The Dissertation Committee for Jared Bryan Shaw Certifies that this is the approved version of the following dissertation:

Development of Tandem Mass Spectrometric Methods for Proteome Analysis Utilizing Photodissociation and Ion/Ion Reactions

Committee:

Jennifer S. Brodbelt, Supervisor

Richard M. Crooks

Lauren J. Webb

Eric V. Anslyn

George Georgiou

Development of Tandem Mass Spectrometric Methods for Proteome Analysis Utilizing Photodissociation and Ion/Ion Reactions

by

Jared Bryan Shaw, B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2013

For my parents, Roger and Sheila

Acknowledgements

I would like to thank my advisor Professor Jennifer Brodbelt for the freedom and opportunity to pursue my own research interests. Thank you to all of the members of the Brodbelt research group. Your friendship, support, enthusiasm and fruitful discussion have been indispensable. I am forever grateful.

Much of the work in this dissertation would not have been possible without the collaborative efforts of many people. Many thanks are owed to Jae Schwartz, John Syka, Jens Griep-Raming, Alexander Makarov, Eugen Damoc and Eduard Denisov at Thermo Fisher Scientific for fruitful discussion, a wonderful collaborative spirit and a once in a lifetime internship experience with the research and development group in Bremen, Germany. Thanks to Desmond Kaplan for the opportunity to collaborate with Bruker Daltonics and help with development of NETD and AI-NETD methods. Thanks to Hua Xu for his work with MassMatrix. Thanks to Neil Kelleher and his research group at Northwestern University for developing software to support UVPD for top-down proteomics.

Finally, I would like to thank my undergraduate research advisor Professor Dimitri Pappas for his mentorship and inspiring my curiosity and enthusiasm for the sciences.

Development of Tandem Mass Spectrometric Methods for Proteome Analysis Utilizing Photodissociation and Ion/Ion Reactions

Jared Bryan Shaw, Ph.D.

The University of Texas at Austin, 2013

Supervisor: Jennifer S. Brodbelt

The utility of 193 nm ultraviolet photodissociation (UVPD) and negative electron transfer dissociation (NETD) for the characterization of peptide anions was systematically evaluated. UVPD outperformed NETD in nearly all metrics; however, both methods provided complementary information to traditional collision induced dissociation (CID) of peptide cations in high throughput analyses. In order to enhance the performance of NETD, activated ion negative electron transfer dissociation (AI-NETD) methods were developed and characterized. The use of low-level infrared photoactivation or collisional activation during the NETD reaction period significantly improved peptide anion sequencing capabilities compared to NETD alone. Tyrosine deprotonation was shown to yield preferential electron detachment upon NETD or UVPD, resulting in N – Cα bond cleavage N-terminal to the tyrosine residue. LC-MS/MS analysis of a tryptic digest of BSA demonstrated that these cleavages were regularly observed under high pH conditions.

Transmission mode desorption electrospray ionization (TM-DESI) was coupled with 193 nm UVPD and CID for the rapid analysis and identification of protein digests. Comparative results are presented for TM-DESI-MS/CID and TM-DESI-MS/UVPD

analyses of five proteolyzed model proteins. In some cases TM-DESI/UVPD outperformed TM-DESI-MS/CID due to the production of an extensive array of sequence ions and the ability to detect low m/z product ions.

193 nm UVPD was implemented in an Orbitrap mass spectrometer for characterization of intact proteins. Near-complete fragmentation of proteins up to 29 kDa was achieved. The high-energy activation afforded by UVPD exhibited far less precursor ion charge state dependence than conventional methods, and the viability of 193 nm UVPD for high throughput top-down proteomics analyses was demonstrated for the less 30 kDa protein from a fractionated yeast cell lysate.

The use of helium instead of nitrogen as the C-trap and HCD cell bath gas and trapping ions in the HCD cell prior to high resolution mass analysis significantly reduced the signal decay rate for large protein ions. As a result, monoclonal IgG1 antibody was isotopically resolved and mass accurately determined. A new high mass record for which accurate mass and isotopic resolution has been achieved (148,706.3391 Da ± 3.1 ppm) was established.

Table of Contents

Chapter 1: Introduction 1
1.1 Introduction 1
1.2 Mass Spectrometric Approaches to Proteomics 3
1.3 The Role of Tandem Mass Spectrometry in Proteomics 6
1.3.1 Polypeptide Fragmentation Nomenclature 8
1.3.2 Collision Induced Dissociation 9
1.3.3 Electron-Based Dissociation Methods 11
1.3.4 Vacuum Ultraviolet Photodissociation 13
1.4 Overview of Chapters 17
1.5 References 20

Chapter 2: Experimental Methods 23
2.1 Overview 23
2.2 Mass Spectrometry 23
2.2.1 Electrospray Ionization 24
2.2.2 Quadrupole Ion Traps 24
2.2.3 Orbitrap Mass Analyzer 26
2.3 Liquid Chromatography 28
2.3.1 Dionex Ultimate 3000 HPLC 28
2.3.2 Eksigent NanoLC-Ultra 29
2.4 Chemicals 29
2.5 Peptide and Protein Preparation 30
2.6 Automated Tandem Mass Spectrometry Data Interpretation 31
2.6.1 SEQUEST 32
2.6.2 MassMatrix 32
2.6.3 ProSightPC 3.0 33
2.7 References 35

Chapter 3: Comparison of Ultraviolet Photodissociation and Negative Electron Transfer Dissociation for the Characterization of Multiply Charged Peptide Anions 37

3.1 Overview 37

3.2 Introduction 38

3.3 Experimental 40

3.3.1 Materials and Sample Preparation 40

3.3.2 Instrumentation 41

3.3.3 Mass Spectrometry and Liquid Chromatography 42

3.4 Results and Discussion 44

3.5 Conclusions 56

3.6 References 57

Chapter 4: Activated Ion Negative Electron Transfer Dissociation of Multiply Charged Peptide Anions 60

4.1 Overview 60

4.2 Introduction 60

4.3 Experimental 64

4.3.1 Materials 64

4.3.2 Instrumentation 65

4.3.3 Mass Spectrometry 66

4.4 Results and Discussion 66

4.4.1 Production of NETD reagent ions: EI vs. CI 66

4.4.2 NETD of Doubly and Triply Deprotonated Peptides 68

4.4.3 AI-NETD of Doubly Deprotonated Peptides 70

4.4.4 AI-NETD of Triply Deprotonated Peptides 80

4.4.5 Hydrogen Content of Product Ions 83

4.4.6 LC-MS/MS and MassMatrix Database Search Results 84

4.5 Conclusions 86

4.6 References 87

Chapter 5: Tyrosine Deprotonation Yields Abundant and Selective Backbone Cleavage in Peptide Anions upon Negative Electron Transfer Dissociation and Ultraviolet Photodissociation 90

5.1 Overview 90

5.2 Introduction 90

5.3 Experimental 92

5.3.1 Materials and Sample Preparation 92

5.3.2 Instrumentation 93

5.3.3 Mass Spectrometry and Liquid Chromatography 94

5.4 Results and Discussion 95

5.5 Conclusions 106

5.6 References 107

Chapter 6: Analysis of Protein Digests by Transmission-mode Desorption Electrospray Ionization Mass Spectrometry with Ultraviolet Photodissociation 110

6.1 Overview 110

6.2 Introduction 110

6.3 Experimental 113

6.3.1 Materials 113

6.3.2 Sample Preparation 114

6.3.3 Transmission Mode DESI, Mass Spectrometry and Ultraviolet Photodissociation 114

6.3.4 Background Subtraction and Protein Database Searching 116

6.4 Results and Discussion 117

6.5 Conclusions 124

6.6 References 125

Chapter 7: Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation 127

7.1 Overview 127

7.2 Introduction 127

7.3 Experimental 129

7.3.1 Materials 129
7.3.2 Sample Preparation 129
7.3.3 Implementation of UVPD 130
7.3.4 Mass Spectrometry 131
7.3.5 Yeast Whole Cell Lysate Preparation 132
7.3.6 Liquid Chromatography 133
7.3.7 Data Processing 133
7.4 Results and Discussion 134
7.5 Conclusions 151
7.6 References 152

Chapter 8: Extending the Isotopically Resolved Mass Range of Orbitrap Mass Spectrometers 155
8.1 Overview 155
8.2 Introduction 155
8.3 Experimental 158
8.3.1 Materials 158
8.3.2 Sample Preparation 158
8.3.4 Mass Spectrometry 159
8.4 Results and Discussion 160
8.5 Conclusions 173
8.6 References 175

Chapter 9: Conclusions 177
References 183

Chapter 1

Introduction

1.1 Introduction

Whole-genome sequencing has made the study of whole proteomes possible. The field of proteomics offers highly complementary information to genomics, as it is the proteome (i.e. the entire set of proteins expressed at a particular time in a cell) that ultimately determines cellular function and phenotype. DNA and RNA sequencing data is used to predict gene products (e.g. protein sequences) and expression levels in cells; however, protein sequence and abundances can only be inferred base on genome and transcriptome data, and this is one of the major driving factors for the development of methods to directly analyze proteins. The field of proteomics and its continued development is critical to the understanding of disease states, the development of new diagnostic and prognostic assays and the development of effective therapeutics. Unlike the finite genome (i.e. the complete complement of genetic information in a cell), the proteome is essentially boundless and constantly changing in response to internal and external stimuli. The human genome is estimated to contain approximately 20,000 protein-coding genes[1]; however, processes such as alternative splicing and post-translational modifications have been estimated to yield more than one million unique protein forms.[2]

The field of proteomics is continually becoming more and more synonymous with high-throughput mass spectrometry based characterization of proteomes. Mass spectrometry has become the method of choice for protein identification[3,4], elucidating protein post-translational modifications[5,6] and mapping of protein-protein interactions[7].

As a result, mass spectrometry is an essential tool used by biologists and biochemists to understand molecular systems regulating cellular pathways. The rapid growth of mass spectrometry based proteomics in the past decade is largely due to significant improvements in experimental strategy, the sensitivity, speed and mass accuracy/resolution of mass spectrometric instrumentation, and tandem mass spectrometry methods.[8] This dissertation focuses on the implementation and development of tandem mass spectrometric methods utilizing ion/ion reactions and ultraviolet photodissociation for both the bottom-up and top-down proteomic approaches.

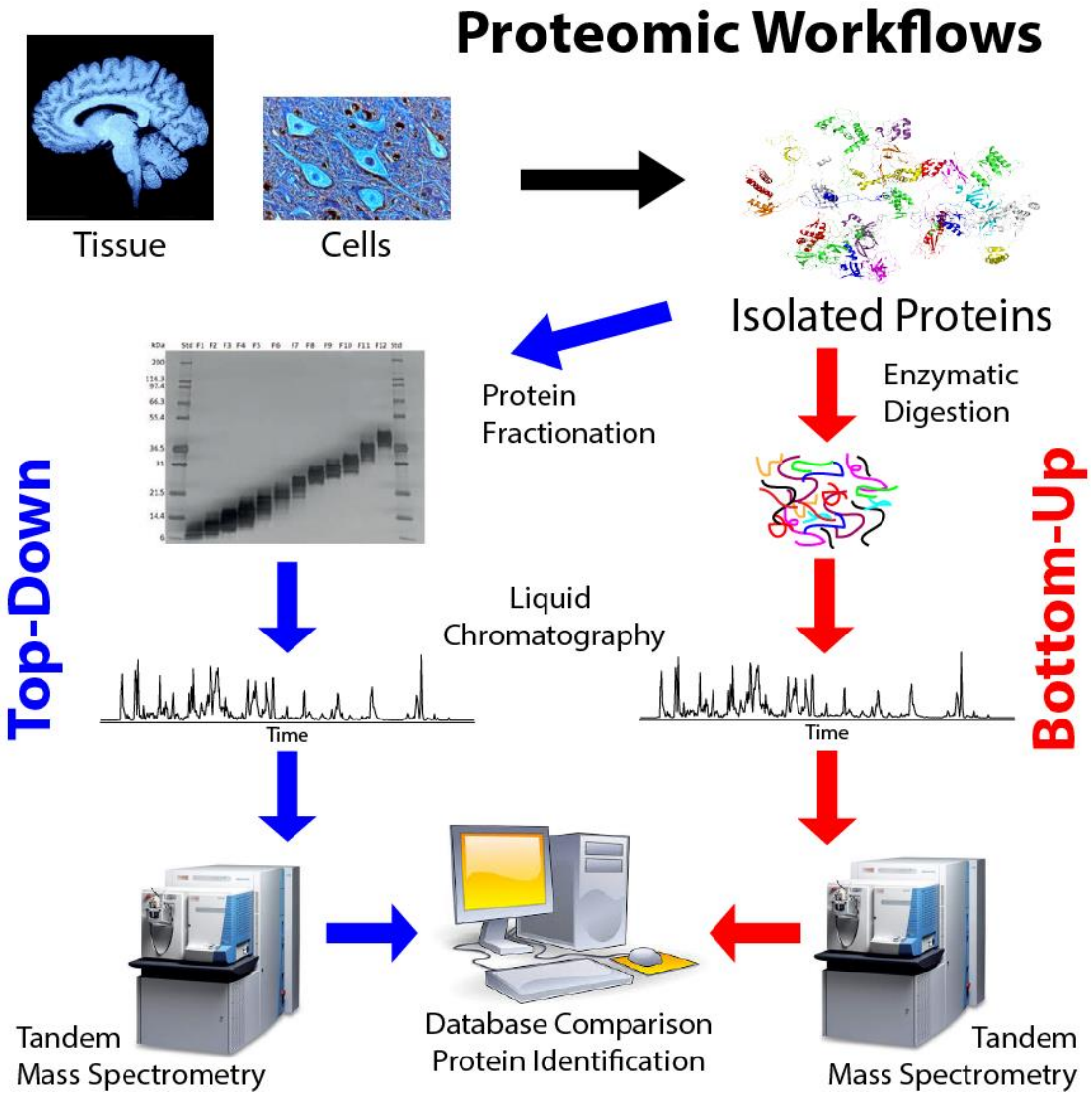

Figure 1.1 Typical workflows used for mass spectrometry based proteomic analyses.

1.2 Mass Spectrometric Approaches to Proteomics

There are two main experimental approaches to mass spectrometry based proteomics, the bottom-up approach and the top-down approach (**Figure 1.1**). The bottom-up approach is the most popular method for tackling large-scale complex mixtures such as whole proteomes.[8,9] In this approach, proteins, which have been isolated

from cells or tissue, are subject to proteolytic digestion to produce peptides.[10] The peptide mixture is separated using reversed phase liquid chromatography or multiple orthogonal dimensions of chromatography. The eluting peptides are then sampled by a mass spectrometer and the ensuing peptide masses and sequences are used to identify the corresponding proteins. Most applications require the use of tandem mass spectrometry (MS/MS) during which the peptides are fragmented in a reproducible manner, and the fragmentation patterns are used to identify the peptide sequences. The most common method for the identification of MS/MS data is through comparison to *in-silico* generated fragmentation patterns of peptide sequences generated from protein sequence databases. The bottom-up method has been very successful because (i) peptides are relatively small and more uniform biochemically than intact proteins, (ii) robust front-end separation methods are available for peptides, (iii) peptide masses are more easily determined with high accuracy and (iv) peptides are more easily fragmented during MS/MS. These characteristics have enabled researchers utilizing the bottom-up approach to identify and quantify thousands of proteins in a single study as wells as identify thousands of sites of post-translational modifications.[11]

Alternatively, the top-down approach bypasses protein enzymatic digestion and utilizes intact protein mass measurements and MS/MS to identify and characterized proteins. The potential major benefits of the top-down approach mainly relate to the characterization of post-translational modifications (PTMs). PTMs specifically modify protein structure and/or function and serve as mechanisms for precise control of cellular pathways to generate response to changes in physiological conditions.[12] Often these mechanisms involve multiple inter-dependent PTMs at multiple sites within a single protein molecule. In other words, the attachment of a PTM at one site triggers modification at another site. As a result, in order to completely understand protein

regulation by PTM, additional information beyond that of peptide-level PTM identification and quantification is required, and concentrated effort to understanding the interconnectivity of PTMs is of the upmost importance to deciphering cellular pathways. The enzymatic digestion of proteins (bottom-up approach) destroys any information on the interconnectivity of PTMs; however, interrogation of PTMs at the proteins-level (top-down approach) provides information on the combinatorial nature of PTMs.

Despite the obvious advantages of the top-down approach, its adoption for routine proteome analysis has been limited due to the analytical challenges related to the analysis of intact proteins. First, intact protein separations are far more difficult due to the vast range of molecular weight and biochemical makeup found in complex mixtures such as cellular proteomes. Effective separation is critical for reduction of sample complexity and increasing dynamic range of detection for mass spectrometric analysis. For example, traditional methods for protein separation, such as two-dimensional polyacrylamide gel electrophoresis, offer unrivaled peak capacity, but low recovery upon extraction of intact proteins from the gel is a major problem. Perhaps the most significant hurdle to routine adoption of the top-down approach is the requirement of a very high performance mass spectrometer. Intact mass determination and MS/MS is much more demanding for intact proteins than for peptides. High mass accuracy and resolution is mandatory for intact mass and fragment ion mass determination. High protein molecular weight and charge of protein ions produced by electrospray ionization (ESI) yield isotopic distributions with very small mass to charge ratio (m/z) increments thus requiring very high resolving power (Figure 1.2). Also, many of the fragment ion isotopic distributions overlap due to the wide distribution of charge and mass of the fragment ions produced during MS/MS. High mass accuracy and resolution (isotopic resolution) is required to distinguish and accurately identify the fragment ions. Additionally, intact protein ions are more difficult

to fragment efficiently and extensively during MS/MS. Well characterized fragmentation pathways observed in peptides are not always observed in protein ions, and the degree of fragmentation which can be achieved by each fragmentation method is largely dependent on features of the protein such as molecular weight, charge, charge density and gas-phase conformation.[12–14]

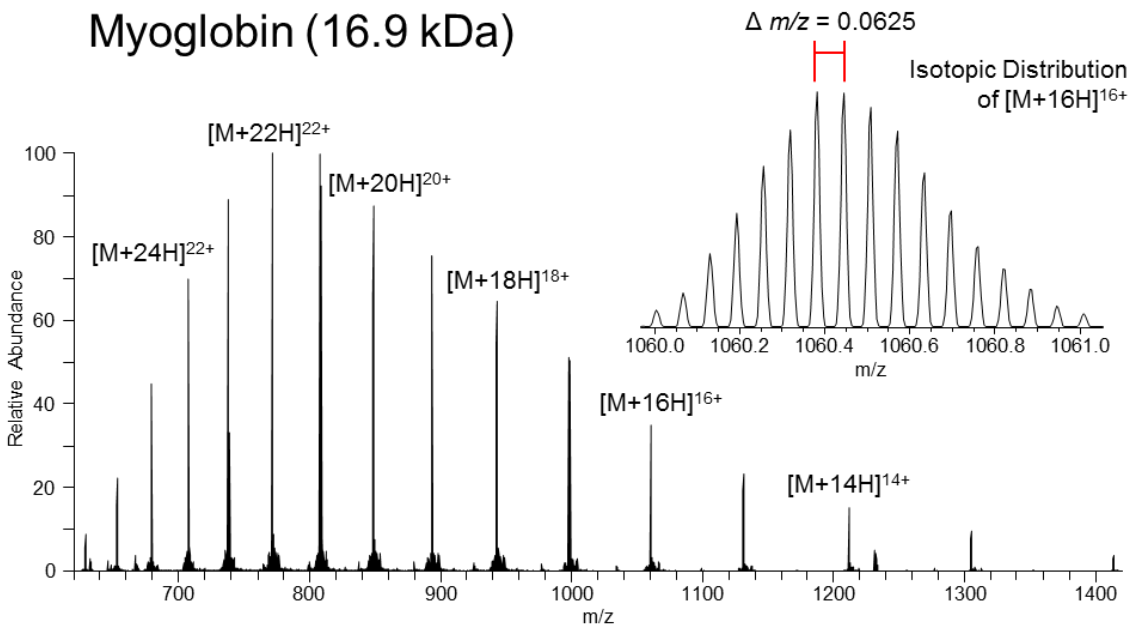

Figure 1.2 Electrospray ionization mass spectrum showing the charge state distribution observed for the protein myoglobin and insert showing the isotopic distribution of the 16+ charge state.

1.3 The Role of Tandem Mass Spectrometry in Proteomics

Determination of the elemental composition of peptide ions is possible at sufficiently high mass measurement accuracy and knowledge of peptide elemental composition allows determination of possible amino acid compositions. Zubarev et al.

have shown mass measurement accuracy of 1 ppm is sufficient to determine unique amino acid compositions for peptides less than approximately 600 Da in mass.[15] However, the average peptide mass for the yeast proteome when digested using the protease trypsin (cleavage C-terminal to lysine and arginine) is approximately 2100 Da (allowing up to two missed cleavages).[16] Additionally, this level of mass accuracy is only achievable with the highest performance (most expensive) mass spectrometers such as Fourier transform ion cyclotron resonance (FT-ICR) and Orbitrap mass spectrometers. Ion-trap mass spectrometers (less expensive) are widely used for proteomic analyses, but are only capable of mass accuracy in the range of 1000 – 2000 ppm.

Furthermore, high mass accuracy alone is insufficient to differentiate peptide isomers which have the same elemental and amino acid compositions and only differ in the amino acid residue order within the peptide sequence. Since accurate intact peptide mass is not sufficient to uniquely identify most peptides, additional information is needed. This is when sequence specific information provided by MS/MS becomes indispensable. MS/MS is a collective term for a number of methods in which precursor ions are selected and fragmented yielding diagnostic product ions that are subsequently mass analyzed. For proteomic applications, tandem mass spectrometry provides a means for fragmenting polypeptide ions along the backbone in a reproducible and consistent manner producing diagnostic product ions that allow determination of the peptide/protein amino acid composition and sequence. In the following sections of this chapter, polypeptide fragmentation nomenclature and the major tandem mass spectrometry methods are discussed.

1.3.1 Polypeptide Fragmentation Nomenclature

In order to determine the primary structure (i.e. the amino acid sequence) of peptides and proteins, MS/MS methods must yield bond cleavage along the polypeptide backbone. Cleavage along the backbone yields product ions with diagnostic mass differences that are characteristic of each amino acid residue. Roepstorff originally proposed nomenclature for polypeptide fragmentation pathways.[17] A graphical illustration of this nomenclature is shown in Figure 1.3.

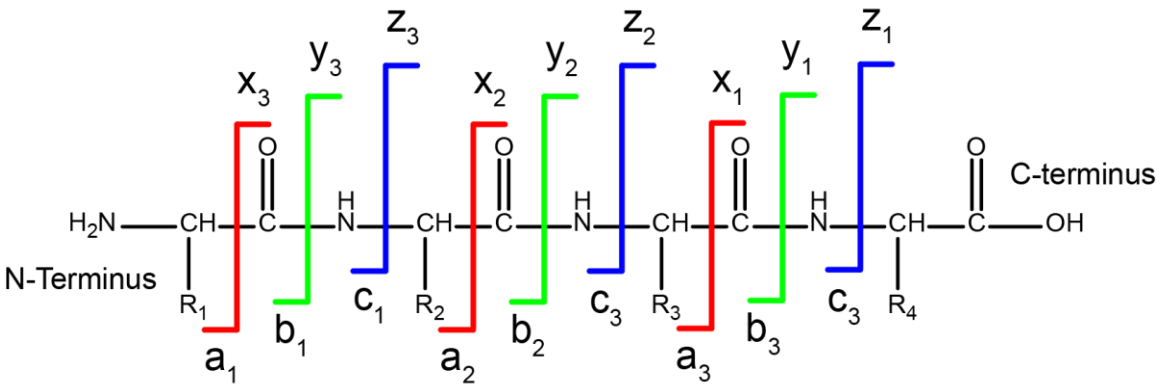

Figure 1.3 Polypeptide fragmentation nomenclature.

Product ions that are produced by cleavage of the polypeptide backbone and retain the N-terminus of the polypeptide are referred to as *a, b* and *c*-ions, whereas product ions that retain the C-terminus of the polypeptide are *x, y* and *z*-ions. Complementary ions pairs are *a/x, b/y* and *c/z* ions. Cleavage of the C_α – C, C – N and N – C_α bonds yield *a/x, b/y* and *c/z* ion pairs, respectively. Subscripts in the product ion name indicate the number of amino acid residues of which the product ions is composed. For example, a_2 and x_2 ions would be produced if fragmentation were to occur at the C_α – C bond of the second residue in the tetrapeptide shown in Figure 1.3.

1.3.2 Collision Induced Dissociation

Collision induced dissociation (CID) or collision activated dissociation (CAD) was the first ion activation method implemented and has played an essential role in the development of MS/MS. CID is still the most widely available and commonly utilized ion activation method, and its fragmentation pathways have been extensively studied and characterized. The two major forms of CID are resonant excitation and beam-type CID.[18] Both forms of CID produce ion activation through energetic collisions between precursor ions and an inert bath gas such as helium or nitrogen. Resonant excitation CID as it is performed in ion trap mass spectrometer involves the application of an AC waveform with frequency equal to the secular frequency of the precursor ion motion in the ion trap. The AC waveform increases the amplitude/velocity of the precursor ion motion producing relatively low energy collisions with the bath gas in the ion trap.[19,20] When an ion with high translational energy undergoes inelastic collisions with a neutral, some of the translational energy is converted to internal energy. Many collisions are required for the precursor ion to accumulate enough internal energy for dissociation to take place. Beam- type CID involves the acceleration of precursor ions into a pressurized collision cell where they undergo multiple collisions with an inert bath gas and dissociate. For the scope of this dissertation, only low-energy forms of CID will be considered, which generally yield collisions in the range of 1 – 100 eV.[18] These methods are referred to as slow heating or threshold dissociation processes, because activation proceeds through small incremental increases in vibration excitation occurring over tens of milliseconds. The energetics of the collisions can be tuned somewhat through manipulation of the amplitude of AC waveform or acceleration voltage. For beam-type CID, energetics can

be tuned through the uses of more of less massive bath gases. The center of mass collision energy (E_{com}) can be represented as follows:

$$E_{com} = \left(\frac{N}{m_p + N} \right) KE$$

where KE is the ion's kinetic energy and N and m_p are the masses of the neutral bath gas and the precursor ion. Increasing the mass of the neutral bath gas allows more of the ion's kinetic energy to be converted into internal energy. Nitrogen and argon are commonly used bath gases for beam-type CID; however, in ion trap mass spectrometers helium is most common because more massive gases degrade performance during mass analysis.

For peptide cation characterization, CID is the standard to which all other ion activation methods are compared. The mobile proton model has served as solid foundation for understanding fragmentation pathways observed for gas phase protonated peptides upon CID.[21–23] The model assumes that ionizing protons are initially located at basic sites such as the N-terminus or side chains of lysine, arginine and histidine. Upon ion activation the ionizing proton(s) can migrate from less basic sites on the polypeptide triggering various charge-site-initiated mechanisms yielding backbone cleavage. More specifically, a mobile proton can produce N-protonation of the amide bond yielding cleavage of the amide bond. Transfer of a proton to the amide nitrogen eliminates resonance stabilization between the carbonyl and nitrogen lone pair electrons. Loss of resonance stabilization greatly weakens the amide bond permitting cleavage that yields diagnostic *b*- and *y*-ions.

There are a number of drawbacks associated with CID that have driven the development of other activations methods. Preferential sequence specific fragmentation can severely limit the information obtained from peptides containing proline and aspartic acid residues.[24,25] Proline containing peptides and proteins exhibit very selective and enhance cleavage of the amide bond N-terminal to the proline residues. This effect is exacerbated in higher charge state ions (more mobile protons). Peptides and proteins also exhibit enhanced cleavage of the amide bond immediately C-terminal to acidic residues in the absence of a mobile proton. Without a mobile proton, cleavage initiated by the acidic hydrogen of aspartic becomes significant. The combination of these effects can severely limit the extent of fragmentation in protein cations which commonly contain multiple proline and aspartic acid residues. Another limitation of CID is related to the analysis of labile PTMs such as phosphorylation. CID is a slow heating/threshold dissociation method; therefore, energy is slowly deposited into the precursor and cleavage of the weakest bonds occurs first. As a result, CID of peptides and proteins containing labile PTMs yield abundant neutral loss of the PTM and limited sequence informative fragmentation. This characteristic limits the utility of CID for peptide identification and site localization of PTMs.

1.3.3 Electron-Based Dissociation Methods

Electron-based dissociation methods, such as electron capture dissociation (ECD)[26,27] and electron transfer dissociation (ETD)[28,29] are attractive alternatives to CID due to the production of more random and extensive fragmentation in peptide/protein cations as well as the preservation of labile PTMs[30–35] during the dissociation process. Electron capture dissociation (ECD) entails the irradiation of multiply charged

peptide/protein cations with low-energy electrons generate by a heated filament.[26,27] ECD is typically performed within the magnetic field of FT-ICR mass spectrometers where electrons and analyte cations can be trapped simultaneously. ETD is the ion trap mass spectrometer equivalent of ECD, and involves the reaction of radical anions with multiply charged peptide/protein cations. Capture of a low-energy electron (ECD) or the transfer of an electron during an ion/ion reaction (ETD) to multiply charge peptide and protein cations is thought to proceed through the capture/transfer of an electron at the site of an ionizing proton and inducing backbone cleavage through the migration of a hydrogen radical.[26,28,36] In contrast to CID, dissociation occurs prior to energy randomization throughout vibrational modes (non-ergodic dissociation) thus labile PTMs are predominantly retained on the product ions. This characteristic allows precise localization of labile PTMs such as phosphorylation, sulfation and glycosylation as well as determination of the polypeptide sequence. The non-ergodic nature of ECD and ETD has greatly extended the utility of the top-down approach due to more random and extensive backbone cleavage and the production of far fewer amino acid sequence specific cleavages.

The major limitations of ECD and ETD stem from a strong dependency on charge density and thus the gas-phase structure of the ions selected for dissociation. Greater charge density yields more coulombic repulsion within the polypeptide ions which results in a more elongated and unfolded gas-phase structure. Low charge density ions generally have a higher degree of intramolecular hydrogen bonding and ionic interactions resulting in a more compact conformation. These interactions can prevent separation of the product ions after backbone cleavage and thus the individual product ions are not detected. The use of low level collisional activation or infrared photoactivation (i.e. vibrational excitation) prior to or during the ECD/ETD process, termed activated-ion ECD/ETD,[37–39]

has been implemented to disrupt peptide/protein secondary and tertiary structure and allows more effective generation and detection of ECD/ETD products of in low charge density peptides and proteins. These methods have greatly extended the mass range of peptides and proteins amenable to MS/MS.

1.3.4 Vacuum Ultraviolet Photodissociation

Photon-based ion activation methods offer a number of potential benefits over conventional activation methods such as CID. First, photodissociation provides very clearly defined and tunable excitation energy. Photodissociation also allows the use of trapping conditions suitable for the detection of low *m/z* ions in ion trap mass spectrometers, because precursor ion motion is not manipulated during dissociation as with CID.[40,41] Secondary dissociation of product ions is also a potential advantage of photodissociation, because uninformative product ions have the potential to absorb additional photons and dissociate into informative product ions.

Photon energies ranging from 0.1 eV to 8 eV (10.6 μm to 157 nm) are commonly used in photodissociation experiments. Activation using low energy infrared photons requires absorption of multiple photons (infrared multiphoton dissociation, IRMPD) for precursor ions to accumulate sufficient vibrational excitation to access fragmentation pathways and generally yields fragmentation similar to CID. On the hand, absorption of an ultraviolet photon yields high energy activation and dissociation from excited electronic states. Ultraviolet photodissociation (UVPD) is typically preformed using 157 nm (7.9 eV) or 193 nm (6.4 eV) photons generated by F_2 or ArF excimer lasers. Of these two wavelengths, 193 nm has been far more prevalent, because there is only a small loss when transmitting through air, common fused silica optics can be used and ArF lasers

produce higher pulse energies.[42] 193 nm UVPD was first utilized for peptide fragmentation in 1984 by Bowers et al.[43], and was subsequently utilized by a number of other groups for various biomolecules.[44–51]

Absorption spectra of polyalanine in the vacuum ultraviolet region show three distinct bands centered at approximately 190 nm, 160 nm and 130 nm. These bands have been attributed to transitions involving the amide bond of the polypeptide backbone.[52] This is the fundamental reasoning behind the use of 157 nm and 193 nm photons for fragmentation of the polypeptides. The high-energy activation afforded by 193 nm UVPD enables dissociation through broad array of pathways not generally accessible by most activation methods.[47,48,51,53–58] 193 nm UVPD promotes homolytic cleavage of the C_α – C bond producing radical *a* and *x*-ions (i.e. a+1 and x+1) which can eliminate a hydrogen atom to from a and *x* ions (Figure 1.4). *v, d* and *w* even electron ions are also produced by loss small molecules, such as CO or CONH, and/or partial amino acid side chains from a+1 and x+1 ions as shown in Figure 1.5.

