techniques for manipulating protein structure.[19] However, it is also possible that changing these parameters might influence the noncovalent attachment of small molecules in ESI-MS experiments. For example, the addition of organic solvents may alter the solution phase binding constants or the addition of acid may lead to competitive binding with lysine side chains.

To examine the influence of pH and organic content on the complexation of 18C6 to lysine in the absence of structural changes, we conducted several control experiments with pentalysine. Previous work relying on circular dichroism (CD) has demonstrated that pentalysine is typically disordered.[20] In Figure 5.1, the results for experiments designed to test the influence of organic solvents are shown. The data in Figure 5.1a reveals that the ratio of naked to complexed peptide does not increase with increasing methanol content. Because the 18C6-lysine interaction is hydrogen bond and ion-dipole mediated, it is anticipated that the binding constant will increase as the content of methanol is increased

due to reduction of the dielectric constant. If this effect were directly correlated with the observations in ESI-MS, the amount of naked peptide would decrease with increasing methanol. The data in Figure 5.1a illustrates that this is not the case. Therefore, the introduction of organic solvents does not influence the number of 18C6s that attach to the peptide in a manner consistent with the expected change in binding constant. Although there are some changes in relative intensities in the spectra, all of the major peaks remain present as the methanol percentage is varied from 0 to 95%. The two extremes are shown in Figure 5.1b and c. The decline of $[K_5 + 418C6]^{4+}$ appears to be connected with the introduction of sodium (present as a contaminant in the methanol). Smaller changes may be related to structural rearrangement of the peptide or to differences in the electrospray process.

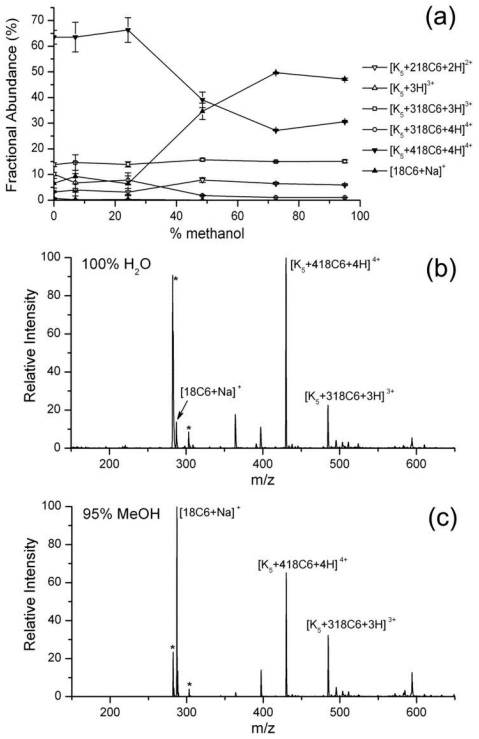


Figure 5.1 (a) The fractional abundance of several species sampled from a solution of pentalysine and 18C6 is plotted as a function of the percentage methanol in solution (with the remainder being water). As the organic content increases, the number of 18C6 adducts does not change significantly, suggesting that small changes in binding constant do not strongly influence the observed attachment of 18C6. The spectra representing the two extreme cases are shown in (b) and (c). Asterisks denote chemical noise from NH_4^+, Na^+, and K^+ present in the solvents that attach to 18C6.

Similarly, experiments with varying amounts of acid were performed to test the extent to which acidic molecules might competitively interfere with complexation of pentalysine by 18C6. The addition of acid leads to several changes in the resulting spectra. The total ion intensity increases significantly, the intensity of naked peptide peaks increases, and the relative intensity of some 18C6 adduct peaks is reduced. It is unclear whether the increased abundance of

naked peptides is a result of reduced 18C6 attachment or simply represents an enhancement of the intensity of previously unobserved ions in the presence of acid. In either case, the net results suggest that some interference due to the addition of acid may occur and reduce the amount of 18C6 complexes that are formed. However, the number of 18C6 adducts does not change (only the relative intensity of some contributing peaks). Furthermore, an experiment resulting in an increase in the number of 18C6 adducts after the addition of acid would unambiguously indicate a structural rearrangement for an experiment involving a macromolecule.

5.3.2 Protein Structure in Methanol

Many studies have shown that organic solvents can influence the equilibrium structures that proteins adopt.[21] It is also well known that the addition of methanol or trifluoro-ethanol can be used to stabilize and even increase the helical content of a protein.[22] Cyt c is an important heme containing electron-transfer protein with a predominantly helical structure[23] and no disulfide bonds. The data in Figure 5.2 show that the addition of 50% methanol does not lead to a significant shift in charge state distribution for cyt c, in agreement with previous results.[24] However, from spectroscopic data, clearly cyt c undergoes a significant structural rearrangement in 50:50 water/methanol.[25]

The spectra for cyt c with 18C6 in water and 50:50 water/methanol are shown in Figure 5.2a and b. It is clear that the number and distribution of crown adducts is very different, suggesting that at least two different structures are being sampled from solution. In Figure 5.2c–f, the 18C6 adducts associated with a particular charge state from Figure 5.1a and b are isolated. The intensities are shown relative to the base peak for easy comparison between the two distributions. The data in Figure 5.1c–f demonstrate that in 50% methanol more 18C6s attach to cyt c both in number and in relative intensity, suggesting a significant structural shift.

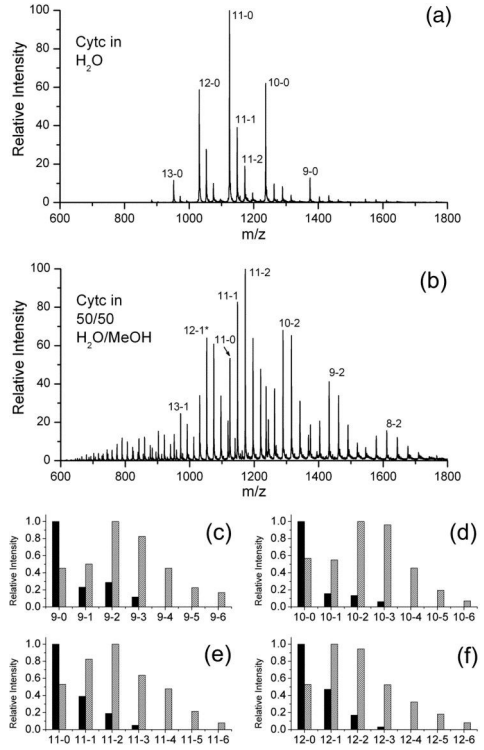


Figure 5.2 ESI-MS spectra of cyt c and 18C6 with no acid in (a) water and (b) 50/50 water/methanol. Several peaks are identified by two numbers (charge state-#18C6). The resulting spectra are noticeably different. In (c)–(f), the data from (a) and (b) are shown for individual charge states for easy comparison. The black bars are data extracted from (a) while the gray bars represent data from (b). Intensities are shown relative to the base peak for each distribution. Asterisk denotes that the peak overlaps with 13–5.

The number of 18C6 adducts that attach to cyt c in Figure 5.2a is qualitatively consistent with the number of lysines that are not in close proximity to a strong hydrogen bond partner in the known crystal structure.[26] The contrast with the results shown in Figure 5.2b are particularly intriguing in light of the fact that the charge state distribution does not shift substantially, which typically indicates that the protein is still in a compact form. However, from the dramatic increase in the number of 18C6s that attach in Figure 5.2b, clearly the observed

structure is not the native structure. This idea is further supported by previous work showing a change in heme coordination in 50% methanol.[27] A substantial portion of the protein in the crystal structure of cyt c is in the random coil configuration. The addition of methanol may induce organization of this region into more helical arrangements.

It is clear from the results in Figure 5.2 that structural rearrangement not involving the complete unfolding of a protein can still lead to substantial modification of the availability of lysine residues for binding by 18C6. Several potential factors may influence lysine availability. For example, intramolecular side-chain–side-chain interactions, such as salt bridges, could potentially interfere. However, there are 19 lysines in cyt c but only a few of these are involved in salt bridges in the crystal form. Other intramolecular hydrogen bonding interactions between side chains or with the peptide backbone may play a significant role in determining lysine availability for cyt c. If these interfering interactions are optimized in the native state, then lysine availability should typically increase if the native structure is perturbed. In the case of cyt c, non-native helical regions induced by the addition of methanol may disrupt intramolecular interactions that otherwise inhibit attachment of 18C6. All of the above mentioned interfering interactions are individually weak in solvated

environments; however, 18C6 is also a weak binder. This allows it to be an efficient, noninterfering probe. In fact, 18C6 does not interfere with and can be used advantageously in conjunction with enzymatic catalysis.[28]

Under identical experimental conditions, ubiquitin yields very different results as shown in Figure 5.3. Ubiquitin is a 76 residue protein with mixed helical and β sheet structure and no disulfide bonds.[29] Comparison of the spectra obtained in water (Figure 5.3a) and 50/50 water/methanol (Figure 5.3b) reveals very similar results, suggesting that no significant structural reorganization takes place in the case of ubiquitin. The data shown in Figure 5.3c–f confirm that the 18C6 adduct distributions associated with each charge state are similar in both shape and relative intensities. The similarity of the results in both solvent systems may be related to the overall stability of the structure of ubiquitin.[30]

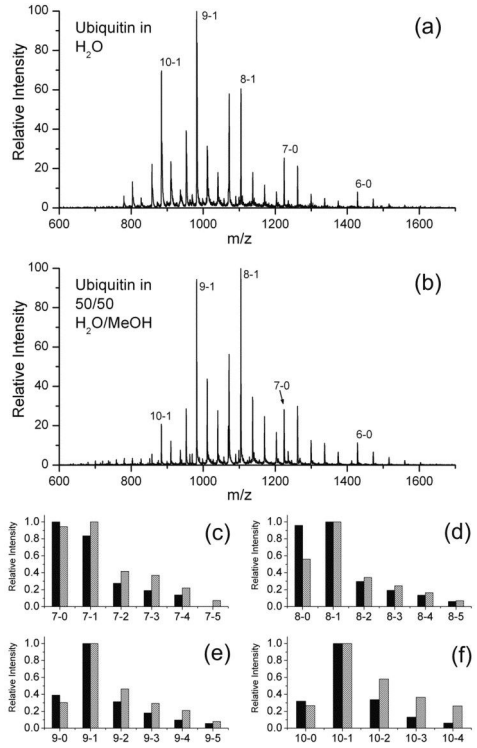


Figure 5.3 ESI-MS spectra of ubiquitin and 18C6 with no acid in (a) water and (b) 50/50 water/methanol. Several peaks are identified by two numbers (charge state-#18C6). The resulting spectra are fairly similar, indicating similar structures are sampled in both experiments. In (c)–(f), the data from (a) and (b) are shown for individual charge states for easy comparison. The black bars are data extracted from (a) while the gray bars represent data from (b). Intensities are shown relative to the base peak for each distribution.

However, closer inspection suggests the presence of two structures in both systems. In water, the first structure is present in the +7 and +8 charge states, as noted by the similar relative intensities of the 7-0, 7-1 and 8-0, 8-1 peaks. In 50/50 water/methanol, the first structure is present only in the +7 charge state. Nearly identical 18C6 distributions are observed in both solvent conditions for the +9

and +10 charge states. These results suggest an equilibrium between the native structure and a very slightly more accessible structure under both solvent conditions, with the more accessible structure being more favored in methanol. It has been shown previously by the addition of a non-native disulfide bond that ubiquitin must be flexible to retain functionality.[31] This observation is in excellent agreement with our observation of two "native" structures. The shift in the 18C6 distribution to include one additional 18C6 suggests that the more accessible form is the result of a structural rearrangement exposing one additional lysine. Previous results suggest that the C-terminal domain (Gly53-Gly76) is less stable than the N-terminal portion of the structure.[26] There is a single lysine residue (Lys63) in this region near the C-terminus, which may be the lysine that 18C6 attaches to in the more accessible structure.

5.3.3 Acid Induced Unfolding

The addition of acid is frequently used to denature proteins.[27] As discussed above, this type of denaturation is commonly accompanied by a shift to higher charges states. The spectrum for ubiquitin and 18C6 in 50/50 water/methanol with 0.5% acetic acid added is shown in Figure 5.4a. Comparison with Figure 5.3b reveals that the charge state distribution has shifted significantly. The data shown in Figure 5.4b–e compares the 18C6 distributions for several charge states

in the absence and presence of acid. The +6 charge state is still present in very small abundance with the addition of acid, yet the structure appears to be unchanged when compared with the +6 charge state in the absence of acid as shown in Figure 5.4b. This observation is in accord with previous studies where it was found that the native structure was still present under acidic conditions.[32] The +8 charge state shown in Figure 5.4c also appears to have a similar distribution with or without acid, although the structure is potentially different from that observed in the +6 charge state. The +10 and +13 charge states clearly reveal the presence of a more open structure which attaches 18C6 in higher abundance. This is most likely the a-state of ubiquitin, which has been studied previously.[27,29] All of the charges states over +10 yield very similar distributions to that shown in Figure 5.4d. It can also be clearly observed that the structure yielding the 10+ charge state is different in the presence and absence of acid, a distinction that cannot be made without the addition of 18C6.

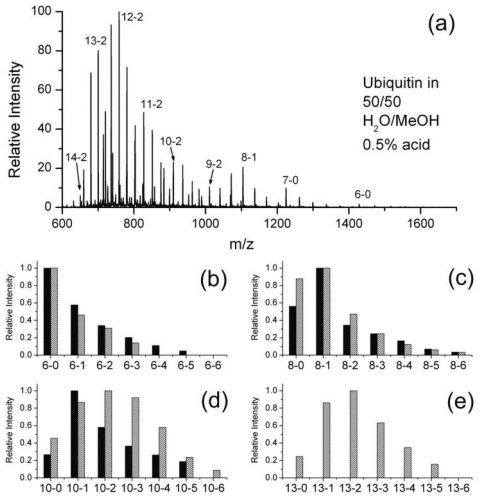


Figure 5.4 ESI-MS spectra of ubiquitin and 18C6 with 0.5% acetic acid (pH 3.6) in 50/50 water/methanol. Several peaks are identified (charge state-#18C6). The spectrum is compared with data from Figure 5.3b in (b)–(e). The black bars are data extracted from Figure 5.3b while the gray bars represent data from (a). Intensities are shown relative to the base peak for each distribution. Acid induces a noticeable structural shift, increasing the number of attached crowns.

The shift in charge state implies a more open structure, which is confirmed by the addition of more 18C6 adducts. The distribution of 18C6 also shifts at the point that would be predicted by the bimodal charge state distribution, in agreement with previous predictions. Furthermore, three different structures are predicted by comparison of the 18C6 distributions in Figure 5.4b, c, and e. The presence of three structures is in agreement with previous results obtained by combined H/D exchange and charge state distribution analysis.[8] In this study, the three structures were detected by comparing results from both techniques. With the current approach, all three structures are observed with a single method.[8]

5.3.4 Melittin and Secondary Structure

The results obtained with cyt c and ubiquitin suggest that the attachment of 18C6 is sensitive to changes in both the overall tertiary structure (i.e., compact or extended) and to more subtle changes in the secondary structure. To probe this possibility more thoroughly, we conducted experiments with melittin, which is known to undergo a transition from a random coil to an α helix under appropriate conditions. Melittin is a 26 residue peptide with an amidated c-terminus that comprises the principle component of honey bee venom. Melittin has a random coil structure in water, which becomes primarily α helical in 25% TFE or >50% methanol.[33,34] The spectra for melittin and 18C6 in water and 20:80 water/methanol are shown in Figure 5.5. Acid (0.5% acetic acid) was added to both samples to enhance the signal, but melittin has no acidic groups and is not sensitive to structural changes under pH 7.[35] Examination of the 18C6 distributions shown in Figure 5.5c–f reveals that lysine availability changes significantly with the transition to an α helical structure.

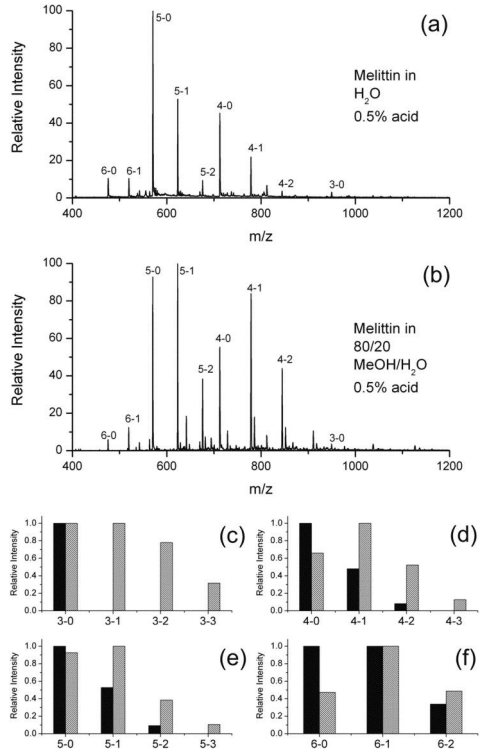


Figure 5.5 ESI-MS spectra of melittin and 18C6 with 0.5% acetic acid in (a) water (pH 3.05) and (b) 20/80 water/methanol (pH 2.97). Several peaks are identified (charge state-#18C6). The 18C6 distributions shift significantly as shown in (c)–(f), where the data from (a) and (b) are extracted for individual charge states. The black bars are data extracted from (a) while the gray bars represent data from (b). Intensities are shown relative to the base peak for each distribution.

This observation can be rationalized in terms of hydrogen bonding with the peptide backbone. In the absence of acidic groups, the backbone carbonyls offer the best potential intramolecular sites for hydrogen bonding. The backbone carbonyls are available to interact with the side chains if the peptide is in a random coil conformation. On the other hand, the carbonyls become largely unavailable if the peptide adopts an α helical structure, which should lead to enhanced availability of the side chains. Indeed, it is observed in Figure 5.5 that

lysine availability increases substantially with the shift to the helical structure. Both the number and relative abundance of 18C6 adducts increases for the helical structure as seen in Figure 5.5c–f. Our data show the formation of three 18C6 adducts suggesting that all three lysines are available, in agreement with the predicted availability based on the X-ray structure.[35] It is important to note that the charge state distribution does not change significantly, which implies that evidence for the structural transition cannot be obtained through charge state analysis alone. However, with the addition of 18C6, the transition is easily detected, confirming that the present technique is sensitive to changes in secondary structure.

5.4 Conclusions

A new method for probing protein structure has been described. The technique relies on ESI-MS for the detection of noncovalent reporter molecules that reveal the folding state of the protein to which they attach. In the present work, 18C6 is used to recognize the side chain of lysine, but nothing prevents the technique from being applied with other small molecules that also interact specifically with amino acid side chains. It is possible to probe changes in both

secondary and tertiary structure, while retaining experimental simplicity. The application of this technique to examine the folding state of proteins will provide a complimentary method for addressing the complex problems of protein folding.