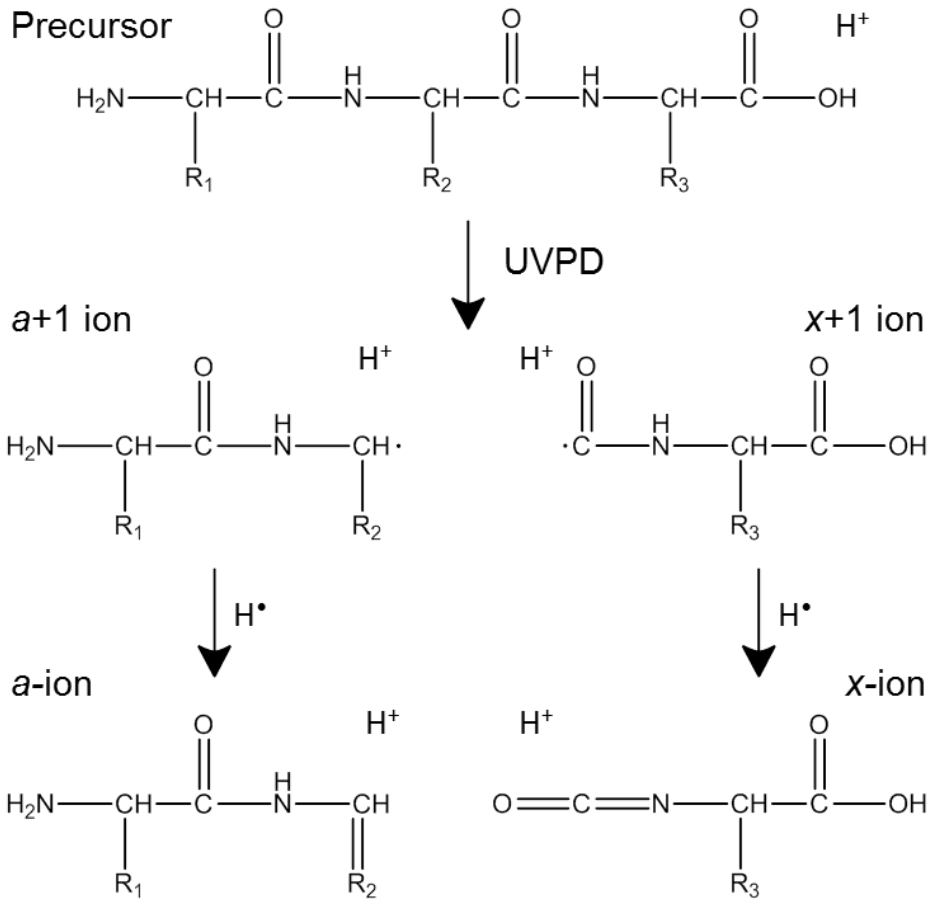

Figure 1.4 Primary product ions produced by 193 nm UVPD of polypeptides

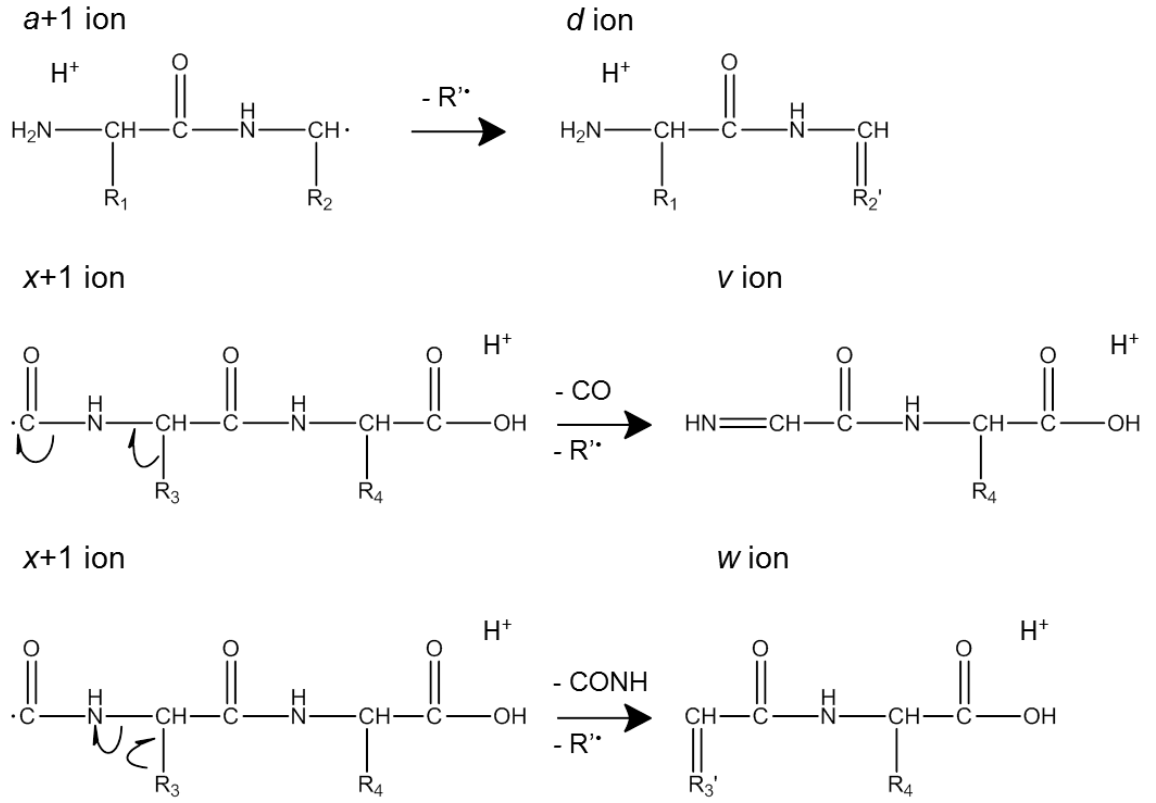

Figure 1.5 Generation of secondary fragment ions from *a*+1 and *x*+1 ions. R' indicates partial side chain loss.

Photoexcited ions may dissociate directly or decay by radiative and/or nonradiative channels. Nonradiative decay is a much higher probability process due to the lack of a solvent system around gas-phase ions to which to transfer energy and the large number of vibrational mode present in polypeptides. From excited electronic states the polypeptide may partition the electronic energy into vibrational modes and dissociate. Many peptide UVPD studies focused on singly charged species with N- or C-terminal arginine residues (sequestered ionizing protons) to mimic peptides generated by trypsin digestion. These experiments yielded almost exclusively ions from the terminus that contained the charge suggesting that the charge is only necessary for detection of the product ions. It was also found that when the peptides contained a lysine residue instead

of an arginine, the proton was more easily mobilized which led to other pathways competing (i.e. c/z and b/y ion formation as is observed with ECD/ETD and CID) with homolytic cleavage of the C_α – C bond to from a/x ions.[48,51,59] The broad array of product ions produced by UVPD has been shown to yield extensive sequence coverage and confident identification of peptides while maintaining labile PTMs. These characteristics were our motivation for pursuing UVPD for high throughput characterization of peptides and proteins.

1.4 Overview of Chapters

The research presented in this dissertation is aimed at the development of tandem mass spectrometry methods involving ion/ion reactions and/or photodissociation to expand the current capabilities of mass spectrometry based proteomics by providing more complete characterization of peptide and protein ions.

CID of peptide anions yields very complicated spectra riddled with neutral losses and unexplainable fragmentation patterns. Chapter 3 presents a systematic comparison of UVPD and negative electron transfer dissociation (NETD) for the characterization of multiply charged peptide anions. UVPD is shown to outperform NETD in terms of peptide sequence coverage and other metrics; however, when combined with traditional CID of peptide cations, both methods yielded complementary information.

In order to enhance the performance of NETD, activated ion negative electron transfer dissociation (AI-NETD) methods were developed and characterized (Chapter 4). The use of low level infrared photoactivation or collisional activation during the NETD reaction period significantly improved peptide anion sequencing capabilities as is

evidenced by greater product ion abundances and peptide sequence coverage compared to NETD alone.

In Chapter 5, tyrosine deprotonation is shown to yield preferential electron detachment upon NETD or UVPD, resulting in prominent N – Cα bond cleavage N-terminal to the tyrosine residue. UVPD of iodo-tyrosine modified peptides was used to generate localized radicals on neutral tyrosine side chains by homolytic cleavage of the C – I bond. Subsequent collisional activation of the radical species yielded the same preferential cleavage of the adjacent N-terminal N – Cα bond. LC-MS/MS analysis of a tryptic digest of BSA demonstrated that these cleavages are regularly observed for peptides when using high pH mobile phases.

Chapter 6 utilized transmission mode desorption electrospray ionization (TM-DESI) coupled with 193 nm UVPD and CID for the rapid analysis and identification of protein digests. Comparative results are presented for TM-DESI-MS/collision induced dissociation (CID) and TM-DESI-MS/UVPD analyses of five proteolyzed model proteins ranging in molecular weight from 8.5 kDa (ubiquitin) to 66 kDa (bovine serum albumin, BSA). In some cases TM-DESI/UVPD yielded greater confidence in database correlation scores for peptides and comparable protein identification compared to TM-DESI-MS/CID due to the production of an extensive array of sequence ions and the ability to detect low m/z terminal sequence ions and immonium ions.

In chapter 7, 193 nm ultraviolet photodissociation (UVPD) is implemented in an Orbitrap mass spectrometer for characterization of intact proteins. Near-complete fragmentation of proteins up to 29 kDa is achieved. The 5 nanosecond, high-energy activation afforded by UVPD exhibits far less precursor ion charge state dependence than conventional collision-based and electron-based dissociation methods. The viability of

193 nm UVPD for high throughput top-down proteomics analyses is demonstrated for the less than 30 kDa proteins from a fractionated yeast cell lysate.

A simple yet effective method for significantly improving the performance of Orbitrap mass spectrometers for the analysis of large biomolecules is presented in Chapter 8. The use of helium instead of nitrogen as the C-trap and HCD cell bath gas and trapping ions in the HCD cell prior to high resolution mass analysis significantly reduced the signal decay rate for large protein ions and allowed the acquisition of longer transients. As a result, the intact masses of the major glycoforms of a monoclonal IgG1 antibody were isotopically resolved and accurately determined. A new high mass record for which accurate mass and isotopic resolution has been achieved (148,706.3391 Da ± 3.1 ppm) was established.

1.5 References

(1) Clamp, M.; Fry, B.; Kamal, M.; Xie, X.; Cuff, J.; Lin, M. F.; Kellis, M.; Lindblad-Toh, K.; Lander, E. S. *Proc. Natl. Acad. Sci.* **2007**, *104*, 19428–19433.

(2) Nørregaard Jensen, O. *Curr. Opin. Chem. Biol.* **2004**, *8*, 33–41.

(3) Chait, B. T. *Annu. Rev. Biochem.* **2011**, *80*, 239–246.

(4) Cox, J.; Mann, M. *Annu. Rev. Biochem.* **2011**, *80*, 273–299.

(5) Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villén, J.; Haas, W.; Sowa, M. E.; Gygi, S. P. *Cell* **2010**, *143*, 1174–1189.

(6) Lemeer, S.; Heck, A. J. *Curr. Opin. Chem. Biol.* **2009**, *13*, 414–420.

(7) Ewing, R. M.; Chu, P.; Elisma, F.; Li, H.; Taylor, P.; Climie, S.; McBroom-Cerajewski, L.; Robinson, M. D.; O'Connor, L.; Li, M.; Taylor, R.; Dharsee, M.; Ho, Y.; Heilbut, A.; Moore, L.; Zhang, S.; Ornatsky, O.; Bukhman, Y. V.; Ethier, M.; Sheng, Y.; Vasilescu, J.; Abu-Farha, M.; Lambert, J.-P.; Duewel, H. S.; Stewart, I. I.; Kuehl, B.; Hogue, K.; Colwill, K.; Gladwish, K.; Muskat, B.; Kinach, R.; Adams, S.-L.; Moran, M. F.; Morin, G. B.; Topaloglou, T.; Figeys, D. *Mol. Syst. Biol.* **2007**, *3*.

(8) Yates, J. R.; Ruse, C. I.; Nakorchevsky, A. *Annu. Rev. Biomed. Eng.* **2009**, *11*, 49–79.

(9) Yates, J. R. *J. Am. Chem. Soc.* **2013**, *135*, 1629–1640.

(10) Switzar, L.; Giera, M.; Niessen, W. M. A. *J. Proteome Res.* **2013**, *12*, 1067–1077.

(11) Geiger, T.; Velic, A.; Macek, B.; Lundberg, E.; Kampf, C.; Nagaraj, N.; Uhlen, M.; Cox, J.; Mann, M. *Mol. Cell. Proteomics* **2013**, *12*, 1709–1722.

(12) Lanucara, F.; Eyers, C. E. *Mass Spectrom. Rev.* **2013**, *32*, 27–42.

(13) Zhou, H.; Ning, Z.; E. Starr, A.; Abu-Farha, M.; Figeys, D. *Anal. Chem.* **2011**, *84*, 720–734.

(14) Armirotti, A.; Damonte, G. *Proteomics* **2010**, *10*, 3566–3576.

(15) He, F.; Emmett, M. R.; Håkansson, K.; Hendrickson, C. L.; Marshall, A. G. *J. Proteome Res.* **2004**, *3*, 61–67.

(16) Cagney, G.; Amiri, S.; Premawaradena, T.; Lindo, M.; Emili, A. *Proteome Sci* **2003**, *1*.

(17) Roepstorff, P. *Biomed. Mass Spectrom.* **1984**, *11*, 601.

(18) Sleno, L.; Volmer, D. A. *J. Mass Spectrom.* **2004**, *39*, 1091–1112.

(19) Louris, J. N.; Cooks, R. G.; Syka, J. E.; Kelley, P. E.; Stafford Jr, G. C.; Todd, J. F. *Anal. Chem.* **1987**, *59*, 1677–1685.

(20) McLuckey, S. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599–614.

(21) Dongre, A. R.; Jones, J. L.; Somogyi, Á.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365–8374.

(22) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35*, 1399–1406.

(23) Paizs, B.; Suhai, S. *Mass Spectrom. Rev.* **2004**, *24*, 508–548.

(24) Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1996**, *31*, 1185–1187.

(25) Gu, C.; Tsaprailis, G.; Breci, L.; Wysocki, V. H. *Anal Chem* **2000**, *72*, 5804–5813.

(26) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.

(27) 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.

(28) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9528 –9533.

(29) Coon, J. J.; Syka, J. E. P.; Schwartz, J. C.; Shabanowitz, J.; Hunt, D. F. *Int. J. Mass Spectrom.* **2004**, *236*, 33–42.

(30) Kelleher, N. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250–4253.

(31) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431–4436.

(32) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793–1800.

(33) Shi, S. D.-H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2000**, *73*, 19–22.

(34) Kleinnijenhuis, A. J.; Kjeldsen, F.; Kallipolitis, B.; Haselmann, K. F.; Jensen, O. N. *Anal. Chem.* **2007**, *79*, 7450–7456.

(35) Kweon, H. K.; Håkansson, K. *J. Proteome Res.* **2008**, *7*, 749–755.

(36) Breuker, K.; Oh, H.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 14011 –14016.

(37) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778–4784.

(38) Håkansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256–3262.

(39) Ledvina, A. R.; Beauchene, N. A.; McAlister, G. C.; Syka, J. E. P.; Schwartz, J. C.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J. *Anal. Chem.* **2010**, *82*, 10068–10074.

(40) Colorado, A.; Shen, J. X.; Vartanian, V. H.; Brodbelt, J. *Anal. Chem.* **1996**, *68*, 4033–4043.

(41) Brodbelt, J. S.; Wilson, J. J. *Mass Spectrom. Rev.* **2009**, *28*, 390–424.

(42) Reilly, J. P. *Mass Spectrom. Rev.* **2009**, *28*, 425–447.

(43) Bowers, W. D.; Delbert, S. S.; Hunter, R. L.; McIver, R. T. *J. Am. Chem. Soc.* **1984**, *106*, 7288–7289.

(44) Hunt, D. F.; Shabanowitz, J.; Yates, J. R. *J. Chem. Soc. Chem. Commun.* **1987**, *0*, 548–550.

(45) Martin, S. A.; Hill, J. A.; Kittrell, C.; Biemann, K. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 107–109.

(46) Williams, E. R.; Furlong, J. J. P.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 288–294.

(47) Cui, W.; Thompson, M. S.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1384–1398.

(48) Morgan, J. W.; Russell, D. H. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 721–729.

(49) Moon, J. H.; Shin, Y. S.; Cha, H. J.; Kim, M. S. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 359–368.

(50) Yoon, S. H.; Kim, M. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1729–1739.

(51) Yoon, S. H.; Chung, Y. J.; Kim, M. S. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 645–655.

(52) Woody, R. W.; Koslowski, A. *Biophys. Chem.* **2002**, *101–102*, 535–551.

(53) Fung, Y. M. E.; Kjeldsen, F.; Silivra, O. A.; Chan, T. W. D.; Zubarev, R. A. *Angew. Chem.* **2005**, *117*, 6557–6561.

(54) Thompson, M. S.; Cui, W.; Reilly, J. P. *Angew. Chem. Int. Ed.* **2004**.

(55) Kjeldsen, F.; Silivra, O. A.; Zubarev, R. A. *Chem.- Eur. J.* **2006**, *12*, 7920–7928.

(56) Thompson, M. S.; Cui, W.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1439–1452.

(57) Kim, T.-Y.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2334–2341.

(58) He, Y.; Webber, N.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 675–683.

(59) Kim, T. Y.; Thompson, M. S.; Reilly, J. P. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 1657–1665.

Chapter 2

Experimental Methods

2.1 Overview

The work presented herein focuses on the development of tandem mass spectrometry methods (MS/MS) that yield more extensive fragmentation and greater numbers of diagnostic product ions. All experiments were performed using quadrupole ion trap mass spectrometers or a hybrid linear quadrupole ion trap/Orbitrap mass spectrometer equipped with electrospray ionization sources. The mass spectrometric instrumentation, chromatography, peptide/protein sample preparation procedures and data analysis methods utilized in this dissertation are discussed in the subsequent sections of this chapter.

2.2 Mass Spectrometry

All mass spectrometric analyses were performed using a three-dimensional quadrupole ion trap (QIT), a linear quadrupole ion trap (LIT) mass spectrometer or a hybrid linear quadrupole ion trap/Orbitrap mass spectrometer. The instruments utilized in this body of work include a Bruker Daltonics HCTultra ETDII three-dimensional quadrupole ion trap modified to enable photodissociation and negative electron transfer reactions, a Thermo Fisher Scientific LTQ XL linear quadrupole ion trap mass spectrometer modified to enable photodissociation and a Thermo Fisher Scientific Orbitrap Elite mass spectrometer also modified to enable photodissociation.

2.2.1 Electrospray Ionization

Electrospray ionization (ESI) has revolutionized the field of biological mass spectrometry since its introduction in 1989.[1] ESI is a soft ionization method as it does not induce fragmentation during the ionization process and greatly extended the mass range and nature of ions capable of being analyzed by mass spectrometry. ESI is performed by the application of a high voltage (typically 3-4 kV) to a conductive needle through which an analyte solution is infused. A Taylor cone is produced at the end of the needle and a fine aerosol of charged droplets containing analyte is produced. The droplets undergo evaporation and coulombic explosions until desolvated analyte ions are produced. Ionization is accomplished through protonation or deprotonation of the analyte molecules. For peptides and proteins, multiply charged ions are typically observed. The degree of charging is related to the number of acidic and basic sites on the molecule and the initial pH of the solution electrosprayed. A nitrogen sheath gas is typically used to help desolvation of ions during the transfer of ions from the ion source to the mass spectrometer inlet.

2.2.2 Quadrupole Ion Traps

A QIT consists of three electrodes, a ring electrode which is sandwiched by two end cap electrodes. The profile of the inner surface of these electrodes is hyperbolic in shape. The end caps electrodes are typically held at ground potential and a radio frequency (RF) potential is applied to the ring electrode which creates a quadrupolar field between the electrodes. It is this field that allows trapping of ions within the ion trap. It is important to note that both cations and anions can be trapped simultaneously, since the trapping field is created by an RF potential only. This factor greatly simplifies the implementation of ion/ion reactions. Ions are injected into the QIT through holes in the

end cap electrodes. Trapping efficiency and resolution are increased through the use of an inert bath gas (typically ~1 of helium) which helps condense the ion cloud to the center of the QIT by collisional cooling. Mass spectra are produced by a linear ramp of the RF potential on the ring electrode from a low potential to a higher potential. Ions of increasing mass are sequentially ejected (mass selective instability) from the trap through holes in the end cap electrodes to a detector (typically an electron multiplier).[2] Resolution in quadrupole ion traps is inversely proportional to the slope of the RF potential ramp (i.e. scan speed).[3] MS/MS is performed in QIT through the application of auxiliary AC waveforms to the end cap electrodes of the QIT. Application of an auxiliary waveform of the same frequency as an ion's secular motion within the QIT causes the amplitude of motion and thus the velocity of the ion to increase. This phenomenon can be used for ion isolation and ion activation (resonant collisional activation).

Standalone linear quadrupole ion trap mass spectrometers were introduced by Thermo Finnigan in 2002, as a high capacity alternative to three dimensional quadrupole ion traps.[4] The claimed advantages of LITs over QITs include increased ion storage volume, enhanced sensitivity and higher trapping efficiency. A LIT as implemented by Thermo Finnigan consists of a quadrupole with rods split into three sections and front and back lenses (Figure 2.1). Opposite RF phases are applied to the two pairs of quadrupole rods to confine ion in the radial directions, and DC potentials applied to the separate sections of the rods confine ions in the axial direction. As with QITs, a bath gas provides collisional cooling of the ion cloud. Opposite phases of an auxiliary AC waveform are added to one pair of opposite rods for ion isolation and ion activation. Mass analysis is accomplished by ejection of ions (mass selective instability) through slots in the center section of the quadrupole rods. DC potentials on the front and back lenses gate the flow of ions into the ion trap. To enable ion/ion reactions, AC waveforms are also added to the

front and back lenses of the LIT. This provides an axial trapping field for simultaneous storage of cations and anions.

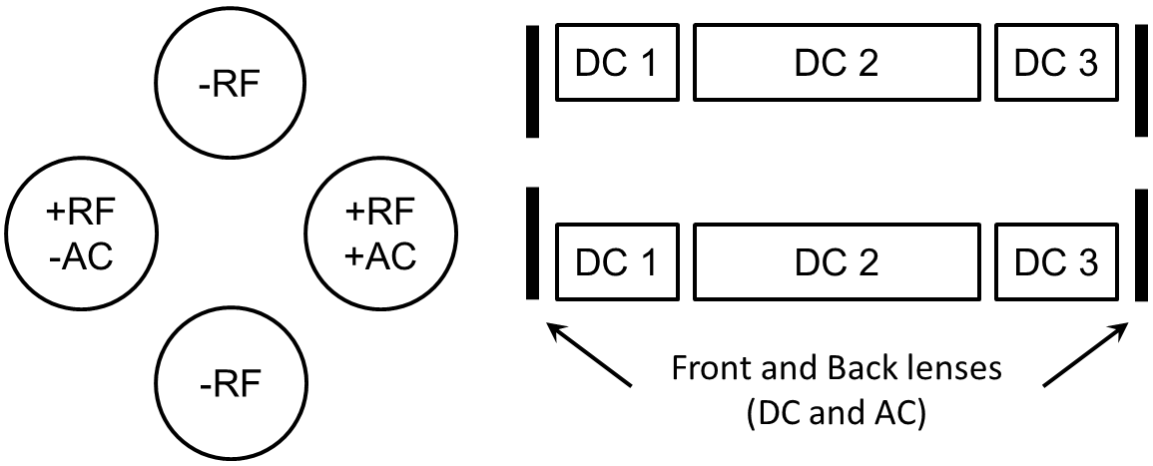

Figure 2.1 Schematic showing the application of RF, DC and AC potentials need for operation of a linear quadrupole ion trap.

2.2.3 Orbitrap Mass Analyzer

The Orbitrap mass analyzer was first described in 2000[5] and commercialized in the form of a hybrid linear ion trap/Orbitrap mass spectrometer in 2005.[6] Orbitrap mass spectrometers have rapidly gained popularity for a wide variety of applications due to its ultrahigh resolving power, high mass accuracy, short acquisition times and robust operation.[7–14]

The trap consists of an outer barrel-like electrode and a spindle shaped center electrode along the axis (Figure 2.2). The Orbitrap mass analyzer employs trapping in an electrostatic field in which ions orbit around the center electrode. Ions are injected into the Orbitrap mass analyzer from a curve linear ion trap (C-trap) as coherent packets of ions (Figure 2.2). Ion injection is offset along the axis of the Orbitrap mass analyzer and

causes the ions to accelerate towards the opposite end of the trap. The field shape (quadrologarithmic) causes the ions undergo harmonic oscillations along the length of the center electrode, and the frequency of ion oscillation is proportional to $(m/z)^{-1/2}$. The frequency of ion oscillation is detected by image current induced on the outer electrodes. Resolution in excess of one million and less than 1 ppm mass accuracy is achievable with Orbitrap mass analyzer.

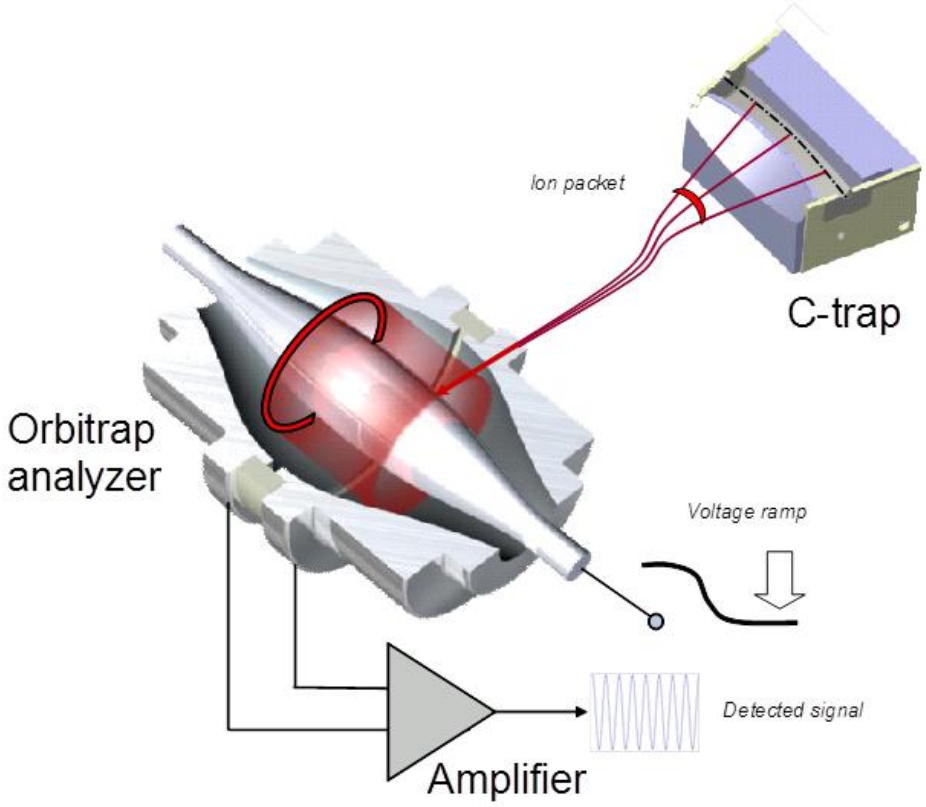

Figure 2.2 Cut-away view of the Orbitrap mass analyzer and C-trap illustrating ion injection into the Orbitrap mass analyzer. Artwork by Thermo Fisher Scientific. For license terms see http://creativecommons.org/licenses/by-sa/3.0/deed.en

2.3 Liquid Chromatography

Prior separation of complex proteomic samples is essential for successful mass spectrometric analysis. Frontend separations enhance the dynamic range of detection, minimize ion suppression effect during the electrospray ionization process and greatly increase the depth of proteome analysis. Peptide mixtures are generally separated by high performance liquid chromatography (HPLC) using C18 reverse phase columns and C4 or C8 reverse phase columns are used with intact proteins. Flow rates in the nL/min to μL/min range are commonly used for proteomic analyses. In general, less sample is required and ionization is more efficient for experiments conduct at lower flow rates.

2.3.1 Dionex Ultimate 3000 HPLC

Peptide separations were perform using a Dionex Ultimate 3000 HPLC equipped with a capillary flow splitter and an Agilent ZORBAX 300Extend-C18 column (Santa Clara, CA) (150 × 0.3 mm, 3.5 μm particle size). The column temperature was held at 25°C. For peptide anion experiments, eluents A and B consisted of 0.05% ammonium hydroxide in water and acetonitrile (~pH 10.5), respectively. For peptide cation experiments, eluents A and B were 0.1% formic acid in water and acetonitrile, respectively. A 120 minute gradient consisting of the first five minutes at 3% eluent B and a linear increase to 35% eluent B over 120 minutes at a flow rate of 4 uL/min was used to separate the protein digests. 2 – 5 μg of protein digest was injected on column.

2.3.2 Eksigent NanoLC-Ultra

Intact protein separations were carried out on an Eksigent nano flow HPLC equipped with self-packed 15 cm analytical and 2 cm trap columns. The packing media was C4, 300 angstrom pore, 5 micron Magic stationary phase (Michrom Bioresources, Inc.). For protein mixtures, eluents A and B were 0.1% formic acid in water and acetonitrile, respectively. A 90 minute gradient consisting of a linear increase to 10% eluent B in the first 10 minutes followed by a linear increase to 60% eluent B at 90 minutes. A flow rate of 300 nL/min was used and approximately 5 µg of protein was injected.

2.4 Chemicals

All model peptides and Fmoc-protected amino acids were purchased from Anaspec (San Jose, CA), and all model proteins were purchased from Sigma Aldrich (St. Louis, MO) except for chicken lysozyme C which was purchased from MP Biomedicals, LLC (Solon, OH). Monoclonal IgG1 (product number: 186006552) was purchased from Waters (Milford, MA). Sodium iodide, chloramine-T, and sodium metabisulfite, proteomics grade trypsin, acetic acid, formic acid, ammonium bicarbonate, ammonium hydroxide, dithiothreitol, iodoacetamide, isopropylsilane (TIPS) were purchased from Sigma Aldrich. β-d_2-Tyrosine was purchased from Cambridge Isotope Laboratories (Andover, MA). Wang resin, fluorenylmethyloxycarbonyl-O-succinimide and tetramethylammonium hexafluorophosphate (HCTU) (ChemPep Inc., Miami, FL) were also purchased. Sep-Pak® 1cc C18 cartridges were obtained from Waters (Milford, MA). HPLC grade and LCMS grade solvents were purchased from Fisher Scientific.

2.5 Peptide and Protein Preparation

All protein digestions were performed as follows. 200 μg of protein was reduced with dithiothreitol (DTT) for 30 minutes then incubated with iodoacetamide at 40°C for one hour in the dark. Excess iodoacetamide was quenched with additional DTT. 4 μL of 1 μg/ μL trypsin in 1 mM HCl was added, giving a protein to enzyme ratio of 50:1. The digestion solution was incubated overnight at 37°C. The digest was desalted using C18 cartridges then dried and reconstituted in water to a concentration of 1 μg/μL.

The method used for tyrosine iodination was performed as previously described.[15] All reagents were prepared in water. The mole ratio between peptide, sodium iodide, chloramine-T, and sodium metabisulfite was 1:1:2:2. To 100 nmol of peptide in 100 μL of water, sodium iodide was added and chloramine-T immediately after and reacted for 3 minutes at room temperature. Sodium metabisulfite was added to quench the reaction.

Human Pin1 expression, purification and oxidation was carried out as follows. The human Pin1 gene was subcloned in a pHIS8 vector, a derivative of pET28a vector (Novagene).[16] The Pin1 R14A mutant was produced using the QuikChange Site-Directed Mutagenesis Kit (Stratagene, CA). The purification of Pin1 R14A mutant was similar to previous reported procedure.[17] Briefly, the Pin1 R14A was overexpressed in using E. coli BL21(DE3) strain at 16°C overnight induced by isopropyl-β-D-thiogalactopyranoside (IPTG). The N-terminal polyhistidine tag was removed by thrombin protease (Novegen Germany) during the overnight dialysis after eluted from Ni-NTA (Invitrogen NY) purification. The protein was further purified by gel filtration superdex75 (GE Healthcare) in 20 mM HEPES 7.5 and 50 mM NaCl.

Purified Pin1 was treated with mild oxidizing condition similarly to previous published protocol.[18] 2 mg/ml purified Pin1 R14A protein was incubated with 30 μM $FeSO_4$/1 mM H_2O_2 for 3hours. The oxidation was stopped by removing hydroxyl free

radicals using dialysis in 20 mM HEPES, 50 mM NaCl overnight. Samples were spun down for 20 min at 20,000 x g to remove possible aggregation particles then concentrated and desalted using a 10 kDa centrifugal filter (Millipore Corporation) prior to mass spectrometric analysis.

Yeast whole cell lysate for top-down analysis was prepared as follows. *S. Cerevisiae* was grown to an OD600 of 0.7, harvested, and lysed in SDS lysis buffer. 500 ug of total protein was loaded to a 12%-T commercial GELFrEE devise and proteins were resolved by SDS-PAGE and eluted off the gel. Fractions were collected and 10% of each fraction was loaded to a conventional SDS-PAGE slab gel for visualization by silver nitrate staining (shown above). The remainder of each fraction was subjected to a methanol/chloroform/water cleanup procedure to remove the SDS from the protein. The SDS-free samples were dried down, and the pellets were stored at -80 C until the time of use. Each pellet was reconstituted and pipetted vigorously 20 – 100 times, and then sonicated in a bath sonicator on ice for 20 minutes. Finally, the reconstituted samples were spun down at high speed (21,000 x g) for 10 minutes.

2.6 Automated Tandem Mass Spectrometry Data Interpretation

With the vast amounts of data acquired during high throughput proteomic analyses, automated interpretation of large proteomic datasets is essential. The most common method for the identification of MS/MS data is through comparison to *in-silico* generated fragmentation patterns of peptide sequences generated from protein sequence databases. Various statistical methods are used to validate the candidate peptide and protein matches, and false discovery rates can be estimated based results from searching against a decoy database containing reversed protein sequences. The database search

algorithms, SEQUEST, MassMatrix and ProSightPC, utilized in this work are discussed in the subsequent section of this chapter.