---

[1] Wales, T. E.; Engen, J. R. *Mass Spectrom. Rev.* **2006**, *25*, 158-170.
[2] Kaltashov, I. A.; Eyles, S. *J. Mass Spectrom. Rev.* **2002**, *21*, 37-71.
[3] Englander, S. W. *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *29*, 213-238.
[4] Mandell, J. G.; Falick, A. M.; Komives, E. A. *Anal. Chem.* **1998**, *70*, 3987-3995.
[5] Garcia, R. A.; Pantazatos, D.; Villareal, F. J. *Assay Drug Dev. Technol.* **2004**, *2*, 81-91.
[6] Zhu, M. M.; Chitta, R.; Gross, M. L. *Int. J. Mass Spectrom.* **2005**, *240*, 213-220.
[7] Roulhac, P. L.; Powell, K. D.; Dhungana, S.; Weaver ,K. D.; Mietzner, T. A.; Crumbliss, A. L.; Fitzgerald, M. C. *Biochemistry* **2004**, *43*, 15767-15774.
[8] Pan, J.; Wilson, D. J.; Konermann, L. *Biochemistry* **2005**, *44*, 8627-8633.
[9] Yan, X.; Watson, J.; Ho, P. S.; Deinzer, M. L. *Mol. Cell. Proteomics* **2003**, *3*, 10-23.
[10] Grandori, R. *J. Mass Spectrom.* **2003**, *38*, 11-15.
[11] Dobo, A.; Kaltashov, I. A. *Anal. Chem.* **2001**, *38*, 4763-4773.
[12] Fenn, J. B. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 524-535.
[13] Konermann, L.; Douglas, D. J. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 435-442.
[14] Konermann, L.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1248-1254.
[15] Julian, R. R.; Beauchamp, J. L. *Int. J. Mass Spectrom.* **2001**, *210*, 61, 613-623.
[16] Cunniff, J. B. and Vouros, P. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1175-1182.
[17] Sproch, N.; Kruger, T. L. *Proceedings 41st ASMS Conference; San Francisco, CA* **1993**, 904a-904b.
[18] Bosch, E.; Bou, P. Allemann, H.; Roses, M. *Anal. Chem.* **1996**, *68*, 3651-3657.
[19] Fink. A. L.; Calciano, L. J.; Goto, Y.; Kurotsu, T.; Palleros, D. R. *Biochemistry* **1994**, *33*, 12504-12511.
[20] Mattice, W. L.; Harrison, W. H. *Biopolymers* **1975**, *14*, 2025-2033.
[21] Buck, M. *Q. Rev. Biophys.* **1998**, *31*, 297-355.
[22] Hirota-Nakaoka, N.; Goto, Y. *Bioorg. Med. Chem.* **1999**, *7*, 67-73.
[23] Bushnell, G. W.; Louie, G. V.; Brayer, G. D. *J. Mol. Biol.* **1990**, *214*, 585.
[24] Konerman, L.; Douglas, D. J. *Biochemistry* **1997**, *36*, 12296-12302.
[25] Suzumura, A.; Paul, D.; Sugimoto, H.; Shinoda, S.; Julian, R. R. Beauchamp, J. L.; Teraoka, J.; Tsukube, H. *Inorg. Chem.* **2005**, *44*, 904-910.
[26] Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531-544.
[27] Brutscher, B.; Bruschweiler, R.; Ernst, R. R. *Biochemistry* **1997**, *36*, 13043-13053.
[28] Ecker, D. J.; Butt, T. R.; Marsh, J.; Sternberg, E.; Shatzman, A.; Dixon, J. S.; Weber, P. L.; Crooke, S. T. *J. Biol. Chem.* **1989**, *264*, 1887-1893.
[29] Hoerner, J. K.; Xiao, H.; Kaltashov, I. A. *Biochemistry* **2005**, *44*, 11286-11294.
[30] Dill, K. A.; Shortle, D. *Annu. Rev. Biochem.* **1991**, *60*, 795-825.

[31] Mohimen, A.; Dobo, A.; Hoerner, J. K.; Kaltashov, I. A. *Anal. Chem.* **2003**, *75*, 4139-4147.
[32] Wang, F.; Polavarapu, P. L. *Biopolymers* **2003**, *70*, 614-619.
[33] Hirota, N.; Mizuno, K.; Goto, Y. *J. Mol. Biol.* **1998**, *275*, 365-378.
[34] Bazzo, R.; Tappin, M. J.; Pastore, A.; Harvey, T. S.; Carver, J. A.; Campbell, I. D. *Eur. J. Biochem.* **1988**, *173*, 139-146.
[35] Eisenberg, D.; Gribskov, M.; Terwillger, T. C. PDB ID: 2 MLT; unpublished.

Chapter 6

PROTEIN-METAL INTERACTIONS OF CALMODULIN AND ALPHA-SYNUCLEIN MONITORED BY SELECTIVE NONCOVALENT ADDUCT PROTEIN PROBING

6.1 Introduction

Proper protein function is determined by a variety of factors including primary sequence, three-dimensional structure, and binding of essential ligands. The interdependence of structure and ligand binding is illustrated by numerous instances in biochemistry where ligand binding confers biological function. For example, association of calcium ions with calmodulin induces a structural rearrangement that reveals a key hydrophobic α helix which can recognize and bind to other proteins, ultimately signaling enzymatic phosphorylation of the substrate.[1,2] Thus, calmodulin acts as an intracellular Ca^{2+} sensor, marking proteins for phosphorylation only when Ca^{2+} is bound.

Ligand binding can also induce structural transitions in proteins that are undesirable and may eventually lead to the formation of protein aggregates, which are associated with several diseases. The pathologies of Alzheimer's disease,[3] Parkinson's disease,[4] and Creutzfeldt-Jakob disease[5] are linked to extracellular aggregation of otherwise harmless neural proteins. One of the key

questions in the study of these neurodegenerative diseases is what causes normally functioning proteins to aggregate? There are several factors that have been associated with protein aggregation, including site specific mutations,[6-7] protein oxidation,[8] and ligand binding.[9-11] For example, in Alzheimer's disease it has been shown that increased levels of metals such as Cu^{2+} and Zn^{2+} are linked to aggregation of Aβ protein *in vitro*.[12] The theory of metal induced aggregation has gained credence following numerous studies tying metal concentrations in the brain with Alzheimer's disease *in vivo*.[9]

Recently, there has been significant interest in α-synuclein, a protein that is suspected to play an essential role in the pathology of Parkinson's disease.[13] α-synuclein is the major component of Lewy bodies,[14] which are protein aggregates found in the neuronal and glial cytoplasm of patients affected with Parkinson's disease,[15] dementia,[16] and a specific variant of Alzheimer's disease.[17] Perhaps the most convincing evidence to date that links α-synuclein with the onset of Parkinson's disease is the discovery that single point mutations in the gene that codes for α-synuclein increases the likelihood of manifesting Parkinson's disease to 85%.[18,19] α-synuclein is found naturally in the cytosol and presynaptic terminals of neurons. A complete understanding of the biological function of α-synuclein remains a significant challenge. It has been shown that synelfin, an α-

synuclein homolog found in zebra finches, is up-regulated during song-learning, which suggests that the protein is involved in neuronal plasticity.[20] Additionally, α-synuclein has been implicated in biochemical pathways involving neuronal cell apoptosis,[21] regulation of dopamine,[22] and molecular chaperoning.[23]

α-synuclein is a natively disordered protein and adopts a wide distribution of conformations.[24] Commonly used techniques to determine protein structure, such as x-ray crystallography, are inappropriate for disordered proteins like α-synuclein. Therefore, studies of α-synuclein have typically relied on techniques that report data averaged over the entire ensemble of conformations. For example, using fluorescence, UV/Vis spectroscopy and far-UV circular dichroism, Fink et al. found that metal cations bind to α-synuclein and accelerate fibril formation *in vitro*.[11] Among the most effective metals at accelerating aggregation were Al^{3+} and Cu^{2+}. In a separate study by Fernandez et al., the putative Cu^{2+} binding sites were localized to near residue-specific detail using nuclear magnetic resonance spectroscopy.[25] Recently, it was shown that time-resolved fluorescence resonance energy transfer experiments with α-synuclein can effectively separate and identify unique conformations in the ensemble due to differences in inter-residue distances and contact rates.[26] Results from these experiments have demonstrated that the protein is highly dynamic. Fast

interchange between conformations occurs even among α-synuclein mutants where the rate of protein aggregation is significantly accelerated. Thus, it is very difficult to examine the different conformational states independently without sophisticated instrumentation and introduction of non-native fluorescence energy transfer donor-acceptor pairs. Fortunately, separation by conformation occurs automatically in the course of electrospray (ESI)[27] MS experiments, as is discussed in more detail below.

A number of approaches utilizing mass spectrometry (MS) have been developed to examine protein structure.[28] Typically, these techniques use chemical reactions to label specific functional groups on the protein with covalently attached mass tags. The number of functional groups available for covalent modification is dependent on the protein conformation and the chemical environment surrounding each reactive group. This allows mass spectrometry to distinguish between different protein conformations. The simplest covalent modification observed in conjunction with MS to examine protein structure is protonation. Basic residues such as arginine, lysine and histidine are expected to be protonated first, followed by protonation at other sites. The basicity of these secondary protonation sites may be enhanced by multi-dentate interactions with the protein backbone, which stabilizes the

charge.[29] A number of reports have been published on the dependency of protein charge state distribution on conformations found in solution using ESI-MS.[30-32] The number of charges that can be placed on a protein is dependent on its conformation and is limited by Coulombic repulsion. Indeed, recently it was reported that the charge state distribution correlates well to the surface area of a protein.[33] Nevertheless, charge state distributions report only the approximate size of a protein in an indirect fashion, meaning that only dramatic changes can be reliably ascribed to structural transitions.

Covalent chemistry can also be tailored to target specific functional groups of a protein. These reactions have been applied extensively in cross-linking studies,[34-38] where length constraints imposed by the cross-linking reagent can provide approximate distances between reactive functional groups. Hydrogen-deuterium exchange (HDX) MS is another approach to examine conformational changes in proteins.[39-42] For example, Gross et al. found that the number of exchangeable hydrogens decreases as the concentration of Ca^{2+} is increased up to approximately 0.25 mM, which they interpret as a tightening of the Ca^{2+} binding domains of calmodulin.[43,44] Reactive labeling groups can also be used to monitor protein surface structure.[45] Specific, noncovalent interactions with the sidechains of proteins can also be used to probe protein structure.[46-50]

Selective noncovalent adduct protein probing mass spectrometry (SNAPP-MS) was recently developed to examine the structure of proteins.[48-50] The technique relies on the selective association of 18-crown-6 ether (18C6) with the side-chain of lysine.[51] The noncovalent interaction between 18C6 and lysine is weak in solution,[52] permitting sensitive sampling of the chemical microenvironment surrounding each lysine. Lysine sidechains involved in intramolecular interactions, like hydrogen bonds and salt-bridges, do not bind well to 18C6.[49] Modulation of these interactions, due to structural rearrangements, typically change the overall lysine availability for the protein. Once 18C6-protein complexes are transferred into the gas phase by ESI, 18C6 strongly associates with the side-chain of lysine, effectively preserving equilibrium solution phase information for subsequent MS detection.[51] The resulting mass spectra reveal a distribution of protein-18C6 complexes (or SNAPP distribution) for each charge state of the protein. Previous work suggests that the intensities and shapes of these SNAPP distributions are the result of statistical averaging of 18C6 attachment to many potential binding sites. Modulation of the lysine availability at any of these sites due to structural shifts or ligand binding will be easily observed by comparison with the unperturbed protein. In addition, sample preparation for a typical SNAPP-MS experiment is

simple, requiring only addition of 18C6 and protein to a suitable solvent, followed by ESI-MS.

Herein, we demonstrate that SNAPP-MS is sensitive to conformational changes due to protein-metal interactions in calmodulin and α-synuclein. We have examined conformational changes in calmodulin due to Ca^{2+} binding, and found that our SNAPP-MS results are in agreement with the known structures for apo and holo calmodulin. For α-synuclein, where the effects of metal binding are not completely understood, we have examined the effect of biologically relevant (μM) concentrations of copper and aluminum. We find that both copper and aluminum bind to α-synuclein, but only aluminum causes a significant change in the lysine availability. This change in lysine availability may reflect a shift in the conformational dynamics of α-synuclein to aggregate-prone structures. Comparison of results with both Al^{3+} and Cu^{2+} reveal that the metals accelerate α-synuclein aggregation by independent mechanisms.

6.2 Materials and Methods

Calcium acetate (Mallinckrodt Baker, Phillipsburg, NJ), acetic acid (EMD Biosciences, Gibbstown NJ), aluminum chloride (Sigma Aldrich, St. Louis, MO) and copper(II) chloride (Sigma Aldrich, St. Louis, MO) were of analytical grade and used without further purification. Water was purified to 18.2 MΩ resistivity

using a Millipore Direct-Q (Billerica, MA). Porcine calmodulin was purchased from Ocean Biologics (Corvallis, OR). Recombinant human α-synuclein (rPeptide, Bogart, GA) was further purified by dialysis against H_2O and lyophilized. Stock solutions of calcium acetate were neutralized with acetic acid before addition to protein. Samples contained 10 μM and 3 μM protein for calmodulin and α-synuclein experiments, respectively. For SNAPP experiments, a concentration of 18C6 equal to two times the number of lysines was used to ensure stoichiometric excess of 18C6. For calmodulin, which has 8 lysine residues, the concentration of 18C6 used was 160 μM. For α-synuclein, which has 15 lysine residues, the concentration of 18C6 was 90 μM. Replicate sample solutions were spiked with the appropriate metal cation to test for protein-ligand binding. To form holocalmodulin, the calmodulin-18C6 solutions mentioned above were spiked with $Ca(OAc)_2$ to give a final concentration of 25 μM Ca^{2+}. To test the effects of metals on α-synuclein, solutions containing both α-synuclein-18C6 were spiked with either $AlCl_3$ or $CuCl_2$ to give final concentrations of 3 μM $AlCl_3$ or 100 μM $CuCl_2$, respectively. All samples were of neutral pH by litmus paper.

Mass spectra were obtained using an LTQ linear ion trap mass spectrometer (Thermo Fisher, Waltham, MA) equipped with a standard electrospray source.

For each protein, voltages, sheath gas flow rates and temperatures in the source region were optimized for observation of noncovalent complexes, and are similar to parameters described previously.[48] Once optimized, these parameters remained unchanged for all experiments for that protein. The following are typical source parameters: spray voltage 4.5 kV, tube lens 150 V, capillary temperature 215 °C, capillary voltage 40V. For calmodulin, the capillary temperature was raised to 285 °C.

6.3 Results and Discussion

6.3.1 Calmodulin

The structure of calmodulin has been described as a "dumbbell" shape because it contains two globular domains bridged by an α helical linker.[1] Each globular domain has two EF-hand structural motifs, which are high affinity binding sites for Ca^{2+} ions.[53-55] Figure 6.1a shows the mass spectrum obtained from a solution containing apocalmodulin and 18C6. A statistical distribution of 18C6 adducts is observed for each charge state. The number of 18C6s that bind to calmodulin decreases with increasing charge state. This observation is highly unusual and would seem to indicate that more extended structures lead to reduced lysine availability. Examination of the sequence of calmodulin suggests a possible explanation. Six out of the eight lysine residues of calmodulin are

adjacent in sequence to an acidic residue. This means that as the protein unfolds, each lysine remains in close proximity to an acidic residue which can interfere with attachment of 18C6.

The most abundant peak observed for any charge state is calmodulin with no 18C6 attached even though 18C6 is present in significant excess in solution. This result is consistent with previous observations[48] and is due to the presence of interfering intramolecular interactions restricting lysine availability. A zoomed-in view of the +12 charge state is shown in Figure 6.1b. The asterisked peaks are 18C6 adducts of calmodulin dimer, which is present in minor abundance. Noncovalent dimers of calmodulin have been observed previously in ESI mass spectra at similar concentrations of protein.[56,57] These dimers have been shown to exist in solution, and are not simply an experimental artifact due to the ionization process.[57]

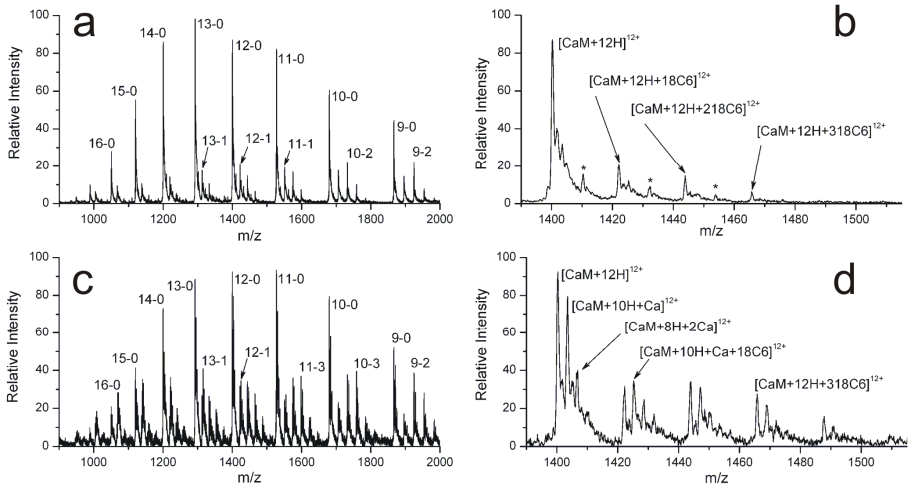


Figure 6.1 (a) ESI-MS of solution of calmodulin (CaM) and 18C6 in H_2O. (b) Close-up of the +12 charge state region of Figure 6.1a. Asterisks indicate CaM dimer-18C6 adduct peaks. (c) ESI-MS of calmodulin, 18C6 and 25 μM Ca^{2+} in H_2O. Comparison between apo and holo solution show an increase in 18C6 attachment in holocalmodulin indicating a structural change has occurred. (d) Close-up of the +12 charge state region of the spectra shown in Figure 6.1c, revealing peaks corresponding to both attachment of 18C6 and Ca^{2+}. The ESI-MS results do not match the Ca^{2+} binding stoichiometry in solution, suggesting loss of Ca^{2+} ions occur during ESI. Labels correspond to (charge state) – (# 18C6s attached).

To test whether SNAPP-MS can detect a conformational change due to metal binding, a 2.5-fold excess of Ca^{2+} was added. Calmodulin binds up to four calcium ions with the following association constants: 1.0×10^7 M^{-1}, 3.98×10^7 M^{-1}, 3.16×10^6 M^{-1}, and 2.51×10^6 M^{-1}.[58] In comparison, association constants determined for 18C6 and Ca^{2+} ($K = 3.39$ M^{-1} to 5.01 M^{-1}) are much smaller and indicate that competitive binding of Ca^{2+} by 18C6 is insignificant.[52] The fractional abundance of each calcium-bound species was calculated at the concentration of

total calcium and total calmodulin used in our experiments.[43] The predominant species expected in solution is calmodulin + $2Ca^{2+}$ (46.7%), followed by calmodulin + $3Ca^{2+}$ (29.6%), calmodulin + $4Ca^{2+}$ (15.0%), calmodulin + $1Ca^{2+}$ (5.8%) and apocalmodulin (2.9%). Of the total calcium in solution, only 0.8% is not bound to calmodulin. In theory, free calcium might hinder 18C6 binding to lysine; however, the concentration of free calcium in solution is over two orders of magnitude lower than the concentration of 18C6 and is therefore not expected to interfere.

Figure 6.1b shows the ESI mass spectrum for the solution containing calmodulin, Ca^{2+} and 18C6. Even cursory comparison with Figure 6.1a reveals significant differences between the two spectra. Closer examination, as shown in Figure 6.1c, reveals additional peaks corresponding to Ca^{2+} adducts. In contrast to what is expected in solution, apocalmodulin remains the largest peak in the spectrum with a maximum observed binding stoichiometry of calmodulin + $3Ca^{2+}$. This discrepancy has been observed previously in ESI-MS experiments,[56,] and is due to loss of Ca^{2+} during the final stages of the desolvation process. Water is an essential ligand in the Ca^{2+} coordination shell when bound to the protein;[1] therefore, it is unsurprising that desolvation in the ESI source destabilizes Ca^{2+} binding. Importantly, the loss of Ca^{2+} does not lead to a substantial shift in the

SNAPP distributions because there is not enough time for the protein to unfold and 18C6 to re-equilibrate before fully desolvated ions are formed. Once the protein is largely desolvated, the 18C6/lysine interaction becomes very strong and the SNAPP distribution is not likely to undergo further changes.