2.6.1 SEQUEST

Thermo Fisher Scientific Proteome Discoverer 1.0 software package was used to perform SEQUEST database searches for interpretation of MS/MS data. SEQUEST search parameters included a signal:noise ratio of 3, a precursor mass tolerance of 3 Da, fragment mass tolerance of 2 Da, 1 missed cleavage by trypsin, methionine oxidation as a dynamic side chain modification and carbamidomethyl as a static side chain modification of cysteine. Product ion series for CID spectra included *a*, *b* and *y* ions and UVPD product ion series were *a, b, c, x, y* and *z* ions. UVPD photoionization background peaks possibly arising from pump oil and hydrocarbon impurities in the ion trap helium bath gas were subtracted as previously described.[19] Nonredundant bovine and chicken protein databases from the NCBI were used for proteins form the respective species. Peptide matches were filtered based on charge state and XCorr scores, and protein matches were rejected if they did not receive more than two peptide matches. MS/MS spectra that passed the filters but received probability scores less than 1.00 were manually verified.

2.6.2 MassMatrix

MassMatrix, originally developed for positive mode CID, was modified to allow searching of negative mode MS/MS data by both UVPD and NETD. The methods were optimized individually based on the specific fragmentation behavior associated with each method. For example, NETD searches were performed with several forms of *a*- and *x*-type ions (e.g. $a/x \pm 1$ due to hydrogen migration), and therefore searching scores are

elevated compared to UVPD, which uses only one form of *a*- and *x*- ions. Preferential cleavages at proline, glutamic acid, and aspartic acid that yield specific Y- and y-type ions, and *c*/*z* ions are also included in searches for UVPD data.[20]

Using these modified MassMatrix algorithms, LC-MS/MS data sets were searched against the cow International Protein Index (IPI v3.73, Sept 27, 2011) modified to include the chicken lysozyme C sequence using the MassMatrix database search engine v2.4.0. LC-MS/MS were performed in triplicate, and the data from the replicate experiments were merged prior to database searching. Search parameters utilized were trypsin, 2 missed cleavages, phosphorylation of S and T as variable modifications, iodoacetamide/carbamidomethylation of C as a fix modification, precursor ion tolerance of 2 Da and a product ion tolerance of 1 Da. Peptide hits were filtered by pp and pp_2 scores less than 5 and pp_{tag} scores less than 0.5[21] for which the pp and pp2 scores are a statistical measure of the number of matched product ions and the total abundance of matched product ions, respectively. All suspect peptide spectral matches (PSMs) were manually verified.

2.6.3 ProSightPC 3.0

Automated database searches for top-down LC-MS/MS data was performed using ProSightPC 3.0 that was modified to incorporate a, b, c, x, y and z ions produced by UVPD. Spectra were converted to neutral monoisotopic masses using the Xtract algorithm considering product ions of S/N 3 and greater and a maximum charge of 25. Spectra were searched against the "top_down_complex" Saccharomyces cerevisiae database downloaded from the ProSightPC database website (ftp://prosightpc.northwestern.edu/). The database contained 23564 basic sequences and

1196890 protein forms. Searches were performed in absolute mass mode with a precursor mass tolerance of 500 Da, 10 ppm fragment ion tolerance, a minimum of 10 fragment ions and with the Δm option. Hits were subsequently filtered to remove one with a P score greater than 1E-4 and mass difference greater than ± 3 Da.

2.7 References

(1) Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. *Science* **1989**, *246*, 64–71.

(2) Stafford Jr., G. C.; Kelley, P. E.; Syka, J. E. P.; Reynolds, W. E.; Todd, J. F. J. *Int. J. Mass Spectrom. Ion Process.* **1984**, *60*, 85–98.

(3) Schwartz, J. C.; Syka, J. E.; Jardine, I. *J. Am. Soc. Mass Spectrom.* **1991**, *2*, 198–204.

(4) Schwartz, J. C.; Senko, M. W.; Syka, J. E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 659–669.

(5) Makarov, A. *Anal. Chem.* **2000**, *72*, 1156–1162.

(6) Makarov, A.; Denisov, E.; Kholomeev, A.; Balschun, W.; Lange, O.; Strupat, K.; Horning, S. *Anal Chem* **2006**, *78*, 2113–2120.

(7) Olsen, J. V.; de Godoy, L. M. F.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. *Mol. Cell. Proteomics* **2005**, *4*, 2010 –2021.

(8) Makarov, A.; Denisov, E.; Lange, O.; Horning, S. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 977–982.

(9) Perry, R. H.; Cooks, R. G.; Noll, R. J. *Mass Spectrom. Rev.* **2008**, *27*, 661–699.

(10) Scigelova, M.; Makarov, A. *Proteomics* **2006**, *6*, 16–21.

(11) Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. *Nat. Methods* **2007**, *4*, 709–712.

(12) McAlister, G. C.; Phanstiel, D.; Good, D. M.; Berggren, W. T.; Coon, J. J. *Anal. Chem.* **2007**, *79*, 3525–3534.

(13) Michalski, A.; Damoc, E.; Hauschild, J.-P.; Lange, O.; Wieghaus, A.; Makarov, A.; Nagaraj, N.; Cox, J.; Mann, M.; Horning, S. *Mol. Cell. Proteomics* **2011**, *10*, 10.1074/mcp.M111.011015.

(14) Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Müller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M.; Dunyach, J.-J.; Cox, J.; Horning, S.; Mann, M.; Makarov, A. *Mol. Cell. Proteomics* **2012**, *11*, 10.1074/mcp.O111.013698.

(15) Liu, Z.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 965–971.

(16) Jez, J. M.; Ferrer, J.-L.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Biochemistry (Mosc.)* **2000**, *39*, 890–902.

(17) Zhang, Y.; Daum, S.; Wildemann, D.; Zhou, X. Z.; Verdecia, M. A.; Bowman, M. E.; Lücke, C.; Hunter, T.; Lu, K. P.; Fischer, G. *Acs Chem. Biol.* **2007**, *2*, 320–328.

(18) Sultana, R.; Boyd-Kimball, D.; Poon, H. F.; Cai, J.; Pierce, W. M.; Klein, J. B.; Markesbery, W. R.; Zhou, X. Z.; Lu, K. P.; Butterfield, D. A. *Neurobiol. Aging* **2006**, *27*, 918–925.

(19) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9*, 4205–4214.

(20) Madsen, J. A.; Kaoud, T. S.; Dalby, K. N.; Brodbelt, J. S. *Proteomics* **2011**, *11*, 1329–1334.

(21) Xu, H.; Freitas, M. *BMC Bioinformatics* **2007**, *8*, 133.

Chapter 3

Comparison of Ultraviolet Photodissociation and Negative Electron Transfer Dissociation for the Characterization of Multiply Charged Peptide Anions

3.1 Overview

Ultraviolet photodissociation at 193 nm (UVPD) and negative electron transfer dissociation (NETD) were compared to establish their utility for characterizing acidic proteomes with respect to sequence coverage distributions (a measure of product ion signals across the peptide backbone), sequence coverage percentages, backbone cleavage preferences, and fragmentation differences relative to precursor charge state. UVPD yielded significantly more diagnostic information compared to NETD for lower charge states ($n \leq 2$), but both methods were comparable for higher charged species. While UVPD often generated a more heterogeneous array of sequence-specific products (*b-,y-, c-, z-, Y-, d-,* and *w*-type ions in addition to *a*- and *x*- type ions), NETD usually created simpler sets of *a*/*x*-type ions. LC-MS/UVPD and LC-MS/NETD analysis of protein digests utilizing high pH mobile phases coupled with automated database searching via modified versions of the MassMatrix algorithm was undertaken. UVPD generally outperformed NETD in stand-alone searches due to its ability to efficiently sequence both lower and higher charge states with rapid activation times. However, when combined with traditional positive mode CID, both methods yielded complementary information with significantly increased sequence coverage percentages and unique peptide identifications over that of just CID alone.

3.2 Introduction

Standard proteomics experiments typically involve a "bottom-up" strategy consisting of protein digestion, liquid chromatographic separation using acidic eluents, ionization in the positive mode, tandem mass spectrometry (MS) via collision induced dissociation (CID),[1,2] and automated database searching.[3,4] This general protocol consistently works well for the analysis of peptides cations.[5] In addition, alternative dissociation methods, such as electron capture/transfer dissociation[6–11] as well as infrared photodissociation (IRMPD)[12–15] and ultraviolet photodissociation (UVPD[16–27] have been developed to supplement CID for the analysis peptide cations. However, this general protocol often falls short for the characterization of more acidic species, such as peptides containing several acidic amino acids and/or acidic post translational modifications (PTM), due to low ion signals and poor MS/MS patterns. Acidic peptides ionize readily in the negative mode via deprotonation;[28,29] however, CID of peptide anions generally yields low sequence coverage, high abundances of uninformative neutral losses, and overall unpredictable fragmentation patterns.[30–32] It is now well recognized that many important modifications such as phosphorylation, sulfonation, acetylation, and glycosylation render proteins more acidic after translation.[33] Moreover, all proteomes exhibit a bimodal distributions of protein isoelectric points (pI).[34–37] In other words, every proteome contains acidic and basic subsets of proteins with very few proteins possessing a pI close to cellular pH.[34–37] As a result of these two factors, there is a need for improved methods for the characterization of the acidic regions of proteins as well as the acidic subset of proteins, including those approaches specifically adapted for analysis of peptide anions.

Several MS/MS and MS^n methods have recently been developed for analysis of peptide anions to combat the downfalls of conventional CID.[38–44] The majority of these

techniques have utilized electrons or photons to promote primarily C_α-C(O) bond cleavages, resulting in *a*- and *x*-type sequence ions. For example, the interaction of peptides with high-energy electrons, termed electron detachment dissociation (EDD), has been shown to be useful for sequencing peptide anions.[38,39] Activated-electron photodetachment dissociation (a-EPD), an MS anion activation method, uses UV irradiation at 262 nm in the first step to produce charge-reduced peptide radicals which are subsequently dissociated by collisional activation in the second step.[40,45] EDD and a-EPD were compared recently for characterization of peptide anions, and while both techniques yielded mostly the same types of ions (*a*/*x*-type), the abundances were significantly different.[40] Cleavages next to aromatic and histidine residues were favored for a-EPD, whereas cleavages next to negative charge solvation sites on the backbone were preferential for EDD.[40] Recently, a method involving high kinetic energy metastable ions (MAD-MS) was used to dissociate phosphorylated and sulfonated peptide anions; C_α-C(O) bond cleavages were again the main dissociation channels.[41] The development of EDD, a-EPD, and MAD-MS all represent promising advances for peptide anion characterization, but these methods generally afford low dissociation efficiencies and thus require long averaging times to generate interpretable spectra.[40,41] Also, implementing these methods on liquid chromatographic timescales is problematic because the required activation times often range from half a second up to multiple seconds for peptide dissociation.[39,45] Two alternative anion dissociation techniques that require less averaging and lower activation times to produce reasonable spectra, therefore allowing more facile integration with LC-MS/MS workflows, are negative electron transfer dissociation (NETD)[42,43] and 193 nm ultraviolet photodissociation (UVPD).[44] The former involves the reaction of peptide anions with electron-deficient radical cations, causing electron transfer from the peptide to the reagent that results in backbone

dissociation to create mainly *a* and *x* ions.[42,43] This method has shown great potential in sequencing peptides in high-throughput experiments.[46] 193 nm UVPD delivers high-energy (6.4 eV) photons to peptides and has been shown to yield diagnostic spectra dominated by *a* and *x* ions and utilizing activation times as little as 5 ns (in combination with LC-MS/MS).[44] UVPD has more recently been used in conjunction with automated database searching and sequencing of peptide anions for the analysis of complex protein samples.[47] Since both NETD and UVPD have been promising for the sensitive and high-throughput analysis of acidic proteomes, we have undertaken a comparison of NETD and UVPD for the characterization of a series of modified (e.g., phosphorylated) and unmodified peptide anions. The sequence coverage distribution (a measure of product ion signals across the peptide backbone), sequence coverage percentages, backbone cleavage preferences, and fragmentation differences with respect to precursor charge state were assessed between the methods. Results from the analysis of a seven protein mixture via automated MassMatrix[48–50] interpretation and LC-MS/MS are also presented.

3.3 Experimental

3.3.1 Materials and Sample Preparation

All peptides were purchased from AnaSpec (Fremont, CA), and all proteins were purchased from Sigma Aldrich (St. Louis, MO) except for chicken lysozyme C which was purchased from MP Biomedicals, LLC (Solon, OH). All peptides and proteins were used without further purification. HPLC grade methanol and water (Fisher Scientific) were used for preparing peptide working solutions. HPLC grade ammonium hydroxide

and argon of ≥99.998% purity were purchased from Sigma Aldrich (St. Louis, MO). All optical components were purchased from Edmund Optics Inc. (Barrington, NJ).

An equimolar mixture of bovine α-casein, β-casein, β-lactoglobulin A, BSA, ribonuclease A and chicken lysozyme C was made in 100 mM sodium bicarbonate. 200 μg of total protein was reduced with dithiothreitol (DTT) for 30 minutes then incubated with iodoacetamide at 40°C for one hour in the dark. Excess iodoacetamide was quenched with additional DTT. 4 μL of 1 μg/ μL trypsin in 1 mM HCl was added, giving a protein to enzyme ratio of 50:1. The digestion solution was incubated overnight at 37°C. The digest was desalted using C18 cartridges then dried and reconstituted in water to a concentration of 1 μg/μL.

3.3.2 Instrumentation

A Bruker Daltonics (Billerica, MA) HCT Ultra ETD II ion trap mass spectrometer was modified to allow the introduction of 193 nm photons into the trapping volume. Briefly, two 25 mm CaF_2 windows were mounted in the vacuum manifold lid above the ion trap. The laser beam was passed through 2 mm holes in the ring electrode (oriented vertically) of the quadrupole ion trap. Broadband deep UV enhanced 12.5 mm mirrors were placed in 45 degree mounts under and next to the ion trap to reflect the laser beam back out of the mass spectrometer. This configuration allowed facile alignment of the laser through the system by monitoring the intensity and profile of the beam exiting the system. Ultraviolet photons (193 nm, 5 ns pulse, 500 Hz, 8 mJ/pulse max) were generated by an ExciStar XS 500 laser (Coherent Inc., Santa Clara, CA) and used without focusing or collimating the beam. An external trigger supplied through the auxiliary interface port gated a pulse/delay generator (Model 505, Berkely Nucleonics Corporation, San Rafael,

CA) which triggered the laser during the fragmentation period of the ion trap scan function (two pulses total). An override in the instrument user interface and replacement of the methane carrier gas with argon allowed efficient production of fluoranthene radical cations by argon chemical ionization and transmission of the radical cations to the ion trap. A reagent crucible temperature of 60 °C, a filament current of 1 μA, and electron energy of 70 eV were used.

3.3.3 Mass Spectrometry and Liquid Chromatography

10 μM peptide working solutions composed of 70/30 (v/v) methanol/water with 0.1% ammonium hydroxide were infused at 1.5 μL/min for electrospray ionization. NETD experiments were performed with a low mass cutoff of *m/z* 160, 10 to 15 ms of reagent accumulation, and 150 to 250 ms reaction time. UPVD experiments were performed with 4 ms activation periods allowing two 4 mJ laser pulses and a q_z-value of 0.110 which allowed detection of low *m/z* product ions.

LC-MS/MS experiments were performed with a DionexUltiMate 3000 system (Sunnyvale, CA) using a capillary flowsplitter. An Agilent ZORBAX 300Extend-C18 column (Santa Clara, CA) (150 × 0.3 mm, 3.5 μm particle size) was used for all separations. The column temperature was held at 25°C. For negative ion mode experiments (NETD and UVPD), eluents A and B consisted of 0.05% ammonium hydroxide in water and acetonitrile, respectively. A 120 minute gradient consisting of the first five minutes at 3% eluent B and a linear increase to 35% eluent B over 120 minutes at a flow rate of 4 uL/min was used to separate the protein mixture digest. CID experiments were performed in positive mode and utilized 0.1% formic acid in water and acetonitrile for eluents A and B, respectively. The same gradient were used for positive

mode except the gradient ended at 40% eluent B. 5 µL of 1 µg/µL protein mixture digest was injected on column. Data-dependent acquisition consisted of a survey scan of *m/z* 400 - 2000 followed by five MS/MS scan events for all activation methods. For LC-MS/NETD, the precursor charge state preference was set to exclude singly charged ions. No precursor charge state preference was specified for LC-MS/UVPD. A preference for doubly charged ions was utilized for LC-MS/CID experiment. For all LC-MS/MS experiments, eight averages were used for all MS/MS events and dynamic exclusion was set to one repeat and exclude for 60 seconds.

MassMatrix was used for all database searching.[48–50] This algorithm, which was originally developed for positive mode CID, was modified to allow searching of negative mode MS/MS data by both UVPD and NETD. The methods were optimized individually based on the specific fragmentation behavior associated with each method. For example, NETD searches were performed with several forms of *a*- and *x*-type ions (e.g. $a/x \pm 1$ due to hydrogen migration), and therefore searching scores are elevated compared to UVPD, which uses only one form of *a*- and *x*- ions. Preferential cleavages at proline, glutamic acid, and aspartic acid that yield specific Y- and y-type ions, and *c*/*z* ions are also included in searches for UVPD.[44]

Using these modified MassMatrix algorithms, LC-MS/MS data sets were searched against the cow International Protein Index (IPI v3.73, Sept 27, 2011) modified to include the chicken lysozyme C sequence using the MassMatrix database search engine v2.4.0. LC-MS/MS were performed in triplicate, and the data from the replicate experiments were merged prior to database searching. Search parameters utilized were trypsin, 2 missed cleavages, phosphorylation of S and T as variable modifications, iodoacetamide/carbamidomethylation of C as a fix modification, precursor ion tolerance of 2 Da and a product ion tolerance of 1 Da. Peptide hits were filtered by pp and pp_2

scores less than 5 and pp_{tag} scores less than 0.5[48] for which the pp and pp2 scores are a statistical measure of the number of matched product ions and the total abundance of matched product ions, respectively. All suspect peptide spectral matches (PSMs) were manually verified.

3.4 Results and Discussion

Figure 3.1 shows NETD (A and C) and UVPD (B and D) spectra of doubly and triply deprotonated KPLLIIAEDVEGEY. NETD of the doubly charged peptide results in production of eight *x*-ions and two *a*-ions, corresponding to ten out of thirteen possible backbone cleavage sites. The dominant dissociation pathways include the neutral loss of CO_2 and non-dissociative electron transfer (NETnoD), the latter yielding intact charge-reduced precursor ions (labeled as $[M-2H]^{-\bullet}$). The most abundant sequence-informative products are the x_{10}, x_{11}, x_{12} and x_{13} ions. A fluoranthene adduct of the intact precursor is also observed in all NETD spectra and is denoted by the label Δ in Figures 3.1A and C. The *m/z* value of the fluoranthene adduct species can be accurately predicted based on the consistent mass shift of 202 Da and thus can be subtracted from the spectra prior to the database searches. UVPD of the doubly charged species yields an extensive array of sequence-informative product ions. Whereas NETD exclusively produced *a*- and *x*-type product ions, UVPD yields *b*-, *y*-, *c*-, *z*-, *Y*-, *d*- and *w*-type product ions in addition to the dominant *a*- and *x*-ions. The large array of product ions upon UVPD yielded 100% sequence coverage as well as complementary N- and C-terminal product ions (i.e., *a*/*x* product ion pairs) for nine of the thirteen backbone cleavage sites.

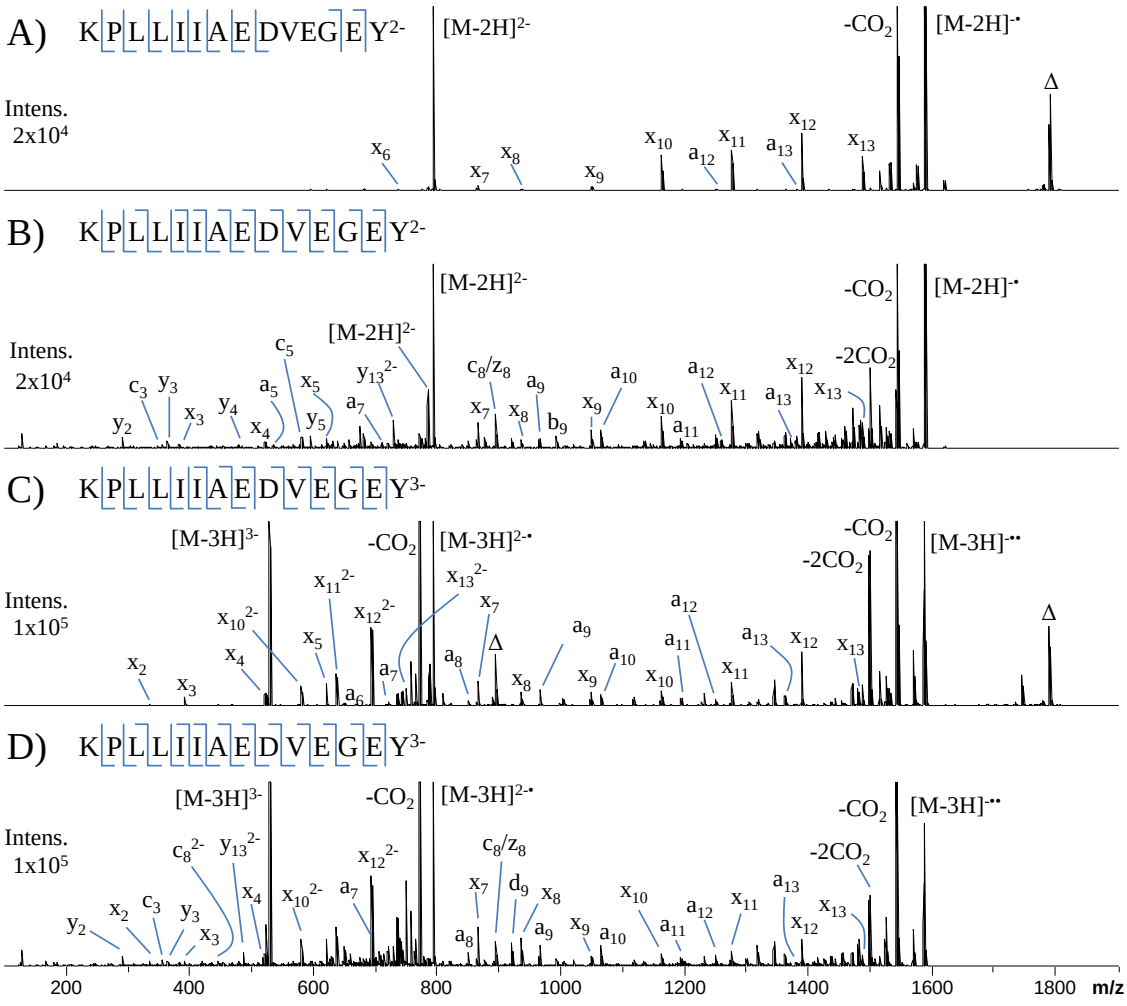

Figure 3.1 NETD (A and C) and UVPD (B and D) spectra of doubly and triply charged KPLLIIAEDVEGEY. Δ indicates fluoranthene adduction to the precursor.

UVPD also promotes the loss of CO_2 from the precursor as well as charge-reduction by photo-induced electron detachment, a pathway noted in our previous UVPD study of peptide anions.[44] NETD of the triply charged peptide in Figure 3.1C provides complete sequence coverage and complementary *a*- and *x*-ions for seven of the thirteen peptide backbone cleavage sites. The triply charged peptide yielded five additional *a*-ions and four *x*-ions relative to the doubly charged peptide. Although the *a/x* ion series has

higher abundance and affords greater coverage than observed for the corresponding doubly charged peptide, the loss of CO_2 and NETnoD remain the most dominant pathways. The fragmentation pattern produced by UVPD of the triply charged species is similar to that of the doubly charged species. Complete sequence coverage is achieved, and complementary *a*/*x* ion pairs are observed for nine of the thirteen peptide backbone cleavage sites.

For the peptide KPLLIIAEDVEGEY, charge reduction accounts for 59% and 42% of product ions (based on peak area) produced by NETD of the 2- and 3- species, respectively. Similarly, UVPD resulted in 31% and 30% charge reduction for the 2- and 3- species, respectively. Neutral loss of CO_2 from the precursor accounts for 24% and 34% of the product ions upon NETD and 15% and 25% upon UVPD for the 2- and 3- species, respectively. UVPD produced less extensive charge reduction and CO_2 loss from the precursor than NETD; however small molecule neutral losses (CO_2, NH_3 and H_2O) from sequence ions were more notable upon UPVD. NETD produced only one detectable neutral loss sequence ion (x_7 - CO_2) for the doubly charged peptide. The loss of CO_2 loss was observed from the x_7, x_9, x_{10}, x_{11}, x_{12} and x_{13} ions upon NETD of the triply charged peptide, as well as ammonia loss from a_{13}. CO_2, NH_3 and H_2O losses occurred for a majority of the sequence ions upon UVPD, with even multiple neutral losses in some cases. This outcome is attributed to the energetics of NETD versus UVPD. Polfer and coworkers estimated the recombination energy of electron transfer with fluoranthene radical cations to be 2.4 – 4.5 eV depending on the site of electron abstraction.[42] In the same study they also reported that the more energetic electron abstraction by Xe radical cations (estimated recombination energy of 6.7-8.7 eV) vastly increased the number of fragment ions observed.[42] However, the additional fragment ions mostly corresponded to sequential neutral losses from $a^{\bullet}$ and *x*-ions instead of new diagnostic products arising

from different fragmentation channels (i.e. b, c, y, z-ions).[42] The greater energy deposition upon absorption of a single 193 nm photon (6.4 eV) and wider array of fragmentation channels accessible compared to electron abstraction by fluoranthene radical cations accounts for the larger array of product ion types as well as the greater extent of neutral losses. UVPD product ions may also possess greater internal energy upon formation, thus contributing to more extensive secondary dissociation and thus elevated levels of neutral losses compared to NETD.

The NETD (A and C) and UVPD (B and D) spectra of the doubly and triply deprotonated phosphopeptide TSTEPQyQPGNEL are illustrated in **Figure 3.2**. The phosphorylation site can be pinpointed to the tyrosine residue based on the NETD and UVPD spectra for both the doubly and triply charged species. However, for the doubly charged species, UVPD affords greater confidence in the identification of the sequence and the localization of the phosphorylation site as a result of the production of peptide backbone cleavages both N-terminal and C-terminal to the phosphorylated residue and complete sequence coverage of the rest of the peptide. NETD did not result in the loss of the phosphate group from the precursor nor from any of the sequence ions. Importantly, UVPD produced $H_2PO_4^-$ and PO_3^- reporter ions that can be used confirm peptide phosphorylation in bottom-up proteomic experiments. The UPVD spectrum of singly charged TSTEPQyQPGNEL is shown in **Figure 3.3**. UVPD of the singly charged species resulted in nearly complete series of *a* and *x* ions as well as confident assignment of the site of phosphorylation. In this case, the phosphate reporter ions fall below the low mass cutoff and are not detected, but there is an abundant phosphate neutral loss from the precursor. In general the UVPD spectrum for the singly deprotonated phosphopeptide is more congested with internal ions and neutral losses compared to the UVPD mass spectra obtained for the corresponding doubly and triply charged species.

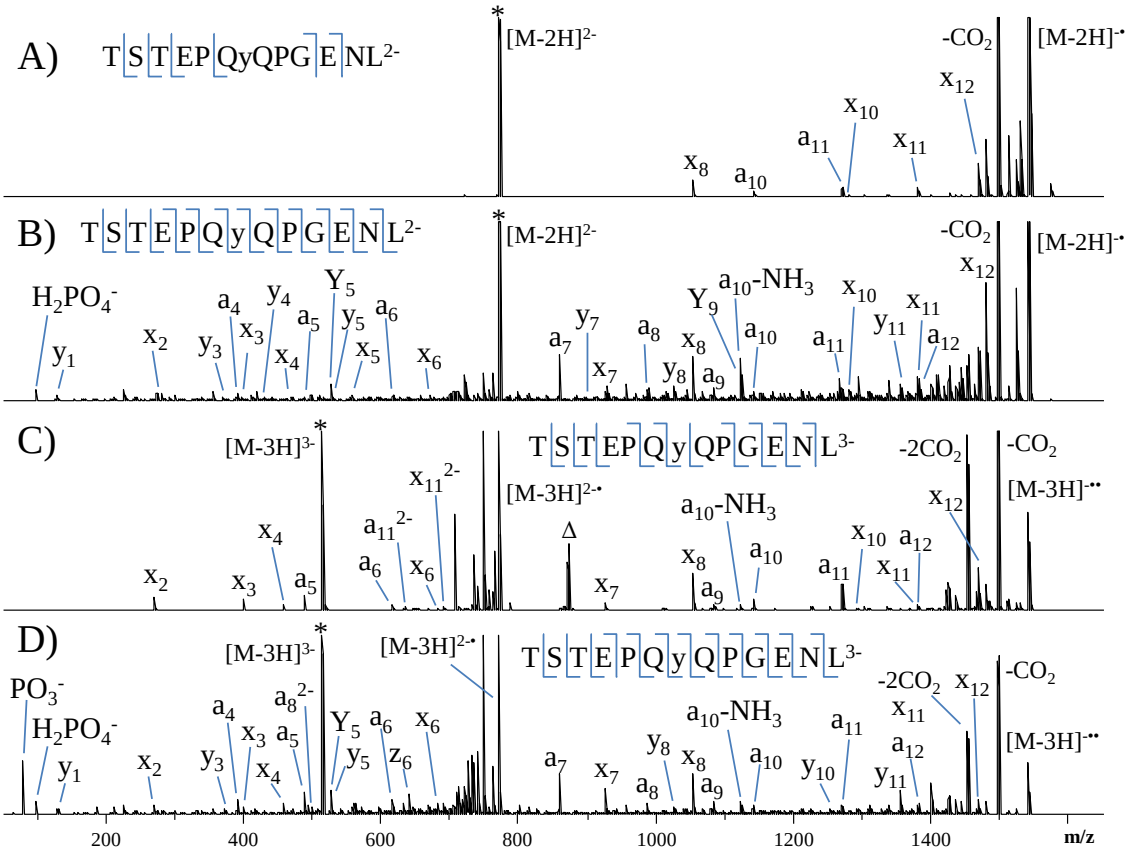

Figure 3.2 NETD (A and C) and UVPD (B and D) spectra of doubly and triply charged TSTEPQyQPGENL (y indicates phosphorylated tyrosine). * indicated the precursor ion.

Since activation/dissociation by NETD is more highly dependent on the peptide charge state than UVPD, a comparison of NETD and UVPD was undertaken for six doubly and triply deprotonated peptides. Sequence coverage distributions were calculated for each peptide and charge state based on the summed peak areas of complementary *a/x* ion pairs normalized to the most abundant *a/x* ion pair.

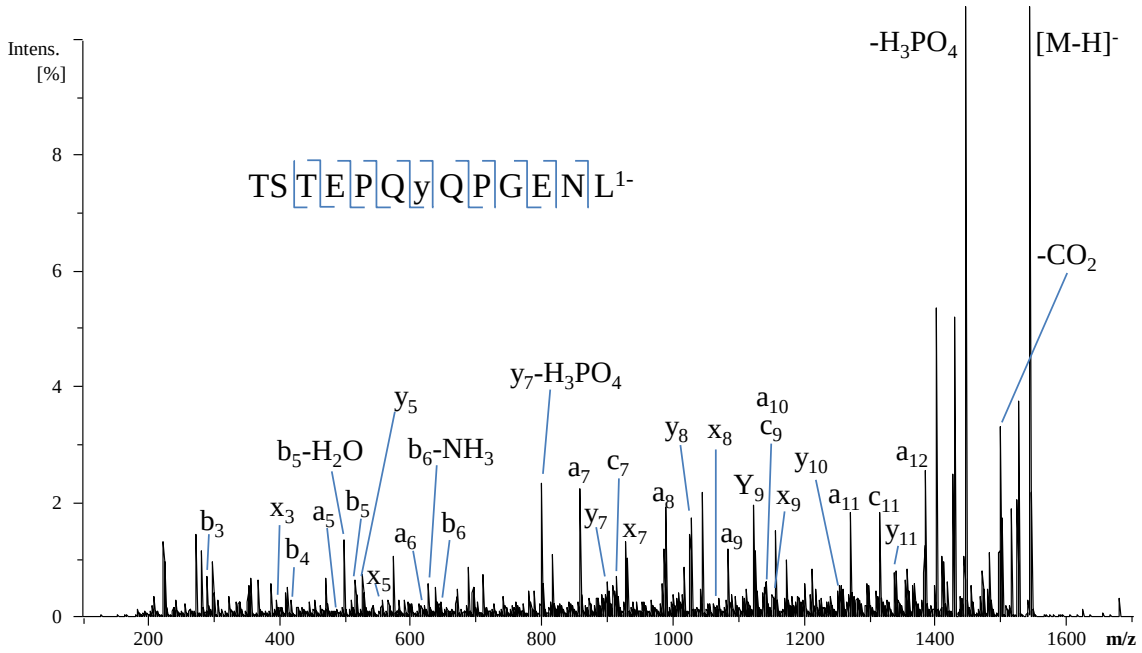

Figure 3.3 UVPD spectrum of singly deprotonated TSTEPQyQPGENL.

Only product ions with S/N ≥ 3 were included in the sequence coverage distributions. The results are shown in **Figure 3.4**. In general, NETD performs poorly for the doubly deprotonated peptides as witnessed by the rather sparse bar graphs in **Figure 3.4**. For these cases, the most dominant pathways produced *x*-ions corresponding to cleavage near the N-terminus with very few low abundance *a*-ions. In fact, there are zero complementary *a*- and *x*-ions pairs detected for the doubly charged precursors, signifying notable localization of the charge sites upon NETD. NETD of the triply deprotonated species yields greater abundances and broader distributions of backbone cleavages compared to the doubly charged species, although still occurring with notable gaps in coverage. UVPD produced substantially broader distributions of peptide backbone cleavages for both the doubly and triply charged species of all the peptides analyzed. For the doubly charged species, UVPD produced greater product ion abundances for nearly every possible peptide backbone cleavage site compared to NETD. This is partially

attributed to the production of complementary *a*- and *x*-ion pairs for most backbone cleavage sites upon UVPD, whereas NETD predominantly produced *x*-ions. **Table 3.1** summarizes the sequence coverages obtained by NETD and UVPD for the same six peptides in the 2- and 3- charge states. UVPD produced 100% sequence coverage for all the peptides (both doubly and triply charged species). NETD produced average sequence coverage of 37% for the doubly charged species and 83% for the triply charged species.

To enable a more extensive comparison of the performance of NETD and UVPD for protemoics-type experiments, as well as benchmarking these MS/MS methods relative to conventional positive mode CID, LC-MS/NETD and LC-MS/UVPD in addition to positive ion mode LC-MS/CID experiments were undertaken for tryptic digests of a model protein mixture containing bovine α-casein, β-casein, β-lactoglobulin A, BSA, ribonuclease A and chicken lysozyme C with automated database searching via MassMatrix. These proteins were chosen for three reasons: they are well characterized model systems, α-casein and β-casein are highly phosphorylated, and the theoretical isoelectric points (based on amino acid content) of these proteins range from ~4.8 for β-lactoglobulin to ~9.3 for lysozyme C and thus should not bias results for either positive or negative mode analyses. **Table 3.2** summarizes the results obtained from the NETD, UVPD and CID data sets. UVPD produced overall higher sequence coverage percentages and greater numbers of unique peptides compared to that of NETD. This outcome can be attributed to the ability of UVPD to efficiently sequence both lower and higher charge states with rapid activation times (4 ms). In contrast, NETD in general yields the best results for higher charged states ($n \geq 3$), and uses fairly modest activation times (250 ms), which can limit the number of MS/MS spectra collected.

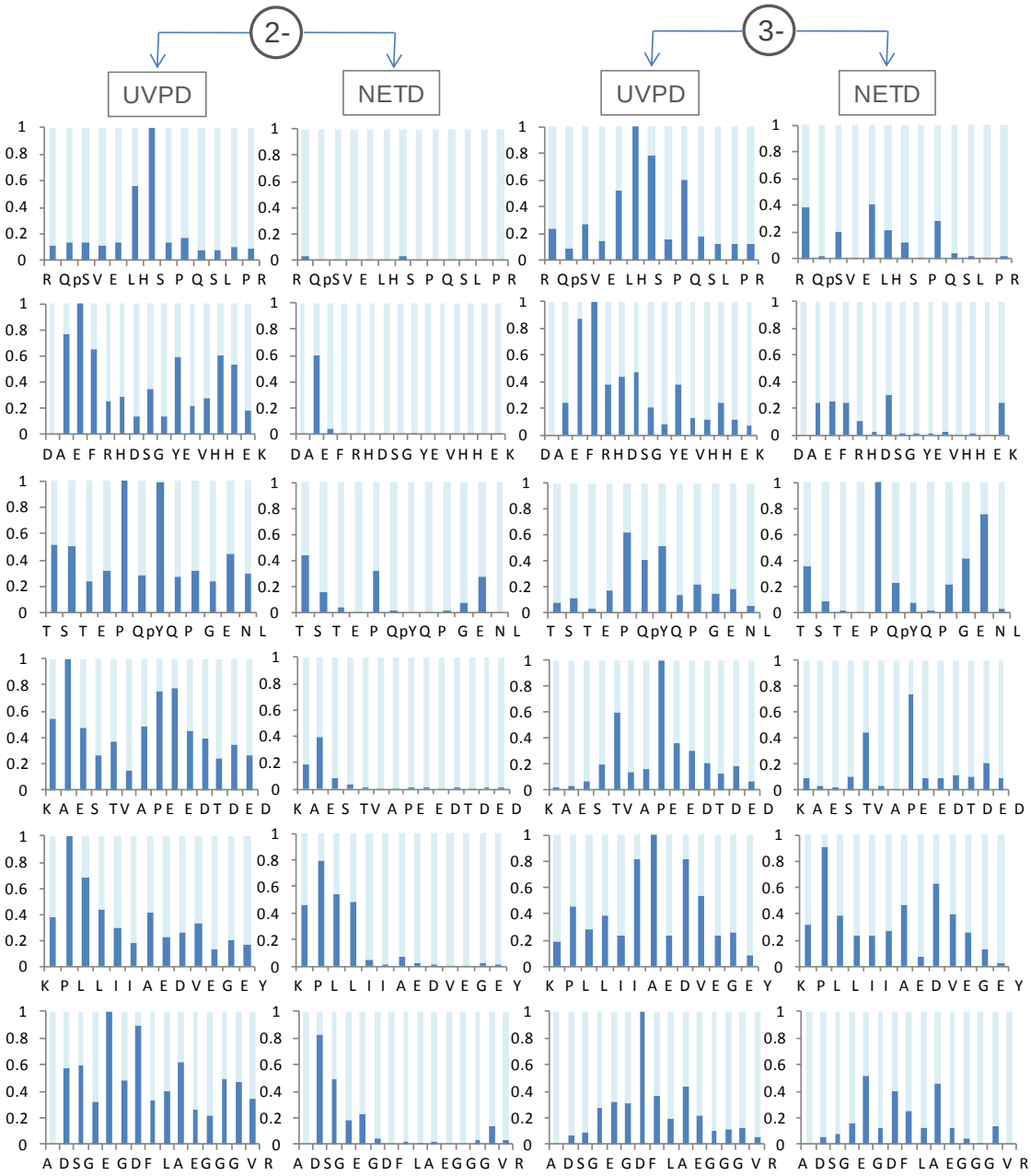

Figure 3.4 Sequence coverage distribution comparison between UVPD and NETD of doubly and triply deprotonated peptide anions. The product ion signals were plotted for summed *a*/*x* pairs per backbone cleavage. For each peptide, ion signals were normalized to the highest *a*/*x* pair between the UVPD and NETD methods.

Peptide Sequence	% Sequence Coverage			
	$[M - 2H]^{2-}$		$[M - 3H]^{3-}$	
	NETD	UVPD	NETD	UVPD
RQsVELHSPQSLPR	15	100	77	100
DAEFRHDSGYEVHHEK	13	100	60	100
TSTEPQyQPGENL	50	100	75	100
KAESTVAPEEDTDED	43	100	93	100
KPLLIIAEDVEGEY	62	100	100	100
ADSGEGDFLAEGGGVR	40	100	93	100

Table 3.1 Sequence coverage comparison of NETD and UVPD of peptide anions. The percentages were calculated from *a*- and *x*-type ions with S/N ≥ 3, and represent the number of observed backbone cleavages divided by the total possible backbone cleavages per peptide.

Protein	% Sequence			Unique Peptides		
	NETD	UVPD	CID	NETD	UVPD	CID
BSA	51	58	70	26	33	51
αS1-casein	35	43	32	9 (4/1)*	12 (4/1)*	9 (2/1)*
αS2-casein	38	32	31	6 (4/4)*	9 (4/4)*	12 (3/2)*
β-lactoglobulin A	64	56	48	7	8	11
β-casein	15	24	25	2 (1/1)*	6 (1/1)*	6 (2/1)*
Lysozyme C	24	43	81	2	4	10
Ribonuclease A	13	42	52	1	3	5

* unique peptides (phosphopeptides/phosphorylation sites)

Table 3.2 MassMatrix comparison of negative mode NETD and UVPD and positive mode CID LC-MS/MS experiments.

As expected due to higher signal intensities of tryptic peptides (ones that contain at least two protonation sites) in the positive mode, CID generally yielded the most unique peptide identifications and sequence coverage compared to both negative mode MS/MS methods, except for three of the model proteins, αS1-casein, αS2-casein, and β-

lactoglobulin A. Both UVPD and NETD, however, often out-performed CID for sequencing and identifying acidic peptide species. UVPD identified a total of nine phosphopeptides corresponding to six unique phosphorylation sites, and NETD also identified nine phosphopeptides corresponding to six unique phosphorylation sites. CID identified seven phosphopeptides corresponding to four unique phosphorylation sites. The Venn diagram in **Figure 3.5** shows the overlap in unique peptides identified in the negative mode by NETD and UVPD and in the positive mode by CID for the protein mixture. A significant overlap in peptide matches was observed between NETD and UVPD with 83% of those identified by NETD also identified by UVPD. UVPD and NETD produced 13 and 12 unique peptides, respectively, that were not identified by CID in positive mode.

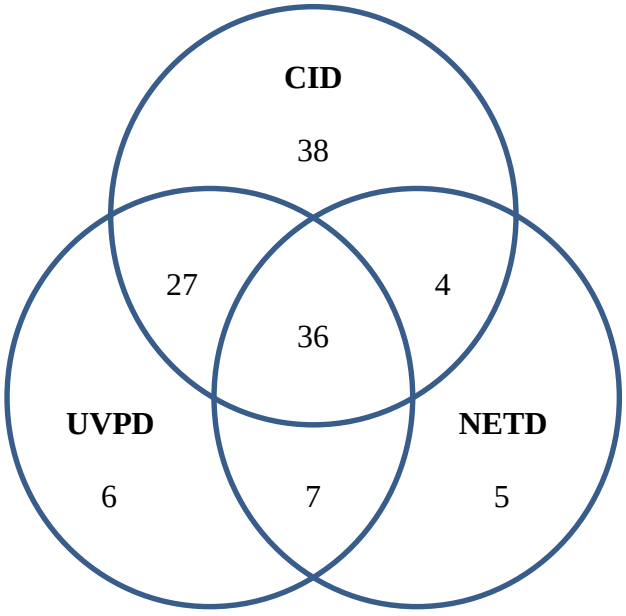

Figure 3.5 Venn diagram comparing the overlap in unique peptides identified in the negative mode by NETD and UVPD and in the positive mode by CID.

Lastly, database searches were undertaken by combining the spectral files for UVPD and positive CID as well as NETD and positive CID to investigate in more detail the overlap between traditional CID and the negative mode methods - these results are illustrated in **Table 3.3**. In all cases except for lysozyme C and ribonuclease A, the combined searches significantly increased the sequence coverage percentages and unique peptide identifications over that obtained by CID alone. The combined NETD/CID searches outperformed the UVPD/CID searches for BSA, but underperformed for β-casein. However, even these search outcomes were fairly similar (i.e., the number of unique peptides and sequence coverage percentages were close if not identical). The integrated NETD/CID and UVPD/CID results were similar for most of the other proteins. The variations that were seen between the two different searching strategies are attributed to the differences in optimum ion charge state for each MS/MS method and the corresponding population of charge states generated for the tryptic peptides of each protein (as described in previous sections). To investigate this hypothesis in more quantifiable detail, the charge state distributions of the peptide spectral matches (PSMs) produced by NETD and UVPD are tabulated in **Figure 3.6**. For UVPD, the charge state distribution ($\bar{x}$ = -2.1, s = 0.72, n = 578) is centered about the 2- charge state which corresponded to 54% of the PSMs. In contrast, the charge state distribution for NETD ($\bar{x}$ = -2.9, s = 0.60, n = 259) is centered about the 3- charge state, which corresponded to 63% of the PSMs. These differences in peptide preference highlight the complementary nature of the two MS/MS methods for peptide anions and are responsible for many of the observed variations in the acidic peptide identifications.

Protein	% Sequence		Unique Peptides	
	NETD/CID	UVPD/CID	NETD/CID	UVPD/CID
BSA	77	73	54	54
αS1-casein	47	46	13 (4/1)*	13 (4/1)*
αS2-casein	46	36	14 (4/4)*	15 (5/4)*
β-lactoglobulin A	63	63	14	13
β-casein	25	29	6 (2/1)*	7 (2/1)*
Lysozyme C	81	81	10	10
Ribonuclease A	52	52	5	5

* unique peptides (phosphopeptides/phosphorylation sites)

Table 3.3 MassMatrix comparison of NETD/CID and UVPD/CID combined positive and negative mode data sets.

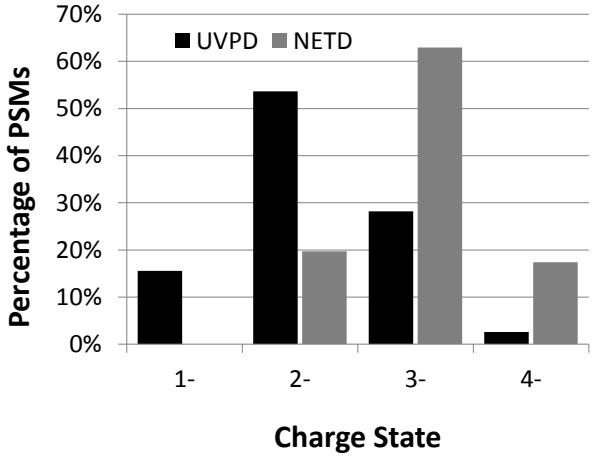

Figure 3.6 Charge state distribution of peptide spectral matches (PSMs) produced by NETD and UVPD for a protein mixture. The total number of PSMs for NETD and UVPD are 259 and 578, respectively.

3.5 Conclusions

UVPD and NETD were compared with respect to sequence coverage distributions (a measure of product ion signals across the peptide backbone), sequence coverage percentages, backbone cleavage specificity, and both LC-MS/MS and automated database searching performance metrics for analysis of peptide anions. The substantially shorter activation times and compatibility with a wider range of charge states allowed the collection of a greater number of MS/MS spectra during a typical a LC gradient for UVPD compared to NETD. However, the spectra that were collected for NETD were generally from peptide species that contained several acidic residues and/or acidic post translational modifications – ones that are more difficult for traditional positive mode CID to efficiently and accurately elucidate. Importantly, when combined with positive mode CID, both UVPD and NETD significantly increased the sequence coverage percentages and unique peptide identifications over that of just CID alone, demonstrating their utility and potential for characterizing acidic proteomes.

3.6 References

(1) Hunt, D. F.; Buko, A. M.; Ballard, J. M.; Shabanowitz, J.; Giordani, A. B. *Biol. Mass Spectrom.* **1981**, *8*, 397–408.

(2) McLuckey, S. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599–614.

(3) Eng, J. K.; McCormack, A. L.; Yates III, J. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976–989.

(4) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. *ELECTROPHORESIS* **1999**, *20*, 3551–3567.

(5) De Godoy, L. M. F.; Olsen, J. V.; Cox, J.; Nielsen, M. L.; Hubner, N. C.; Frohlich, F.; Walther, T. C.; Mann, M. *Nature* **2008**, *455*, 1251–1254.

(6) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.

(7) Kelleher, N. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250–4253.

(8) 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.

(9) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528 –9533.

(10) Coon, J. J.; Syka, J. E. P.; Schwartz, J. C.; Shabanowitz, J.; Hunt, D. F. *Int. J. Mass Spectrom.* **2004**, *236*, 33–42.

(11) Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. *Mol. Cell. Proteomics* **2007**, *6*, 1942 –1951.

(12) Bowers, W. D.; Delbert, S. S.; Hunter, R. L.; McIver, R. T. *J. Am. Chem. Soc.* **1984**, *106*, 7288–7289.

(13) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809–2815.

(14) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 2067–2070.

(15) Wilson, J. J.; Brodbelt, J. S. *Anal. Chem.* **2006**, *78*, 6855–6862.

(16) Morgan, J. W.; Hettick, J. M.; Russell, D. H. In *Biological Mass Spectrometry*; Academic Press, 2005; Vol. Volume 402, pp. 186–209.

(17) Kim, T. Y.; Thompson, M. S.; Reilly, J. P. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 1657–1665.

(18) Ly, T.; Julian, R. R. *Angew. Chem. Int. Edit.* **2009**, *48*, 7130–7137.

(19) Reilly, J. P. *Mass Spectrom. Rev.* **2009**, *28*, 425–447.

(20) Ly, T.; Julian, R. R. *J. Am. Chem. Soc.* **2008**, *130*, 351–358.

(21) Morgan, J. W.; Russell, D. H. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 721–729.

(22) Thompson, M. S.; Cui, W.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1439–1452.

(23) Kim, T.-Y.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2334–2341.

(24) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9*, 4205–4214.

(25) Shin, Y. S.; Moon, J. H.; Kim, M. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 53–59.

(26) Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2001**, *22*, 197–206.

(27) Yoon, S. H.; Chung, Y. J.; Kim, M. S. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 645–655.

(28) Janek, K.; Wenschuh, H.; Bienert, M.; Krause, E. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1593–1599.

(29) Gunawardena, H. P.; Emory, J. F.; McLuckey, S. A. *Anal. Chem.* **2006**, *78*, 3788–3793.

(30) Ewing, N.; Cassady, C. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 105–116.

(31) Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Mass Spectrom. Rev.* **2002**, *21*, 87–107.

(32) Steinborner, S. T.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 253–258.

(33) Jensen, O. N. *Nat. Rev. Mol. Cell. Biol.* **2006**, *7*, 391–403.

(34) Schwartz, R.; Ting, C. S.; King, J. *Genome Res.* **2001**, *11*, 703 –709.

(35) Knight, C. G.; Kassen, R.; Hebestreit, H.; Rainey, P. B. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8390 –8395.

(36) Weiller, G. F.; Caraux, G.; Sylvester, N. *Proteomics* **2004**, *4*, 943–949.

(37) Kiraga, J.; Mackiewicz, P.; Mackiewicz, D.; Kowalczuk, M.; Biecek, P.; Polak, N.; Smolarczyk, K.; Dudek, M.; Cebrat, S. *BMC Genomics* **2007**, *8*, 163.

(38) Kjeldsen, F.; Silivra, O. A.; Ivonin, I. A.; Haselmann, K. F.; Gorshkov, M.; Zubarev, R. A. *Chem. Eur. J.* **2005**, *11*, 1803–1812.

(39) Kjeldsen, F.; Horning, O. B.; Jensen, S. S.; Giessing, A. M. B.; Jensen, O. N. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1156–1162.

(40) Larraillet, V.; Vorobyev, A.; Brunet, C.; Lemoine, J.; Tsybin, Y. O.; Antoine, R.; Dugourd, P. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 670–680.

(41) Cook, S.; Jackson, G. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 1088–1099.

(42) Huzarska, M.; Ugalde, I.; Kaplan, D. A.; Hartmer, R.; Easterling, M. L.; Polfer, N. C. *Anal. Chem.* **2010**, *82*, 2873–2878.

(43) Coon, J. J.; Shabanowitz, J.; Hunt, D. F.; Syka, J. E. P. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 880–882.

(44) Madsen, J. A.; Kaoud, T. S.; Dalby, K. N.; Brodbelt, J. S. *Proteomics* **2011**, *11*, 1329–1334.

(45) Antoine, R.; Joly, L.; Tabarin, T.; Broyer, M.; Dugourd, P.; Lemoine, J. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 265–268.

(46) McAlister, G. C.; Russell, J. D.; Rumachik, N. G.; Hebert, A. S.; Syka, J. E. P.; Geer, L. Y.; Westphall, M. S.; Pagliarini, D. J.; Coon, J. J. *Anal. Chem.* **2012**, *84*, 2875–2882.

(47) Madsen, J. A.; Xu, H.; Horton, A. P.; Boutz, D. R.; Kaoud, T. S.; Dalby, K. N.; Brodbelt, J. S. *submitted*.

(48) Xu, H.; Freitas, M. *BMC Bioinformatics* **2007**, *8*, 133.

(49) Xu, H.; Zhang, L.; Freitas, M. A. *J. Proteome Res.* **2008**, *7*, 138–144.

(50) Xu, H.; Hsu, P.-H.; Zhang, L.; Tsai, M.-D.; Freitas, M. A. *J. Proteome Res.* **2010**, *9*, 3384–3393.

Chapter 4

Activated Ion Negative Electron Transfer Dissociation of Multiply Charged Peptide Anions

4.1 Overview

The implementation and evaluation of activated ion negative electron transfer dissociation (AI-NETD) in order to enhance the analytical capabilities of NETD for the elucidation of multiply deprotonated peptide anions is described. The analytical figures-of-merit and fragmentation characteristics are compared for NETD alone and with supplemental collisional activation of the charge reduced precursors or infrared photoactivation of the entire ion population during the NETD reaction period. The addition of supplemental collisional activation of charge reduced precursor ions or infrared photoactivation of the entire ion population concomitant with the NETD reaction period significantly improves sequencing capabilities for peptide anions as evidenced by the greater abundances of product ions and overall sequence coverage. Neither of these two AI-NETD methods significantly alters the net fragmentation efficiencies relative to NETD; however, the sequence ion conversion percentages with respect to formation of diagnostic product ions are notably higher. Supplemental infrared photoactivation outperforms collisional activation for most of the peptide fragmentation metrics evaluated.

4.2 Introduction

Recent bioinformatics studies have shown that all proteomes contain a bimodal distribution of protein isoelectric points (pI) composed of acidic and basic subsets, and

very few proteins poses pIs near physiological pH.[1,2] Post-translational modifications (PTMs), including phosphorylation, acetylation, sulfation, nitrosylation and glycosylation (sialic acid), as well as the presence of acidic amino acid residues (Asp,Glu) greatly increases the acidity of peptides and proteins. Consequently these species readily ionize in negative ion mode, and in some cases prove difficult to analyze in the positive ion mode by traditional tandem mass spectrometric strategies. In fact, it has been shown previously that almost all peptides, regardless of acidity, produce satisfactory ion abundances in the negative ion mode, whereas acidic peptides are not as easily ionized in the positive ion mode.[3] However, the negative ion mode is rarely utilized for structural determination of peptides and proteins due to the lack of adequate ion activation/dissociation methods.

Although collision induced dissociation (CID) remains the most popular MS/MS method for proteomic applications due to its widespread availability on commercial mass spectrometers and the extensive development of database search algorithms compatible with CID mass spectra, CID is less effective in general for characterization of peptide anions because the resulting fragmentation patterns are dominated by neutral losses (e.g., side-chain groups, NH_3, H_2O, and PTMs) as well as internal product ions.[4,5] The low abundance product ions that are produced by CID of peptide anions are mainly comprised of *c* and *y* ions, in contrast to the conventional *b* and *y* ions derived from peptide cations, along with lower levels of *a, b, x, and z* ions. More substantial sequence information can be obtained for multi-charged peptide anions rather than singly charged anions.[6–10] The ability to generate more informative fragmentation patterns of peptide anions would further expand the capabilities of tandem mass spectrometry for addressing a greater array of proteomic problems, such as those involving the acidic proteome or proteins decorated with acidic post-translational modifications.

Electron-based dissociation methods, specifically electron capture dissociation (ECD)[11–14] and electron transfer dissociation (ETD),[15,16] have gained wide acceptance for their broad applications in both bottom-up and top-down proteomic approaches. In ECD, a subthermal (<0.2 eV) electron is captured by a multiply charged peptide or protein cation. Alternatively, ETD utilizes a radical reagent anion (e.g., fluoranthene) to transfer an electron to a multiply charged peptide or protein cation. Each of these electron-based methods promotes exothermic charge reduction that may occur in conjunction with bond cleavages to yield *c* and *z* sequence ions. In comparison to CID, ECD and ETD show far less dependence on amino acid composition as they are non-ergodic processes.[11,17–19] Perhaps the most attractive characteristic of these electron-based activation methods is that labile post-translation modifications are largely left intact during the dissociation process, making these methods effective for mapping post-translational modifications.[20–24] However, as precursor *m/z* increases and/or charge state decreases, nondissociative electron capture/transfer processes becomes the dominant pathway, resulting in charge-reduced ions in which the backbone is cleaved but non-covalent bonds hold the product ions together.[25,26] Consequently, the secondary and tertiary structure of peptides and proteins greatly influences electron capture/transfer efficiencies and product ion yields.[27] A number of methods have been introduced to disrupt the secondary and tertiary structure of peptides and proteins and increase product ion yields from ECD and ETD. Ion activation prior to or during the ECD/ETD reaction period by low level collisional activation, infrared irradiation or thermal activation have been shown to substantially increase product ion yield for both ECD and ETD.[28–34] These methods are collectively termed “activated ion” ECD/ETD (AI-ECD and AI-ETD). An additional benefit of some of the activated ion methods is decreased hydrogen atom migration within the long-lived product ion complexes after the peptide backbone has

been cleaved, thus leading to lower abundances of fragment ions that contain extra or fewer hydrogen atoms. Low-level resonant collisional excitation of the charge-reduced species of doubly protonated peptides concomitant to ETD (also termed ETcaD) has been shown to increase H atom migration,[32] whereas concomitant infrared irradiation has been shown to reduce H atom migration and yielded substantially more peptide spectral matches for the analysis of complex mixtures compared to ETD and ETcaD.[33,34]

The electron-based dissociation methods have been less widely adapted for the analysis of multiply charged peptide and protein anions. Zubarev and coworkers introduced electron detachment dissociation (EDD) in which irradiation of peptide poly-anions with fast (>10 eV) electrons induced electron detachment and subsequent preferential cleavage of the C_α – C bonds, thus producing radical $a^\bullet$ and even electron *x* ions as well as *c*- and *z*- type ions.[35,36] Most importantly, like ECD and ETD, EDD preserves post-translational modifications and has allowed location of the sulfonation sites of the synthetic peptide caerulein[35] and later the phosphorylation site of another synthetic peptide.[36] More than a decade ago McLuckey et al. first demonstrated negative electron transfer dissociation (NETD) through the reaction of multiply deprotonated nucleic acids with rare gas cations.[37,38] Hunt and coworkers applied NETD to the sequence analysis of multiply deprotonated phosphopeptides.[39] In these latter experiments electron abstraction by xenon cations from the peptide anions produced *a*- and *x*-type sequence ions exclusively, but neutral losses of carbon dioxide and phosphoric acid from the precursor and numerous product ions complicated the spectra. The radical cation fluoranthene was compared to the xenon cation as a NETD reagent for sequencing mono- and diphosphopeptides in a recent study by Polfer and colleagues.[40] It was found that fluoranthene radical cations produced far fewer neutral losses and the resulting NETD spectra were thus easier to interpret. These findings were rationalized by the lower

ionization potential (IP) of fluoranthene (IP = 7.9 eV) compared to xenon (IP = 12.1 eV). The recombination energies, based on the difference in the electron affinity of the peptide anion and the ionization potential of the cation, were estimated to be 2.5 – 4.5 eV for fluoranthene and 6.7 – 8.7 eV for xenon. In two recent studies, the Coon group has systematically evaluated the propensities for side-chain losses upon NETD of deprotonated peptides[41] and implemented NETD on an orbitrap mass spectrometer to facilitate the comprehensive mapping of acidic regions of the *S. cerevisiae* proteome.[42]

Based on prior results for ECD and ETD of positively charged ions, it is anticipated that the secondary and tertiary structure of peptide and protein anions should likewise have a significant impact on NETD efficiencies and overall product ion yield. In this study, we present activated ion negative electron transfer dissociation (AI-NETD) with a focus on improving its practical capabilities for the analysis of doubly deprotonated peptide anions. Various analytical figures-of-merit and fragmentation characteristics are compared for NETD alone and with supplemental collisional activation of the charge reduced precursors or infrared photoactivation of the entire ion population during the NETD reaction period.

4.3 Experimental

4.3.1 Materials

All peptides were purchased from AnaSpec (Fremont, CA) and used without further purification. HPLC grade ammonium hydroxide, methanol and water (Fisher Scientific) were used for preparing peptide working solutions. Argon of ≥99.998% purity was purchased from Sigma Aldrich (St. Louis, MO).

4.3.2 Instrumentation

All experiments were performed on a Bruker Daltonics HCT Ultra ETD II ion trap mass spectrometer. An IR beam was introduced via a 25 mm ZnSe window mounted in the vacuum manifold above the ion trap and transmitted through a 2 mm hole in the ring electrode of the quadrupole ion trap. Deep UV enhanced 12.5 mm mirrors (greater than 85% reflectance from 170 nm - 11μm) were placed in 45 degree mounts under and next to the ion trap to reflect the laser beam back out of the mass spectrometer. This configuration allowed facile alignment of the laser through the system by monitoring the intensity and profile of an alignment HeNe diode laser beam exiting the system. Infrared photons were generated by a 50 W CO_2 laser (Synrad, Mukilteo, WA) which was used without focusing or collimating the beam. All optical components were purchased from Edmund Optics (Barrington, NJ). The laser power was controlled by a pulse width modulation signal generated by an UC-2000 laser controller (Synrad, Mukilteo, WA). An external trigger supplied through the auxiliary interface port was passed through an optocoupler to reduce noise, and then amplified to an appropriate level to gate the laser.

The negative chemical ionization (nCI) source normally generates fluoranthene anions by methane chemical ionization. An override in the instrument user interface and replacement of the methane carrier gas with argon allowed efficient production and transmission of fluoranthene radical cations to the ion trap during NETD experiments. Other reagent ionization conditions were the same as for generating fluoranthene anions for ETD. A reagent crucible temperature of 60 °C, a filament current of 1 μA, and an electron energy of 70 eV were used.

4.3.3 Mass Spectrometry

Anions of model peptides were produced by electrospray ionization and were directly infused at a flow rate of 1.5 μL/min. Peptide working solutions were prepared at 10 μM in 70/30 (v/v) of methanol/water with 0.1% ammonium hydroxide. NETD and AI-NETD experiments were performed using the same conditions except for the addition of supplemental activation during the AI-NETD experiments. Peptide anion accumulation was typically less than 5 ms, and precursor isolation was performed with an isolation width of 2 *m/z*, and the low mass cutoff was set at *m/z* 150. Reagent cations were accumulated for 10-20 ms, and an NETD reaction time of 250 ms was utilized unless otherwise noted. These parameters usually reduced the precursor to approximately 50% relative abundance. For AI-NETD experiments utilizing collisional activation, a parameter called "smart decomposition" was enabled. Smart decomposition entails application of gentle collisional activation to the charge reduced precursor during the NETD reaction period. For AI-NETD experiments utilizing IR photoirradiation, low-level photoactivation (3-18 W) was applied during the NETD reaction period.

4.4 Results and Discussion

4.4.1 Production of NETD reagent ions: EI vs. CI

There are two common strategies for generating fluoranthene radical cations utilized for the NETD experiments. One option is to generate the fluoranthene radical cations by electron ionization (EI) as previously reported.[40] Briefly, the carrier gas (typically methane) for the CI source is turned off to suppress the formation of protonated fluoranthene (*m/z* 203) and promote the formation of the radical cation (*m/z* 202). A

second option is to replace the methane carrier gas with argon and generate the fluoranthene radical cations by CI. **Figure 4.1** contains reagent ion spectra generated by EI (A) and argon CI (B). Approximately equal ion abundances of the radical cation species are achieved for both methods, however the ion accumulation time for argon CI was fifty times shorter than for EI (0.3 ms compared to 15 ms). It is also worth noting that fragmentation of fluoranthene during the ionization process is an order of magnitude lower in the argon CI spectrum. This is important because the reagent ion is not mass selected in this experimental set-up. All subsequent experiments were performed using argon CI to generate fluoranthene radical cations.

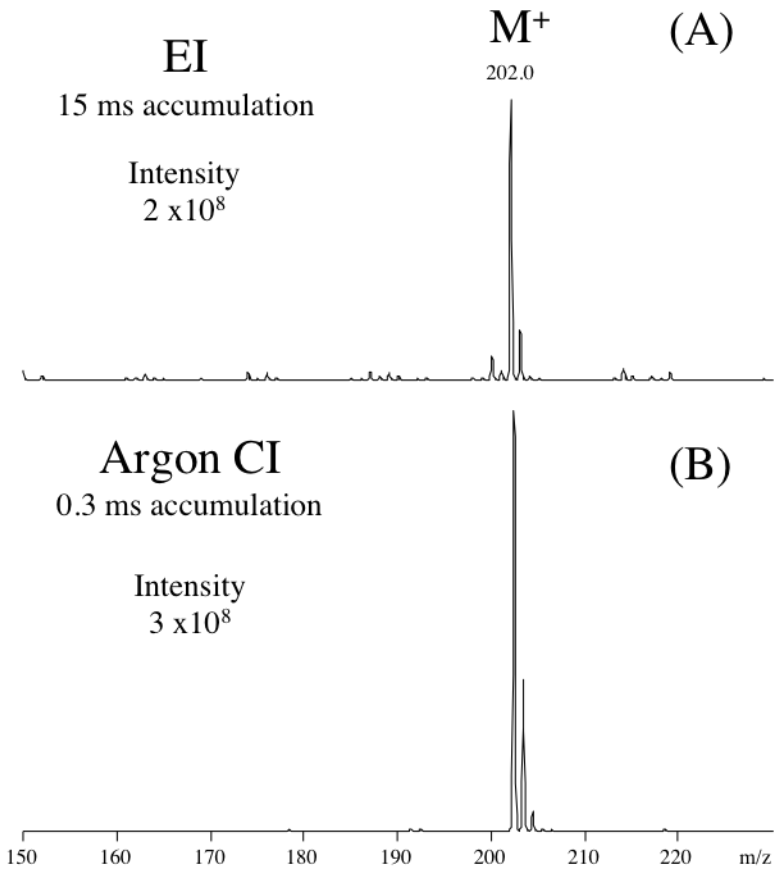

Figure 4.1 Spectra comparing EI (A) and Argon CI (B) of fluoranthene to generate reagent radical cations for NETD.