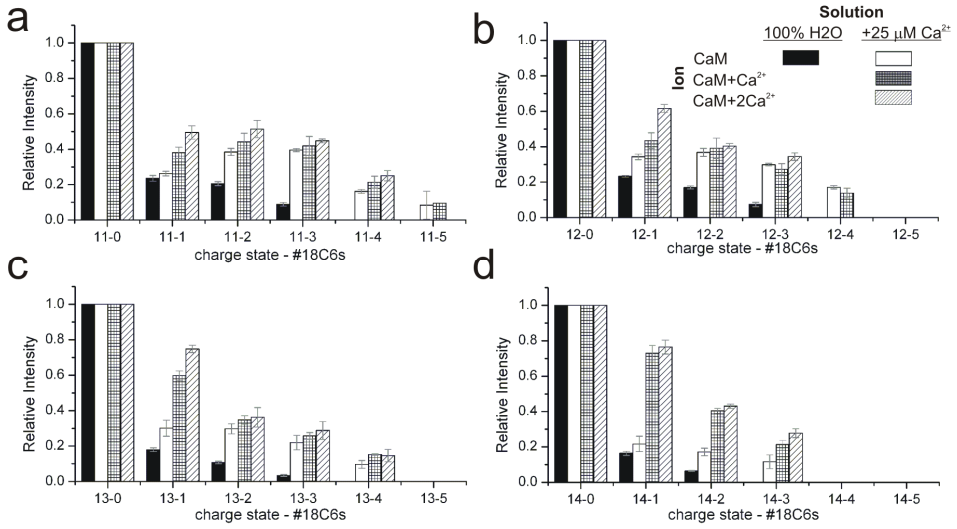


Figure 6.2 Extracted 18C6 distributions for +11 (a), +12 (b), +13 (c) and +14(d) from spectra shown in Figure 6.1. Addition of Ca^{2+} in solution increases the number and intensity of 18C6-calmodulin complexes. Ions retaining Ca^{2+} in the gas phase have 18C6 distributions shifted towards more 18C6 adducts attached compared to apocalmodulin. The significant change in 18C6 distributions permits detection of protein-metal ligand interaction. X-axis labels correspond to (charge state) – (# 18C6s attached).

Additional insight can be extracted by examining adduct distributions for individual charge states. The black bars in Figure 6.2 show SNAPP distributions taken from the mass spectrum for the +11 through +14 charge states for apocalmodulin sampled from water. For straightforward comparison, adduct distributions sampled from Ca^{2+} solution (white and hatched bars) are shown

alongside the distributions sampled from water. The appearance of Ca^{2+}-protein complexes in the mass spectra provides an additional level of information because now there are superimposed adduct distributions involving protein and both 18C6 and Ca^{2+}.

The gas phase $[Calmodulin + nH]^{n+}$ ions sampled from Ca^{2+} solution are produced from a mixture of apocalmodulin, and calmodulin-Ca^{2+} complexes where Ca^{2+} has been selectively lost during ESI. By mass spectrometry alone, these isobaric ions are indistinguishable. However, addition of 18C6 reveals intriguing differences between $[Calmodulin + nH]^{n+}$ ions sampled from water and Ca^{2+} solution. As shown in Figure 6.2, the SNAPP distributions for $[Calmodulin + nH]^{n+}$ from Ca^{2+} solution (white bars) show an increase both in intensity and number of 18C6 adducts compared to data obtained from water (black bars). These results suggest that the solution phase conformations of calmodulin-Ca^{2+} complexes have increased lysine availability compared to apocalmodulin. Furthermore, observation of the protein-metal complex is not required to observe a conformational change using SNAPP-MS, as is easily confirmed by inspection of the isobaric black and white bars in Figure 6.2, which both correspond by mass to apocalmodulin. This is a significant advantage in using SNAPP for monitoring ligand-binding events because many protein-ligand complexes may be

destabilized during desolvation, resulting in ions that are apoprotein by mass, but structurally originate from holoprotein conformations in solution.

The number of 18C6 adducts also increases with additional calcium adducts (see hatched bars in Figure 6.2). These increases are due to the greater proportion of holocalmodulin vs. apocalmodulin contributing to the SNAPP distributions. For example, the [Calmodulin + $1Ca^{2+}$] ion may arise from ESI of Calmodulin + $1Ca^{2+}$ in solution and also loss of Ca^{2+} from [Calmodulin + nCa^{2+}] (n > 1), where greater structural rearrangement is expected. Thus, the peak assignment provides only a lower limit on the Ca^{2+} stoichiometry of the calmodulin-Ca^{2+} complex sampled from solution.

Table 6.1 Distance to Nearest Heteroatom Neighbor for Each Lysine in CaM[a]

	apoCaM 1QX5[55]	**holoCaM** 3CLN[1]
LYS13	3.66	5.18
LYS21	2.91	2.60
LYS30	3.70	2.91
LYS75	3.27	4.84
LYS77	4.51	5.46
LYS94	3.14	5.66
LYS115	6.26	6.74
LYS148	n/a[b]	n/a[b]

[a] 1QX5 and 3CLN are Protein Data Bank IDs

[b] No electron density found.

Acidic residues in the EF-hand motifs of calmodulin provide six out of the seven essential oxygen ligands to bind Ca^{2+}. In the absence of Ca^{2+}, these acidic residues may be hydrogen bonded or salt-bridged to nearby positively charged residues, such as lysine and arginine. To examine whether addition of Ca^{2+} changes intramolecular interactions involving lysine, we have compared the x-ray crystal structures of apocalmodulin and holocalmodulin. Table 6.1 lists the average nearest heteroatom distance measured for each lysine. These distances are a coarse measure of the local environment surrounding each lysine. If a hydrogen bond acceptor is within ~3Å of a lysine, then 18C6 attachment will likely be hindered. In particular, it has been demonstrated previously that lysines in salt-bridges are unavailable for 18C6 attachment.[49] The average nearest heteroatom distance for every lysine either increases or remains unchanged upon addition of Ca^{2+}, except for Lys30, which decreases from 3.70 A to 2.91 A. Based on this analysis, an increase in lysine availability is expected for holocalmodulin compared to apocalmodulin (in agreement with the experimental results). Closer inspection of the crystal structure for apocalmodulin reveals that two lysines, Lys21 and Lys94 are salt-bridged to E31 and E104, respectively, which are residues involved in Ca^{2+} binding. Addition of Ca^{2+} displaces lysine, which should increase lysine availability. In fact, closer examination reveals that the

intramolecular interactions for Lys21 and Lys94 change significantly. Zoomed-in views of Lys94 in the crystal structures for apo and holocalmodulin are shown in Figure 6.3 and highlight the changes in sidechain interactions due to calcium binding. Instead of being salt-bridged to E104 in the apo structure as seen in Figure 6.3a, Lys94 is freely available for 18C6 attachment in the holo structure (Figure 6.3b). Lys21 is no longer salt-bridged to E31 in holocalmodulin. Instead, Lys21 is hydrogen-bonded to the sidechain of Thr29, which is also expected to be less competitive towards 18C6 binding. The changes observed in these x-ray structures are consistent with our observation that lysine availability increases upon Ca^{2+} binding and is further evidence that SNAPP-MS samples solution phase structure. Increases in lysine availability due to the presence of metal may be an indication that the metal binding chemistry involves acidic sidechains. However, further work would be required to confirm this hypothesis.

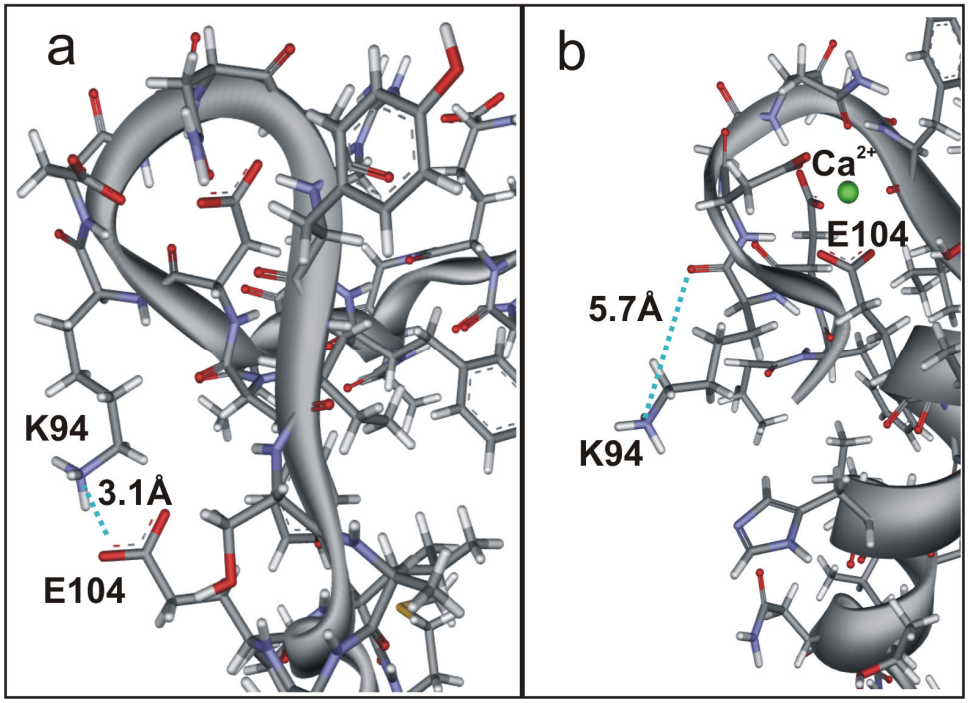


Figure 6.3 Partial view of the crystal structure of apocalmodulin (a), PDB ID: 1QX5, highlighting the sidechain interactions of Lys94. Note that the sidechain of Lys94 is salt-bridged to E104. In the corresponding crystal structure of holocalmodulin (b), PDB ID: 3CLN, the calcium (in green) displaces Lys94 from the intramolecular salt-bridge to coordinate with E104. Lys94 is then freely available to attach 18C6.

6.3.2 α-synuclein

An important question in the pathology of Parkinson's disease is how normally benign α-synuclein proteins aggregate. An existing theory paralleled in other amyloidogenic diseases such as Alzheimer's disease,[12] is that a conformational change occurs due to metal ligand binding.[59,60] Fink et al. demonstrated that out of fifteen metal cations studied, Al^{3+} accelerated fibril formation the most and caused conformational changes detectable by intrinsic protein fluorescence and far UV-circular dichroism.[11] The acidic C-terminal

domain was proposed as the binding site for Al^{3+}. Metal binding is thought to interfere with autoinhibitory interactions between the N- and C- termini, which protect the central hydrophobic domain from interacting with other α-synuclein proteins to form oligomers. [10,61] Although this result is a significant step in understanding the metallochemistry of α-synuclein, the biological relevance of the high Al^{3+} concentrations (0.1 – 1 mM) required to induce observable changes in conformation and fibril formation is questionable.[62]

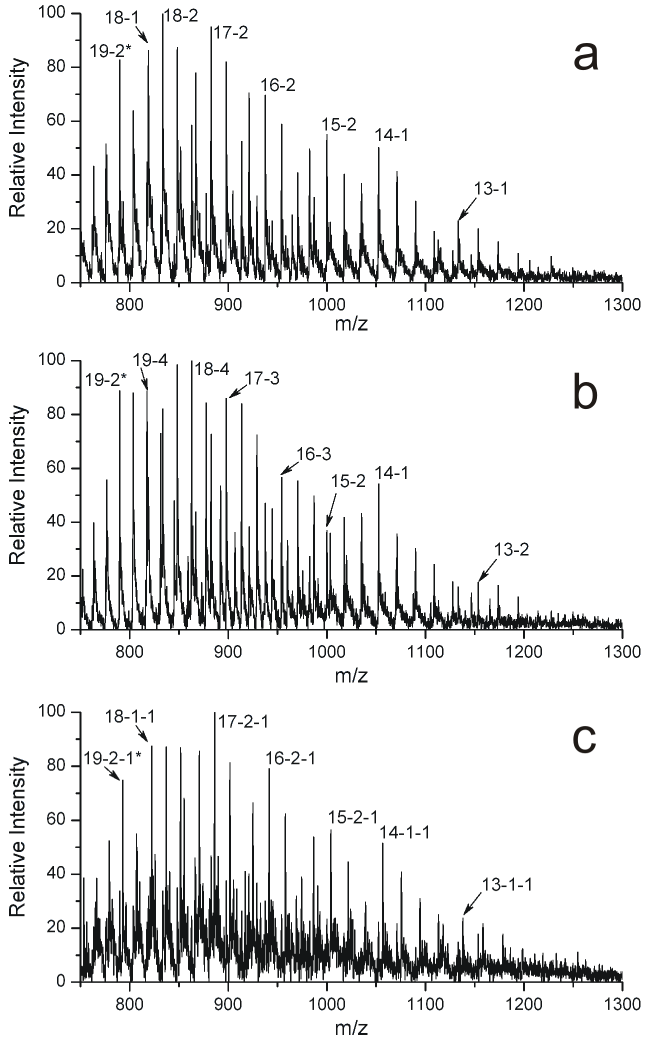


Figure 6.4 Mass spectra of electrosprayed solutions of α-synuclein in water (a), in 3 μM $AlCl_3$ (b), and in 100 μM $CuCl_2$ (c). The charge state distribution for α-synuclein sampled from each solution is broad, which is consistent with a natively disordered protein. The 18C6 distributions for α-synuclein in the presence of aluminum are significantly different from water and indicates a shift in the protein dynamics. In contrast, the charge state distribution is insensitive to the effects of aluminum or copper on α-synuclein. The labels in (a) and (b) correspond to (charge state) – (# 18C6s attached). Cu^{2+}-protein adducts are observed in (c), thus peaks are labeled (charge state) – (# 18C6s attached) – (# Cu^{2+} attached). An asterisk indicates that the peak assignment is ambiguous due to an m/z overlap with the 18C6 distribution of another charge state.

We have examined whether α-synuclein binds Al^{3+} by comparing 18C6 attachment to α-synuclein in the presence and absence of Al^{3+} in solution. Figure 6.4a shows the mass spectrum of α-synuclein and 18C6 in water. SNAPP distributions for the +13, +16, +17, and +18 charge states of α-synuclein in water are shown in Figure 6.5 (black bars). The maximum number of 18C6s attached is eight, about half (8/15) of the lysines present in the amino acid sequence of α-synuclein. Cytochrome c, which is similar to α-synuclein in molecular weight and number of lysines, binds only three 18C6s under similar conditions,[48] presumably due to higher order structure which prevents 18C6 attachment. The 18C6 distribution for the +13 charge state has maximal intensity at 13-2, or protein with two adducts attached. At higher charge states, the distributions shift to higher numbers of 18C6s. We can quantitatively compare the distributions by calculating the average number of 18C6s bound. For the +13 charge state, the average number of 18C6s attached is 2.16 ± 0.04. Higher charge states have a higher number of 18C6s bound, which are given in parentheses: +16 (2.81 ± 0.05), +17 (2.66 ± 0.03), +18 (2.50 ± 0.05). Interestingly, the 18C6 binding reaches a maximum at the +16 charge state and then decreases slightly with increasing charge. These data suggest that the conformations populated in these charge states have different lysine availabilities, and imply the existence of a

heterogeneous ensemble of conformations in solution. Assuming that conformations have varying propensities towards protonation due to differences in repulsive Coulombic interactions, then the 18C6 distributions that are resolvable by charge state represent different dynamic states of the protein. The wide charge state distribution observed for α-synuclein (from +11 to +20) is also consistent with a heterogeneous population of conformers. However, changes to the 18C6 distributions upon metal binding shown below reveal that these 18C6 distributions are more sensitive structural probes than examining charge states alone. Indeed, addition of metal cations does not significantly change the charge state distribution of the protein, but does significantly affect the 18C6 distributions.

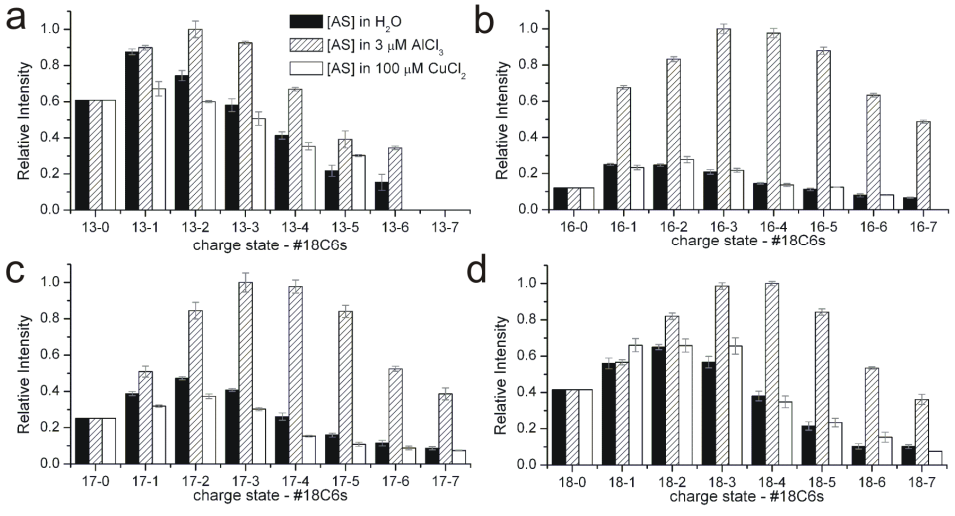


Figure 6.5 18C6 distributions for α-synuclein are shown for the +13 (a), +16 (b), +17 (c), and +18 (d) charge states in water (black bars) and in 3 μM Al^{3+} (hatched bars). A shift

towards increased 18C6 attachment occurs for the α-synuclein ion upon addition Al^{3+} to solution. Addition of 100 μM Cu^{2+} produces protein-Cu^{2+} adducts, with [α-synuclein + Cu^{2+}] being dominant. 18C6 distributions are shown for the [α-synuclein + Cu^{2+}] (white bars), which show only minor changes relative to α-synuclein sampled from water.

ESI-MS of α-synuclein and 18C6 in the presence of 3 μM Al^{3+} does not produce any observable Al^{3+} adducts in the gas phase, as seen in Figure 6.4b. The absence of protein-metal complexes in the gas phase does not preclude their presence in solution, due to destabilization of protein-metal complexes during desolvation. However, we are still able to monitor the effects of Al^{3+} by examining the 18C6 distributions for the apo α-synuclein ion. At 3 μM Al^{3+}, we observe that the 18C6 distributions for α-synuclein dramatically shift towards more adducts, as shown in the hatched bars in Figure 6.5. However, the magnitude of the increase is charge state dependent. The average number of 18C6s that bind to α-synuclein in the presence of Al^{3+} are as follows: +13 (2.56 ± 0.04), +16 (3.72 ± 0.03), +17 (3.59 ± 0.01), +18 (3.46 ± 0.01). Like in water, 18C6 binding reaches a maximum at the +16 charge state. For the +16, +17, and +18 charge states, addition of Al^{3+} increases the average number of 18C6s by approximately one. The +13 charge state is affected less by Al^{3+}, having an increase of only ~ 0.4.