4.4.2 NETD of Doubly and Triply Deprotonated Peptides

To illustrate the difference in NETD efficiencies between doubly and triply deprotonated peptides, the model peptide ADSGEGDFLAEGGGVR was reacted with fluoranthene radical cations for 250 ms, and the product ion spectra are displayed **Figure 4.2**. The triply charged species produced 25 of the 30 possible *a*- and *x*-type product ions, whereas the doubly charged species produced only five *a*- and *x*-type product ions and yielded very limited sequence coverage. NETD fragmentation efficiencies as defined by:

(sum of product ion peak areas / sum of product ions and surviving precursor peak areas) for the triply and doubly charged species are 90% and 84%, respectively. However, these efficiencies are misleading in terms of sequence informative product ion yields because the most abundant pathways are charge reduction and the neutral loss of 44 Da. The loss of 44 Da from the precursor presumably arises from the loss of CO_2 from the C-terminus or D and E side chains.[39] When the peak areas corresponding to the charge-reduced precursor, loss of CO_2, and fluoranthene adducts are excluded from the product ion sum in the numerator, fragmentation efficiencies in terms of conversion of the precursor to sequence informative product ions for the triply and doubly charged species drop to 21% and 5%, respectively. Charge reduction accounts for 27% and 42% of the product ion current for the triply and doubly charged species, respectively. The loss of CO_2 from the precursor accounts for 45% and 44% of the product ion current for the triply and doubly charged species, respectively. The abundance of charge reduction and loss of CO_2 from the precursor greatly limits the sensitivity and utility of NETD for proteomic analyses. In order to improve the analytical merits of NETD, the amount of charge reduction and neutral loss of CO_2 must be decreased and the number and abundances of peptide backbone cleavage product ions (*a*- and *x*-type ions) must be increased.

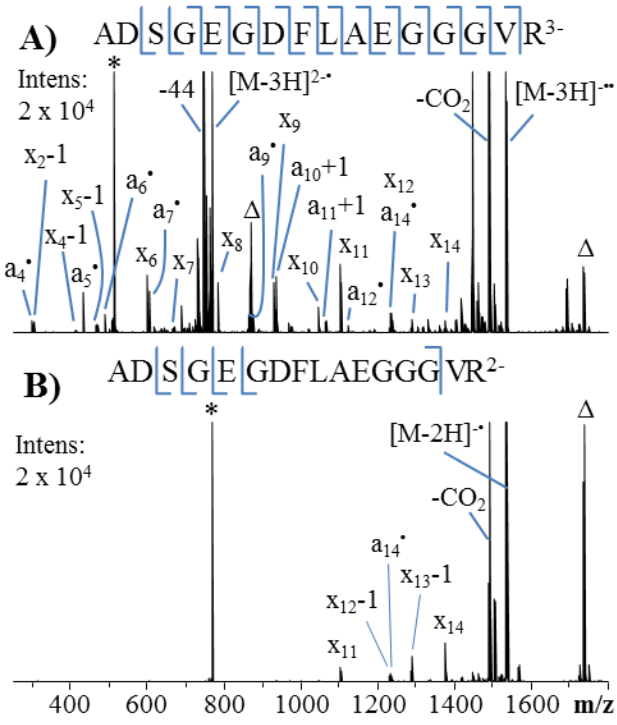

Figure 4.2 NETD spectra of A) triply and B) doubly deprotonated ADSGEGDFLAEGGGVR. 250 ms reaction with fluoranthene radical cations. * and Δ represent the precursor and fluoranthene adducted precursor, respectively.

4.4.3 AI-NETD of Doubly Deprotonated Peptides

The doubly deprotonated peptides EGVNDNEEGFFSAR and TSTEPQyQPGENL were subjected to 250 ms reactions with fluoranthene radical cations with and without simultaneous collisional activation of the charge reduced precursor or infrared photoactivation of the entire ion population. The resulting NETD product ion spectra are displayed in **Figure 4.3 A** and **D** along with the corresponding AI-ETD spectra with simultaneous collisional activation of the charge reduced precursor (**Figure 4.3 B** and **D**) or with 12 watts infrared photoactivation (**Figure 4.3 C** and **F**). Although

the sequence coverage obtained for each peptide upon NETD alone is limited, both AI-NETD modes provide substantially improved sequence coverage. Backbone cleavages are observed between every residue of EGVNDNEEGFFSAR for both AI-NETD methods. Excluding sites N-terminal to proline which are not cleavable by most electron based dissociation methods due to the cyclic structure formed between the side chain and the amide nitrogen, there are three missed cleavages for TSTEPQyQPGENL upon AI-NETD with collisional activation and only a single missed cleavage using AI-NETD with infrared photoactivation.

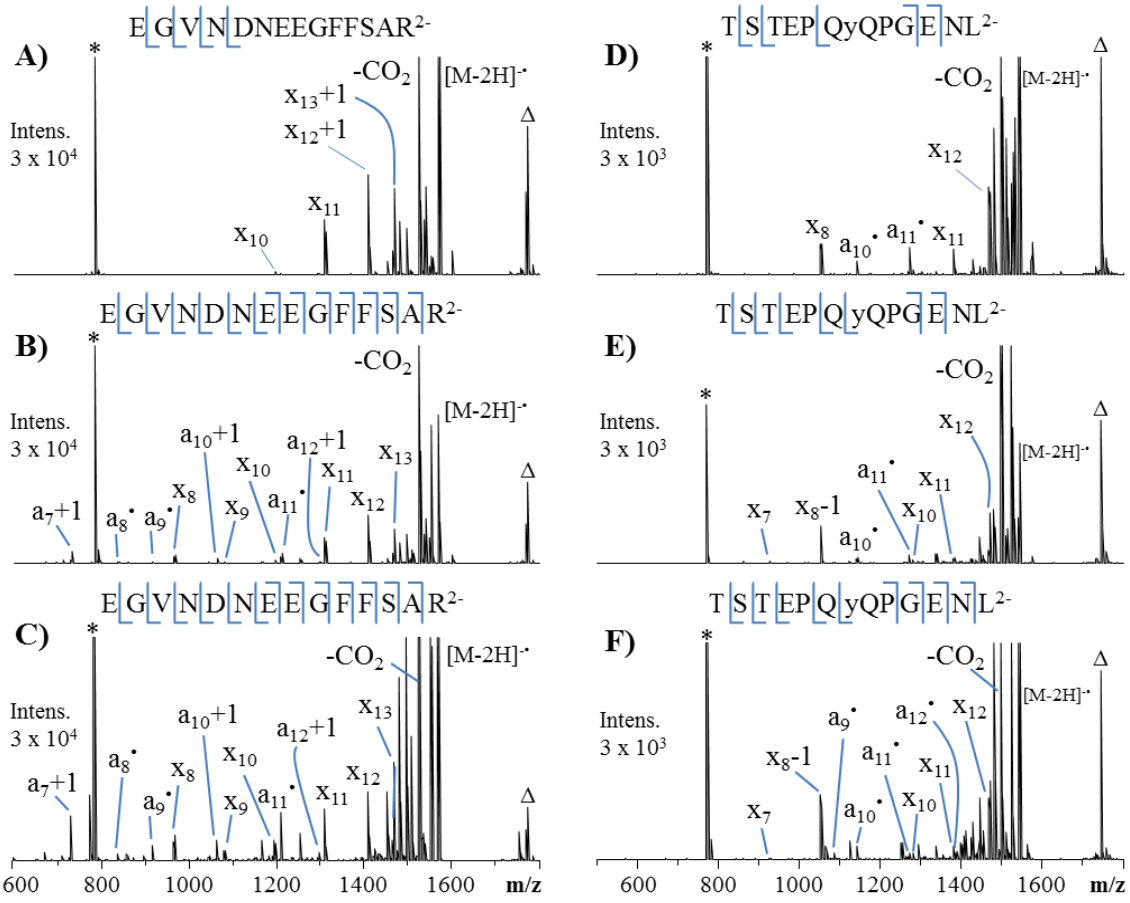

Figure 4.3 NETD (A and D), AI-NETD with collisional activation of the charge-reduced precursor (B and E) and AI-NETD with infrared photoactivation (C and F) product ion spectra of doubly deprotonated EGVNDNEEGFFSAR and TSTEPQyQPGENL. * represent the precursor and Δ represents $[M-2H+Fluoranthene]^{-\bullet}$.

In general the abundances of the diagnostic *a*- and *x*-type product ions are greatest for AI-NETD utilizing infrared photoactivation. This point is illustrated in more detail in **Figure 4.4** in which the sum of the peak areas of *a*- and *x*- ions corresponding to each inter-residue cleavage normalized to the most abundant *a*•/x ion pair are compared for each method for the peptides RQsVELHSPQSLPR and EGVNDNEEGFFSAR. Of the

nine cleavage sites observed for RQsVELHSPQSLR by AI-NETD with collisional activation, AI-NETD with IR photoactivation yielded greater product ion abundances for seven of the sites and promoted cleavages at two additional sites not observed upon AI-NETD with collisional activation. For the peptide EGVNDNEEGFFSAR, AI-NETD with IR photoactivation resulted in greater product ion abundances at every observed cleavage site.

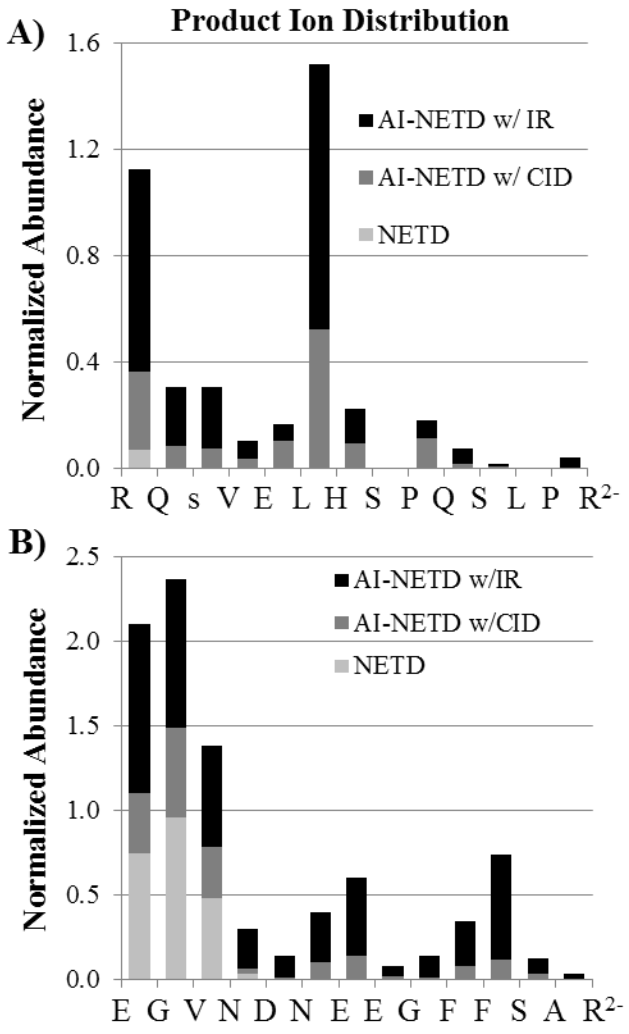

Figure 4.4 Sum of the complementary *a*- and *x*-ion peak areas corresponding to each inter-residue cleavage normalized to the most abundant *a*- and *x*-ion pair for doubly deprotonated peptides RQsVELHSPQSLPR (A) and EGVNDNEEGFFSAR (B). s indicates a phosphorlyated serine residue.

The primary objective of the AI-NETD strategy is to provide the peptide ions with sufficient internal energy to disrupt noncovalent bonds that maintain secondary structure, thus ensuring more efficient conversion of non-dissociated charged reduced species (NETnoD) into diagnostic *a*- and *x*-type product ions. The infrared laser power has a significant impact on the abundances and types of product ions observed upon AI-

NETD. The laser power must be sufficiently high to promote unfolding of the peptides but not so greater as to directly cause dissociation and production of *b*- and *y*-type ions. Energy variable experiments were performed to determine the optimal laser power. Doubly deprotonated RQsVELHSPQSLR and EGVNDNEEGFFSAR were reacted with fluoranthene radical cations for 250 ms during which the ion cloud was irradiated with IR photons. Spectra were acquired using laser powers between 0 and 18 watts. **Figures 4.5 A** and **B** show the distributions of selected (most abundant) product ions as a function of laser power for doubly deprotonated RQsVELHSPQSLR and EGVNDNEEGFFSAR, respectively. From these plots it is evident that the range of optimal laser power falls between approximately 9 and 15 watts with *a*- and *x*-type product ions reaching maximum abundance. This range will vary between instrument types due to differences in operating pressure leading to varying rates of collisional cooling, but the same trend is expected. Above approximately 15 watts laser power, secondary dissociation (predominantly losses of CO_2, H_2O and NH_3) of product ions becomes more significant (**Figure 4.5**), thus accounting for the decay in ion abundances for many of the products. Interestingly, some of the product ions are only observed at laser powers greater than 6-7 watts (i.e. x_{11} and x_{10} of RQsV0ELHSPQSLPR and a_{10} of EGVNDNEEGFFSAR). For both peptides the appearance of the y_{10} ion signifies the onset of photodissociation at approximately 10 watts. Moreover, the loss of the phosphate moiety from the doubly charged precursor of RQsVELHSPQSLPR is observed in low abundance without IR irradiation and up to approximately 13 watts. Above 13 watts of IR irradiation the loss of the phosphate moiety becomes one of the dominant dissociation pathways.

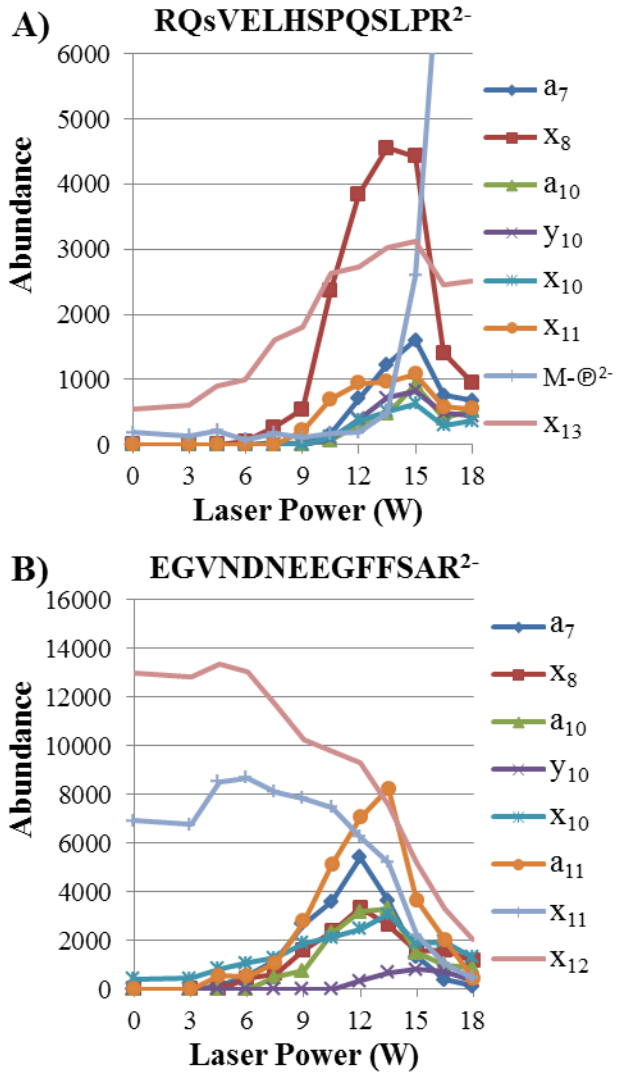

Figure 4.5 Plot of product ion abundance as a function of IR laser power during 250 ms NETD reaction period for doubly deprotonated peptides RQsVELHSPQSLPR (A) and EGVNDNEEGFFSAR (B). M-Ⓟ$^{2-}$ represents the neutral loss of phosphoric acid and lower case s indicates a phosphorylated serine residue.

Once optimized, the AI-NETD methods were applied to a larger set of nine peptides. AI-NETD with IR photoactivation was performed with a laser power of 12 watts, and 250 ms reaction times were used for all experiments. The peptides sequences

and sequence coverage obtain for each peptide analyzed by NETD, AI-NETD with collisional activation of the charge reduced precursor, and AI-NETD with IR photoactivation are summarized in **Table 4.1**. Sequence coverages are greater across the board for both AI-NETD methods compared to NETD alone, and the performance of AI-NETD with IR photoirradiation equals or exceeds AI-NETD with collisional activation for all nine peptides. The mean sequence coverages for these nine peptides were 27%, 79% and 91% for NETD, AI-NETD with collisional activation and AI-NETD with IR photoactivation, respectively. Sequence ion conversion, percent charge reduction and percent neutral loss of CO_2 are shown for each peptide in **Table 4.2**. The overall fragmentation efficiencies are essentially equal for NETD and both AI-NETD methods (81% ± 2.5). However, sequence ion conversion (in which sequence ions are all product ions except the charge reduced precursor, neutral loss of CO_2 from the precursor, and peptide/fluoranthene adducts) is significantly increased for the activated ion methods with supplemental collisional activation yielding 24% conversion and with IR photoactivation yielding 31% conversion compared to only 8% conversion for NETD alone. Overall the amount of charge reduction observed is decreased by more than 50% for AI-NETD with IR photoactivation compared to NETD. This particular metric is not applicable for AI-NETD with collisional activation because the charge reduced precursor is being directly dissociated. The percentage of product ion current attributed to neutral loss of CO_2 is slightly increased upon AI-NETD with IR photoactivation, but is nearly doubled upon collisional activation of the charge reduced precursor.

Peptide Sequence	% Sequence Coverage		
	NETD	AI-NETD w/ CID	AI-NETD w/ IR
ADSGEGDFLAEGGGVR	33	93	93
DAEFRHDSGYQVHHQK	13	60	80
DAFLGSFLYEYSR	33	75	83
DRVYIHPFHL	13	75	100
DRVYIHPFHLVIHN	8	58	75
EGVNDNEEGFFSAR	31	100	100
KAESTVAPEEDTDED	54	92	100
RQsVELHSPQSLPR	8	91	100
TSTEPQyQPGENL	50	70	90
Average	**27 ± 16**	**79 ± 14**	**91 ± 9**

Table 4.1 Sequence Coverage for model peptides produced by NETD, AI-NETD with Collisional Activation and AI-NETD with IR Photoactivation

Peptide Sequence	% Sequence Ion Conversion			% Charge Reduction			% Neutral Loss of CO_2		
	NETD	AI-NETD w/ CID	AI-NETD w/ IR	NETD	AI-NETD w/ CID	AI-NETD w/ IR	NETD	AI-NETD w/ CID	AI-NETD w/ IR
ADSGEGDFLAEGGGVR	16	24	31	26	NA	16	44	56	37
DAEFRHDSGYQVHHQK	5	16	25	71	NA	17	18	64	45
DAFLGSFLYEYSR	8	15	22	65	NA	25	21	65	44
DRVYIHPFHL	3	23	18	79	NA	39	15	64	36
DRVYIHPFHLVIHN	3	13	24	82	NA	27	12	62	40
TSTEPQyQPGENL	9	29	36	44	NA	27	37	44	21
RQsVELHSPQSLPR	6	46	34	67	NA	45	19	28	12
KAESTVAPEEDTDED	14	25	46	22	NA	12	47	52	23
EGVNDNEEGFFSAR	9	27	46	47	NA	19	35	48	18
Average	**8 ± 5**	**24 ± 10**	**31 ± 10**	**56 ± 22**		**25 ± 11**	**28 ± 13**	**54 ± 12**	**31 ± 12**

Table 4.2 Sequence Ion Conversion, Percent Charge Reduction and Percent Neutral Loss of CO_2 for NETD, AI-NETD with Collisional Activation and AI-NETD with IR Photoactivation

As noted above, the neutral loss of CO_2 from the precursors, corresponding to the loss of carbon dioxide from the C-terminus or the side chains of aspartic and glutamic acid residues, is a major dissociation pathway for both NETD and AI-NETD methods. Other side chain and neutral losses are also observed to a lesser extent. **Figure 4.6** shows the expanded regions of the NETD spectra near the charge reduced precursor of DRVYIHPFL upon NETD **(A)**, AI-NETD with collisional activation of the charge reduced precursor **(B)** and AI-NETD with IR photoactivation **(C)**. Low levels of partial or entire side chain losses of arginine, histidine, isoleucine and leucine occur upon NETD. These results are consistent with previously reported observations by Coon and coworkers which included accurate mass determinations and proposed mechanisms for neutral losses by NETD.[41] As exemplified by the spectra in Figure 5, AI-NETD with collisional activation of the charge reduced precursor results in significantly elevated abundances of side chain losses compared to NETD. In contrast, AI-NETD with IR photoactivation produces only slightly elevated levels of side chain losses compared to NETD. Both AI-NETD methods yield tyrosine side chain loss products which are not observed in the corresponding NETD spectrum. This indicates that tyrosine forms a stable radical that is delocalized on the tyrosine side chain upon NETD which does not lead to fragmentation. The additional vibrational excitation of the AI-NETD methods allows hydrogen abstraction and subsequent β-scission yielding side chain loss.[43]

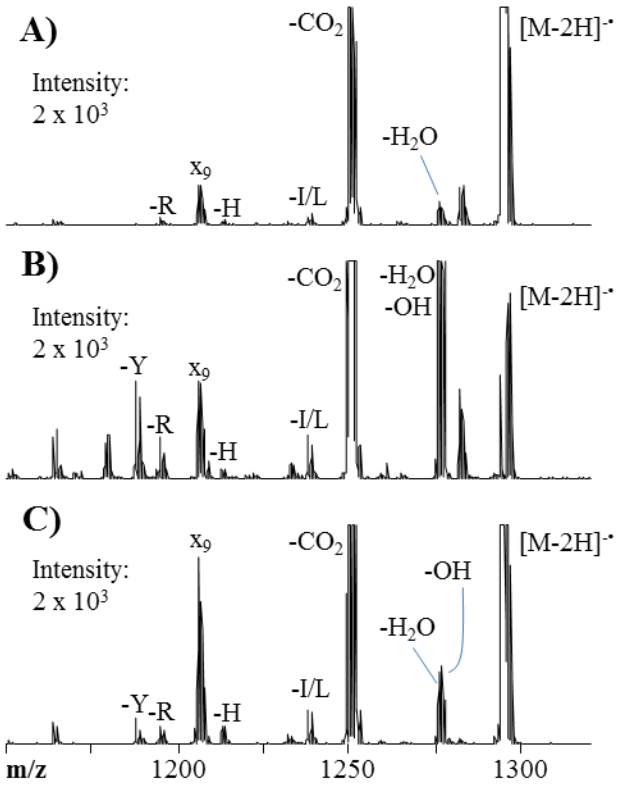

Figure 4.6 Neutral losses from doubly deprotonated DRVYIHPFHL produced by NETD (A), AI-NETD with collisional activation of the charge reduced precursor (B) and AI-NETD with 12W infrared photoactivation (C).

4.4.4 AI-NETD of Triply Deprotonated Peptides

Figure 4.7 shows NETD (**A**) and AI-NETD with IR photoactivation (**B**) spectra of triply deprotonated EGVNDNEEGFFSAR. In contrast to the doubly deprotonated species, NETD of the triply deprotonated species yielded complete sequence coverage and complementary *a*- and *x*-ions for nine of the 13 peptide backbone cleavage sites. AI-NETD with IR photoactivation also yielded the complete sequence coverage and complementary *a*- and *x*-ions for nine of the 13 peptide backbone cleavage sites. Product ion distributions are shown in **Figure 4.8** for triply charged RQsVELHSPQSLPR and

EGVNDNEEGFFSAR. The product ion abundances are greater for 11 of the 13 cleavage sites with AI-NETD with IR photoactivation for EGVNDNEEGFFSAR. For the triply charged species, the abundances of product ions corresponding to each cleavage site are more evenly distributed throughout the peptide compared to the doubly charged species in which fragmentation is most abundant near the N-terminus. AI-NETD with IR photoactivation of RQsVELHSPQSLPR produced greater product ion abundances for all but one cleavage site compared to NETD, and both methods produced complete sequence coverage (excluding N-terminal to proline). In contrast to EGVNDNEEGFFSAR, the product ion abundances for the 3- charge state of RQsVELHSPQSLPR are very similar to the 2- charge state with x_8 and x_{13} ions by far the most abundant product ions. The comparison was extended to the same set of peptides analyzed in the 2- charge state, and in general the AI-NETD methods provide greater product ion abundances, more complementary *a*- and *x*-ion pairs, and slightly better sequence coverage (data not shown).

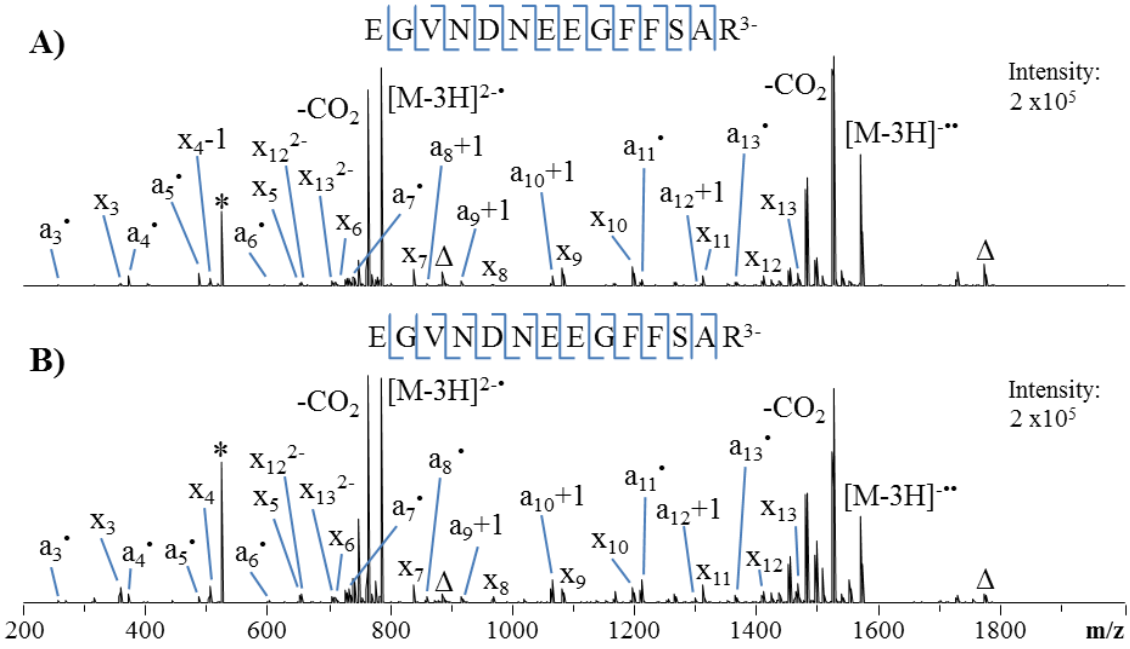

Figure 4.7 NETD (A) and AI-NETD with infrared photoactivation (B) product ion spectra of triply deprotonated EGVNDNEEGFFSAR. * represent the precursor and Δ represents a fluoranthene adduct to the precursor.

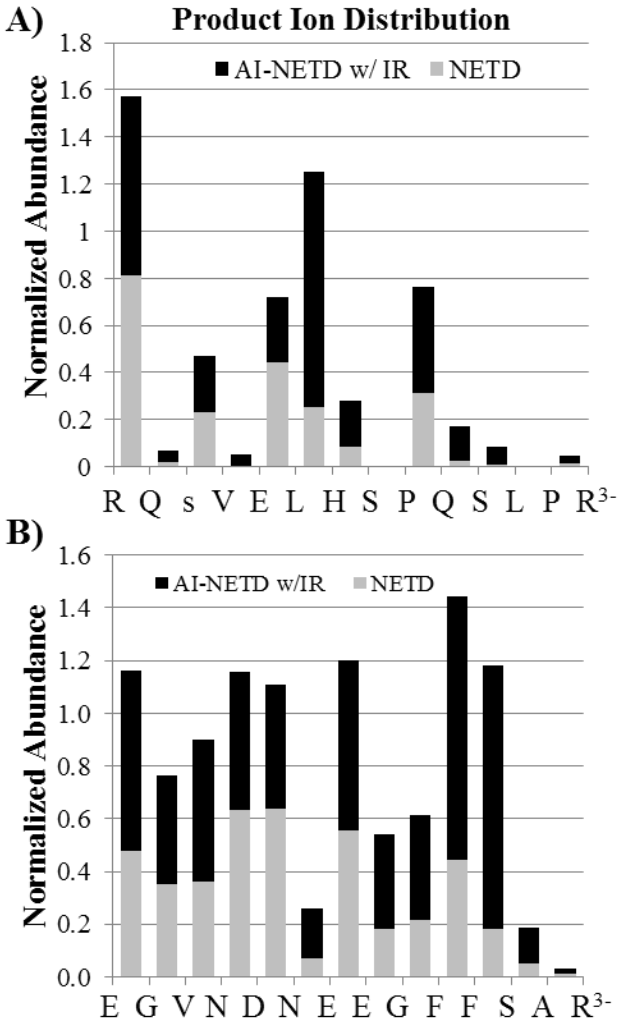

Figure 4.8 Sum of the complementary *a*- and *x*-ion peak areas corresponding to each inter-residue cleavage normalized to the most abundant *a*- and *x*-ion pair for triply deprotonated peptides RQsVELHSPQSLPR (A) and EGVNDNEEGFFSAR (B). s indicates a phosphorlyated serine residue.

4.4.5 Hydrogen Content of Product Ions

One of the key advantages of AI-ETD with IR photoactivation is the reduction of hydrogen atom migration which yields product ion isotopic distributions that more

closely resemble theoretical product ion distributions when compared to ETD or ETcaD. This phenomenon is attributed to the reduction in gas-phase secondary structure which limits intramolecular interactions that would otherwise allow hydrogen atom migration to readily occur.[32–34] However, this trend does not appear to hold true or is not as pronounced for the NETD and AI-NETD methods. For all the peptides analyzed, the amount of hydrogen atom migration appears to be nearly equal for AI-NETD with collisional activation and AI-NETD with IR photoactivation. The *a*- and *x*-type product ions in **Figure 2** that are labeled as either n + 1 or n - 1 have gained or lost a hydrogen atom and are thus shifted one Da relative to the theoretical masses of the radical *a*-ion or even electron *x*-ion. Although we have observed consistent results for the peptides analyzed, a larger dataset is needed to draw broader conclusions on this matter.

4.4.6 LC-MS/MS and MassMatrix Database Search Results

To evaluate each method for a standard proteomics experiment, we performed triplicate LC-MS/MS of a *Halobacterium salinarium* NRC-1 lysate using a 180 min gradient with 0.1% ammonium hydroxide mobile phases. The results from MassMatrix database searches are summarized in **Figure 4.9.** The total number of unique peptides identified by AI-NETD with IR photoactivation , AI-NETD with collisional activation and NETD were 1494, 1046 and 1230, respectively. The number of peptide spectral matches (PSMs) produced by AI-NETD with IR photoactivation were 54, 15 and 27% greater than that produced by NETD for the 2-, 3- and 4- charge states, respectively. AI-NETD with simultaneous collisional activation of the charge reduced precursor produce 16% more PSMs for the 2- charge state compared to NETD; however, AI-NETD with collisional activation produced fewer PSMs than NETD for the 3- and 4- charge states.

Space charge effects could explain why AI-NETD with collisional activation under-performed for the 3- and 4- charge states. Collisional activation of the charge reduced precursor requires charge state determination of the precursor from the initial survey mass spectrum. For intense chromatographic peaks, injection charge control (ICC) may not be able to adequately control the number of charges in the ion trap leading to space charging and poor isotopic resolution for the most abundant species in the survey mass spectrum. At this point we do not have a means to control the number of reagent ions injected (beyond a fixed injection time) during the ion/ion reaction period. Space charging during the ion/ion reaction period can alter ion frequencies, thus resulting in shifts in the frequencies of the ions that are selected for resonant excitation by the collisional activation waveform. Both of these phenomena would cause collisional activation to be ineffective.

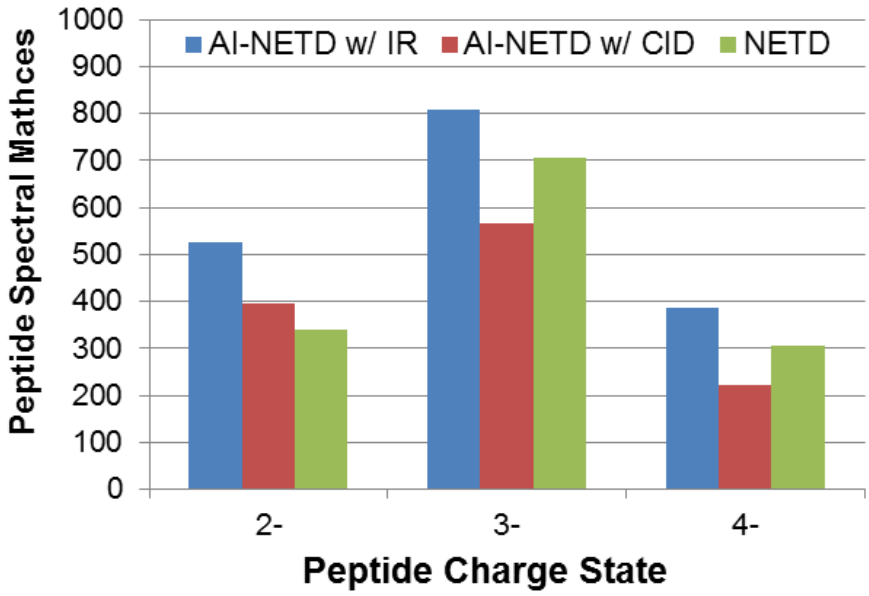

Figure 4.9 Number of peptide spectral matches produced by AI-NETD with IR photoactivation, AI-NETD with collisional activation and NETD for the analysis to H. salinarium NCR-1cell lysate.

4.5 Conclusions

The use of supplemental collisional activation of charge reduced precursors or infrared photoactivation of the entire ion population performed concomitant with the NETD reaction period greatly improves sequencing capabilities for peptide anions by increasing product ion abundances and sequence coverage. The two AI-NETD methods do not substantially change the fragmentation efficiencies relative to NETD, but the sequence ion conversion percentages, in terms of formation of diagnostic product ions, are substantially increased. Supplemental infrared photoactivation outpaces collisional activation for most of the metrics used to assess fragmentation performance.

4.6 References

(1) Schwartz, R.; Ting, C. S.; King, J. *Genome Res.* **2001**, *11*, 703 –709.

(2) Knight, C. G.; Kassen, R.; Hebestreit, H.; Rainey, P. B. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8390 –8395.

(3) Clipston, N. L.; Jai-nhuknan, J.; Cassady, C. J. *Int. J. Mass Spectrom.* **2003**, *222*, 363–381.

(4) Ewing, N.; Cassady, C. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 105–116.

(5) Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Mass Spectrom. Rev.* **2002**, *21*, 87–107.

(6) Steinborner, S. T.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 253–258.

(7) O'Hair, R. A. J.; Blanksby, S.; Styles, M.; Bowie, J. H. *Int. J. Mass Spectrom.* **1999**, *182-183*, 203–211.

(8) Brinkworth, C. S.; Dua, S.; McAnoy, A. M.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1965–1973.

(9) Chass, G. A.; Marai, C. N. J.; Harrison, A. G.; Csizmadia, I. G. *J. Phys. Chem. A* **2002**, *106*, 9695–9704.

(10) Steinborner, S. T.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1243–1247.

(11) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.

(12) 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.

(13) 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.

(14) Simons, J. *Chem. Phys. Lett.* **2010**, *484*, 81–95.

(15) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528 –9533.

(16) Coon, J. J.; Syka, J. E. P.; Schwartz, J. C.; Shabanowitz, J.; Hunt, D. F. *Int. J. Mass Spectrom.* **2004**, *236*, 33–42.

(17) Breuker, K.; Oh, H.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14011 –14016.

(18) Good, D. M.; Coon, J. J. *Biotechniques* **2006**, *40*, 783–789.

(19) Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. *Mol. Cell. Proteomics* **2007**, *6*, 1942 –1951.

(20) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793–1800.

(21) Håkansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530–4536.

(22) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431–4436.

(23) Jones, A.; Mikhailov, V.; Iniesta, J.; Cooper, H. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 268–277–277.

(24) Kelleher, N. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250–4253.

(25) Pitteri, S. J.; Chrisman, P. A.; McLuckey, S. A. *Anal. Chem.* **2005**, *77*, 5662–5669.

(26) Pitteri, S. J.; Chrisman, P. A.; Hogan, J. M.; McLuckey, S. A. *Anal. Chem.* **2005**, *77*, 1831–1839.

(27) Robinson, E.; Leib, R.; Williams, E. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1470–1479–1479.

(28) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778–4784.

(29) Håkansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256–3262.

(30) Zabrouskov, V.; Whitelegge, J. P. *J. Proteome Res.* **2007**, *6*, 2205–2210.

(31) Lin, C.; Cournoyer, J. J.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 780–789.

(32) Swaney, D. L.; McAlister, G. C.; Wirtala, M.; Schwartz, J. C.; Syka, J. E. P.; Coon, J. J. *Anal. Chem.* **2007**, *79*, 477–485.

(33) Ledvina, A. R.; McAlister, G. C.; Gardner, M. W.; Smith, S. I.; Madsen, J. A.; Schwartz, J. C.; Stafford, G. C.; Syka, J. E. P.; Brodbelt, J. S.; Coon, J. J. *Angew. Chem. Int. Edit.* **2009**, *48*, 8526–8528.

(34) Ledvina, A. R.; Beauchene, N. A.; McAlister, G. C.; Syka, J. E. P.; Schwartz, J. C.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J. *Anal. Chem.* **2010**, *82*, 10068–10074.

(35) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299–302.

(36) Kjeldsen, F.; Silivra, O. A.; Ivonin, I. A.; Haselmann, K. F.; Gorshkov, M.; Zubarev, R. A. *Chem. Eur. J.* **2005**, *11*, 1803–1812.

(37) Herron, W. J.; Goeringer, D. E.; McLuckey, S. A. *J. Am. Chem. Soc.* **1995**, *117*, 11555–11562.

(38) Stephenson, J. L.; McLuckey, S. A. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 875–880.

(39) Coon, J. J.; Shabanowitz, J.; Hunt, D. F.; Syka, J. E. P. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 880–882.

(40) Huzarska, M.; Ugalde, I.; Kaplan, D. A.; Hartmer, R.; Easterling, M. L.; Polfer, N. C. *Anal. Chem.* **2010**, *82*, 2873–2878.

(41) Rumachik, N. G.; McAlister, G. C.; Russell, J. D.; Bailey, D. J.; Wenger, C. D.; Coon, J. J. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 718–727.

(42) McAlister, G. C.; Russell, J. D.; Rumachik, N. G.; Hebert, A. S.; Syka, J. E. P.; Geer, L. Y.; Westphall, M. S.; Pagliarini, D. J.; Coon, J. J. *Anal. Chem.* **2012**, *84*, 2875–2882.

(43) Ly, T.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1148–1158.

Chapter 5

Tyrosine Deprotonation Yields Abundant and Selective Backbone Cleavage in Peptide Anions upon Negative Electron Transfer Dissociation and Ultraviolet Photodissociation

5.1 Overview

Tyrosine deprotonation in peptides yields preferential electron detachment upon NETD or UVPD, resulting in prominent N – Cα bond cleavage N-terminal to the tyrosine residue. UVPD of iodo-tyrosine modified peptides was used to generate localized radicals on neutral tyrosine side chains by homolytic cleavage of the C – I bond. Subsequent collisional activation of the radical species yielded the same preferential cleavage of the adjacent N-terminal N – Cα bond. LC-MS/MS analysis of a tryptic digest of BSA demonstrated that these cleavages are regularly observed for peptides when using high pH mobile phases.

5.2 Introduction

Bottom-up workflows for qualitative and quantitative protein identification utilizing liquid chromatography tandem mass spectrometry (LC-MS/MS) have emerged as one of the top technologies of choice for proteome analyses.[1] The success of these strategies hinges upon the ability to accurately predict fragmentation patterns of theoretical peptide sequences generated *in silico* from protein sequence databases.[2,3] The benchmark for peptide cation dissociation is collision induced dissociation (CID),[4,5] in addition to electron capture dissociation (ECD)[6,7] and electron transfer dissociation (ETD)[8,9] which afford complementary information to CID and enhanced identification of labile post-translational modifications (PTMs).[10] The dissociation mechanisms and

characteristic product ion types of these methods have been investigated in depth and are fairly well understood. A number of preferential cleavage types have been reported for peptide cations. For example, cleavages at protonated histidine and N-terminal to proline residues are commonly observed upon CID of protonated peptides, and in the absence of a mobile proton, enhanced cleavage also occurs at the amide bond immediately C-terminal to acidic residues.[11,12] Disulfide bond cleavage in multiply charged peptide and protein cations upon ECD and ETD is readily observed and for lower charge states is more favored than peptide/protein backbone cleavage.[13–15]

Exploration of the negative ion mode for protein identification affords an attractive alternative due to the fact that a large portion of biologically significant PTMs, such as phosphorylation, sulfonation, nitration, and glycosylation with acidic glycans, as well as many naturally occurring peptides and proteins, are acidic and thus readily form anions. Negative ion mode analyses[16,17] provide complementary information to that obtained in the positive ion mode, and the combined data sets allow more comprehensive proteomics analyses by tandem mass spectrometry. However, CID of peptide anions is generally ineffective due to extensive and uninformative neutral losses.[18,19] Recently, a number of dissociation methods including electron detachment dissociation (EDD),[20,21] negative electron transfer dissociation (NETD),[22,23] ultraviolet photodissociation (UVPD)[24] at 193 nm, negative-ion electron capture dissociation (niECD)[25], and electron photodetachment dissociation (EPD)[26–28] have been shown to be viable alternatives to CID for peptide anion characterization. Higher pH conditions using basic mobile phases for LCMS enhance deprotonation of peptide sites not routinely deprotonated (e.g. tyrosine side chain with pKa of ~10). Deprotonation of alternative sites may promote alternative fragmentation pathways as described herein.

5.3 Experimental

5.3.1 Materials and Sample Preparation

DRVYIHPFHLVIHN was purchased from Anaspec Inc. (Fremont, CA). Bovine serum albumin (BSA), trypsin and ≥99.998% pure argon were purchased from Sigma Aldrich (St. Louis, MO). RGYALG peptide was purchased from American Peptide Company (Sunneyvale, CA). β-d_2-Tyrosine was purchased from Cambridge Isotope Laboratories (Andover, MA). Fmoc-protected amino acids were purchased from Anaspec (San Jose, CA) or LC Sciences (Houston, TX). Wang resin, fluorenylmethyloxycarbonyl-O-succinimide (Fmoc-oSu) and tetramethylammonium hexafluorophosphate(HCTU) (ChemPep Inc., Miami, FL), isopropylsilane(TIPS) (Sigma Aldrich) were also purchased.

All optics were purchased from Edmund Optics (Barrington, NJ). Sodium iodide, chloramine-T, and sodium metabisulfite used for peptide iodination along with HPLC grade water and acetonitrile were purchased from Fisher Scientific.

Synthesis of deuterium labeled RGY(d2)ALG Fmoc protected β-d_2-Tyrosine was synthesized by reaction with 1 equivalent Fmoc-oSu and 1 equivalent $NaHCO_3$in 1:1 dioxane: H_2O solution overnight. After dissolving the reaction mixture in 5% HCl, the product was taken up using ethyl acetate and washed with 0.1M HCl and water. Following drying with anhydrous Na_2SO_4, organic solvent was evaporated away and the protected amino acid was crystallized. The peptide was synthesized by solid phase reaction using Wang-resin and Fmoc chemistry. HCTU and piperidine were used for carboxyl group activation and amine group deprotection, respectively. The peptide was cleaved from resin with 95:2.5:2.5 TFA:H_2O:TIPS.

The method used for tyrosine iodination was performed as previously described.[29] All reagents were prepared in water. The mole ratio between peptide, sodium iodide,

chloramine-T, and sodium metabisulfite was 1:1:2:2. To 100 nmol of peptide in 100 μL of water, sodium iodide was added and chloramine-T immediately after and reacted for 3 minutes at room temperature. Sodium metabisulfite was added to quench the reaction. The reaction mixture was then diluted to give 10 μM peptide and analyzed.

200 μg of BSA was reduced with dithiothreitol (DTT) for 30 minutes then incubated with iodoacetamide for at 55°C for one hour in the dark. Excess iodoacetamide was quenched with additional DTT. 4 μL of 1 μg/μL trypsin in 1 mM HCl was added giving a protein to enzyme ratio of 50:1. The digestion solution was incubated over night at 37°C. The tryptic peptides were desalted using C18 cartridges then dried and reconstituted in water to a concentration of 1 μg/μL.

5.3.2 Instrumentation

NETD and AI-NETD experiments were perform on a Bruker Daltonics (Billerica, MA) ETD equipped quadrupole ion trap mass spectrometer. Briefly, the methane carrier gas for the negative chemical ionization (nCI) source was replaced with argon. An override in the instrument user interface allowed manipulation of the nCI source and ion optic potentials which allowed generation of fluoranthene radical cations by argon chemical ionization and transmission to the ion trap. Two 2 mm holes were drilled in the ring electrode to allow irradiation of the trapping volume with 10.6 μm IR photons generated by a 50 W CW CO_2 laser (Synrad, Mukilteo, WA). Laser power was modulated by a UC-2000 laser controller (Synrad, Mukilteo, WA). A flange mounted ZnSe window allowed transmission of the laser beam into the vacuum chamber.

UVPD experiments were performed on a Velos dual cell linear ion trap mass spectrometer or an LTQ linear ion trap mass spectrometer (Thermo Scientific, San Jose,

CA). The back flange of the instrument was modified to allow introduction of the laser beam coaxial to the dual cell linear ion trap through a 25 mm calcium fluoride window and a stainless steel plate with a 2 mm aperture. For the UVPD experiments on the dual linear ion trap, 193 nm photons were generated by a 500 Hz ArF excimer laser (Coherent Inc., Santa Clara, CA). A TTL signal supplied during the ion activation period of the ion trap scan function gated a pulse/delay generator (Model 505, Berkely Nucleonics Corporation, San Rafael, CA) which triggered the laser every 2 ms during the activation period. The firmware of the mass spectrometer was modified to allow UVPD to be performed in the low pressure cell of the dual cell linear ion trap. For the experiments on the linear ion trap, 266 nm photons were generated from the 4^{th} harmonic generation of a Nd:YAG laser (Continuum, Santa Clara, CA). The isolation window for precursor ions in collision induced dissociation was set as 4 m/z.

5.3.3 Mass Spectrometry and Liquid Chromatography

For the experiments on the 3D-quadrupole ion trap or dual linear ion trap peptide anions were produced by negative mode electrospray ionization from 10 μM working solutions containing 70% acetonitrile, 30% water and 0.1% ammonium hydroxide. NETD and AI-NETD experiments were performed using 250 ms ion/ion reaction periods, and AI-NETD experiments were performed with 12 W IR irradiation during the entire ion/ion reaction period. UVPD experiments were performed with a q_z-value of 0.100 and an activation time of 2 ms (allowing one laser pulse) and a laser power of 3 mJ/pulse. For experiments on the LTQ linear ion trap, peptides (10 μM) were dissolved in 40:60 H_2O:ACN with 30mM NH_4Ac and electrosprayed into the mass spectrometer.

LC-MS/MS experiments were performed using a Dionex Ultimate 3000 HPLC (Sunnyvale, CA) equipped with a capillary flow splitter and a Agilent ZORBAX 300Extend-C18 column (Santa Clara, CA) (150 × 0.3 mm, 3.5 μm particle size). Eluent A was water with 0.05% ammonium hydroxide (~pH 10.5) and eluent B was acetonitrile with 0.05% ammonium hydroxide. A linear gradient over 60 minutes from 3% eluent B to 35% eluent B was used. 2 μg of tryptic peptides were injected on column. Data-dependent acquisition consisted of an MS^1 analysis from which MS/MS was performed on the 10 most abundant ions. Dynamic exclusion parameters included 1 repeat and exclude for 60 seconds.

Databases searches were performed using MassMatrix and search parameters included a precursor and fragment mass tolerance of 2 and 1 Da, respectively, and carbamidomethylation of C as a fixed modification.

5.4 Results and Discussion

In this chapter, preferential cleavage N-terminal to deprotonated tyrosine residues in peptide anions upon UVPD at 193 nm or upon negative electron transfer dissociation (NETD) is demonstrated and mechanistic details are elucidated. The dominant backbone cleavages of peptide anions upon NETD[21-22] typically result in $a^{\bullet}$- and x-type product ions and upon UPVD[23] *a*- and *x*-type with lower levels of *b, c, y, Y,* and *z* ions. Conversely, NETD (**Figure 5.1A**), NETD with simultaneous infrared photoactivation (a process termed AI-NETD) (**Figure 5.1B**), and UVPD (**Figure 5.1C**) of triply deprotonated DRVYIHPFHLVIHN yield abundant c_3 and $z_{11}^{\bullet}$ ions which arise uniquely from cleavage N-terminal to the tyrosine residue. This process is likewise notable upon NETD and UVPD of the 3- charge states of peptides NEKYAQAYPNVS, Ac-

DRVYIHPFHLVIHN, DRVYIHPFHLLVYS and KTMTESSFYSNMLA (not shown). The formation of high abundance *c*- and *z*-type product ions has not been reported previously for NETD or UVPD at 193 nm with evidence suggesting this process arises from a highly site-selective, radical-directed cleavage process that occurs upon electron detachment of tyrosine-containing peptides.

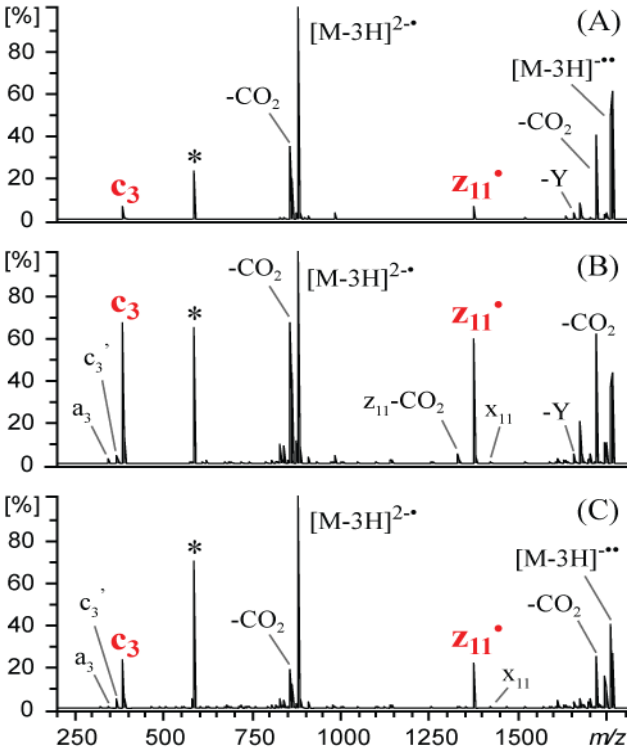

Figure 5.1 NETD (A), NETD with simultaneous infrared photoactivation (B) and UVPD (C) spectra of triply deprotonated DRVYIHPFHLVIHN. * and ' represent the precursor and loss of water, respectively.

The electron hole recombination energy upon electron transfer from peptide anions to fluoranthene radical cations has been previously estimated as 2.5 - 4.5 eV based on the difference in ionization energy (IE) of fluoranthene (IE = 7.9 eV) and the electron affinities (EA) of carboxylate (EA = 3.4 eV) and phosphate (EA = 5.4 eV) groups in

phosphopeptides.[23] The energy of a 193 nm photon is 6.4 eV. In a recent top-down EDD study, Breuker et al. tabulated the electron affinities (EA) of radical functional groups in peptides and proteins.[30] The EA of each radical functional group is equivalent to the ionization energy of the corresponding deprotonated functional group, and all are less than 3.5 eV. In contrast, ionization energies of neutral amino acids are between 7 and 10 eV.[31–33] Therefore, electron detachment, a key step in the selective backbone cleavage, likely occurs at sites of deprotonation.

The predicted sites of deprotonation in DRVYIHPFHLVIHN include the side-chains of aspartic acid and tyrosine, and the C-terminus, and in the triply charged species all three sites are expected to be deprotonated. Delocalization of the negative charges over backbone amides may also occur as a means of charge solvation.[21] The IEs of deprotonated aspartic acid and the C-terminus have been estimated to be 3.34 eV, and the IE of deprotonated tyrosine as 2.17 eV.[30] Due to the lower IE of deprotonated tyrosine, electron detachment is most favored at the deprotonated phenol of tyrosine over the carboxylates of aspartic acid or the C-terminus. In the doubly deprotonated species a majority of the ions are expected to be deprotonated at the carboxylates of aspartic acid and the C-terminus, not the tyrosine side-chain. In fact, NETD and UVPD of doubly deprotonated DRVIHPFHLVIHN (**Figure 5.2A and B**) yielded very low abundance c_3 and $z_{11}^{\bullet}$ product ions. This outcome is expected if it is tyrosine deprotonation that specifically promotes the preferential cleavage.

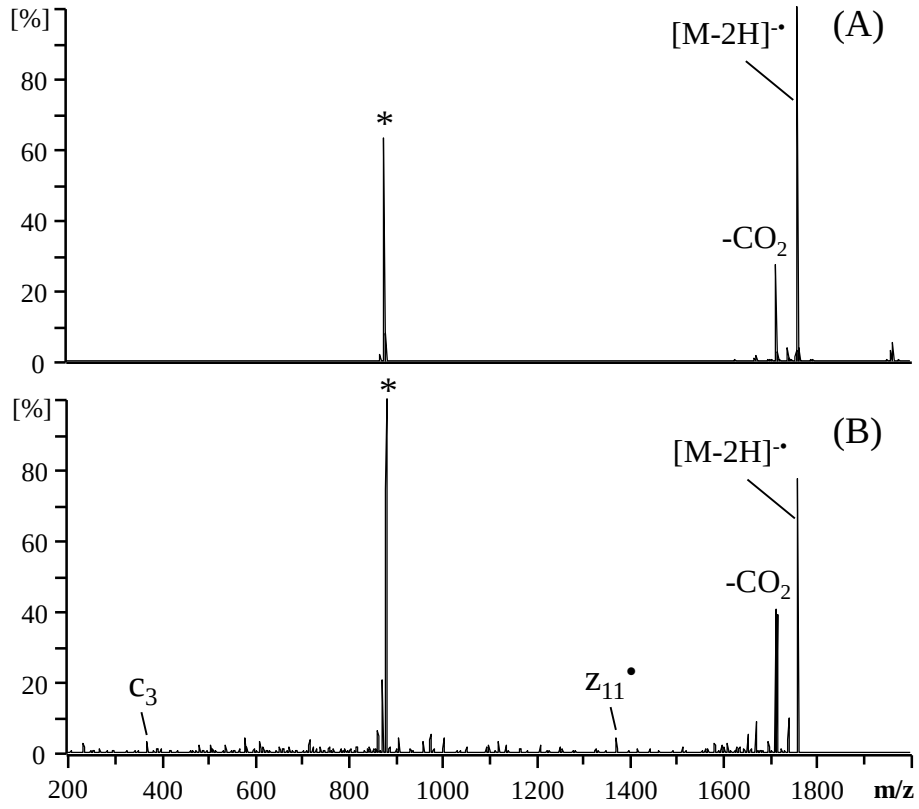

Figure 5.2 NETD mass spectrum (A) and 193 nm UVPD mass spectrum (B) of doubly deprotonated DRVYIHPFHLVIHN. * indicates precursor ions.

Breuker et al. proposed a mechanism for the formation of *c* and *z* ions N-terminal to tyrosine residues upon EDD (see Ref 29, Scheme S2), and it is reasonable to consider a similar pathway during NETD and 193 nm UVPD. In fact, deuterium labeling of the beta carbon atom yields results entirely consistent with the previously proposed mechanism. CID of $[RGYALG - H]^{-\cdot}$ (**Figure 5.3 top**) yields $[z_4 - H]^{-\cdot}$ ions (expected cleavage N-terminal to tyrosine); however, CID of the tyrosine beta position deuterium labeled derivative $[RGY(d2)ALG - H]^{-\cdot}$ yields $[z_4(d1) - H]^{-\cdot}$ ions (**Figure 5.3 bottom**), in which one of the beta position deuterium atoms is transferred to the N-terminal fragment. In the context of the tyrosine-specific cleavage in the present study, we presume that the initial

radical ion formed upon electron detachment from the phenol oxygen undergoes radical directed dissociation (RDD) to yield N – C_α bond cleavage. The substantially increased abundances of the c_3 and $z_{11}^\bullet$ ions upon NETD with simultaneous IR photoactivation (**Figure 5.1B**) compared to NETD alone (**Figure 5.1A**) suggests that the additional vibrational energy provided by IR photoactivation allows a greater portion of these tyrosyl radical ions to undergo the radical migration/N – C_α backbone cleavage process.

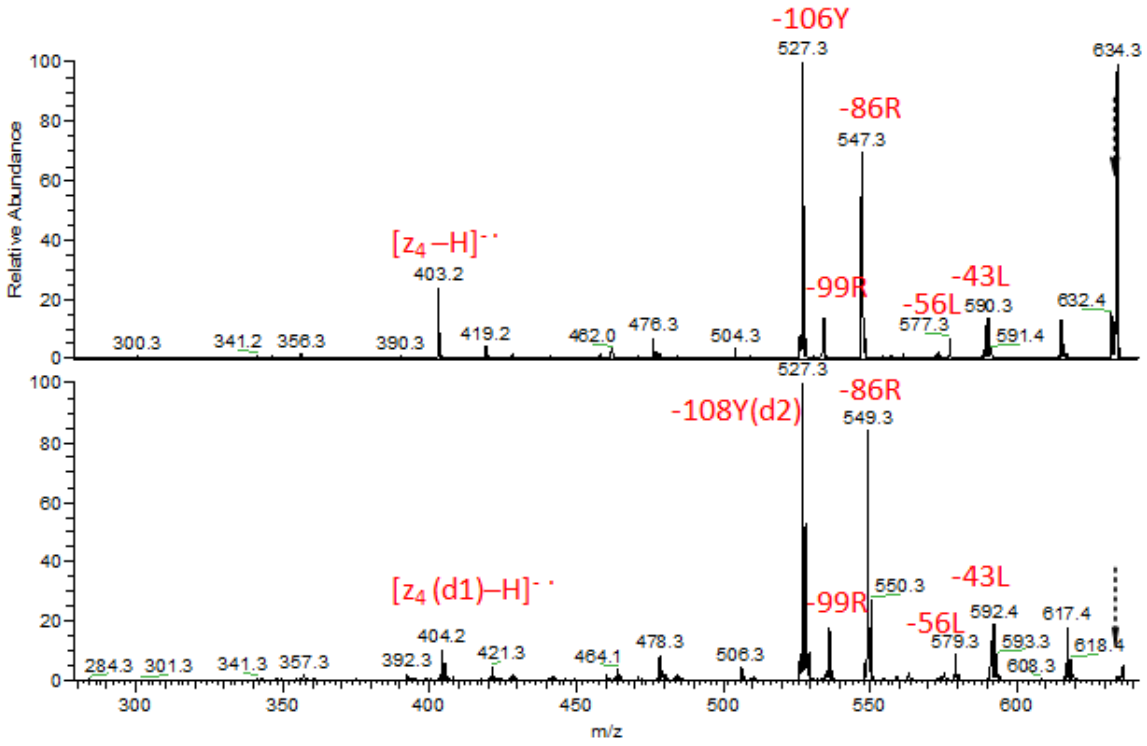

Figure 5.3 CID mass spectra of (top) [RGYALG - H]$^{\cdot}$ and (bottom) [RGY(d2)ALG - H]$^{\cdot\cdot}$ radical anions generated by 266 nm UVPD of iodo-tyrosine.

CID of the charge reduced precursor ([M - 3H]$^{2-\bullet}$) produced by NETD (**Figure 5.4A**) or by UVPD (**Figure 5.4B**) of triply deprotonated DRVYIHPFHLVIHN yields the same abundant c_3 and $z_{11}^\bullet$ ions along with neutral losses of water, carbon dioxide and tyrosine side chain. This indicates that a portion of the [M - 3H]$^{2-\bullet}$ ions formed upon

electron detachment from the triply deprotonated peptide precursors contain stable radicals localized on the tyrosine side chain or *c* and *z* ions held together through non-covalent interactions. Subsequent collisional activation of these species provided enough internal energy to surpass the activation barriers for radical migration or non-covalent bond cleavage.

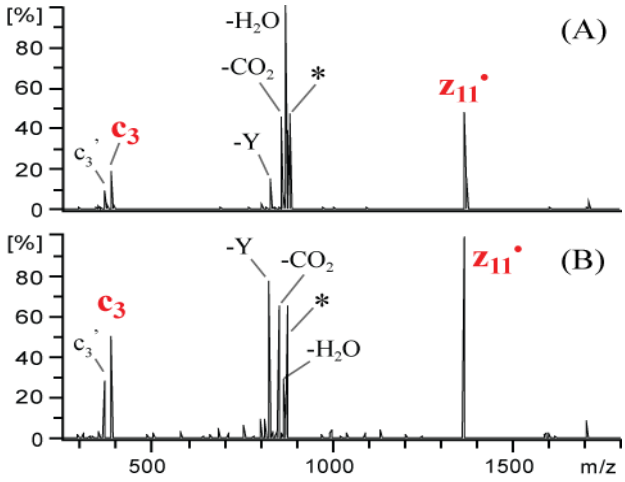

Figure 5.4 MS^3 spectrum upon CID of the charge-reduced precursor ($[M - 3H]^{2-\cdot}$) produced by (A) NETD and (B) UVPD of triply deprotonated DRVYIHPFHLVIHN . * and ' represent the precursor and loss of water, respectively.

Previous work has shown that 266 nm UVPD of iodo-tyrosine-containing peptide and protein cations results in homolytic cleavage of the C – I bond, thus yielding a localized radical on the tyrosine residue.[29,34] Subsequent CID of these peptide/protein radical cations led to dominant backbone fragmentation in the form of *a* ions C-terminal to the tyrosine residue.[29,34] This process is illustrated in the equation below.

$$[\mathrm{M + nH + I}]^{\mathrm{n+}} \overset{h\nu}{\Rightarrow} [\mathrm{M} + n\mathrm{H}]^{n+\bullet} \overset{CID}{\Longrightarrow} \text{fragments}$$

In addition to selective *a* ion formation adjacent to the modified tyrosine residue, abundant secondary cleavages producing complementary *a* and *y* ions were observed N-

terminal to proline residues in the vicinity of the modified tyrosine residue. These experiments were recently extended to peptide anions,[35] thus providing a convenient method for creating radicals on tyrosine side chains in peptide anions in the present study. The resulting ions allow determination of whether the selective formation of *c* and *z* ions N-terminal to tyrosine upon NETD and 193 nm UVPD proceeds through a tyrosyl radical.

UVPD at 193 nm (**Figure 5.5A**) of doubly deprotonated DRVY^{I}IHPFHLVIHN (where Y^{I} represents iodo-tyrosine) yields predominantly an intact charge reduction product ($[M - 2H + I]^{2-\cdot}$) via electron photodetachment and a wide array of low abundance *a, x, c* and *z* ions as expected. Homolytic cleavage of the C – I bond (i.e. $[M - 2H]^{2-\bullet}$) also occurs at approximately 10% relative abundance. Very low abundance c_3 and $[z_{11} + I]^{\bullet}$ ions are observed and are expected based on the results from the non-iodinated peptide (i.e. in the 2- charge state Y^{I} is not expected to be deprotonated, and thus Y-directed cleavage is not favored.). Based on the iodinated peptide results from Julian et al., collisional activation of $[M - 2H]^{2-\bullet}$ is expected to yield selective *a* ion formation C-terminal to the iodinated tyrosine residue. Instead, CID of $[M - 2H]^{2-\bullet}$ (**Figure 5.5B**) yields dominant c_3 and $z_{11}^{\bullet}$ ions corresponding to cleavage N-terminal to the tyrosine residue. In fact, the a_4 and x_{10} ions which arise from cleavage C-terminal to the tyrosine residue are observed at only 1% relative abundance. In addition, y_7 ions are observed at 90% relative abundance which corresponds to cleavage C-terminal to proline. The UVPD/CID experiments of other iodo-tyrosine peptides similarly reproduced the same tyrosine selective cleavages that occur from non-iodinated peptides containing a deprotonated tyrosine. However, the abundant y_7 ion formation C-terminal to proline is not observed upon NETD or UVPD of the triply deprotonated non-iodinated peptide or upon CID of $[M - 3H]^{2-\bullet}$ ion produced by NETD or UVPD of triply charged

DRVYIHPFHLVIHN. CID of the doubly deprotonated even electron species (**Figure 5.6**) also fails to yield y_7 ions, a results which indicates the y_7 ions are produced by RDD. This is consistent with the analogous results reported by Julian et al. showing that *y* ion formation N-terminal to proline proceeds through a radical mechanism.[34,35]

The fact that the tyrosine residue is not deprotonated in the 2- charge state of the iodo-tyrosine-containing peptide and is deprotonated in the 3- charge state of the non-iodinated peptide likely accounts for the observation of the abundant y_7 ion. UVPD of triply deprotonated DRVY^{I}IHPFHLVIHN (**Figure 5.5C**) shows that the presence of iodine on the tyrosine side chain does not impede the selective cleavage N-terminal to the deprotonated tyrosine residue. Dominant c_3 and $[z_{11}+I]^{\bullet}$ ions are produced. The absence of y_7 ions upon UVPD of the 3- charge state of the iodinated peptide supports our hypothesis that the phenol oxygen needs to be protonated for *y* ion formation. Collectively, these results suggest that the formation of tyrosyl radicals, whether promoted by NETD or UVPD of peptides containing a deprotonated tyrosine residue, is a key step that precedes the selective cleavage N-terminal to the tyrosine residue.