These results suggest that the effect of Al^{3+} binding on the structure of α-synuclein varies as a function of conformational state. It is known that

accelerated fibril formation occurs by favoring a partially folded intermediate that is on-pathway to aggregation.[11] Our results suggest that increased aggregation by Al^{3+} may be due to structural perturbations to only a fraction of α-synuclein conformations. These results are also consistent with attachment to the C-terminal domain, where Al^{3+} binding may displace lysines involved in intramolecular interactions with acidic sidechains. The absence of arginine residues in α-synuclein means that any intramolecular salt-bridge interaction will most likely involve an acidic residue and lysine. Interestingly, 11 out of the 15 lysines of α-synuclein reside on the N-terminal domain, while 15 out of 24 acidic residues are on the C-terminal tail. The SNAPP results with Al^{3+} are consistent with a model where transient interactions occur between the acidic C-terminal tail and lysine sidechains on the N-terminal portion of the protein. These long-range contacts have been proposed previously, and are thought to protect the central hydrophobic domain of α-synuclein from intermolecular interactions that lead to aggregation.[61] Binding of Al^{3+} to the acidic domain prevents these transient salt-bridges, effectively increasing the availability of the lysine sidechains towards 18C6 attachment.

Copper(II) ions also accelerate fibril formation of α-synuclein.[11] Cu^{2+} was found in previous work to interact with the N-terminal domain of α-synuclein.[25]

Specifically, there is evidence that Cu^{2+} binds to His50 and at least one other nitrogen ligand, provided either by the amide backbone or by the N-terminus. α-synuclein also contains multiple non-specific Cu^{2+} binding sites with low affinity located on the C-terminal domain.[25,63]

ESI-MS of protein, 18C6, and 100 μM Cu^{2+} in water produces peaks corresponding to mixed complexes of α-synuclein, 18C6 and Cu^{2+}. The corresponding mass spectrum is shown in Figure 6.4c. Here, the interactions between α-synuclein and metal are retained in the gas phase. α-synuclein + Cu^{2+} is the dominant copper adduct in the mass spectrum, consistent with a single high affinity binding site. The SNAPP distributions for the [α-synuclein + $1Cu^{2+}$] ion are shown in white bars in Figure 6.5. Comparison with the distributions from α-synuclein in water reveals that protein complexation with Cu^{2+} changes the 18C6 adduct distributions slightly, but these differences are within error for most peaks, suggesting that the α-synuclein conformations in both solutions are structurally similar. More information is revealed by examination of the shapes of the 18C6 distributions. For the +13 and +17 charge states, the 18C6 distributions for α-synuclein + $1Cu^{2+}$ are less intense when compared to α-synuclein. The +16 charge state shows little change, while distributions for the +18 charge state increase compared to the distributions for α-synuclein. Lower

concentrations of copper produced SNAPP distributions with no significant deviations from the SNAPP distributions obtained in water. These results suggest that 18C6 binding is unaffected by differences in ionic strength at these concentrations of metal ion. In addition, copper adducts are no longer observed at low μM concentrations.

The limited effect of copper on the SNAPP distributions indicates that copper does not significantly affect the lysine availability of α-synuclein. It is possible that copper causes a substantial structural rearrangement which leads to no net change in lysine availability and is therefore difficult to detect. This explanation is deemed unlikely due to the high sensitivity of SNAPP to structural change. The results are more consistent with a scenario where copper attachment does not play a significant role in altering the structure of monomeric α-synuclein. Since copper is known to facilitate fibril formation, this suggests that copper may serve to stabilize higher order aggregates which might otherwise dissociate back into monomeric form. This hypothesis is consistent with previous observations demonstrating that the hydrodynamic radius of the protein and the chemical shifts of the backbone amide groups do not change upon Cu^{2+} binding.[25,61] Additionally, no changes in secondary structure are observed in the far UV-circular dichroism spectrum.[64]

6.4 Conclusions

The interactions between Ca^{2+} and calmodulin were explored with SNAPP-MS. The SNAPP distributions revealed a significant conformational change due to Ca^{2+} binding, resulting in an increase in the overall intensity and number of 18C6 adducts attached to calmodulin. Inspection of the crystal structures for apo and holo calmodulin suggests that Lys94 and Lys21 are key residues in producing the observed increase in lysine availability. Importantly, structural changes are still detected in the SNAPP distributions even when the calcium adducts are lost during ESI, meaning that gas phase retention of the solution phase ligand is not required. In general, this example reveals that SNAPP-MS is an excellent method for probing ligand binding events which involve structural rearrangement of the host protein.

In addition, SNAPP-MS was used to examine the less established metal binding properties of α-synuclein. Our results demonstrate that α-synuclein undergoes significant structural changes in the presence of biologically relevant concentrations of Al^{3+}.[62] Interestingly, there is a charge state dependence on the magnitude of the structural change, suggesting that select conformations of α-synuclein are affected more substantially than others. These results give rise to

the hypothesis that aluminum-induced acceleration of fibril formation may proceed through only a limited set of conformations. Additionally, experiments with Cu^{2+} suggest that significant structural changes do not occur upon Cu^{2+} binding. These results are consistent with complementary spectroscopic studies showing that binding of Cu^{2+} does not perturb the conformational dynamics of α-synuclein. We propose that Cu^{2+} binding may only stabilize the structure of higher order aggregates, without significant impact on the monomer. However, Al^{3+} clearly causes a significant perturbation in the conformations adopted by monomeric α-synuclein, which may enable aggregation by essentially generating aggregate prone monomers.

---

[1] Babu, Y. S.; Bugg, C. E.; Cook, W. J. *J. Mol. Bio.* **1988**, *204*, 191-204.
[2] Klee, C. B.; Vanaman, T. C. *Prot. Chem.* **1982**, *35*, 213-321.
[3] Walsh, D. M.; Hartley, D. M.; Kusumoto, Y.; Fezoui, Y.; Condron, M. M.; Lomakin, A.; Benedek, G. B.; Selkoe, D. J.; Teplow, D. B. *J. Biol. Chem.* **1999**, *274*, 25945-25952.
[4] Conway, K. A.; Lee, S. J.; Rochet, J. C.; Ding, T. T.; Williamson, R. E.; Lansbury, P. T. *Proc. Natl. Acad. Sci.* **2000**, *97*, 571-576.
[5] Prusiner, S. B. *Proc. Natl. Acad. Sci.* **1998**, *95*, 13363-13383.
[6] Wood, S. J.; Wypych, J.; Stevenson, S.; Louis, J. C.; Citron, M.; Biere, A. L. *J. Biol. Chem.* **1999**, *274*, 19509-19512.
[7] El-Agnaf, O. M. A.; Jakes, R.; Curran, M. D.; Wallace, A. *FEBS Lett.* **1998**, *440*, 67-70.
[8] Krishnan, S.; Chi, E. Y.; Wood, S. J.; Kendrick, B. S.; Li, C.; Garzon-Rodriguez, W.; Wypych, J.; Randolph, T. W.; Narhi, L. O.; Biere, A. L.; Citron, M.; Carpenter, J. F. *Biochemistry* **2003**, *42*, 829-837.
[9] Adlard, P. A.; Bush, A. I. *J. Alzheimer's Disease* **2006**, *10*, 145-163.
[10] Wright, J. A.; Brown, D. R. *J. Neurosci. Res.* **2008**, *86*, 496-503.
[11] Uversky, V. N.; Li, J.; Fink, A. L. *J. Biol. Chem.* **2001**, *47*, 44284-44296.
[12] Bush, A. I.; Pettingell, W. H.; Multhaup, G.; Paradis, M.; Vonsattel, J.; Gusella, J. F.; Beyreuther, K.; Masters, C. L.; Tanzi, R. E. *Science* **1994**, *265*, 1464-1467.
[13] Dawson, T. M.; Dawson, V. L. *Science* **2003**, *302*, 819-822.
[14] For a comprehensive review, see: Shults, C. W. *Proc. Natl. Acad. Sci.* **2006**, *103*, 1661-1668.

[15] Lewy, F. H. in Handbuch der Neurologie ed. Lewandowsky, M. (Springer, Berlin), Vol. III, pp. 920-933.
[16] Gibb, W. R. G.; Esiri, M. M.; Lees, A. J. *Brain* **1987**, *110*, 1131-1153.
[17] Hansen, L.; Salmon, D.; Galasko, D.; Masliah, E.; Katzman, R.; DeTeresa, R.; Thal, L.; Pay, M. M.; Hofstetter, R.; Klauber, M.; Rice, V.; Butters, N.; Alford, M. *Neurology* **1990**, *40*, 1-8.
[18] Polymeropoulos, M. H.; Lavedan, C.; Leroy, E.; Ide, S. E.; Dehajia, 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.; Nussbaum, R. L. *Science* **1997**, *276*, 2045-2047.
[19] Kruger, R.; Kuhn, W.; Muller, T.; Woitalla, D.; Graeber, M.; Kosel, S.; Przuntek, H.; Epplen, J. T.; Schols, L.; Riess, O. *Nat. Genet.* **1998**, *18*, 106-108.
[20] George, J. M.; Jin, H.; Woods, W. S.; Clayton, D. F. *Neuron* **1995**, *15*, 361-372.
[21] Kholodilov, N. G.; Neystat, M.; Oo, T. F.; Lo, S. E.; Larsen, K. E.; Sulzer, D.; Burke, R. E. *J. Neurochem.* **1999,** *73*, 2586-2599.
[22] Lee, M.; Hyun, D.; Halliwell, B.; Jenner, P. *J. Neurochem.* **2001**, *76*, 998-1009.
[23] Rekas, A.; Adda, C. G.; Aquilina, J. A.; Barnham, K. J.; Sunde, M.; Galatis, D.; Williamson, N. A.; Masters, C. L.; Anders, R. F.; Robinson, C. V.; Cappai, R.; Carver, J. A. *J. Mol. Bio.* **2004**, *340*, 1167-1183.
[24] Eliezer, D.; Kutluay, E.; Bussell, Jr., R.; Browne, G. *J. Mol. Biol.* **2001**, *307*, 1061-1073.
[25] Rasia, R. M.; Bertoncini, C. W.; Marsh, D.; Hoyer, W.; Cherny, D.; Zweckstetter, M.; Griesinger, C.; Jovin, T. M.; Fernandez, C. O. *Proc. Natl. Acad. Sci.* **2005**, *102*, 4294-4299.
[26] Lee, J. C.; Langen, R.; Hummel, P. A.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci.* **2004**, *101*, 16466-16471.
[27] Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *46*, 64-71.
[28] For a review: Kaltashov, I. A.; Eyles, S. *Mass Spectrom. Rev.* **2002**, *21*, 37-71.
[29] Abzalimov, R. R.; Frimpong, A. K.; Kaltashov, I. A. *Int. J. Mass Spectrom.* **2006**, *253*, 207-216.
[30] Loo, J. A.; Loo, R. R. O.; Udseth, H. R.; Edmonds, C. G.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 101-105.
[31] Konermann, L.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1248-1254.
[32] Dobo, A.; Kaltashov, I. A. *Anal. Chem.* **2001**, *73*, 4763-4773.
[33] Kaltashov, I. A.; Mohimen, A. *Anal. Chem.* **2005**, *77*, 5370-5379.
[34] Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. *Proc. Natl. Acad. Sci.* **2000**, *97*, 5802-5806.
[35] Kruppa, G. H.; Schoeniger, J.; Young, M. M. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 155-162.
[36] Pearson, K. M.; Pannell, L. K.; Fales, H. M. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 149-159.
[37] Bennett, K. L.; Kussman, M.; Bjork, P.; Godzwon, M.; Mikkelsen, M.; Sorensen, P.; Roepstorff, P. *Prot. Sci.* **2000**, *9*, 1503-1518.
[38] Back, J. W.; de Long, L.; Muijsers, A. O.; de Koster, C. G. *J. Mol. Biol.* **2003**, *331*, 303-313.
[39] Miranker, A.; Robinson, C. V.; Radford, S. E.; Aplin, R. T.; Dobson, C. M. *Science* **1993**, *262*, 896-900.
[40] Wagner, D. S.; Anderegg, R. J. *Anal. Chem.* **1994**, *66*, 706-711.
[41] Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1993**, *115*, 6317-6321.
[42] Mandell, J. G.; Falick, A. M.; Komives, E. A. *Anal. Chem.* **1998**, *70*, 3987-3995.
[43] Nemirovskiy, O.; Giblin, D. E.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 711-718.

[44] Zhu, M. M.; Rempel, D. L.; Zhao, J.; Giblin, D. E.; Gross, M. L. *Biochemistry* **2003**, *42*, 15388-15397.
[45] Mendoza, V. L.; Vachet, R. W. *Anal. Chem.* **2008,** 80, (8), 2895-2904.
[46] Friess, S. D.; Zenobi, R. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 810-818.
[47] Friess, S. D.; Daniel, J. M.; Hartmann, R.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *219*, 269-281.
[48] Ly, T.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1209-1215.
[49] Liu, Z.; Cheng, S.; Gallie, D.R.; Julian, R. R. *Anal. Chem.* **2008**, *80*, 3846-3852.
[50] Ly, T.; Liu, Z.; Pujanauski, B. G.; Sarpong, R.; Julian, R. R. *Anal. Chem.* **2008**, *80*, 4807-5276.
[51] Julian, R. R.; Beauchamp, J. L. *Int. J. Mass Spectrom.* **2001**, *210*, 613-623.
[52] Izatt, R. M.; Pawlak, K.; Bradshaw, J. S. Bruening, R. L. *Chem. Rev.* **1991**, *91*, 1721-2085.
[53] Zhang, M.; Tanaka, T.; Ikura, M. *Nat. Struct. Biol.* **1995**, *2*, 758-767.
[54] Kuboniwa, H.; Tjandra, N.; Grzesiek, S.; Ren, H.; Klee, C. B.; Bax, A. *Nat. Struct. Biol.* **1995**, *2*, 768-776.
[55] Schumacher, M. A.; Crum, M.; Miller, M. C. *Structure* **2004**, *12*, 849-860.
[56] Hu, P. F.; Ye, Q. Z.; Loo, J. A. *Anal. Chem.* **1994**, *66*, 4190-4194.
[57] Lafitte, D.; Heck, A. J. R.; Hill, T. J.; Jumel, K.; Harding, S. E.; Derrick, P. J. *Eur. J. Biochem.* **1999**, *261*, 337-344.
[58] Linse, S.; Helmersson, A.; Forsén, S. *J. Biol. Chem.* **1991**, *266*, 8050-8054.
[59] Paik, S. R.; Shin, H. J.; Lee, J. H.; Chang, C. S.; Kim, J. *Biochem. J.* **1999**, *340*, 821-828.
[60] Paik, S. R.;Lee, J. H.; Kim, D. H.; Chang, C. S.; Kim, J. *Arch. Biochem. Biophys.* **1997**, *344*, 325-334.
[61] Binolfi, A.; Rasia, R. M.; Bertoncini, C. W.; Ceolin, M.; Zweckstetter, M.; Griesinger, C.; Jovin, T. M. Fernandez, C. O. *J. Am. Chem. Soc.* **2006**, *128*, 9893-9901.
[62] Corrigan, F. M.; Reynolds, G. P.; Ward, N. I. *BioMetals* **1993**, *6*, 149-154.
[63] Sheng, Y.; Zabrouskov, V.; Loo, J. A. Top-down Mass Spectrometry of Noncovalent Protein Complexes for Determining Binding Sites. Presented at Lake Arrowhead Ion Chemistry Conference, Lake Arrowhead, CA, 1/11/2008.
[64] Hoyer, W.; Cherny, D.; Subramaniam, V.; Jovin, T. M. *Biochemistry* **2004**, *43*, 16233-16242.

Chapter 7

SURVEYING UBIQUITIN STRUCTURE BY NONCOVALENT ATTACHMENT OF DISTANCE-CONSTRAINED BIS(CROWN) ETHERS

7.1 Introduction

Protein structure is one of the keys to understanding function or malfunction[1] in biological systems at the molecular level; therefore, methods which enable the rapid screening of protein structure are needed. Mass spectrometry (MS) has several inherent advantages over x-ray[2] or NMR[3] in terms of speed and sensitivity, which has led to the development of several tools[4-14] for examining protein structure. This statement requires clarification because the most relevant environment for studying protein structure is obviously in solution, yet MS is conducted exclusively in the gas phase. Despite this fact, cleverly designed experiments can utilize MS to track structurally dependent mass shifts that occur in solution yet are retained and observed in the gas phase. In other words, the mass spectrometer is used as a detector, not a reaction vessel. One of the most common techniques employs hydrogen/deuterium exchange to probe the chemical availability of backbone amide hydrogens.[5-7] The degree of

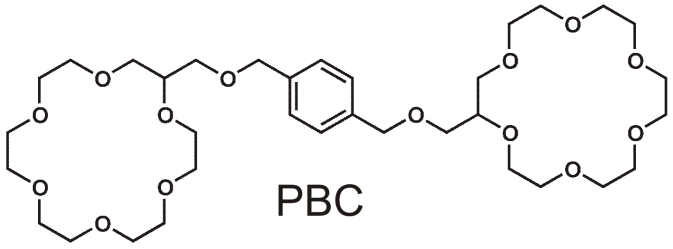
exchange is largely determined by intramolecular interactions (typically hydrogen bonds)[15] within the protein and can be easily quantified by the resulting mass shift for each exchange. In practical terms, faster exchange is usually observed for more exposed portions of the protein. Quenching of the exchange, followed by enzymatic digestion, and further MS analysis of the fragments can be used to localize fast or slow exchanging regions.[6] Three dimensional structures, such as those obtained by x-ray or NMR, are not revealed and neither are tertiary structural interactions.

This shortcoming can be partially addressed by the use of chemical crosslinking studies.[9-11] In these experiments, reactive amino acid side chains are covalently crosslinked to one another. The crosslinking reagents are chosen such that the maximum distance between two amino acids can be determined.[16] Following enzymatic digestion and analysis by MS, the residues which are crosslinked are revealed because of the mass shift caused by the crosslinking agent. In this manner, proximity relationships between different amino acids can be established and provide experimental constraints on computational models of protein structure.[17] These experiments usually only provide information about the proximity of side chains, and the backbone is not directly probed. Caution should be employed while interpreting the results, because factors other than

spatial proximity can influence crosslinking potential including chemical reactivity[18] (which may be a function of intramolecular entanglement) and protein structural dynamics.

A third method for examining protein structure, known as SNAPP (selective noncovalent adduct protein probing) has been developed recently.[12] SNAPP utilizes a selective interaction between 18-crown-6 ether (18C6) and lysine to examine protein structure.[19] 18C6 associates weakly with protonated primary amines (such as the side chain of lysine) in solution via the formation of three hydrogen bonds. The dissociation constant for the 18C6-protonated lysine complex has not been determined, and may be influenced by a variety of factors including local sequence or protein structure. Nevertheless, the binding constant should be similar in magnitude to protonated butylamine, for which thermodynamic data is available. Protonated butylamine binds to 18C6 with a dissociation constant of ~110 mM in water.[20] Therefore, attachment of 18C6 to lysine depends on the degree to which the side chain is occupied by intramolecular interactions such as salt bridges or hydrogen bonds.[21] These potentially interfering intramolecular interactions prevent attachment of 18C6 and are inseparably connected to the structure of the protein, making attachment of 18C6 (or lack thereof) a sensitive probe of protein structure.12 The number of

18C6s that attach to a protein can be easily determined with MS because the 18C6/lysine interaction becomes strong in the gas phase (ΔH ~-150 kJ/mol)[19] and causes an easily detectable mass shift. A typical SNAPP experiment is conducted by electrospraying a solution containing the protein and 18C6 directly into the mass spectrometer. The resulting mass spectrum contains an intensity distribution of protein-18C6 complexes. If the ensemble of conformations is heterogeneous, then the observed 18C6 distributions represent statistical averages of the entire ensemble. Assuming that distinct conformations have different propensities for protonation, then the 18C6 distributions for different structures are resolvable by differences in charge state, and therefore different dynamic states of the protein can be simultaneously detected. Site specific information can be obtained by determining which lysine residues are capable of attaching 18C6.[21] The three dimensional relationship between different lysine side chains, however, is not examined.



The present work extends the utility of SNAPP with the implementation of the bis(crown) probe shown above (PBC). The covalent attachment of two 18C6

moieties[22] introduces an additional distance constraint factor into SNAPP experiments. Molecular modeling suggests that the maximum distance between the centers of the two crown ethers is ~19 Å. Results with ubiquitin demonstrate that PBC is a viable SNAPP reagent. Comparison with SNAPP experiments using 18C6 can be used to easily identify the degree of bidentate binding to a protein. The A-state of ubiquitin contains one abundant PBC adduct which is attached by a bidentate interaction. Determination of the precise binding sites for PBC is more challenging. Collision induced dissociation (CID) experiments can be used to narrow down the possible sites and can also be used to confirm the degree of bidentate binding. Site-directed mutagenesis was used to pinpoint the preferred bidentate binding location to residues Lys29 and Lys33. These results suggest that PBC will be useful for SNAPP experiments to probe changes in protein structure under equilibrium conditions in a general manner; however, the further extraction of site specific information with the noncovalent crosslinking approach may be challenging.

7.2 Materials and Methods

7.2.1 General Protocols and Instrumentation

Unless otherwise noted, all commercial reagents and solvents were used as received. Bovine ubiquitin, oligonucleotide primers, HPLC-grade methanol and

HPLC-grade acetonitrile were all purchased from Sigma-Aldrich (St Louis, MO). Water was purified to 18.2 MΩ resistivity using a Millipore Direct-Q (Millipore, Billerica, MA). *Escherichia coli* BL21 (DE3) cells were purchased from Novagen Inc. (Madison, WI). T4 DNA ligase, pGEM T-vector, *BamH1* and *Nde1* were purchased from New England Biolabs (Beverly, MA); Wizard SV Gel and PCR Clean-up System and Plasmid Miniprep Kit were purchased from Promega (Madison, WI). Gene Jet Plasmid Miniprep Kit was purchased from Fermentas Life Science (Glen Bernie, MA).

All air or moisture sensitive reactions were conducted in oven-dried glassware under an atmosphere of nitrogen using dry, deoxygenated solvents. Tetrahydrofuran (THF) and diethyl ether (Et_2O) were distilled under nitrogen from sodium benzophenone ketyl. 18-crown-6-methanol was purchased from Alfa Aesar and α,α'-dibromo-*p*-xylene was purchased from Aldrich; both were used without further purification. ^1H NMR spectra were recorded on a Bruker AV-500 (at 500 MHz) in chloroform-*d* at 23 ˚C. Chemical shifts were referenced to the residual chloroform-H peak, which was set at 7.26 ppm for ^1H. Data for ^1H NMR are reported as follows: chemical shifts (δ ppm), multiplicity, (s = singlet, d = doublet, t = triplet, q = quartet, dd = doublet of doublet, dt = doublet of triplet, m = multiplet, br = broad resonance), coupling constants (Hz) and integration. IR

spectra were recorded on a Nicolet MAGNA-IR 850 spectrometer and are reported in frequency of absorption (cm^{-1}). All mass spectra were obtained using an LTQ linear ion trap mass spectrometer (Thermo Fisher, Waltham, MA) with a standard electrospray ionization source.

7.2.2 Organic Synthesis

To a stirred solution of 18-crown-6-methanol (28.3 mg, 0.096 mmol), α,α'-dibromo-*p*-xylene (12.7 mg, 0.048 mmol) in THF (500 μL) was added NaH (4.3 mg, 60% dispersion in mineral oil, 0.106 mmol) with stirring. Stirring was continued at room temperature for 24 hours. Diethyl ether (2 mL) was added, and the mixture was filtered through celite, and the solvent was removed by evaporation under reduced pressure to yield 41 mg of crude product (product + NaCl). The crude product was resuspended in water and purified using reverse-phase high performance liquid chromatography (HPLC) on a 150 mm x 30 mm, 10 μm particle size C18 Gemini column (Phenomenex, Torrance, CA). The elution gradient was 0% - 90% acetonitrile over 30 minutes, using water and acetonitrile as co-solvents. The observed retention time for PBC was ~13.9 min (~42% acetonitrile). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.31 (s, 4H), 4.55 (s, 4H), 3.6-3.9 (m, 50H). FTIR (thin film) 2923.1, 2853.4, 1923.0, 1437.9, 1228.4. MS (m/z): $[M+Na]^+$ calc'd for $[C_{34}H_{59}O_{14}Na]^+$, 713.37; found, 713.48.

7.2.3 Selective Noncovalent Adduct Protein Probing Mass Spectrometry (SNAPP-MS)

Solutions were prepared either with 50/50 water/methanol + 1% acetic acid to generate the A-state, or with water alone to sample the native structure. To ensure stoichiometric excess of adduct, solutions were prepared such that the number of crown moieties was twice the number of possible binding sites. For 18C6, the concentrations used were 10 μM protein:160 μM 18C6. For PBC, which contains two crowns, the molar ratio was 10 μM protein:80 μM 18C6. Ions were generated by direct infusion into the electrospray inlet of the LTQ linear ion trap mass spectrometer. Source and ion lens parameters (electrospray voltage, ion transfer capillary temperature, tube lens voltage) were optimized for gentle ionization and maximal adduct intensities and are similar to parameters used previously.[12] Once determined, these instrument parameters remained constant for a given analyte. Ions of interest were then isolated in the trap and subjected to a resonance excitation RF voltage to produce fragmentation after many collisions with He buffer gas. Multiple stages of CID (MS^n) can be achieved in the trap by re-isolating and exciting subsequent dissociation products. MS^n experiments were used to locate the noncovalent attachment of PBC.

7.2.4 Ab initio Calculations

Semi-empirical calculations (PM3)[23] were performed as implemented in Gaussian 03 Version 6.1 Revision D.01. To determine the maximum binding distance for PBC, a ternary complex of PBC and two ammonium adducts was minimized in the gas phase. Coulombic repulsion between the two ammoniums should provide the driving force to maximize separation between the two crown ether centers without significant distortion of PBC. The *Arabidopsis* mutants were prepared as described in detail elsewhere.[21]

7.3 Results and Discussion

Ubiquitin is an important signaling protein that has been well characterized by a variety of mass spectrometric and other techniques.[24] The results for SNAPP experiments with ubiquitin using PBC and 18C6 as the probe reagents are shown in Figure 7.1. The structure of Ubiquitin is quite stable,[25] but can be denatured in the presence of acid and organic co-solvent to yield the "A-state".[25,26] In Figure 7.1a, the natively folded structure is represented in the +6 and +7 charge states. The A-state structure is partially denatured and is observed in higher charges states, such as +9 and +10 shown in Figure 7.1b.[12,27] Comparison of the SNAPP distributions for the two structures reveals that both PBC and 18C6 attachment is structure dependent. This conclusion is drawn from the fact that both the total number of crowns attaching and the most abundant number of adducts shifts to

a higher number of adducts for the A-state with PBC and 18C6. Therefore, both reagents can be used for SNAPP experiments to evaluate changes in protein structure, but is there additional information provided by the distance constraint component of PBC?

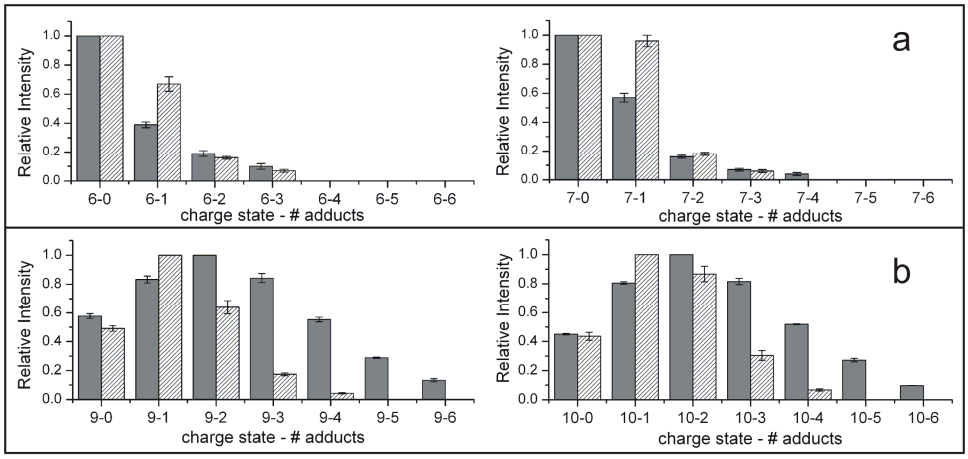


Figure 7.1 Comparisons of 18C6 (solid) and PBC (hatched) distributions for ubiquitin sampled from native conditions (a) and A-state conditions (b). The average numbers of adducts observed (18C6/PBC) by charge state are +6) 0.64/0.63, +7) 0.69/0.68, +9) 2.34/1.27, and +10) 2.23/1.46. The difference between the average number of 18C6 and PBC suggests a bidentate interaction is present in the A-state.

Careful comparison of the 18C6 and PBC distributions reveals that some of the PBC adducts are attached in a bidentate fashion. If two lysines are within ~19Å (the maximum separation as determined by modeling the distance between two ammonium ions attached to PBC at the PM3 semi-empirical level), then one PBC can attach to both residues. If two lysine residues are not within this distance, then two PBC adducts can attach, each in a monodentate fashion. A single bidentate attachment is expected to be more favorable than two

monodentate attachments, due to the lower entropic costs of forming the bidentate crown-lysine complex. Additionally, monodentate attachment of PBC may be somewhat less favorable than attachment of 18C6 due to scavenging interactions between the unoccupied crown of PBC with free ions in solution. The unoccupied crown of PBC acts as a "handle" for competitive processes leading to dissociation of the PBC-protein complex. These scavenging interactions may become more common in the late stages of ionization, i.e. evaporating droplets, where concentrations of free ions are higher than in bulk solution. In contrast, 18C6 can only interact with a single lysine; therefore, comparison of the two adduct distributions should reveal information about the degree of bidentate binding because each bidentate PBC will be replaced by two 18C6s while each monodentate PBC will be matched by a single 18C6. It is not anticipated that there will be an exact quantitative correlation; however, the net result should be a reduction in the number of PBC adducts by approximately one relative to 18C6 for each bidentate PBC.

Returning to the data in Figure 7.1, the average number of adducts can be calculated for each charge state to determine the extent of bidentate attachment for PBC. The average numbers by charge state for 18C6/PBC are 6+, 0.64/0.63; 7+, 0.69/0.68; 9+, 2.34/1.27; 10+, 2.23/1.46. For the folded structure, there is almost no

change in the number of adducts that attach between PBC and 18C6 (although the shape of the distributions do vary). This suggests that there are no lysine-lysine pairs that are available to bind 18C6 and held within ~19Å for the native state structure. Chemical cross-linking studies have found lysine-lysine pairs within this distance.[11,28] There are several experimental differences which may account for this discrepancy. First, noncovalent attachment of 18C6 may be more sensitive to interfering intramolecular interactions than covalent reactions.[29] Second, SNAPP experiments sample equilibrium conditions, whereas covalent crosslinking is not reversible and may trap transient states or exaggerate the importance of minor conformations. In either case, a more dramatic shift occurs in the A-state, suggesting that this structure exhibits a significantly higher degree of bidentate binding. The difference between the average number of 18C6 and PBC adducts in the A-state suggests that one PBC is probably interacting with ubiquitin via bidentate binding. More detailed information, such as the specific binding site for the bidentate PBC, cannot be extracted from the SNAPP distribution alone and requires further experiments.

Collision induced dissociation (CID) experiments were performed to evaluate potential binding sites. Fortuitously, protonated ubiquitin ions that carry a high net charge (≥ +10 charges) have a facile cleavage point at Pro19

(which has been identified in previous work)[30] that facilitates fragmentation of the protein. The results for fragmenting $[Ubi+PBC+11H]^{11+}$ are shown in Figure 7.2a. The base peak results from simple loss of the PBC adduct, as would be expected given the noncovalent attachment. Indeed, CID of noncovalent complexes of PBC and ubiquitin at lower charge states, and of PBC and other proteins, such as cytochrome c, results in loss of PBC as the exclusive dissociative pathway. Surprisingly, several fragments are produced in substantial abundance with PBC still attached. In the comparable experiment with 18C6, no fragments with 18C6 still attached are detected, suggesting that bidentate attachment is required for PBC to be retained following fragmentation of the protein. Following re-isolation and fragmentation of the $[y_{58}+PBC+9H]^{9+}$ fragment, loss of PBC again yields the base peak as shown in Figure 7.2b. However, there is a small amount of a secondary fragment (y_{58}-y_{24}) produced with PBC still attached. In Figure 7.2c, an MS^4 experiment is conducted to verify the assignment of this peak. Indeed, loss of PBC from this fragment yields the base peak, confirming that PBC was attached to the precursor. These results suggest that the possible locations for bidentate binding are most likely between residues 20-52. Unfortunately, there are four lysine residues in this region of the protein, and

further experiments (which are described below) are necessary to pinpoint the location. Fragmentation of higher charge states produced similar product ions.

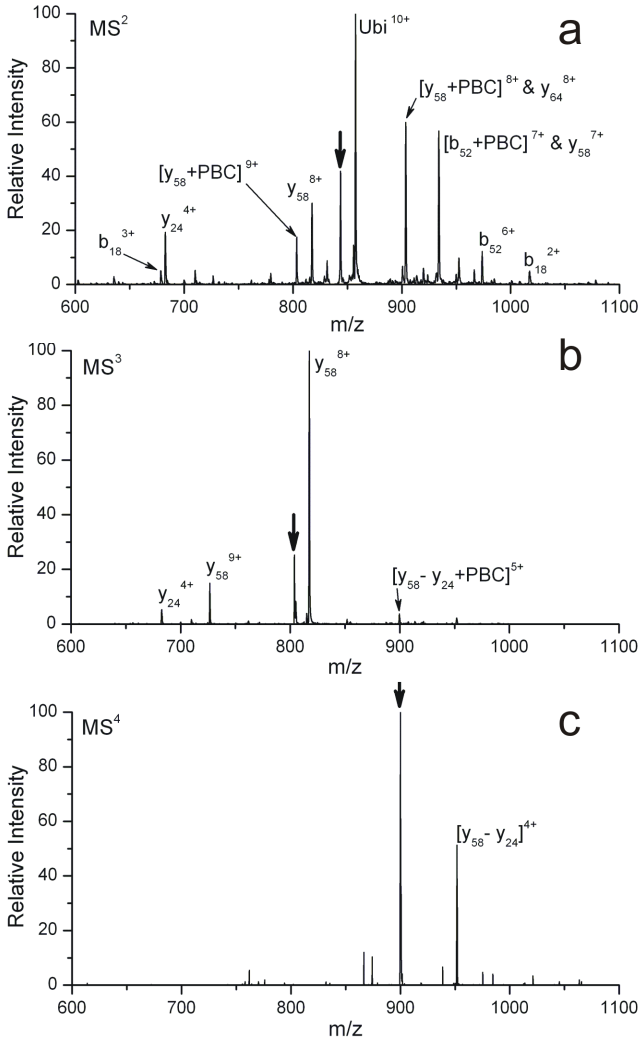


Figure 7.2 CID spectra of (a) [Ubiquitin + PBC + 11H]$^{11+}$, (b) PBC-retaining fragment product [y_{58} + 9H + PBC]$^{9+}$, and (c) [y_{58} - y_{27} + PBC + 5H]$^{5+}$. Observation of [y_{58}-y_{27}+PBC+5H]$^{5+}$ localizes the bidentate attachment to residues 20-52. However more specific information could not be obtained by CID alone, due to competitive noncovalent loss of PBC from [y_{58}-y_{27}+PBC+5H]$^{5+}$. Bold down arrows indicate peaks being subjected to collisional excitation.

A comparison of the adduct distributions and the CID experiments discussed above reveals that there is a single lysine-lysine pair bridged by a

bidentate PBC. To confirm that there is only one bidentate PBC attachment, we performed CID experiments on complexes where the number of PBC molecules attached is greater than one, as shown in Figure 7.3. Figure 7.3a shows the MS^2 of the ternary complex of ubiquitin and two PBC molecules, which yielded primarily loss of one or both PBC adducts. Nevertheless, a very small amount of the y_{58} fragment with two PBC adducts attached was produced. This assignment was confirmed with MS^3, which generated several peaks corresponding to the loss of PBC as shown in Figure 7.3b. No additional fragmentation with retention of both adducts was observed. It is known that the A-state of ubiquitin is dynamic in nature, and it is possible that the second bidentate PBC attaches to one of the less abundant conformations or only attaches to the protein transiently due to a less favorable interaction. In either case, the amount of the second bidentate adduct appears to be minimal from these results and analysis of Figure 7.1. The PBC triple adduct was subjected to CID as shown in Figure 7.3c. In this case, only loss of PBC is observed, confirming that some of the PBC adducts are most likely attached to the protein by a monodentate interaction. Furthermore, this result demonstrates that it is not possible for any arbitrary PBC adduct to adopt a bidentate state in the gas phase during the CID process.

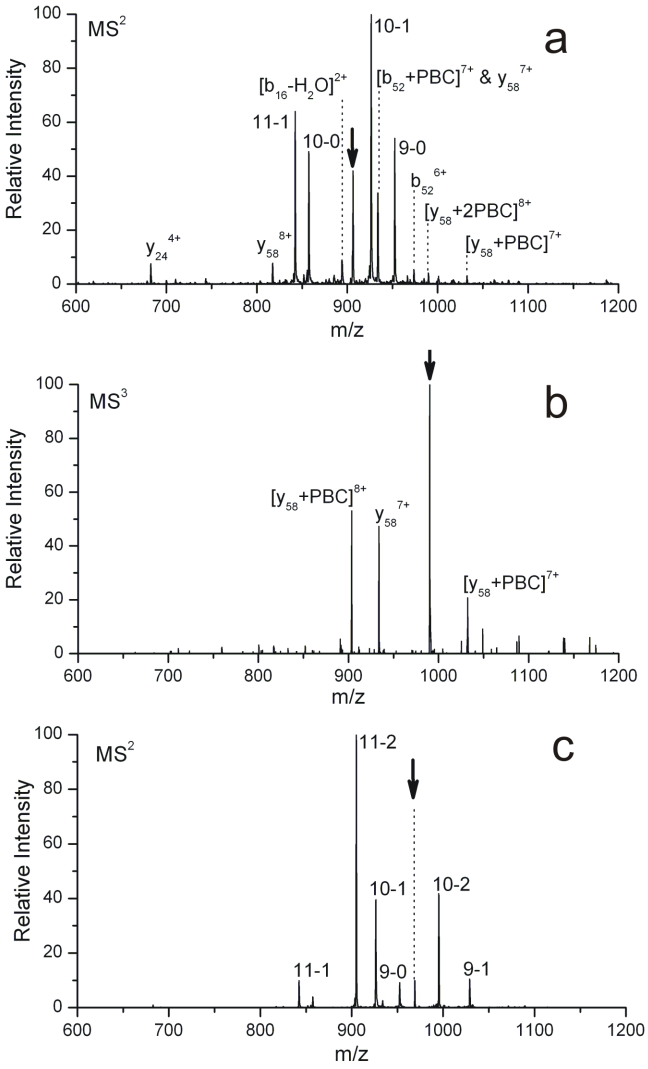


Figure 7.3 (a) CID of [Ubiquitin + 11H + 2PBC]$^{11+}$ produces [y_{58} + 8H + 2PBC]$^{8+}$ in low abundance, suggesting that the second bidentate attachment is more weakly bound than the first. (b) Confirmation of the assignment is provided by CID of the PBC-retaining product [y_{58} + 8H + 2PBC]$^{8+}$, which results in loss of noncovalently bound PBC without backbone fragmentation. (c) CID of [Ubiquitin + 11H + 3PBC]$^{11+}$ and other higher order Ubiquitin-PBC complexes (data not shown) exclusively produces noncovalent losses of PBC and [PBC+H]$^{+}$. Bold down arrows indicate peaks being subjected to collisional excitation.