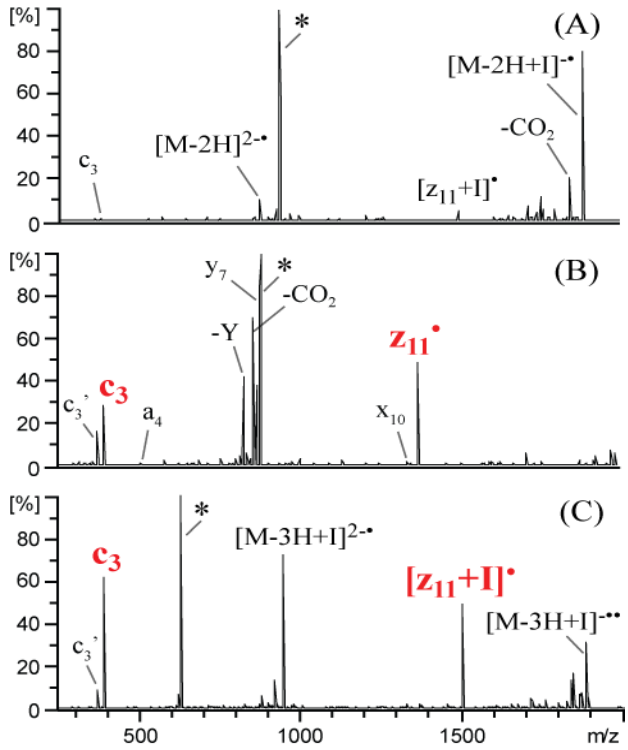

Figure 5.5 193 nm UVPD (A) of doubly deprotonated DRVY^{I}IHPFHLVIHN (Y^{I} represents iodo-tyrosine) and (B) subsequent CID of $[M - 2H]^{2-\bullet}$. (C) 193 nm UVPD of $[M - 3H + I]^{3-}$. *, ' and I represent the precursor, loss of water or iodine, respectively.

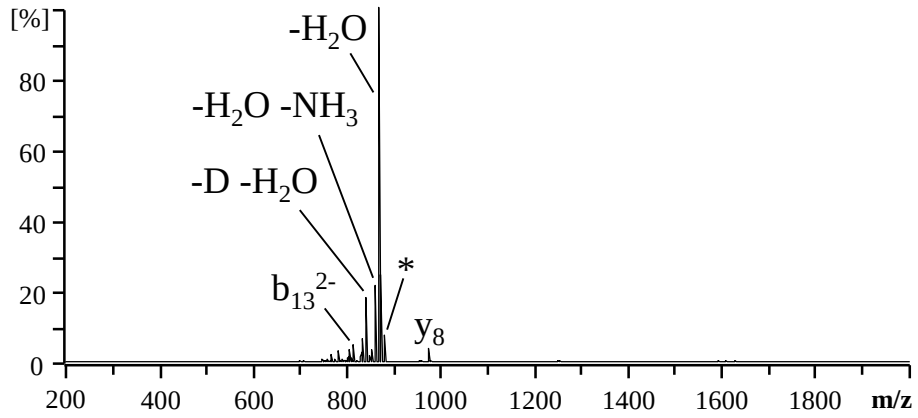

Figure 5.6 CID mass spectrum of doubly deprotonated DRVYIHPFHLVIHN. * represents precursor ion.

To gauge the frequency with which the preferential cleavage N-terminal to tyrosine might occur in a bottom-up proteomics experiment, we performed LC-MS/NETD and LC-MS/UVPD of a bovine serum albumin (BSA) tryptic digest using 0.05% ammonium hydroxide mobile phases (~pH 10.5). LC-MS/MS spectra were interpreted using the MassMatrix database search engine[36,37] which was adapted to search the appropriate ion series for each dissociation method. Tyrosine-containing peptide matches in which the charge state exceeded the number of acid functionalities (i.e. E, D and C-terminus) were manually interpreted to confirm preferential cleavage N-terminal to deprotonated tyrosine residues. The peptide sequence, charge state, and the percentage of total sequence ion peak area for Y-selective *c* and *z* ions for those peptides that produced preferential cleavage upon NETD are shown in **Table 5.1**. **Figure 5.7** shows the NETD spectrum of the 4- charge state of RHPYFYAPELLYYANK from BSA. For peptides with multiple tyrosine residues, *c* and *z* ion formation was observed N-terminal to all tyrosine residues. Of the 16 tyrosine-containing peptides identified by each method, six yielded preferential cleavage N-terminal to tyrosine residues. However, three of the 16

peptides contained N-terminal tyrosine residues, thus *c* and *z* ion formation N-terminal to the tyrosine residue is not possible. The seven remaining tyrosine-containing peptides did not produce charge states consistent with tyrosine deprotonation, and thus preferential cleavage was not observed upon NETD or UVPD.

Peptide Sequence	Charge State[1]
HPYFYAPELLYYANK	3-, 4- (39, 56%)
RHPYFYAPELLYYANK	3-, 4- (69, 88%)
GLVLIAFSQYLQQCPFDEHVK	4- (20%)
RHPEYAVSVLLR	3- (63%)
LGEYGFQNALIVR	3- (23%)
LGEYGFQNALIVRYTR	4- (53%)

[1]The percentage of sequence ion peak area due to the *c* and *z* ions from Y-selective cleavage is given in parenthesis for each charge state.

Table 5.1 BSA tryptic peptides exhibiting preferential cleavage N-terminal to tyrosine upon NETD.

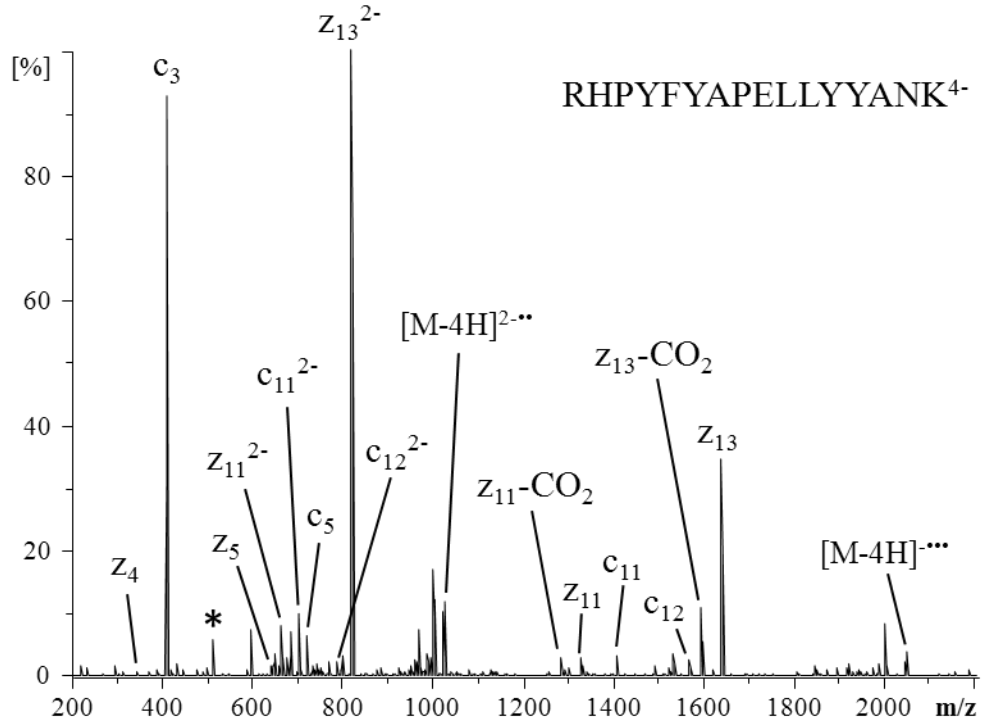

Figure 5.7 NETD mass spectrum of the 4- charge state of BSA tryptic peptide RHPYFYAPELLYYANK.

5.5 Conclusions

In summary, tyrosine deprotonation during negative ion mode analyses yields selective and enhanced *c* and *z* ion formation N-terminal to the tyrosine residue upon NETD and 193 nm UVPD. The lower ionization energy of deprotonated tyrosine phenol compared to the carboxylates of acidic residues and the C-terminus favors preferential electron detachment upon NETD or UVPD. UVPD/CID experiments utilizing iodo-tyrosine derivatives confirmed that selective and enhanced *c* and *z* ion formation proceeds through a tyrosyl radical. LC-MS/MS experiments showed that this cleavage specificity can be expected to occur frequently during bottom-up proteomics experiments and could be readily incorporated into database search algorithms to facilitate peptide anion

identification. This fragmentation pathway may prove to be a valuable diagnostic for the characterization of post-translational modifications of tyrosine, such as nitration which has been shown to lower the pKa of the phenol hydroxyl function.[38]

5.6 References

(1) Mann, M.; Kelleher, N. L. *P. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18132 –18138.

(2) Eng, J. K.; McCormack, A. L.; Yates III, J. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976–989.

(3) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. *ELECTROPHORESIS* **1999**, *20*, 3551–3567.

(4) McLuckey, S. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599–614.

(5) Hunt, D. F.; Buko, A. M.; Ballard, J. M.; Shabanowitz, J.; Giordani, A. B. *Biol. Mass Spectrom.* **1981**, *8*, 397–408.

(6) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.

(7) Breuker, K.; Oh, H.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14011 –14016.

(8) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528 –9533.

(9) Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. *Mol. Cell. Proteomics* **2007**, *6*, 1942 –1951.

(10) Zhou, Y.; Dong, J.; Vachet, R. W. *Curr. Pharm. Biotechno.* **2011**, *12*, 1558–1567.

(11) Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1996**, *31*, 1185–1187.

(12) Gu, C.; Tsaprailis, G.; Breci, L.; Wysocki, V. H. *Anal. Chem.* **2000**, *72*, 5804–5813.

(13) 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.

(14) Chrisman, P. A.; Pitteri, S. J.; Hogan, J. M.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1020–1030.

(15) Simons, J. *Chem. Phys. Lett.* **2010**, *484*, 81–95.

(16) Ewing, N.; Cassady, C. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 105–116.

(17) Janek, K.; Wenschuh, H.; Bienert, M.; Krause, E. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1593–1599.

(18) Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Mass Spectrom. Rev.* **2002**, *21*, 87–107.

(19) Clipston, N. L.; Jai-nhuknan, J.; Cassady, C. J. *Int. J. Mass Spectrom.* **2003**, *222*, 363–381.

(20) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299–302.

(21) Kjeldsen, F.; Silivra, O. A.; Ivonin, I. A.; Haselmann, K. F.; Gorshkov, M.; Zubarev, R. A. *Chem. Eur. J.* **2005**, *11*, 1803–1812.

(22) Coon, J. J.; Shabanowitz, J.; Hunt, D. F.; Syka, J. E. P. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 880–882.

(23) Huzarska, M.; Ugalde, I.; Kaplan, D. A.; Hartmer, R.; Easterling, M. L.; Polfer, N. C. *Anal. Chem.* **2010**, *82*, 2873–2878.

(24) Madsen, J. A.; Kaoud, T. S.; Dalby, K. N.; Brodbelt, J. S. *Proteomics* **2011**, *11*, 1329–1334.

(25) Yoo, H. J.; Wang, N.; Zhuang, S.; Song, H.; Håkansson, K. *J. Am. Chem. Soc.* **2011**, *133*, 16790–16793.

(26) Antoine, R.; Joly, L.; Tabarin, T.; Broyer, M.; Dugourd, P.; Lemoine, J. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 265–268.

(27) Larraillet, V.; Antoine, R.; Dugourd, P.; Lemoine, J. *Anal. Chem.* **2009**, *81*, 8410–8416.

(28) Larraillet, V.; Vorobyev, A.; Brunet, C.; Lemoine, J.; Tsybin, Y. O.; Antoine, R.; Dugourd, P. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 670–680.

(29) Liu, Z.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 965–971.

(30) Ganisl, B.; Valovka, T.; Hartl, M.; Taucher, M.; Bister, K.; Breuker, K. *Chem. Eur. J.* **2011**, *17*, 4460–4469.

(31) Slifkin, M. A.; Allison, A. C. *Nature* **1967**, *215*, 949–950.

(32) Klasinc, L. *J. Electron. Spectrosc. Relat. Phenom.* **1976**, *8*, 161–164.

(33) Dehareng, D.; Dive, G. *Int. J. Mol. Sci.* **2004**, *5*, 301–332.

(34) Ly, T.; Julian, R. R. *J. Am. Chem. Soc.* **2008**, *130*, 351–358.

(35) Moore, B.; Sun, Q.; Hsu, J. C.; Lee, A. H.; Yoo, G. C.; Ly, T.; Julian, R. R. *J. Am. Soc. Mass Spectrom.* **2011**, *23*, 460–468.

(36) Xu, H.; Freitas, M. *BMC Bioinformatics* **2007**, *8*, 133.

(37) Xu, H.; Freitas, M. A. *Proteomics* **2009**, *9*, 1548–1555.

(38) Sokolovsky, M.; Riordan, J. F.; Vallee, B. L. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 20–25.

Chapter 6

Analysis of Protein Digests by Transmission-mode Desorption Electrospray Ionization Mass Spectrometry with Ultraviolet Photodissociation

6.1 Overview

Transmission mode desorption electrospray ionization is coupled with ultraviolet photodissociation at 193 nm for the analysis of protein digests. SEQUEST is utilized for data interpretation and database searches. Comparative results are presented for TM-DESI-MS/collision induced dissociation (CID) and TM-DESI-MS/UVPD analyses of five proteolyzed model proteins ranging in molecular weight from 8.5 kDa (ubiquitin) to 66 kDa (bovine serum albumin, BSA). In some cases TM-DESI/UVPD yielded greater confidence in database correlation scores for peptides and comparable protein identification compared to TM-DESI-MS/CID due to the production of an extensive array of sequence ions and the ability to detect low m/z terminal sequence ions and immonium ions.

6.2 Introduction

The undisputed success of MALDI-MS and ESI-MS methods for proteomic applications has led to increasing emphasis on the adaptation of these methods for higher throughput analysis. There have been significant advances in the development of MALDI-MS for high-throughput applications because of its natural compatibility with the analysis of discrete samples, such as those from a complex 2D-gel separation, robotically deposited on a multi-well plate, with analysis requiring just a few laser shots per sample and detection limits in the sub fmol range.[1,2] For bottom-up proteomics

studies, peptide mass fingerprinting (PMF) and and now primarily sequencing by tandem mass spectrometry (MS/MS) are the most frequently used methods for protein identification.[3,4] Numerous MS/MS-based strategies have emerged as valuable tools for the analysis of both pure protein samples and complex mixtures and are readily integrated with HPLC-ESI-MS platforms.[5–10]

The benefits of high-throughput workflows have motivated the search for new alternatives to MALDI-TOF-MS and HPLC-ESI-MS/MS methods. In this context, ambient mass spectrometry techniques, such as desorption electrospray ionization (DESI),[11] provide simple yet highly efficient methods for analyzing multiple discrete samples, generating multiply charged analyte ions, and allowing analysis in a high-throughput manner. Ambient mass spectrometry has revolutionized the means by which samples are introduced to the mass spectrometer. Since the introduction of DESI,[11] the field of ambient mass spectrometry has grown rapidly due the recognition of the innumerable possible applications that exploit the ability to analyze samples quickly and with minimal sample preparation in a high-throughput manner. As a result, nearly thirty ambient ionization techniques have been introduced in the past six years, as summarized in two recent reviews.[12,13]

In addition to touted benefits for desorption and ionization of small molecules, ESI-like ambient ionization techniques are capable of analyzing large non-volatile biomolecules, and a number of recent studies have been reported for the investigation of peptides, proteins and enzymatically digested proteins.[14–24] In one of the first studies, DESI was coupled to an FTICR mass spectrometer for high resolution detection of small proteins (up to insulin, MW = 5.8 kDa).[14] The rapid analysis of tryptic peptides from bovine serum albumin (BSA) without sample pre-treatment or clean-up by DESI-ion mobility mass spectrometry was reported,[15] allowing the identification of BSA by a

pseudo-peptide mass fingerprint and MASCOT database search strategy. DESI-MS was also utilized to analyze intact proteins up to 18 kDa with ng detection limits with generally less successful detection of much larger proteins (up to 66.4 kDa).[16] The use of electrospray-assisted laser desorption ionization (ELDI) was demonstrated for the analysis of liquid samples of peptides and proteins up to 66 kDa (BSA) and tryptic protein digests, as well as the combination of ELDI with tandem mass spectrometry for effective MS/MS characterization.[17] In a follow-up study the successful analysis of proteins up to 80 kDa by IR-ELDI was reported.[18] Nano-porous alumina surfaces were shown to afford higher ion abundances, greater signal stability, and improved limits of detection compared to other surfaces for DESI analysis of peptides and proteins.[19] Another hybrid method, MALDESI, was demonstrated for the analysis of peptides and proteins ranging from 1 to 17 kDa.[20] In a pair of studies that focused on the use of DESI for direct detection of analytes separated on thin layer chromatography plates, DESI was employed for the analysis of tryptic peptides from cytochrome c and myoglobin with sequence coverages of 81% and 74% for cytochrome c and myoglobin, respectively, based on data dependent MS/MS.[21,22] DESI was also used for the direct analysis of liquid samples in which amino acids, peptides, proteins up to 66 kDa and tryptic digests were successfully ionized from solutions with pmol detection limits.[23] More recently, DESI-FTICR-MS was successfully utilized for protein sequencing of intact myoglobin and cytochrome-c by collision induced dissociation (CID) and electron capture dissociation (ECD), requiring approximately 10 nmol for top-down analysis.[24]

Due to the inherent complexity of biological samples and transient nature of analyte ion signal generated by desorption ionization methods, it is crucial for ion activation to be fast and efficient in order to maximize the depth of analysis for each sample in MS/MS experiments. Recently, a number of studies[25–31] have shown short

wavelength (157 and 193 nm) ultraviolet photons provide fast and efficient activation of peptide ions upon the absorption of a single photon. Ultraviolet photodissociation (UVPD) at 157 and 193 nm produces fragmentation patterns containing *a, b, c, x, y, z* sequence ions as well as immonium and *v* and *w* side-chain loss ions. In a recent study from our group[31] it was shown that UVPD at 193 nm using a single 5 ns, 8 mJ laser pulse provided comparable and often improved peptide identification by *in silico* database searching for biologically relevant samples. The integration of fast UV photoactivation with the minimal sample preparation, high-throughput capabilities of DESI offers a compelling opportunity for proteomics applications. We have previously reported transmission mode desorption electrospray ionization (TM-DESI),[32,33] a method employing a simplified geometry for DESI experiments in which the sample is placed in-line between the ESI source and the mass spectrometer inlet on a mesh substrate, allowing transmission of the ESI plume through the sample. In the present study, the high-throughput analysis of proteins is facilitated by the use of TM-DESI-MS and TM-DESI-MS/UVPD. Confident protein identification is achieved through the use of both peptide mass fingerprinting and MS/MS sequencing methods.

6.3 Experimental

6.3.1 Materials

Bovine model proteins α-casein, bovine serum albumin (BSA), β-lactoglobulin A, carbonic anhydrase, cytochrome c, ubiquitin, proteomics grade trypsin, acetic acid, ammonium bicarbonate, dithiothreitol (DTT) and iodoacetamide (IAM) were purchased from Sigma-Aldrich (St. Louis, MO). Chicken lysozyme was purchased from MP

Biomedicals (Solon, OH). HPLC grade acetonitrile, methanol and water were purchased from Fisher Scientific (Fairlawn, NJ). Sep-Pak® 1cc C18 cartridges were obtained from Waters (Milford, MA). Polypropylene mesh with a 149 μm open space and 106 μm strand diameter (PP149) was purchased from Small Parts Inc. (Miramar, FL).

6.3.2 Sample Preparation

Model proteins stock solutions were prepared at 100 μM in ~ pH 8, 100 mM ammonium bicarbonate for trypsin digestion. Sulfhydryl reduction was performed by adding 10 μL of acetonitrile and 5 μL of 100 mM DTT to 50 μL of protein stock in a 2 mL centrifuge tube, then incubated at 40°C for one hour. Reduced sulfhydryl groups were then alkylated by adding 4 μL of 500 mM IAM and reacted in the dark for one hour at 37°C. Excess IAM was quenched by adding another 15 μL of 500 mM DTT and reacted for 30 minutes at room temperature. 1 mg/mL trypsin in 1 mM HCl was added to give a protein to trypsin ratio of 50:1. The total volume was increased to 100 μL with water and incubated for 16 hours at 37°C. Protein digests were cleaned-up using C18 cartridges then diluted to 10 μM with 90:10 (v/v) of methanol/water with 1% acetic acid.

6.3.3 Transmission Mode DESI, Mass Spectrometry and Ultraviolet Photodissociation

All experiments were performed using an Omni Spray ion source (Prosolia, Inc., Indianapolis, IN) mounted to a Thermo Fisher Scientific LTQ XL mass spectrometer (Waltham, MA). The Omni Spray source was equipped with a sample holder that permits 0° desorption geometry. The mesh substrate was held between a mask and backing plate made from sheet polyetheretherketone (PEEK). The mask is 1.5 mm thick and has six 4

mm square openings with 2 mm spacing to allow transmission of the electrospray plume through the mesh substrate. The openings in the mask are analogous to the array of wells on a MALDI sample plate. The distance from the electrospray emitter to the mesh was 2 mm, and the distance from the mesh to the mass spectrometer inlet was 8 mm. Samples were deposited on PP149 mesh in 2 μL aliquots and analyzed wet. The electrospray solvent was methanol at a flow rate of 5 μL/min with an electrospray voltage of 4.5 kV. The sample holder was scrolled perpendicularly to the electrospray plume at 250 μm/s and yielded an analysis time of 2.4 minutes for six sample wells. Nitrogen at a pressure of 100 psi was used as the nebulizing gas and the heated capillary temperature was 250°C.

193 nm photons were generated by a Coherent Excistar XS excimer laser (Santa Clara, CA). The back plate of the mass spectrometer was modified with a CF viewport flange with a CaF_2 window as described previously.[34] The unfocused laser beam was aligned on axis with the linear ion trap through a 2 mm aperture. XCalibur version 2.2 software package was used to perform data dependent acquisitions. The data dependent acquisition employed the following scan sequence: a full scan mass spectrum over the range of *m/z* 400-2000, then MS/MS spectra for each of the twenty most abundant ions in the full scan mass spectrum. A single 5 ns, 8 mJ pulse and a *q*-value of 0.100 were used during the activation period of each MS/MS scan for UVPD experiments. The commercial LTQ software limited the minimum activation period to 0.03 ms even though the isolated ions were only irradiated for 5 ns. For CID experiments, the standard 30 ms activation period with a *q*-value of 0.250 was used, and the normalized collision energy was set to 35%. An isolation width of 2 *m/z* was used for all MS/MS scan events. For all experiments, automated gain control (AGC) was set to 3×10^4 for MS and 1×10^4 for MS/MS scans. For MS and MS/MS scans the maximum injection times were set to 50 ms and 100 ms respectively. Each MS scan was the average of six microscans and each

MS/MS scan was the average of 4 microscans. MS spectra and CID MS/MS spectra were collected in profile mode. UVPD MS/MS spectra were collected in centroid mode to allow photoionization background subtraction prior to protein database processing. Dynamic exclusion parameters included duration of 45 s, exclusion list size of 50 *m/z* values, and a repeat count of one.

6.3.4 Background Subtraction and Protein Database Searching

UVPD photoionization background peaks possibly arising from pump oil and hydrocarbon impurities in the ion trap helium bath gas were subtracted as previously described.[31] Thermo Fisher Scientific Proteome Discoverer 1.0 software package was used to perform SEQUEST database searches for interpretation of MS/MS data. SEQUEST search parameters included a signal:noise ratio of 3, a precursor mass tolerance of 3 Da, fragment mass tolerance of 2 Da, 1 missed cleavage by trypsin, methionine oxidation as a dynamic side chain modification and carbamidomethyl as a static side chain modification of cysteine. Product ion series for CID spectra included *a*, *b* and *y* ions and UVPD product ion series were *a, b, c, x, y* and *z* ions. Nonredundant bovine and chicken protein databases from the NCBI were used for proteins form the respective species. Peptide matches were filtered based on charge state and XCorr scores, and protein matches were rejected if they did not receive more than two peptide matches. MS/MS spectra that passed the filters but received probability scores less than 1.00 were manually verified.

6.4 Results and Discussion

Protein digests were deposited into the wells of the TM-DESI sample holder, then the sample holder was rastered in front of the ESI plume. Full scan mass spectra and MS/MS spectra were acquired with ion activation by CID or UVPD. MS/MS sequencing experiments were performed using a sample well scan rate of 250 μm/s because this rate provided a reasonable compromise between the number of MS/MS spectra acquired per sample well and total analysis time. The total analysis time was 2.4 min (for six wells), and approximately 20 MS/MS spectra could be acquired per sample well. The same sample was deposited in all six sample wells, and SEQUEST scores were generated from the combined spectra from all six sample wells.

Five model proteins ranging in molecular weight from 8.5 kDa for ubiquitin to 66 kDa for BSA were analyzed by both CID and UVPD. **Figure 6.1** shows UVPD and CID spectra for one representative tryptic peptide from β-lactoglobulin A, VAGTWYSLAMAASDISLLDAQSAPLR. UVPD generated a significantly larger array of sequence-informative ions (e.g., *a*, *b*, *c*, *x*, *y*, *z*) compared to CID (predominately *b* and *y* type ions) as well as greater sequence coverage. The ability to operate at lower *q*-values for UVPD also allowed detection of terminal sequence ions that fall below the low mass cut-off in the corresponding CID spectrum. For some of the peptides analyzed, the greater variety of sequence ions and the ability to detect low mass terminal sequence ions led to greater confidence in the peptide identification as reflected by the higher XCorr and probability scores for UVPD. Ten replicate experiments were performed for each of the five model protein for both UVPD and CID. **Table 6.1** summarizes the overall SEQUEST search results for both UVPD and CID sequencing experiments. The overall results for CID and UVPD are similar with respect to number of unique peptides, protein score, and sequence coverage for the proteins analyzed in this study. CID marginally

outperforms UVPD for three of the five proteins analyzed, an outcome attributed to the decreased sensitivity of UVPD due to secondary dissociation of product ions and lower fragmentation efficiency and product ion abundance compared to CID. It has been shown that fragmentation efficiency for UVPD at 193 nm is greater for peptides with amino acids containing aromatic side chains (e.g., F, W and Y).[31] Peptides containing zero aromatic residues showed fragmentation efficiencies of approximately 50%, and peptides containing multiple aromatic residues yielded fragmentation efficiencies of greater than 90%. However, due to the large number of fragmentation pathways accessible by 193 nm UVPD, the product ion current is spread between a large number of product ions and thus the individual abundances of many of the product ions are low. As a result UVPD may not perform as well for low abundance peptides, leading to fewer peptide spectral matches (PSMs).

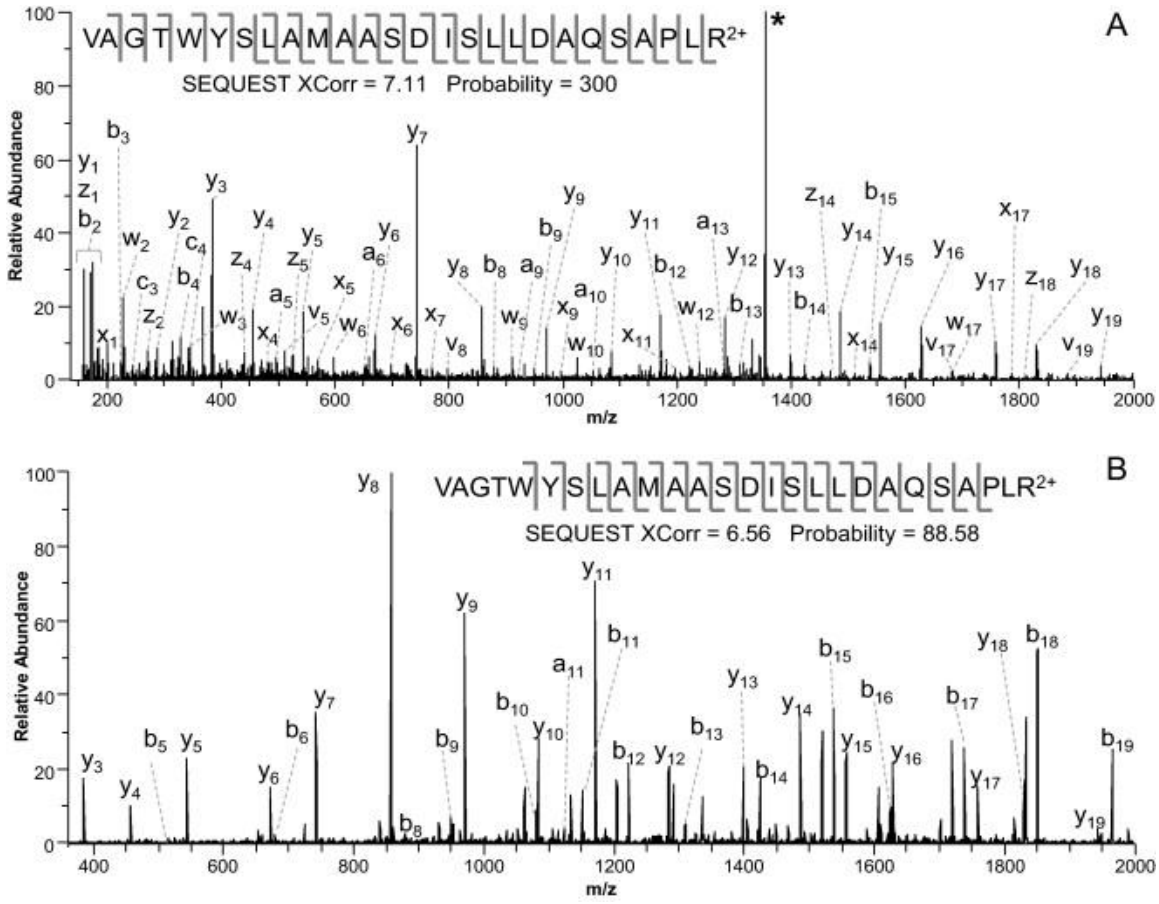

Figure 6.1 (A) UVPD (one 5 ns pulse at 193 nm) and (B) CID spectra for β-lactoglobulin A tryptic peptide VAGTWYSLAMAASDISLLDAQSAPLR with sequence coverage, SEQUEST XCorr and probability scores shown below the sequence. The selected precursor ions are labeled with an asterisk.

Protein	MW (kDa)	SEQUEST Score		Unique peptides		Sequence coverage	
		UVPD	CID	UVPD	CID	UVPD	CID
α-Casein	24.5	95 ± 7	85 ± 12	7 ± 1	7 ± 1	37% ± 1	37 ± 4
β-Lactoglobulin A	20	103 ± 10	88 ± 18	9 ± 1	8 ± 1	54% ± 3	52% ± 3
BSA	66	123 ± 22	189 ± 20	21 ± 3	28 ± 2	43% ± 6	54% ± 5
Carbonic anhydrase	29	57 ± 13	76 ± 12	8 ± 2	9 ± 2	47% ± 9	53% ± 7
Lysozyme	14.7	78 ± 19	84 ± 12	10 ± 1	11 ± 2	69% ± 3	75% ± 7

Table 6.1 Results for SEQUEST database searches for UVPD and CID.

Table 6.2 displays the average number of PSMs obtained by CID and UVPD for each of the model proteins. For α-casein, β-lactoglobulin A, carbonic anhydrase and lysozyme, UVPD and CID perform very similarly. However, in the case of BSA, CID produced on average 21 more PSMs than UVPD which led to the identification of seven more unique peptides on average. These observations can be rationalized by the effects of ion suppression during ESI and the greater average sensitivity of CID compared to UVPD. The molecular weight of BSA is more than twice that of the next largest protein analyzed and produces 46 tryptic peptides (assuming zero missed cleavages and minimum peptide length of six residues). Since the digests are analyzed without prior chromatographic separations, some peptides will experience significant suppression during ESI relative to others.

Protein	PSMs	
	UVPD	CID
α-Casein	23 ± 2	26 ± 3
β-Lactoglobulin A	30 ± 4	23 ± 4
BSA	38 ± 7	59 ± 5
Carbonic anhydrase	19 ± 4	20 ± 3
Lysozyme	24 ± 3	23 ± 5

Table 6.2 Peptide spectral matches for UVPD and CID obtained for five model proteins

Figure 6.2 shows worst case and best case scenarios for UVPD relative to CID. **Figure 6.2A** is a representative DESI-MS survey scan from one BSA analysis. The ion of *m/z* 767 corresponding to doubly charged LKCCDKPLLEK was chosen for MS/MS analysis and representative UVPD and CID are shown in **Figure 6.2B** and C, respectively. The abundance of the LKCCDKPLLEK peptide is low (~10% relative abundance), and this peptide does not contain any aromatic residues, moderating its native absorption cross-section at 193 nm. The XCorr and probability scores for the CID spectrum are significantly higher and produce a confident identification whereas the UVPD spectrum does not pass the database search results filtering. **Figure 6.2D** and **E** are UVPD and CID spectra, respectively, from the same BSA analyses of doubly charged LGEYGFQNALIVR. The peptide LGEYGFQNALIVR of moderate abundance (~40% relative abundance) contains two aromatic residues. In this case UVPD significantly out-performs CID as indicated by the greater XCorr and probability scores. Once again, the diverse array of product ions produced by UVPD at 193 nm and the ability to detect low mass immonium and terminal sequence ions lead to greater confidence in the peptide match.