In order to further pinpoint the primary bidentate binding site, experiments were performed with a series of ubiquitin mutants from *Arabidopsis thaliana*. *Arabidopsis* ubiquitin yields nearly identical results to bovine ubiquitin for every structural probing technique that has been applied to both proteins.[21] Furthermore, the overall backbone structure is highly conserved within the ubiquitin family.[31] In fact, the ubiquitin-related protein, RUB1, has 29 amino acid substitutions (as compared to bovine), yet yields a structure with an essentially identical backbone.[32] The primary sequences of ubiquitin derived from bovine and *Arabidopsis* differ by only three residues (S19P, D24E, and A57S), indicating that results from both proteins should be highly comparable. In the *Arabidopsis* mutants, each lysine is mutated to asparagine one at a time. This enables the contribution of each lysine as a binding site for 18C6 or PBC to be evaluated independently. The results are summarized in Figure 7.4a, where the magnitude of the peak is proportional to the degree of binding at a particular lysine. The results suggest that K29 and K33 are the most likely sites for PBC binding. Since bidentate binding is predicted to be strongly preferred over monodentate binding, these sites are also the most likely pair for the bidentate PBC. This conclusion is in good agreement with the results obtained by CID in Figure 7.3. Furthermore, predictions based on NMR data for the A-state structure identified

organized secondary structural regions, but an absence of stable tertiary structure.[25] The cartoon structure shown in Figure 7.4b represents the secondary structural elements that were predicted by NMR with an absence of any tertiary structure.[25] Lys29 and Lys33 are the same side of the central α-helix. The predicted distance between these residues is in the 10-15Å range, depending on the relative orientation of the side chains. Therefore, a sustainable bidentate interaction is possible between these residues in the absence of tertiary structure. The absence of a strong bidentate PBC bridging any two of the secondary structure elements in Figure 7.4b agrees with the NMR prediction of a highly dynamic structure lacking stable tertiary features. It is interesting to note that Lys29 and Lys33 are on the same helix in the folded structure as well; however, results with 18C6 indicate that these residues are intramolecularly entangled. Analysis of the crystal structure suggests that the sequence remote residues 52 and 14 are the interaction sites. This information also suggests that the A-state lacks any stable tertiary structure.

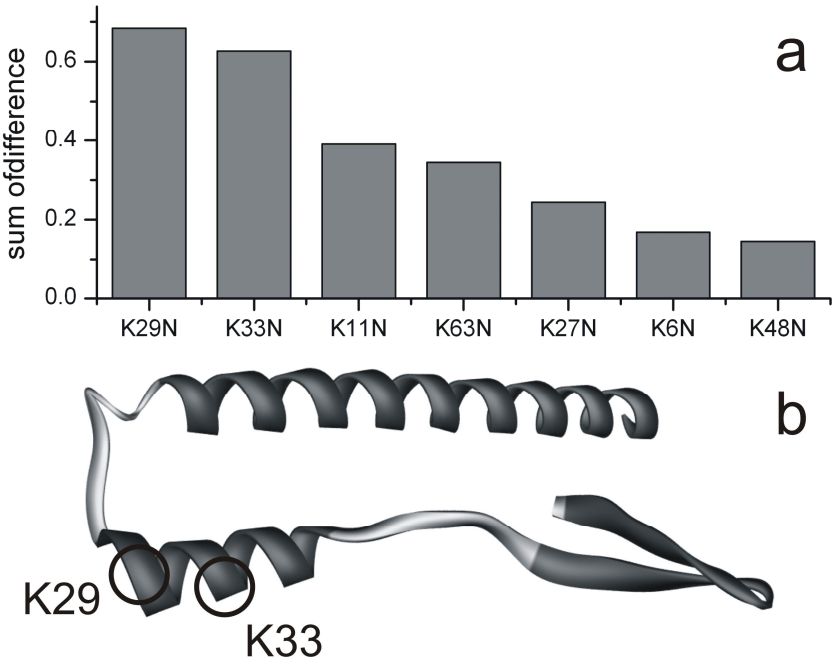


Figure 7.4 (a) Data from seven single residue mutants of *Arabidopsis* ubiquitin. Differences in PBC distributions for the 10+ charge state between wild-type and each mutant were summed. Mutations that affected the PBC distribution most, K29N and K33N, are the putative sites for PBC bidentate attachment. (b) Ribbon cartoon of the A-state of ubiquitin depicting a loose tertiary structure composed of two α helices and a β sheet. Random coil and highly flexible regions are shaded light gray, while distinct secondary structures are shaded dark gray. A single PBC was to found bridge K29 and K33 (circled), which are on the same α helix.

7.4 Conclusions

We have studied whether noncovalent attachment of a novel bis(crown) ether molecule is specific towards lysine pairs within a distance imposed by a semi-rigid linker. This approach is similar to chemical cross-linking in that it potentially provides distance constraints between residues that are spatially close, information which is useful in constructing theoretical models. However, by using noncovalent interactions, this approach avoids kinetic traps that can

potentially occur during covalent modification and provides complementary information. Our SNAPP data for the A-state of ubiquitin suggests a single lysine-lysine pair within ~19 Å. Data from CID experiments and *Arabidopsis* ubiquitin mutants confirm the bidentate attachment and localize the pair to K29 and K33, in agreement with the NMR studies of the A-state.[25] These results also demonstrate that some noncovalent interactions in solution are strong enough in the gas phase to be competitive with protein backbone fragmentation using conventional CID. However, CID experiments may only provide a rough estimate of where the noncovalent adduct is attached, and is limited by the number of facile cleavage sites on a protein. Nevertheless, PBC is shown to be a useful SNAPP reagent which can rapidly and easily probe protein structure. The noncovalent approach for probing distance constraints may find utility in systems where dynamics are important, or where the required pH may prove problematic for traditional covalent crosslinking studies. Both issues are likely important for studying the A-state of ubiquitin.

---

[1] Chiti, F.; Dobson, C. M. *Ann. Rev. Biochem.* **2006**, *75*, 333-366.
[2] Matthews, B. W. *Annu. Rev. Phys. Chem.* **1976**, *27*, 493-523.
[3] Dyson, H. J.; Wright, P. E. *Chem. Rev.* **2004**, *104*, 3607-3622.
[4] Loo, J. A.; Loo, R. R. O.; Udseth, H. R.; Edmonds, C. G.; Smith, R. D. *Rapid Comm. Mass. Spectrom.* **1991**, *5*, 101-105.
[5] Katta, V.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 214-217.
[6] Zhang, Z. Q.; Smith, D. L. *Prot. Sci.* **1993**, *2*, 522-531.
[7] Wales, T. E.; Engen, J. R. *Mass Spectrom. Rev.* **2006**, *25*, 158-170.

[8] Kaltashov, I. A.; Eyles, S. J. *Mass Spectrom. Rev.* **2002**, *21*, 37-71.
[9] Sinz, A. *Mass. Spectrom. Rev.* **2006**, *25*, 663-682.
[10] Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W. Dollinger, G. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5802-5806.
[11] Kruppa, G. H.; Schoeniger, J.; Young, M. M. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 155-162.
[12] Ly, T.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1209-1215.
[13] Friess, S. D.; Daniel, J. M.; Zenobi, R. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2664-2675.
[14] Friess, S. D.; Daniel, J. M.; Hartmann, R.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *219*, 269-281.
[15] Englander, S. W. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 213-238.
[16] Green, N. S.; Reisler, E.; Houk, K. N. *Protein Sci.* **2001**, *10*, 1293-1304.
[17] van Dijk, A. D. J.; Boelens, R.; Bonvin, A. M. J. J. *FEBS J.* **2005**, *272*, 293-312.
[18] Swaim, C. L.; Smith, J. B.; Smith, D. L. *J. Am. Soc. Mass. Spectrom.* **2004**, *15*, 736-749.
[19] Julian, R. R.; Beauchamp, J. L. *Int. J. Mass Spectrom.* **2001**, *210*, 613-623.
[20] Izatt, R. M.; Terry, R. E.; Haymore, B. L.; Hansen, L. D.; Dalley, N. K.; Avondet, A. G.; Christensen, J. J. *J. Am. Chem. Soc.* **1976**, *98*, 7620-7626.
[21] Liu, Z.; Cheng, S.; Gallie, D. R.; Julian, R. R. Accepted in *Anal. Chem.* for publication.
[22] Julian, R. R.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L. *Angew. Chem. Int. Ed.* **2003**, *42*, 1012-1015.
[23] Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209-220.
[24] Hershko, A.; Ciechanover, A. *Annu. Rev. Biochem.* **1998**, *67*, 425-479.
[25] Brutscher, B.; Bruschweiler, R.; Ernst, R. R. *Biochemistry* **1997**, *36*, 13043-13053.
[26] Hoerner, J. K.; Xiao, H.; Kaltashov, I. A. *Biochemistry* **2005**, *44*, 11286-11294.
[27] Pan, J. X.; Wilson, D. J.; Konermann, L. *Biochemistry* **2005**, *44*, 8627-8633.
[28] Novak, P.; Young, M. M.; Schoeniger, J. S.; Kruppa, G. H. *Eur. J. Mass Spectrom.* **2003**, *9*, 623-631.
[29] For example, compare ref 21 and: Novak, P.; Kruppa, G. H.; Young, M. M.; Schoeniger, J. *J. Mass Spectrom.* **2004**, *39*, 322-328.
[30] Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Anal. Chem.* **2001**, *73*, 3274-3281.
[31] Walters, K. J.; Goh, A. M.; Wang, Q. H.; Wagner, G.; Howley, P. M. *Biochim. Biophys. Acta* **2004**, *1695*, 73-87.
[32] Rao-Naik, C.; de la Cruz, W.; Laplaza, J. M.; Tan, S.; Callis, J.; Fisher, A. J. *J. Biol. Chem.* **1998**, *273*, 34976-34982.

Chapter 8

ELUCIDATING THE TERTIARY STRUCTURE OF PROTEIN IONS *IN VACUO* WITH SITE SPECIFIC PHOTO-INITIATED RADICAL REACTIONS

8.1 Introduction

Elucidation of protein structure is a challenging task in any setting and the gas phase, which is the ultimate low dielectric environment, is no exception. Upon desolvation of a protein, Coulombic interactions transition from playing a rather minor part to being a dominant factor controlling structure. As a result, the overall structure of a protein in the gas phase is highly dependent on charge state, with repulsive interactions consistently leading to greater unfolding for higher charge states. This effect has been demonstrated for numerous proteins in experiments utilizing ion mobility, which can easily measure the collision cross section or overall size of a protein.[1-6] Nevertheless, more detailed information about protein structure, for example at the residue level, is very difficult to obtain. Ion molecule reactions, such as gas phase H/D exchange, can yield information about exchangeability in distinct regions of a protein,[7-9] but the translation of such information into structure is typically very difficult. Various dissociation methods have also been explored as means for obtaining

information about protein structure, although most have not succeeded.[10,11] Some success is afforded with electron capture dissociation (ECD), which has the potential to reveal hydrogen-bonding networks.[12-14] In these experiments, it is assumed that facile dissociation of the N-C_α bond occurs without rupturing noncovalent interactions;[15] therefore, regions where ECD fragments are observed as separated products likely do not have significant tertiary contacts to the remaining protein. Unfortunately, further examination of portions of the protein with significant tertiary contact is not possible.

Molecular modeling has also been used to evaluate gas phase protein structure, frequently in conjunction with experiment.[16-18] Due to the constraints imposed by the size of proteins and even large peptides, the level of theory for modeling is typically restricted to molecular mechanics. Modeling has been largely successful in evaluating the structures of peptides, where conformational space is somewhat restricted, when used in conjunction with ion mobility.[19,20] In proteins, however, the larger molecular size leads to an exceedingly vast number of potential structures, many of which have the same nominal cross sections and calculated energies. Given that collisional cross section is the only parameter experimentally measured, distinguishing which (or how many) of these conformers are correct is frequently not possible. Despite these limitations, ion

mobility presently offers the best data for comparison with theory since incorporation of H/D exchange or ECD data into molecular mechanics simulations is not straightforward. Therefore, there remains a significant need for experimental methods that can reveal residue level information that can be unambiguously leveraged to obtain protein structure directly, or in combination with simulations.

Of all gas phase proteins, ubiquitin is one of the most often examined. Ion mobility measurements of protonated ubiquitin have shown that charge states greater than +8 have collisional cross sections that are consistent with elongated conformers whereas the +4 charge state has a collisional cross section that is compact and consistent with the crystal structure.[1] Intermediate charge states are mixtures of compact, partially folded, and elongated structures. These results suggest that for specific charge states, i.e. the +4 through the +6 charge states of ubiquitin, the condensed phase structure may be retained in the gas phase. H/D exchange results are consistent with the ion mobility distributions, indicating that low charge states are heterogeneous in structure whereas higher charge states are nearly homogeneous.[7] ECD of the +5 charge state yields no separated backbone fragments, suggesting that the structure is fully folded with significant tertiary contacts throughout.[12] ECD of the +6 charge state yields fragmentation

between residues 1 through 14, and 74 through 76, suggesting that tertiary contact in these regions has been reduced. Although considerable information about the structure of ubiquitin has been acquired from all of these experiments, the actual gas phase structures for the various charge states remain unknown.

In the present manuscript, we report a new method for examining protein structure in the gas phase. This method relies on site-specific incorporation of a photolabile radical precursor at a tyrosine residue, followed by radical initiation, migration, and dissociation.[21] It is found that radical migration is highly sensitive to protein structure, and consequently, spatial proximity to the initial radical site dictates where radical directed dissociation (RDD) will occur. Experiments with ubiquitin reveal that radical migration to sites that are distant in sequence from Tyr59 (> 5 residues away) decreases significantly upon disruption of the tertiary structure. Moreover, radical transfers from Tyr59 to specific residues are highly structurally dependent. The charge-induced unfolding of ubiquitin proceeds by disruption of contacts between Tyr59 and the N-terminal portion of the protein and a significant rearrangement of the C-terminal half. Radical transfer between Tyr59 and residues that are distant in sequence indicate the presence of specific, through-space contacts. These experimentally derived contacts provide multiple constraints for both generating and validating structures with molecular

mechanics. The structures identified by this methodology are stable towards 10 ns unconstrained molecular dynamics simulations, low in energy, and consistent with both the RDD data and cross sections obtained by ion mobility.

8.2 Materials and Methods

8.2.1 Materials

Bovine ubiquitin was purchased from Sigma Aldrich (St. Louis, MO). HPLC-grade acetonitrile, sodium iodide, chloramine-T hydrate, sodium metabisulfite and glacial acetic acid were purchased from Thermo Fisher Scientific (Waltham, MA). Water was purified to 18.2 MΩ resistivity using a Millipore Direct-Q (Millipore, Billerica, MA). Dialysis membranes and clips (MWCO = 3500) were purchased from Spectra Por (Rancho Dominguez, CA).

8.2.2 Chloramine-T Iodination of Proteins

Milligram quantities of protein are iodinated using sodium iodide as the iodine source, chloramine-T as the oxidant and sodium metabisulfite as the quenching reagent.[21] Iodinated proteins are dialyzed against water and stored frozen in 50 nmol aliquots.

8.2.3 Electrospray Mass Spectrometry and Photodissociation

Solutions containing 10 μM of protein in 50/50 H_2O/acetonitrile + 0.5% acid are infused into the standard electrospray source of a Thermo LTQ linear

quadrupole ion trap mass spectrometer. The back plate of the instrument is modified with a quartz window to transmit fourth-harmonic (266 nm) pulses from a flashlamp-pumped Nd:YAG laser (MiniLite, Continuum, Santa Clara, CA). Laser pulses are synchronized to a typical MS^2 activation step by enabling a diagnostic trigger in the LTQ Tune software, which delivers a 3.3V differential signal, which is converted to be a TTL pulse by a digital delay generator and sent to the laser at the beginning of each MS^2 activation period. Ions are isolated using an isolation width of 3 m/z units for photodissociation (PD). PD products (i.e. loss of iodine atom) are re-isolated using an isolation width of 10 m/z units and subjected to collision induced dissociation (CID). At a minimum, manual charge state assignment is attempted for all peaks >5% relative abundance using the zoom and ultrazoom scan modes of the LTQ, which can typically resolve the isotopic envelope of up to a +8 ion. Charge states for many peaks below the 5% threshold were resolvable and assigned. Peaks in experimental spectra are assigned to backbone fragments with the aid of Fragmentor software (v. 1.0.0.3), which is available on-line at http://faculty.ucr.edu/ryanj/fragmentor.html. We adopted the standard nomenclature for peptide fragments and added the superscripts "*" and "°" to indicate NH_3 loss and H_2O loss, respectively (e.g., b^*_{58}

≡ b_{58} - NH3). Assignments do not distinguish between closed shell vs. radical (e.g., -H·) fragments.

8.2.4 Molecular Modeling

The Maestro and MacroModel computing suite (Schrodinger Inc., San Diego, CA) was used to build models of protonated ubiquitin and perform all molecular modeling simulations. The OPLS atomic force fields is used for all calculations.[22,23] The initial "seed" structure is derived from the atomic coordinates determined by x-ray crystallography (1UBI).[24] All acidic residues are simulated as the charge neutral species. It is likely that the choice of protonation sites significantly affects the molecular modeling results. Therefore, we have referred to previous investigations on the protonation locations in ubiquitin.[12,25] For the +4 charge state of ubiquitin, two of the charges reside on Arg42 and Arg54. The third charge is sequestered by one of the C-terminal arginines (Arg72 and Arg74). The last charge is flexible, and can be located on any lysine residue. Taking into consideration these probabilities, we have chosen Arg42, Arg54, Arg74 and Lys6. For the +6 charge state, we have chosen Lys6, Lys11, Lys33, Arg42, Arg54, and Arg74 as the charged sites.

Tertiary contact information obtained from RDD is implemented as 6 ± 1.5 Å distance constraints with force constants of 100 $kJ \cdot mol^{-1} \cdot Å^2$. The structure is

subjected to 120 rounds of simulated annealing cycles that begin with a heating step to 2500 K (10 ps stochastic dynamics simulation), followed by several cooling steps (10 ps each) to a final bath temperature of 50 K. Additional simulated annealing experiments are performed until the last 120 rounds resulted in conformations with similar calculated energies. This process was repeated at incrementally lower initial temperatures (2000 K, 1500 K, 750 K, 300 K) to obtain lower energy structures.

8.2.5 Calculated Cross Sections

Cross sections are calculated using the exact hard spheres scattering approximation as implemented in the MOBCAL program, which was developed and currently maintained by the Jarrold laboratory at Indiana University.[26,27]

8.3 Results and Discussion

A photolabile carbon-iodine bond is incorporated into Tyr59 of ubiquitin using well-established iodination chemistry.[21] Irradiation of iodoubiquitin with a single pulse of 266 nm photons dissociates the carbon-iodine bond, as shown in Figure 7.1a, to yield a radical localized to a single carbon center on the side chain of tyrosine. The photodissociation yield for the +4 and +5 charge states of ubiquitin are unusually low (~12-17%) compared to higher charge states of ubiquitin, which are typically around ~30%. The CID spectrum for the radical

photoproduct [(Ubiquitin)· + 5H]$^{5+}$ is shown in Figure 8.1b. The key features of this spectrum are the a, c, and z fragment ions, which are absent in the CID spectrum of the even-electron protein.[28] As observed previously for the +6 charge state,[21] the most intense backbone fragmentation (e.g. c_{54}, a_{59}, a_{54}, c_{56}) is due to radical-directed dissociation (RDD) and occurs nearby in sequence to the initial radical site, Tyr59.

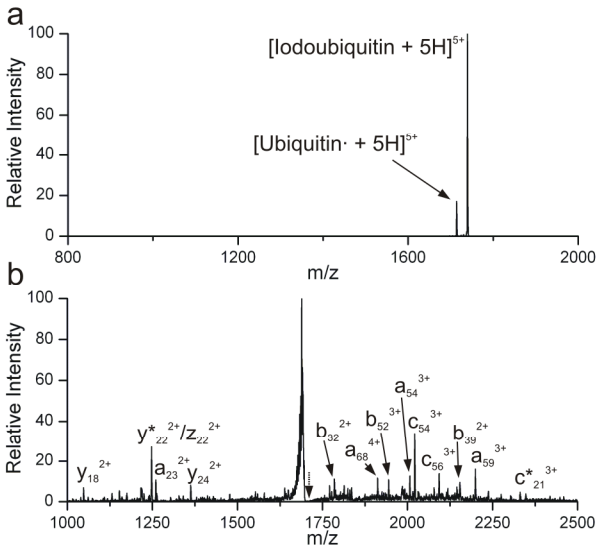


Figure 8.1 a) Photodissociation of the +5 charge state of iodoubiquitin cleaves the iodine atom, yielding a protein radical ion. b) Collision induced dissociation of the radical product from (a), yields prominent radical-directed backbone fragment ions of the a, c, and z varieties. Dotted arrow indicates precursor m/z.

The relationship between amino acid sequence and backbone fragmentation is better illustrated by plotting the RDD backbone fragmentation intensities by residue number. Figure 8.2 shows the RDD fragmentation plots for charge states

+4 through +10. Arginine and tyrosine side chain loss intensities are also included in Figure 8.2 at the residues from which these side chain losses were determined to originate. Although other side chain losses were observed, they are not incorporated into Figure 8.2 because residue specific information was not obtainable. RDD of the +4 charge state (Figure 8.2, top) results in fragmentation that is well-distributed across the ubiquitin sequence. The most prominent fragmentations are cleavage of the backbone bonds connecting Arg42 and Leu43, Asp21 and Thr22, and Val17 and Glu18. Interestingly, backbone fragmentation at these sites, and at other residues distant in sequence from Tyr59, decreases significantly with increasing charge state. In the +6 and +7 charge states prominent arginine side chain loss is observed from one of the C-terminal arginines (Arg72 or Arg74, vide infra) and backbone fragmentation at Tyr59 becomes abundant. It is clear that RDD fragmentation becomes increasingly localized to residues close in sequence to Tyr59 in the higher charge states. In addition, tyrosine side chain loss appears in the +8 charge state and becomes the nearly exclusive fragmentation product by the +10 charge state. These observations are all consistent with the idea that structure strongly influences RDD as will be detailed below.

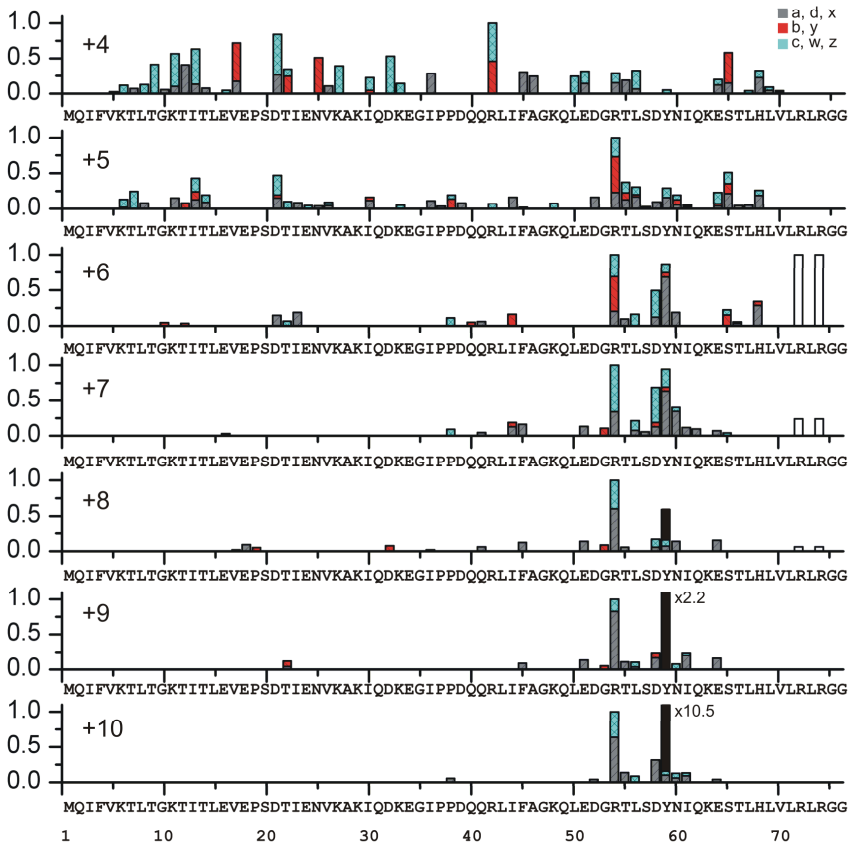


Figure 8.2 Plots of radical-directed fragmentation as a function of sequence for charge states +4 through +10. Arginine and tyrosine side chain loss intensities are shown as white and black bars, respectively. For each charge state, the sum of RDD backbone fragmentation at each residue and side chain loss intensities are normalized to the highest sum.

Close inspection of Figure 8.2 reveals that the RDD fragmentation for the +4 charge state is significantly different from the +5. In contrast, the most abundant fragments in the CID spectra of the even-electron +4 and +5 charge states are identical, suggesting that RDD is sensitive even to modest variations in protein structure which do not influence CID.[11] Changes in RDD fragmentation are most

easily explained by differences in radical migration. In previous work, we have demonstrated that radical migration occurs easily if the reaction is downhill and if the radical donor and acceptor can attain the proper orientation relative to one another.[29] In proteins, both of these factors are functions of three dimensional structure. It is observed generally in Figure 8.2 that sequence remote radical migration is facilitated in compact structures adopted by low charge states, and attenuated by the extended structures found for higher charge states. A quantitative treatment for these observations is provided in Figure 8.3, where the sum of all backbone fragmentation located more than five residues away from Tyr59 is shown as a function of charge state (blue squares). The sums are normalized to the +4 charge state. Distal fragmentation drops steeply from +4 to +7 and then decreases at a lower rate from +7 to +10. These two trends are consistent with the average collisional cross sections measured by ion mobility, which are shown as red triangles.[1] The degree of sequence remote radical migration in ubiquitin is therefore clearly correlated with overall protein size.

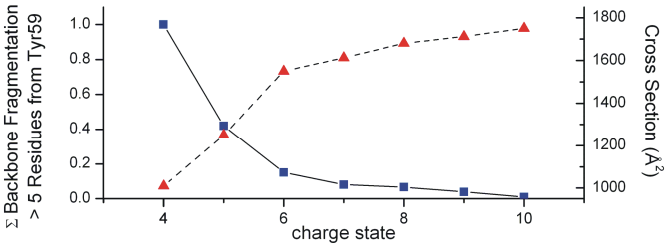


Figure 8.3 Comparison of backbone fragmentation distant from Tyr59 (> 5 residues) summed for each charge state (squares) and the average collisional cross section from ion mobility data (triangles).[1] Fragmentation at distal sites decreases significantly with increasing charge. The inverse relationship with the average collisional cross sections indicates that radical migration to sequence remote residues diminishes as the protein adopts more elongated structures.

8.3.1 Radical migration through-space and through-sequence

Radical transfer proceeds either through-space or through-sequence until the radical reaches a destination where backbone fragmentation or side chain loss is favorable. In proteins, these locations are abundant and likely limit the maximum number of consecutive radical transfers. The theoretical limit can be estimated by calculating the probability that a random hydrogen atom abstraction reaction will occur at a site that will induce fragmentation. Sites that induce fragmentation have been identified previously by examining the mechanisms of backbone fragmentation and side chain losses observed in peptides.[29,30] In ubiquitin, the number of carbon-bound hydrogen atoms capable of being abstracted by the initial radical based on reaction enthalpy is 470. Upon abstraction, 285 of the subsequent carbon-centered radicals are expected from

previous experiments to dissociate into observable products such as backbone fragmentation and side chain loss.[30] Thus, in the absence of structural effects, there is a ~60% probability that the first transfer places the radical in a site susceptible to fragmentation and the probability increases to ~94% by three radical transfers. This simple model does not take into account several important points. First, the number of possible exothermic hydrogen atom abstractions decreases significantly after the first transfer.[30] Moreover, only the first radical transfer from the tyrosyl radical to an α or β carbon is predicted by theory to be potentially barrierless.[29] Consistent with this prediction are ion-molecule experiments performed on peptides that demonstrate radical migration is fast (< ms) and occurs without additional collisional activation.[31] Barriers to subsequent transfers (such as between the α and β carbons) are typically moderate in the absence of structural constraints. In reality, structural effects cannot be ignored and will likely limit migration to donor/acceptor pairs which are in close contact. Therefore, the statistically predicted maximum of three radical transfers is probably an overestimate. Migration to sequence remote residues, for example to Thr66, Thr22, and Thr14 in the +4 and +5 charge states, most likely occurs via through-space migration because an improbably high number of radical transfers would be required otherwise.

8.3.2 Tertiary Contacts Direct Radical Fragmentation of Distal Side chains

Examination of side chain losses reveals additional structural information. Side chain losses in the low charge states (+4 and +5) appear facile and result in peaks corresponding to loss of multiple side chains, shown in the trailing edge of the precursor m/z in Figure 8.1a. In contrast, an intense fragment corresponding to loss of 87 Da is observed for the +6 charge state, as shown in Figure 8.4a. This charged loss occurs from the side chain of arginine. Further tandem MS experiments reveal the location of the arginine from which side chain loss occurs. CID of $[\text{Ubiquitin} + 5\text{H} - 86]^{5+}$, shown in Figure 8.4b, yields $[y_{24} + 2\text{H} - 86]^{2+}$, which limits the side chain loss to three arginines: Arg54, Arg72 and Arg74. Further MS on the y_{24} fragment (inset) yields a pseudo b_6 fragment ion containing the residues Gly53 through Asp57, which eliminates Arg54 as a candidate. Thus, Arg72 and Arg74 are the only remaining candidates. These results indicate that there is significant interaction between Tyr59 and the C-terminus in the +6 charge state.

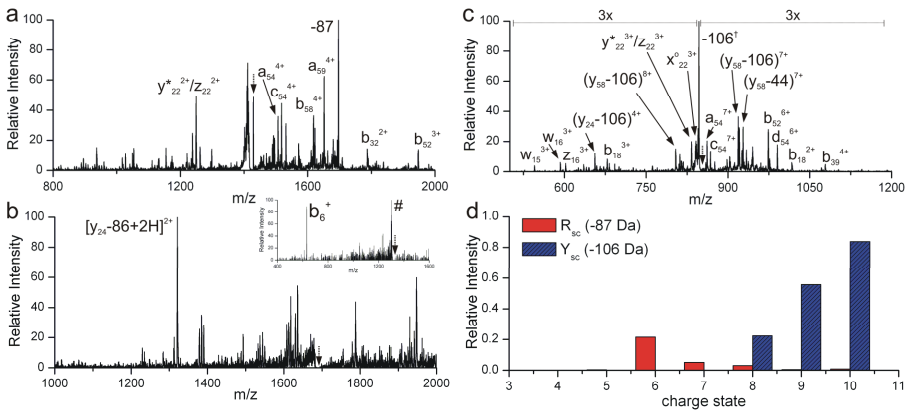


Figure 8.4 a) CID of [(Ubiquitin)· + 6H]$^{6+}$ yields loss of protonated arginine side chain (-87) as the most intense fragment ion. Further MS/MS experiments locate the arginine residue from which the -87 loss originates. b) CID of the 87 loss yields an 86 Da-shifted y_{24} fragment ion, which narrows the loss to Arg54, Arg72, and Arg74. CID of the y_{24} fragment in (b) yields a prominent pseudo b_6 ion containing residues Gly53 through Asp57, which eliminates Arg54 as a candidate (inset). Thus, the side chain loss originates from either Arg72 or Arg74. c) CID of [(Ubiquitin)· +10H]$^{10+}$ yields loss of tyrosine side chain as most abundant fragment ion. d) Plot of side chain losses from tyrosine and arginine as a function of charge state. # neutral loss of H_2O from the molecular ion. † loss of tyrosine sidechain. Dotted arrows indicate precursor m/z.

A tertiary contact between Tyr59 and Arg74 is consistent with ECD studies of the +6 charge state of ubiquitin.[12] Interestingly, separated products due to cleavage between residues 14 and 74 are absent, indicating the existence of tertiary structure in this region. This tertiary interaction disappears with additional charges, as indicated by increased ECD fragmentation in the C-terminal region and a substantial decrease in arginine side chain loss in the RDD. Arginine side chain loss occurs in competition with many other processes. As shown in Figure 8.4d, the highest fractional yield of arginine side chain loss for

any charge state is 20% (+6 charge state). Tyrosine side chain loss, on the other hand, can be much more abundant as will be explained presently.

Fragmentation in high charge states is marked by abundant loss of tyrosine side chain. For example, the most abundant RDD peak for $[(\text{Ubiquitin})\cdot + 10\text{H}]^{10+\cdot}$ is tyrosine side chain loss (-106), as shown in Figure 8.4c. We have previously demonstrated that the 106 loss from tyrosine is due to a strained 1,3 hydrogen atom transfer from the hydroxyl to the *ortho* radical on tyrosine.[29] This strained migration is likely disfavored when alternate exit channels exist. Abundant loss of 106 therefore suggests that tyrosine is isolated in the +10 charge state, and that secondary and tertiary contacts with the remaining protein are minimal. Consistent with this interpretation is the increase of tyrosine side chain loss with increasing charge, as shown in Figure 8.4d. Here, the fractional abundances of tyrosine and protonated arginine side chain loss are shown as a function of charge state. The y-axis represents the fraction of the specific side chain loss relative to the sum of all side chain losses. Although loss of tyrosine side chain is the most abundant fragment in the +10 charge state, retention of backbone fragmentation at Thr55 and Asp58, i.e. the presence of the a_{54} and a_{58} fragment ions in Figure 8.4c, indicate that there remains a low probability for radical transfer from the tyrosine side chain to nearby residues.

8.3.3 Resolution of Conformational Classes

Closer examination of intense RDD fragments as a function of charge state reveals additional trends. Figure 8.5 plots the intensities of key backbone and side chain fragments versus charge state, as extracted from the data in Figure 8.2. It can be observed that the fragmentation intensities at Ile13/Thr14 and Asp21/Thr22 are highest in the +4 charge state and decrease along similar trajectories from +4 to +6, as shown in Figure 8.5a. Backbone fragmentation at these two locations disappears completely above the +7 charge state. These results indicate that low charge states are populated by conformations that favor radical migration to the N-terminal half of ubiquitin. Moreover, the unfolding event that mitigates migration to Thr14 also prevents migration to Thr22. Fragmentation at Ser65/Thr66 also decreases with increasing charge state, but persists through the +7 charge state. Together these results suggest that protonation disrupts tertiary contacts to the N-terminal half of ubiquitin at the same time as the C-terminal half is undergoing structural reorganization. Indeed, the appearance of arginine side chain loss and backbone fragmentation at Asp58, Tyr59, and Asn60 in the +6 and +7 charge states suggests that the C-terminal half has restructured to favor radical migration to these sites. Backbone

fragmentation intensities at Asp58, Tyr59 and Asn60 together follow the same trend, with maxima at the +7 charge state.

In contrast, the intensity of arginine side chain loss, shown in Figure 8.5c does not correlate with these backbone fragments or with any other fragmentation pathway. The lack of correlation suggests that a significant population of the +6 ions are structurally unique and do not appear in other charge states. In contrast, fragmentation at Arg54/Thr55 exhibits a unique charge state dependence which spans across several charge states, as shown in Figure 8.5c. Facile radical migration to Thr55 precludes secondary structures that contain an extended backbone or prevent direct contact between Thr55 and Tyr59, such as the β sheet structural motif. These structures would mitigate through-space and through-sequence transfers to Thr55. It is also unlikely that the gas phase secondary structure mimics the 3_{10} helix found in the condensed phase in this region of the protein, which places the side chain of Tyr59 on the opposite face of the helix as Thr55.[24] Instead, facile radical migration between residues separated by four amino acids in sequence is consistent with an α helix or a random coil. Our results therefore suggest that the secondary structure in this region is distorted from the solution structure and persistent through a number of different tertiary motifs.

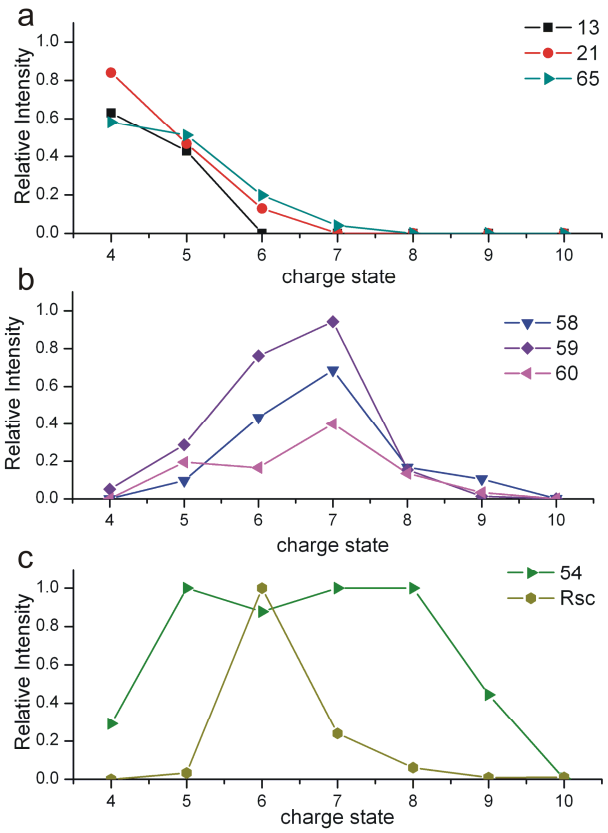


Figure 8.5 Selected backbone and side chain fragmentation products as a function of charge state (a-c). Fragmentation at residues 13, 21, and 65 dominate at lower charge states (compact conformers), and disappear by the +8 charge state (a). Partially unfolded conformers are characterized by enhanced fragmentation nearby Tyr59 (b), and fragmentation at residues 54 and 72/74 (c, see text).

8.3.4 Application to Conformational Searching

Molecular mechanics simulations have been used extensively in conjunction with gas phase techniques to study the structure of proteins.[16-18] Experimentally determined proximities between Tyr59 and other residues revealed by RDD can be implemented in molecular mechanics simulations as distance constraints.

Theory predicts that abstraction of the C_α hydrogen atom from glycine by tyrosyl radical requires that the radical donating carbon atom be within 2.7 Å of the C_α of glycine.[29] Therefore, what is likely measured in RDD experiments is the frequency of short-range contact, which is related to the equilibrium distance separating the two residues and the dynamic motion of the protein.

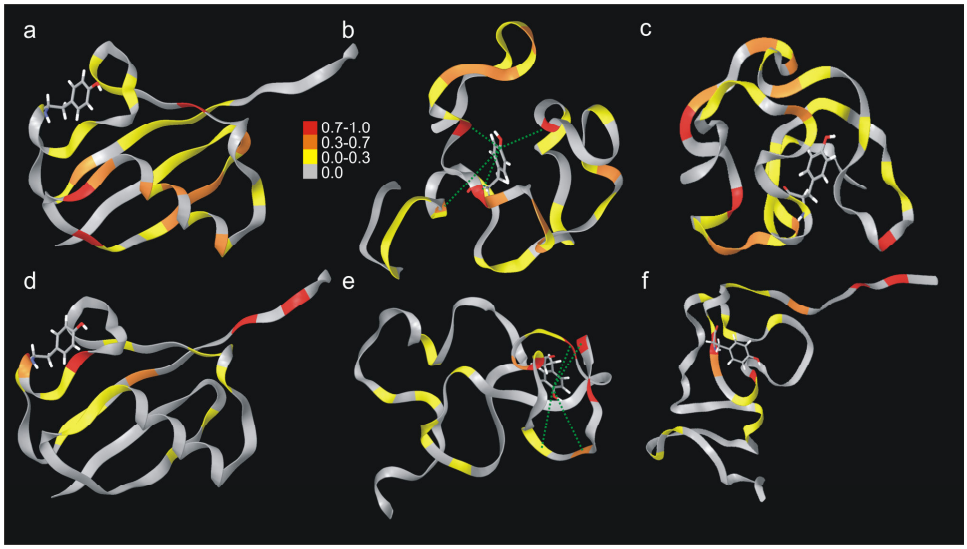


Figure 8.6 a) The crystal structure of ubiquitin with color coding indicating the intensity of RDD backbone fragmentation for the +4 charge state. b) Structure of the +4 charge state of ubiquitin calculated by MDSA using distance constraints derived from experimental data, which are shown as green dotted lines. The coloring is the same as in (a). c) MDSA output without distance constraints. d) The crystal structure of ubiquitin colored by RDD fragment intensities of the +6 charge state. Arg72 and Arg74 are highlighted in red, due to the dominant arginine side chain loss observed. e, f) The output structures of a constrained and unconstrained MDSA calculation, respectively, with identical coloring as (d).

The +4 charge state adopts compact conformations exclusively, as determined by ion mobility. The cross sectional distribution, which has maxima at 1004 and 1059 $Å^2$, is similar to the theoretical cross section calculated for the x-

ray crystal structure (1059 Å^2). Figure 8.6a shows the crystal structure of ubiquitin colored by RDD fragmentation intensity. The Cz on Tyr59 is within 6 Å of the β carbons of Leu50, Glu51, Arg54, and Asp58. Limited RDD fragmentation is observed at these residues, and the most intense RDD fragments are at locations that are distant through-space from Tyr59. Numerous consecutive radical transfers would be required to explain the most intense RDD fragmentation, which as discussed above, are unlikely to occur. The incongruity between the gas phase data and the crystal structure suggests that the +4 charge state is populated by conformers that are structurally distinct from those found in the condensed phase. A molecular dynamics/simulated annealing (MDSA) approach was used to find structures that are consistent with the RDD results. Figure 8.6b shows the output of an MDSA calculation of the +4 charge state of ubiquitin. The distances between Tyr59 and four residues where the most intense RDD fragments occur (Arg42, Thr22, Val17, and Thr66) were initially constrained to 6 Å ± 1.5 Å (see Experimental section). These constraints are represented in Figure 8.6b as green dotted lines. Molecular mechanics (MM) energies and key inter-residue distances for calculated structures are summarized in Table 8.1. The best-fit structure was stable towards a 10 ns stochastic dynamics simulation without distance constraints. Structures were sampled every 10 ps, yielding 1000

structures in total. The low RMSD (1.3 Å) indicates that removal of the distance constraints does not destabilize the best-fit structure.

The secondary and tertiary structural elements that characterize condensed phase ubiquitin are absent in the best-fit structure and the backbone is significantly distorted relative to the crystal structure. Consistent with these qualitative observations is a calculated root mean square deviation (RMSD) of 13.5 Å between the structures shown in Figures 8.6a and b. Interestingly, the N- and C-termini interact in a pincer-like fashion in Figure 8.6b, which is consistent with structural data from ECD.[12] Closer inspection of the best-fit structure indicates that solvation of the charged sites is a dominant factor in this *in vacuo* protein structure. Although the pronounced effect of charge-charge repulsive interactions on structure is well-documented, the combined MDSA and RDD results indicate that favorable electrostatic interactions influence the organization of protein structure at lower charge states significantly. As shown in Table 8.1, the calculated MM energy for the best-fit structure is lower than the crystal structure by 625 kJ/mol. The difference in energy is due to a stabilization of electrostatic interactions (-531 kJ/mol) in the best-fit structure and a modest stabilization of the van der Waals interactions (-100 kJ/mol). The decrease in electrostatic energy is due in part to the significant rearrangement of the

backbone structure surrounding the charged residues in order to better solvate the positive charges. Indeed, the local structure surrounding charged residues are loop-regions, where backbone carbonyls and polar side chains are oriented toward the charges.

Table 8.1 A Comparison between the Crystal and Calculated Structures

Structure	Cross Section[a] Å^2	Energy[b] kJ·mol^{-1}	X_{corr}	Distance from Tyr59 (Å)			
				Val17	Thr22	Arg42	Thr66
+4, Crystal Structure	1059	-11165	0.30	13.1	9.3	12.5	15.1
+4, Constrained[d]	1075	-11792	0.47	7.2	6.4	6.6	7.5
+4, Unconstrained[d]	1067	-11837	0.30	10.2	19.6	14.1	18.2
				Thr55	Thr66	His68	Arg74
+6, Crystal Structure	1059	-10457	0.67	8.9	15.1	13.9	22.4
+6, Constrained[d]	1129	-11359	0.87	6.8	8.2	6.4	7.9
+6, Unconstrained[d]	1294	-11101	0.64	5.5	9.9	14.8	35.2
+6, Rearranged[e]	1053	-10982	0.78	11.7	13.7	12.5	3.8

[a] Exact hard sphere approximation
[b] Gradient-minimized MM energies calculated using the OPLS force field
[c] 1.0 indicates perfect correlation; 0.23 is score for a random structure (see text for details)
[d] MDSA [e] MD

Despite significant differences in structure and calculated MM energies between Figures 8.6a and b, the two conformations would be difficult to distinguish using ion mobility. For example, the calculated cross sections for the best-fit and crystal structures are 1059 Å^2 and 1075 Å^2, respectively. The 1.5% relative difference is within the trial-to-trial error reported previously for ubiquitin cross sections.[1,4,6] Moreover, an MDSA approach without guidance

from residue-resolution experimental constraints would have a significantly smaller chance of successfully finding the best-fit structure. Indeed, the MDSA experiment repeated without distance constraints resulted in a structurally distinct conformer (13 Å RMSD relative to Figure 8.6a) that has a calculated collisional cross section of 1067 $Å^2$. The color-coded calculated structure, shown in Figure 8.6c, indicates that structure is inconsistent with the most intense RDD fragmentation pathways. However, the structure is consistent with numerous less intense RDD fragmentations, including at residues Asp32, Ile36 and Glu64. The distances between Tyr59 and these residues are 5.4 Å, 4.4 Å, and 4.0 Å, respectively. Interestingly, the distances to these residues are much greater in the best-fit structure (12.3 Å, 13.3 Å, and 14.0 Å).

A quantitative model is useful to evaluate which structures are most consistent with the observed RDD fragmentation. The following scoring algorithm (equation 1) was developed to quantify the quality of the fit between the calculated structures and the RDD fragmentation data,

$$\frac{\sum_{i=1}^{76} x_i \cdot I_{RDD}}{\sum_{i=1}^{76} I_{RDD}} = X_{corr} \qquad (1)$$

where x_i is 1 if the i^{th} residue is within 10 Å of Tyr59 in the calculated structure and 0 if not, I_{RDD} is the relative RDD fragmentation intensity at residue *i*, and X_{corr} is the correlation score. Thus, X_{corr} of 1.0 is obtained when all residues where RDD fragmentation occurs are within 10 Å of Tyr59, indicating perfect correlation between the calculated structure and the RDD data. The score for a 'random structure' of ubiquitin was also calculated. Here, Tyr59 was randomly assigned to be near 10 – 25 residues and correlated with the RDD data. The average score of 1 x 10^6 trials is 0.225, with a standard deviation of 0.092. Therefore, only calculated structures with scores greater than ~ 0.32 are significantly better fits than a random structure. The scores for all calculated structures are listed in Table 8.1. The scores for the crystal structure and the unconstrained structures (both 0.30), indicate that these structures are poorly correlated with the RDD fragmentation. The score for the constrained structure is significantly higher than the crystal structure (0.47), but captures less than half of the RDD fragmentations. The low score is not surprising considering the wide sequence distribution of RDD fragmentation, which would be difficult to explain with a single structure. Instead, these results are consistent with heterogeneous tertiary interactions surrounding Tyr59, and imply that multiple structures exist as compact conformers in the gas phase.

According to ion mobility, the +6 charge state adopts a broad distribution of compact and partially unfolded conformations, with slight maxima at 1041 $Å^2$ and 1220 $Å^2$, respectively.[1] Figure 8.6d shows the crystal structure with the RDD data from the +6 charge state superimposed. The most striking difference between the RDD fragmentation and the crystal structure is the large separation between the C-terminus and Tyr59 (Tyr59 to Arg74 distance is 21.5 Å). Interestingly, this difference would easily be resolved if the C-terminal tail pivoted towards Tyr59, which would also require disruption of the cross β-sheet structure between the C-terminal tail and residues 42 through 45. This 'rearranged' structure was modeled computationally and subjected to a 20 ns molecular dynamics simulation. The MM energy for the final, minimized structure is -10982 kJ/mol, which is comparable to the energy calculated for alternative structures (vide infra). Tyr59 and Arg74 are in close proximity (3.8 Å inter-residue distance); however, Thr55, Thr66, and His68 are separated from Tyr59 by 11.7 Å, 13.7 Å and 12.5 Å, respectively. Importantly, the fragmentation at Thr66 and His68 is not easily explained by either the crystal structure, or simple rearrangement of the C-terminal tail.

Figure 8.6e shows the result of the MDSA calculation on the +6 charge state with distance constraints imposed between Tyr59 and Arg74, Thr55, Thr66 and

His68. Arg74 was chosen as the protonation site for the MDSA calculation; however, Arg72 is an equally probable choice.[9] To test whether this choice significantly affects the modeling results, we moved the charge from Arg74 to Arg72 in the output structure and subjected it to a 10 ns MD simulation. The RMSD between the final structure obtained by MD and the structure in Figure 8.6e is 2.4 Å, indicating that small shifts in protonation location (i.e. Arg74 → Arg72) do not significantly affect the stability of the calculated structure.

As observed for the +4 charge state, the +6 structure is significantly distorted relative to the crystal structure. The calculated cross section for the 'best-fit' structure is 1129 $Å^2$. The calculated cross section falls within the distribution measured by ion mobility.[1] Figure 8.6f shows the output of an unconstrained MDSA search. Comparison of the N-terminal domains in Figures 8.6e and f reveal that both are relatively compact and consist of loop regions that facilitate solvation of positive charges by the backbone. However, the specific backbone carbonyls that participate in charge solvation are different and indicate that there is more than one stable permutation of backbone-charge interactions. Inter-conversion between the two structures may require substantial activation energy to shift charge solvation sites, but the relative energies of the conformations (at least at the MM level) suggest that both might exist. Indeed the more open

conformation may represent the partially unfolded structure observed in ion mobility experiments; however, the RDD data suggests that this is not the dominant conformer observed in an ion trap. It is straightforward to envision combined ion mobility/radical migration experiments that would enable examination of such possibilities.

Evaluation of the 'fitting' scores for the calculated structures for the +6 charge state in Table 8.1 indicates that the constrained structure is most correlated with RDD fragmentation (score: 0.86). All other calculated structures score lower and have higher MM energies. Unlike the +4 charge state where multiple conformers are required to explain the RDD fragmentation results, the high correlation between the RDD fragmentation and the constrained structure suggests that tertiary interactions near Tyr59 are nearly homogeneous for the conformer(s) adopted in the +6 charge state.

8.4 Conclusions and Outlook

We have shown that radical migration is a sensitive probe of the gas phase structure of proteins. Different conformations of ubiquitin were sampled in the gas phase by examining different charge states. Low charge states are populated by numerous conformers that contain significant secondary and tertiary structure which facilitates through-space radical migration from Tyr59 to

sequence remote sites. In contrast, radical migration is significantly attenuated in higher charge states, which are known to have extended backbone structures. More specific information is obtained by examining where radical-directed fragmentation occurs, which reveals short-range contacts between Tyr59 and the site of dissociation. Residue-level structural information obtained by radical migration can be used as structural constraints which aid conformational search algorithms. These initial results suggest that radical migration is a promising novel method to examine the gas phase structure of proteins with residue resolution.

Distance-dependent transfer has been used with remarkable success to study protein structure in the condensed phase. Quantitative models of transfer of excitation energy and electrons have been developed and are used in conjunction with experimental methods to examine protein structural dynamics with high resolution. Similarly, the precision of the RDD technique would be significantly improved if a quantitative model of hydrogen atom transfer is determined, which connects fragmentation intensities and rates to the equilibrium distance between two residues. Such a model would permit elucidation of gaseous protein structure with an unprecedented level of detail.

[1] Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 954-961.
[2] Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119*, 2240-2248
[3] Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1995**, *117*, 10141-10142.
[4] Hoaglund, C. S.; Valentine, S. J.; Sporleder, C. R.; Reilly, J. P.; Clemmer, D. E. *Anal. Chem.* **1998**, *70*, 2236-2242.
[5] Purves, R. W.; Barnett, D. A.; Ells, B.; Guevremont, R. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 738-745.
[6] Myung, S.; Badman, E. R.; Lee, Y. J.; Clemmer, D. E. *J. Phys. Chem. A.* **2002**, *106*, 9976-9982.
[7] Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Int. J. Mass Spectrom.* **1999**, *185/186/187*, 565-575.
[8] Suckau, D.; Shi, Y.; Beu, S. C.; Senko, M. W.; Quinn, J. P.; Wampler, F. M.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 790-793.
[9] Gross, D. S.; Schnier, P. D.; Rodriguez-Cruz, S. E.; Fagerquist, C. K.; Williams, E. R. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 3143-3148.
[10] Demmers, J. A. A.; Rijkers, D. T. S.; Haverkamp, J.; Killian, J. A.; Heck, A. J. R. *J. Am. Chem. Soc.* **2002**, *124*, 11191-11198.
[11] Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Anal. Chem.* **2001**, *73*, 3274-3281.
[12] Breuker, K.; Oh, H. B.; Horn, D. M.; Cerda, B. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124*, 6407-6420.
[13] Horn, D. M.; Breuker, K.; Frank, A. J.; McLafferty, F. W. *J. Am. Chem. Soc.* **2001**, *123*, 9792-9799.
[14] Oh, H.; Breuker, K.; Sze, S. K.; Ge, Y.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15863-15868.
[15] McLafferty, F. W.; Horn, D. M.; Breuker, K.; Ge, Y.; Lewis, M. A.; Cerda, B.; Zubarev, R. A.; Carpenter, B. K. *J. Am. Soc. Mass Spec.* **2001**, *12*, 245-249.
[16] Bohrer, B. C.; Merenbloom, S. I.; Koeniger, S. L.; Hilderbrand, A. E.; Clemmer, D. E. *Annu. Rev. Anal. Chem.* **2008**, *1*, 293-327.
[17] Bernstein, S. L.; Liu, D. F.; Wyttenbach, T.; Bowers, M. T.; Lee, J. C.; Gray, H. B.; Winkler, J. R. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1435-1443.
[18] Wu, C.; Murray, M. M.; Bernstein, S. L.; Condron, M. M.; Bitan, G.; Shea, J. E.; Bowers, M. T. *J. Mol. Biol.* **2009**, *387*, 492-501.
[19] Wyttenbach, T.; von Helden, G.; Bowers, M. T. *J. Am. Chem. Soc.* **1996**, *118*, 8355-8364.
[20] Hudgins, R. R.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **1998**, *120*, 12974-12975.
[21] Ly, T.; Julian, R. R. *J. Am. Chem. Soc.* **2008**, *130*, 351-358.
[22] Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657-1666.
[23] Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225-11236.
[24] Vijay-kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531-544.
[25] Schnier, P. D.; Gross, D. S.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1086-1097.
[26] Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Phys. Chem.* **1996**, *100*, 16082-16086.
[27] Shvartsburg, A. A.; Jarrold, M. F. *Chem. Phys. Lett.* **1996**, *261*, 86-91.
[28] Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Anal. Chem.* **2001**, *73*, 3274-3281.

[29] Ly, T.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1148-1158.
[30] Sun, Q.; Nelson, H.; Ly, T.; Stoltz, B. M.; Julian, R. R. *J. Proteome Res.* **2008**, *8*, 958-966.
[31] Moore, B. N.; Blanksby, S. J.; Julian, R. R. *Chem. Comm.* **2009**, 5015-5017.

Chapter 9

CONCLUDING REMARKS

The broader goal of this dissertation is to expand the bioanalytical scope of mass spectrometry beyond measuring the m/z of peptides and proteins. It is easy to understand how this trajectory can quickly bifurcate into numerous branches. Indeed, the analytical applications described here span facilitating identification of protein primary structure to elucidating the tertiary structure of proteins in the gas phase. Perhaps more fundamentally important in advancing our understanding of chemistry is the discovery that photodissociation mass spectrometry provides a unique vantage point to examine and utilize intramolecular radical migration in biomolecules.

Although radical chemistry has been the subject of intense investigation, long-range radical migration is a topic that has received relatively scant attention in the literature. However, long-range radical migration is essential to the activity of key life-sustaining enzymes, such as those involved in photosynthesis and aerobic metabolism. Radical migration also plays an important role in the transmission and repair of free radical damage. There are numerous inherent difficulties with studying radical migration. Radicals are transient due in part to

their high reactivity towards molecular oxygen. Hydroxyl radical, which is the most studied free radical species, attacks amino acids nearly indiscriminately, producing a variety of biomolecular radical species that may differ in their chemistry. At the level of a peptide or protein, the competing radical mechanisms become difficult to disentangle.

The most profound characteristic of the PD technique is the capability of synthesizing radicals on a peptide or protein with exceptionally high control over the radical location in high yield (30-50% from one laser pulse) and selectivity (> 95%). Thus, chemically distinct radical species can be neatly prepared with photodissociation. The subsequent radical migration and chemistry may be investigated using mass spectrometry and other gas phase methods. In the intervening time since the discovery of RDD, it was determined that structure and C-H bond dissociation energies are the most important factors governing radical migration.[1] Interestingly, these findings may have novel implications in the catalytic mechanisms of ribonucleotide reductases. It has been difficult to rationalize the proposed radical migration in these enzymes by comparing the reduction potentials of the isolated amino acids.[2,3] Our results suggest that C-H bond dissociation energies are more appropriate parameters to

evaluate radical migration in situations where charge separation is unfavorable (such as the gas phase or the catalytic pocket of an enzyme).

In summary, this dissertation describes the development of RDD and SNAPP as mass spectrometry-based analytical platforms to analyze protein structure. Several important and seminal applications are offered here. However, the diversity of the applications already shown is a strong indication that these techniques are robust platforms for future investigations into the myriad physical and chemical properties of proteins.

---

[1] See Chapter 3.
[2] Stubbe, J.; Nocera, D. G.; Yee, C. S.; Chang, M. C. Y. *Chem. Rev.* **2003**, *103*, 2167-2201.
[3] Stubbe, J.; van der Donk, W. A. *Chem. Rev.* **1998**, *98*, 705-765.