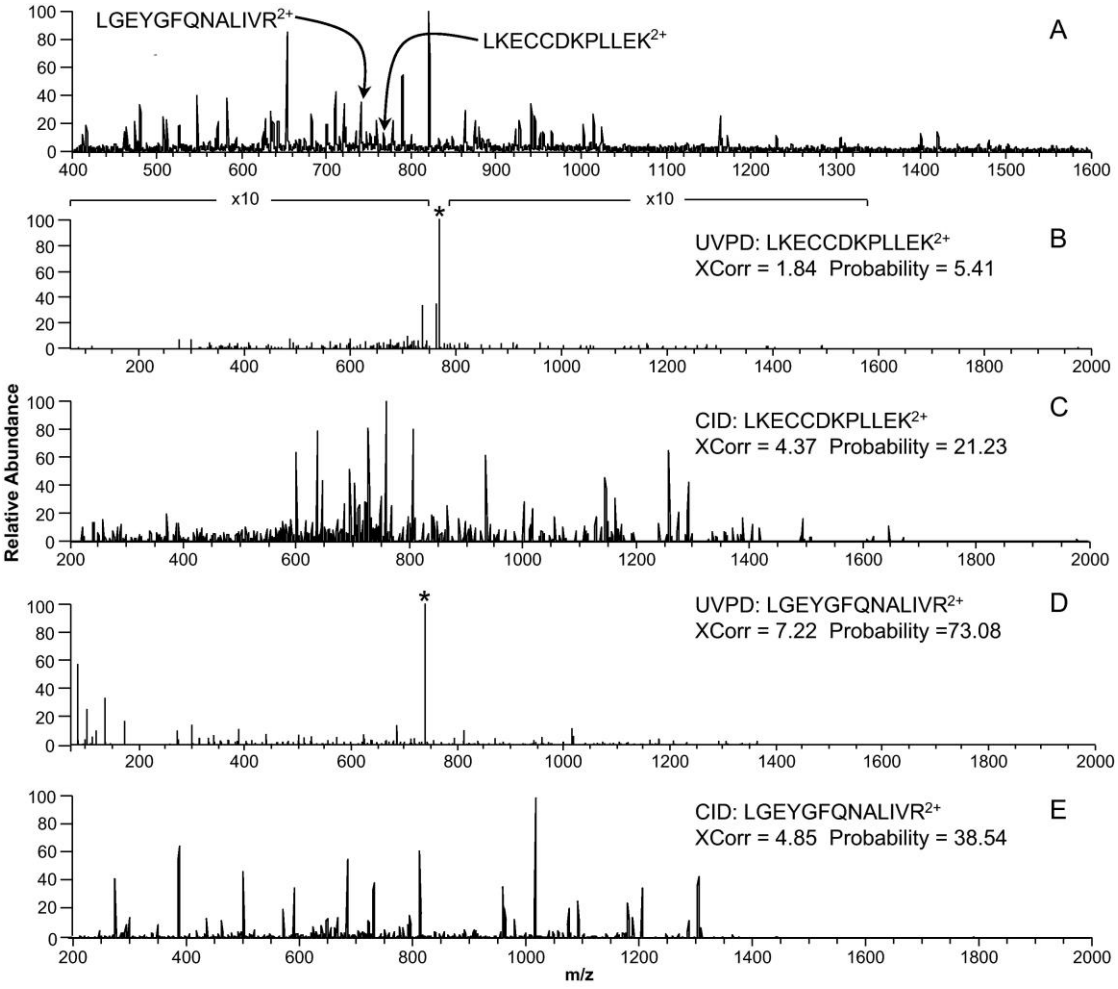

Figure 6.2 A representative survey scan (A) of a BSA tryptic digest, data dependent CID (C and E) and UVPD (B and D) spectra of the peptides LKECCDKPLLEK and LGEYGFQNALIVR. Asterisks indicate the precursor ions.

From the five model proteins, 67 unique peptides were identified by both UVPD and CID. CID identified 10 unique peptides that were not identified by UVPD and seven additional unique peptides were identified by UVPD only. **Figure 6.3** displays the average peptide XCorr (**Top**) and probability (**Bottom**) scores of the top scoring PSM for the unique peptides identified by both CID and UVPD as a function of peptide length. The results are similar with neither method displaying a distinct advantage for any

particular length of peptide. In general, both the XCorr and probability scores increase with peptide length. This trend is due to the additive nature of their calculations, thus spectra of larger peptides will score higher than smaller peptides even for MS/MS spectra of similar quality.[35] The significance is that there are only marginal differences in the peptide XCorr and probability scores produced by CID and UVPD for each peptide length, and overall both methods are very comparable. The key advantage of UVPD is the rich fragmentation information and very short activation time (i.e., a single 5 ns pulse) compared to the standard 30 ms activation period for CID, making the latter less efficient for high-throughput applications.

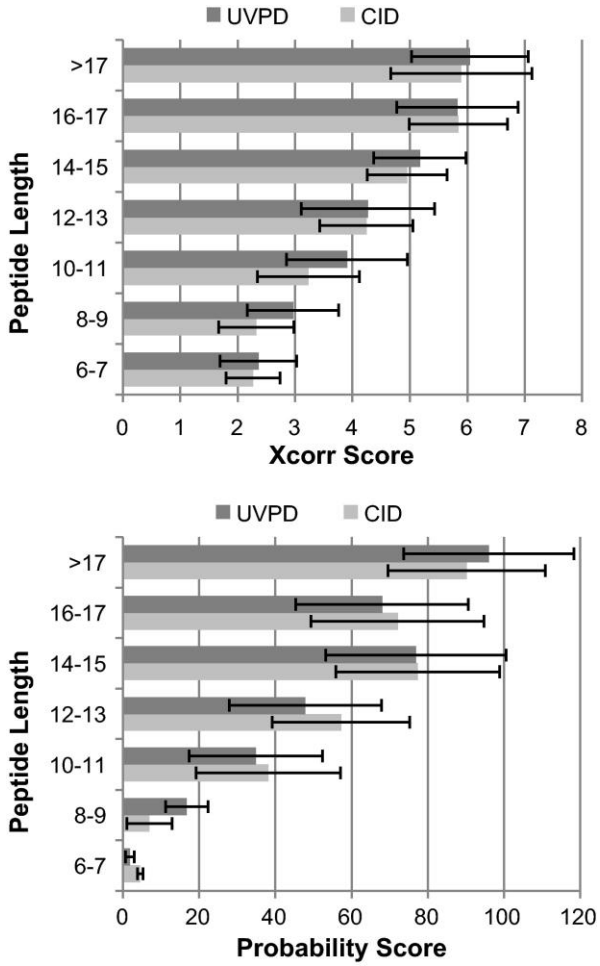

Figure 6.3 Average peptide XCorr (Top) and probability (Bottom) scores of the top scoring PSMs of the peptides identified by both CID and UVPD as a function of peptide length.

6.5 Conclusions

The coupling of TM-DESI with UVPD at 193 nm allowed rapid MS/MS analysis of protein digests and confident protein identification through SEQUEST database searches. In some cases, UVPD at 193 nm provided greater confidence in peptide identification compared to CID due to the extensive array of sequence ion types produced and the ability to detect terminal sequence ions at reduced *q*-values. Design and

production of sample well masks analogous to a 96 well MALDI target would further increase the throughput of TM-DESI experiments for MS/MS sequencing.

6.6 References

(1) Schuerenberg, M.; Luebbert, C.; Eickhoff, H.; Kalkum, M.; Lehrach, H.; Nordhoff, E. Anal. Chem. 2000, 72, 3436–3442.

(2) Mueller, D. R.; Voshol, H.; Waldt, A.; Wiedmann, B.; Van Oostrum, J. In Subcellular Proteomics; Subcellular Biochemistry; Springer Netherlands, 2007; Vol. 43, pp. 355–380–380.

(3) Colinge, J.; Bennett, K. L. PLoS Computational Biology 2007, 3, e114–1160.

(4) Thiede, B.; Hohenwarter, W.; Krah, A.; Mattow, J.; Schmid, M.; Schmidt, F.; Jungblut, P. R. Methods 2005, 35, 237–247.

(5) Cramer, R.; Corless, S. Rapid Commun. Mass Spectrom. 2001, 15, 2058–2066.

(6) Pandey, A.; Mann, M. Nature 2000, 405, 837–846.

(7) Gevaert, K.; Vandekerckhove, J. Electrophoresis 2000, 21, 1145–1154.

(8) Gevaert, K.; Demol, H.; Martens, L.; Hoorelbeke, B.; Puype, M.; Goethals, M.; Damme, J. V.; Boeck, S. D.; Vandekerckhove, J. Electrophoresis 2001, 22, 1645–1651.

(9) Yates, J. R. Electrophoresis 1998, 19, 893–900.

(10) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. ELECTROPHORESIS 1999, 20, 3551–3567.

(11) Takats, Z.; Wiseman, J. M.; Gologan, B.; Cooks, R. G. Science 2004, 306, 471–473.

(12) Weston, D. J. Analyst 2010, 135, 661–668.

(13) Ifa, D. R.; Wu, C.; Ouyang, Z.; Cooks, R. G. Analyst 2010, 135, 669–681.

(14) Bereman, M. S.; Nyadong, L.; Fernandez, F. M.; Muddiman, D. C. Rapid Commun. Mass Spectrom. 2006, 20, 3409–3411.

(15) Kaur-Atwal, G.; Weston, D. J.; Green, P. S.; Crosland, S.; Bonner, P. L. R.; Creaser, C. S. Rapid Commun. Mass Spectrom. 2007, 21, 1131–1138.

(16) Shin, Y.-S.; Drolet, B.; Mayer, R.; Dolence, K.; Basile, F. Anal. Chem. 2007, 79, 3514–3518.

(17) Peng, I. X.; Shiea, J.; Loo, R. R. O.; Loo, J. A. Rapid Commun. Mass Spectrom. 2007, 21, 2541–2546.

(18) Peng, I. X.; Loo, R. R. O.; Margalith, E.; Little, M. W.; Loo, J. A. Analyst 2010, 135, 767–772.

(19) Sen, A.; Nayak, R.; Darabi, J.; Knapp, D. Biomedical Microdevices 2008, 10, 531–538.

(20) Sampson, J. S.; Hawkridge, A. M.; Muddiman, D. C. J. Am. Soc. Mass Spectrom. 2008, 19, 1527–1534.

(21) Pasilis, S.; Kertesz, V.; Van Berkel, G.; Schulz, M.; Schorcht, S. Anal. Bioanal. Chem. 2008, 391, 317–324.

(22) Pasilis, S. P.; Kertesz, V.; Berkel, G. J. V.; Schulz, M.; Schorcht, S. J. Mass Spectrom. 2008, 43, 1627–1635.

(23) Miao, Z.; Chen, H. J. Am. Soc. Mass Spectrom. 2009, 20, 10–19.

(24) Stokes, A. A.; Clarke, D. J.; Weidt, S.; Langridge-Smith, P.; Mackay, C. L. Int. J. Mass Spectrom. 2010, 289, 54–57.

(25) Morgan, J. W.; Russell, D. H. J. Am. Soc. Mass Spectrom. 2006, 17, 721–729.

(26) Thompson, M. S.; Cui, W.; Reilly, J. P. J. Am. Soc. Mass Spectrom. 2007, 18, 1439–1452.

(27) Shin, Y. S.; Moon, J. H.; Kim, M. S. J. Am. Soc. Mass Spectrom. 2010, 21, 53–59.

(28) Zhang, L.; Reilly, J. P. Anal. Chem. 2010, 82, 898–908.

(29) Kim, T. Y.; Thompson, M. S.; Reilly, J. P. Rapid Commun. Mass Spectrom. 2005, 19, 1657–1665.

(30) Kim, T.-Y.; Reilly, J. P. J. Am. Soc. Mass Spectrom. 2009, 20, 2334–2341.

(31) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. J. Proteome Res. 2010, 9, 4205–4214.

(32) Chipuk, J. E.; Brodbelt, J. S. J. Am. Soc. Mass Spectrom. 2008, 19, 1612–1620.

(33) Chipuk, J. E.; Brodbelt, J. S. J. Am. Soc. Mass Spectrom. 2009, 20, 584–592.

(34) Gardner, M. W.; Vasicek, L. A.; Shabbir, S.; Anslyn, E. V.; Brodbelt, J. S. Anal. Chem. 2008, 80, 4807–4819.

(35) Sadygov, R. G.; Cociorva, D.; Yates, J. R. Nat. Methods 2004, 1, 195–202.

Chapter 7

Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation

7.1 Overview

The top-down approach to proteomics offers compelling advantages due to the potential to provide complete characterization of protein sequence and post-translational modifications. Here we describe the implementation of 193 nm ultraviolet photodissociation (UVPD) in an Orbitrap mass spectrometer for characterization of intact proteins. Near-complete fragmentation of proteins up to 29 kDa is achieved, and unambiguous localization of a single residue mutation and protein modifications are demonstrated for Human Pin1, a protein implicated in the development of Alzheimer's disease and cancer pathogenesis. The 5 nanosecond, high-energy activation afforded by UVPD exhibits far less precursor ion charge state dependence than conventional collision-based and electron-based dissociation methods.

7.2 Introduction

Recent advances in instrumentation and experimental design have propelled mass spectrometry to the forefront of proteome research. Conventional bottom-up proteome analysis is based on the ability to sequence the constituent peptides of an enzymatically digested protein mixture. By this approach, the identification of many proteins and post-translation modifications (PTMs) is possible; however, complete characterization of any protein is rare.[1] This limits the ability to determine sequence truncations, single nucleotide polymorphisms (SNPs) and the combinatorial nature of PTMs. The top-down

approach,[2–4] an alternative that involves the interrogation of intact proteins, provides both intact protein and fragment mass measurements and has the potential to overcome the analysis deficiencies noted above (splice variants, PTMs, mutations, etc.).

Despite the promise of the top-down approach, complete characterization (in terms of sequence coverage) has generally been limited to proteins less than 10 kDa due to the inability to efficiently dissociate larger proteins and thus unambiguously characterize whole proteoforms. Slow heating methods, such as collision induced dissociation (CID) and infrared multiphoton dissociation (IRMPD), often yield selective cleavage of the most labile bonds (e.g. non-covalent interactions and amide bonds N-terminal to proline residues) resulting in limited sequence coverage.[5,6] The advent of electron capture dissociation (ECD)[7] and electron transfer dissociation (ETD)[8] has extended the utility of the top-down approach as a result of dissociation occurring prior to energy randomization. This attribute greatly decreases the degree of amino acid-specific cleavages seen with slow heating methods, yielding more random and extensive fragmentation. The major limitations of ECD and ETD stem from the strong dependency on charge density and thus the gas-phase structure of the ions selected for dissociation. The use of collisional activation or infrared photoactivation prior to or during the ECD/ETD process, termed activated-ion ECD/ETD,[9–11] has been implemented to disrupt protein tertiary structure and allows effective generation and detection of ECD/ETD products of larger proteins. In addition, nozzle-skimmer dissociation[12,13] has extended the mass range of proteins that can be dissociated to greater than 200 kDa; however, nozzle-skimmer dissociation is inherently non-specific because it occurs prior to selection of a precursor. The challenge of protein separations remains a hurdle associated with top-down proteomics, but this factor primarily limits throughput rather than impeding sequence coverage.[14]

193 nm ultraviolet photodissociation (UVPD) has been shown to be a viable dissociation method for the characterization of peptides[15–20] and small proteins cations.[21–23] UVPD at 193 nm of peptide cations yields extensive fragmentation producing *a, b, c, x, y,* Y*, z, v, w* and *d* ions.[17,20,21,23,19] Here we present the implementation of 193 nm UVPD in a hybrid linear ion trap Orbitrap mass spectrometer (Thermo Scientific Orbitrap Elite[24]) and the application of UVPD to top-down proteomics.

7.3 Experimental

7.3.1 Materials

Ubiquitin from bovine erythrocytes, carbonic anhydrase from bovine erythrocytes, and myoglobin from equine skeletal muscle were purchased from Sigma-Aldrich and used without further purification. All other chemicals were from Thermo Fisher Scientific.

7.3.2 Sample Preparation

Human Pin1 expression, purification and oxidation. The human Pin1 gene was subcloned in a pHIS8 vector, a derivative of pET28a vector (Novagene).[25] The Pin1 R14A mutant was produced using the QuikChange Site-Directed Mutagenesis Kit (Stratagene, CA). The purification of Pin1 R14A mutant was similar to previous reported procedure.[26] Briefly, the Pin1 R14A was overexpressed in using E. coli BL21(DE3) strain at 16°C overnight induced by isopropyl-β-D-thiogalactopyranoside (IPTG). The N-terminal polyhistidine tag was removed by thrombin protease (Novegen Germany) during the overnight dialysis after eluted from Ni-NTA (Invitrogen NY) purification. The

protein was further purified by gel filtration superdex75 (GE Healthcare) in 20 mM HEPES 7.5 and 50 mM NaCl.

Purified Pin1 was treated with mild oxidizing condition similarly to previous published protocol.[27] 2 mg/ml purified Pin1 R14A protein was incubated with 30 μM $FeSO_4$/1 mM H_2O_2 for 3hours. The oxidation was stopped by removing hydroxyl free radicals using dialysis in 20 mM HEPES, 50 mM NaCl overnight. Samples were spun down for 20 min at 20,000 x g to remove possible aggregation particles then concentrated and desalted using a 10 kDa centrifugal filter (Millipore Corporation) prior to mass spectrometric analysis.

7.3.3 Implementation of UVPD

UVPD was implemented in a manner similar to what we have previously described for infrared photodissociation.[28] Briefly, Experiments were performed using a modified Thermo Scientific Orbitrap Elite mass spectrometer (Thermo Fisher Scientific) equipped with ETD (**Figure 7.1**). The portion of the vacuum manifold containing the higher energy collision dissociation (HCD) cell and a portion of the ETD reagent transfer ion optics was modified to incorporate an optical window (quartz with antireflective coating) concentric with the HCD cell. An ETD reagent transfer bent quadrupole with rods rotated 45 degrees was used, so the laser beam could be transmitted between the rods. An optical periscope (Thorlabs) consisting of two mirrors (25 mm, Edmund Optics) was used to elevate and align the laser beam coaxial to the HCD cell. The excimer laser was mounted with the beam output concentric with the bottom mirror of the optical periscope. The instrument firmware was modified to allow trapping of ions in the HCD cell for a user specified amount of time and to output a TTL trigger during the HCD scan

events. The TTL trigger was used gate a pulse/delay generator (Berkeley Nucleonics model 505) which triggered each pulse of an ArF excimer laser (Coherent Excistar XS). Direct measurements of the collision gas pressure were made via a flange-mounted pirani gauge connected to the HCD cell housing by 6 mm tubing.

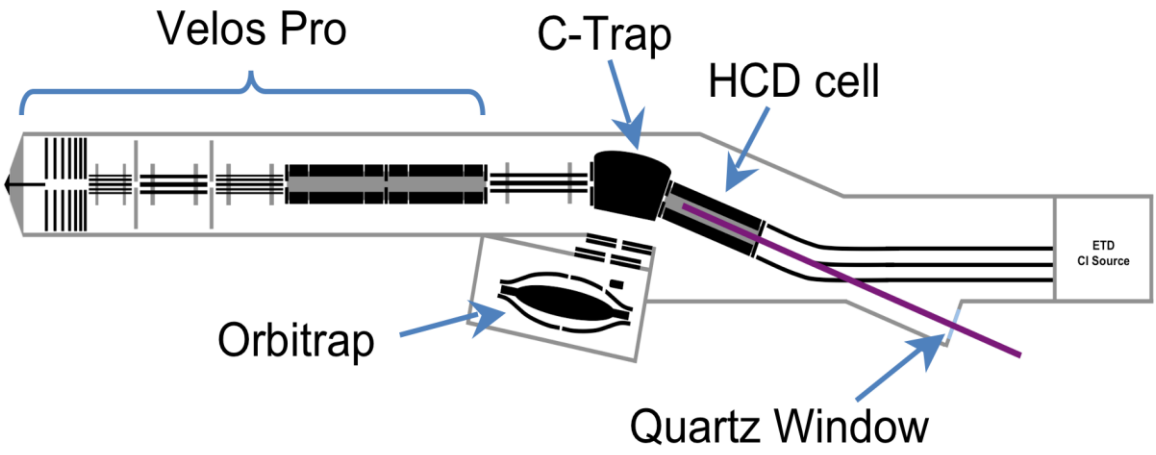

Figure 7.1 Schematic showing modifications made to the Thermo Scientific Orbitrap Elite mass spectrometer to allow photodissociation in the HCD cell.

7.3.4 Mass Spectrometry

Protein solutions were prepared at 10 μM in 49.5:49.5:1 (v/v/v) water:acetonitrile:formic acid, and were analyzed by direct infusion at a flow rate of 3 μL/min. Spectra for all dissociation methods were acquired using a mass range of 200-2000 *m/z* and resolving power of 240,000 at *m/z* 400. The AGC target for MS2 was set to one million, and isolation widths of 8 *m/z* and 25 *m/z* were used for model proteins and Pin1, respectively. Normalized collision energies for CID and HCD were 15-30%, and ETD reaction times were 5-15 ms with a reagent AGC target of three hundred thousand. For UVPD experiments, precursor ions were transferred to HCD cell with normalized

collision energy of 1% (no collisional induced dissociation occurs at this setting), and spectra were acquired using one laser pulse at 1-4 mJ/pulse. The HCD collision gas pressure was reduced to a pressure measured as a delta of 0.1E-10 Torr in the UHV portion of the vacuum chamber containing the Orbitrap analyzer (5 mTorr collision gas pressure). To afford the most meaningful comparisons, the same amount of spectral averaging was performed for each dissociation method; however, the extent of averaging varied with protein molecular weight. 50, 200, and 500 scans were averaged for ubiquitin, myoglobin and carbonic anhydrase, respectively. This amount of averaging corresponded to acquisition times of approximately 45 sec, 2.5 min. and 6 min. for ubiquitin, myoglobin and carbonic anhydrase, respectively.

LC-MS/UVPD data was collected using a data dependent top three acquisition strategy. MS^1 survey scans with 3 microscans at a resolving power 120,000 at *m/z* 400 were used. MS/MS acquisitions consisted of an 8 *m/z* isolation width, six microscans, a resolving power of 60,000 at *m/z* 400, NCE of 1% and an activation time of 1 ms. An AGC target value of one million was used for MS^1 and MS/MS acquisitions.

7.3.5 Yeast Whole Cell Lysate Preparation

S. Cerevisiae was grown to an OD600 of 0.7, harvested, and lysed in SDS lysis buffer. 500 ug of total protein was loaded to a 12%-T commercial GELFrEE device and proteins were resolved by SDS-PAGE and eluted off the gel. Fractions were collected and 10% of each fraction was loaded to a conventional SDS-PAGE slab gel for visualization by silver nitrate staining. The remainder of each fraction was subjected to a methanol/chloroform/water cleanup procedure to remove the SDS from the protein. The SDS-free samples were dried down and the pellets were stored at -80 C until the of use.

Each pellet was reconstituted and pipetted vigorously 20 – 100 times, and then sonicated in a bath sonicator on ice for 20 minutes. Finally, the reconstituted samples were spun down at high speed (21,000 x g) for 10 minutes.

7.3.6 Liquid Chromatography

Separations were performed using an Eksigent nanoLC-ultra HPLC equipped with self-packed 15 cm analytical and 2 cm trap columns. The packing media was C4, 300 angstrom pore, 5 micron Magic stationary phase (Michrom Bioresources, Inc.). Eluents A and B were 0.1% formic acid in water and acetonitrile, respectively. Reconstituted fractions were injected (5 μL) onto the trap column and eluted using a 90 minute gradient consisting of a linear increase to 15% eluent B in the first 10 minutes followed by a linear increase to 60% eluent B at 90 minutes. A flow rate of 300 nL/min was used.

7.3.7 Data Processing

MS^2 spectra were deconvolved using the Xtract algorithm (Thermo Fisher Scientific) with S/N threshold of 3. CID, HCD, ETD and UVPD spectra were converted to either neutral (ProSightPC) or protonated (manual interpretation) monoisotopic masses. The monoisotopic spectra for all dissociation methods were interpreted using a custom version of ProSightPC 3.0 (Thermo Fisher Scientific) which included the ability to search *a/x*, *b/y* and *c/z* ion pairs in a single search in order to accommodate the complete array of ions observed in UVPD spectra. Single protein mode with fragment mass tolerance set to 10 ppm was used for all fragmentation methods. Manual interpretation of UVPD data was performed using the combination of unprocessed raw

spectra and deconvolved spectra by comparison to theoretical fragment masses generated by ProteinProspector v5.10.3 (http://prospector.ucsf.edu). Sequence coverage was calculated as a percentage value based on the total number of observed inter-residue cleavages (considering all ion series) divided by the number of total number of inter-residue bonds in the protein.

Automated database searches for LC-MS/MS data were performed using ProSightPC. Spectra were converted to neutral monoisotopic masses using the Xtract algorithm considering product ions of S/N 3 and greater and a maximum charge of 25. Spectra were searched against the "top_down_complex" Saccharomyces cerevisiae database downloaded from the ProSightPC database website (ftp://prosightpc.northwestern.edu/). The database contained 23564 basic sequences and 1196890 protein forms. Searches were performed in absolute mass mode with a precursor mass tolerance of 500 Da, 10 ppm fragment ion tolerance, a minimum of 10 fragment ions and with the Δm option. Hits were subsequently filtered to remove one with a P score greater than 1E-4 and mass difference greater than ± 3 Da.

7.4 Results and Discussion

Precursor ions were mass selected in the linear ion trap then transferred to and stored in the HCD cell where UVPD was performed. After a single 5 ns laser pulse, product ions were transferred to the Orbitrap analyzer for high resolution and mass accuracy analysis. Reduction of the collision gas pressure in the HCD cell was not required as collisional cooling is not a major competing pathway to dissociation by UVPD at 193 nm; however, reduction of the collision gas pressure does yield lower pressure in the Orbitrap analyzer and better data quality. Reduction of the collision gas

pressure allowed detection of low abundance and larger product ions not observed at the standard collision gas pressure (**Figure 7.2**).

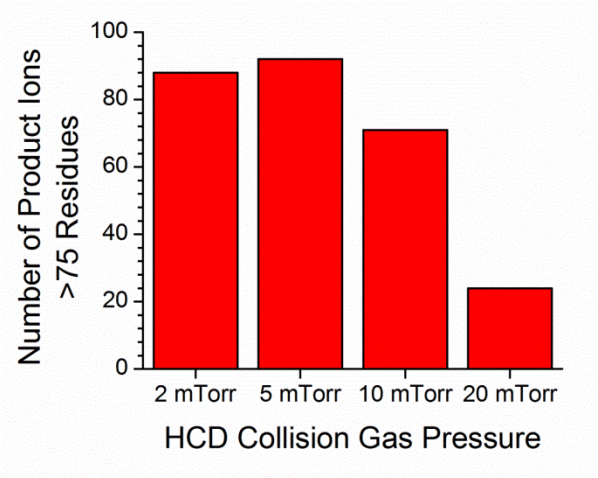

Figure 7.2 The number of UVPD product ions observed containing greater than 75 amino acids for the 20+ charge state of myoglobin as a function of HCD collision gas pressure. 10 mTorr is normal operating pressure and 5 mTorr was used for all other experiments.

First, the small protein ubiquitin (76 residues, 8.6 kDa) was analyzed to benchmark UVPD fragmentation characteristics against other dissociation methods. A representative UVPD mass spectrum is shown in **Figure 7.3A** for the 11+ charge state of ubiquitin. The inset, which illustrates an expanded *m/z* region in **Figure 7.3A,** shows that even low abundance product ions are highly resolved and detected with high S/N. For comparison, UVPD, ETD, HCD and CID were performed for the 7+ to 13+ charge states of ubiquitin. The sequence coverage produced (**Figure 7.4A**) by ETD, HCD and CID varied substantially with ETD exhibiting significantly greater sequence coverage for high charge states and HCD and CID performing better for lower charge states. Unlike these well-studied activation methods, UVPD exhibited very little charge state dependence, producing 100% sequence coverage for the 7+ through 12+ charge states and only one

missed cleavage (99% sequence coverage) for the 13+ charge state by manual interpretation. Manual interpretation of the UVPD spectra yielded similar results to those obtained by ProSightPC, albeit with some differences attributed to incorrect deconvolution of low mass and low charge state product ions. This discrepancy arises from low S/N of ^{13}C isotopic peaks of low molecular weight ions and the overlap of isotopic distributions of low charge state ions with the isotopic distributions of more abundant and highly charged ions.[29] Increasing the S/N threshold for deconvolution from 1.5 to 10 yielded less than a seven percent change in the sequence coverage produced by UVPD for all charge states of ubiquitin (**Figure 7.5**).

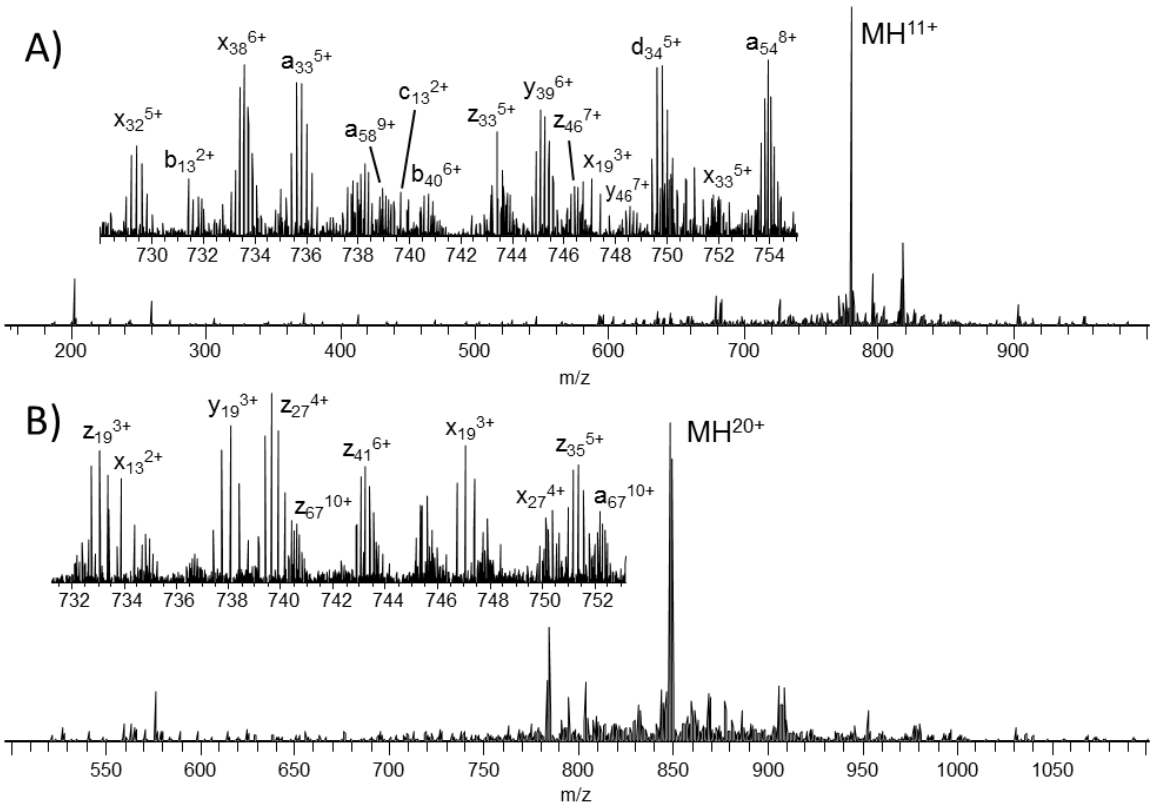

Figure 7.3 UVPD spectra of the 11+ charge state of (A) ubiquitin and (B) the 20+ charge state of myoglobin.

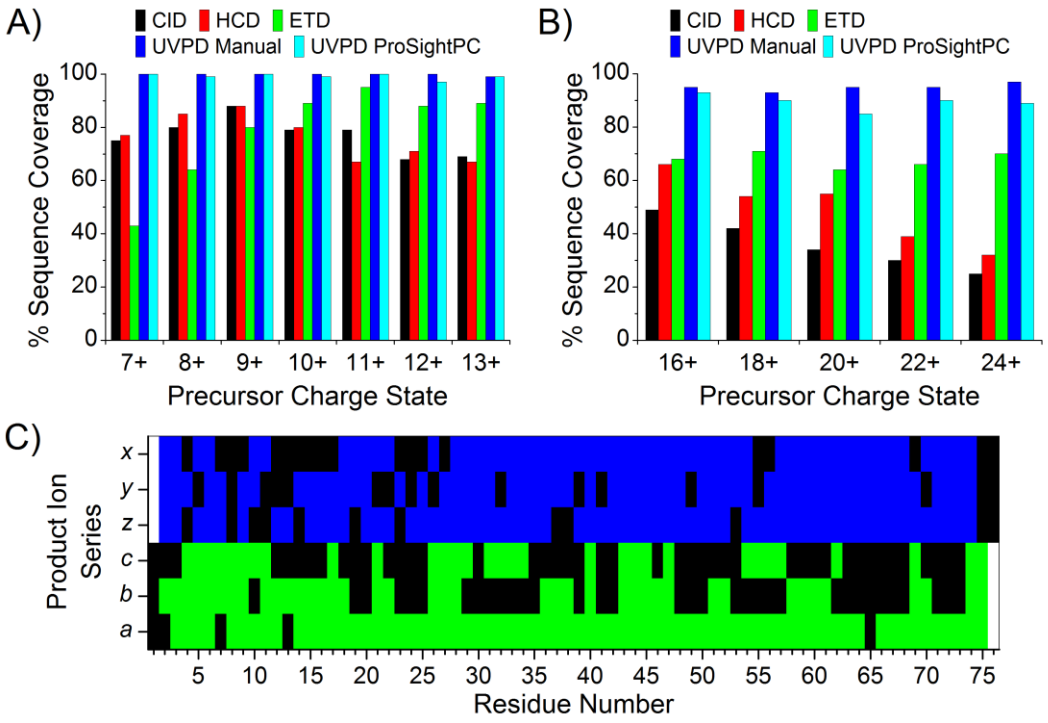

Figure 7.4 Sequence coverage produced by CID, HCD, ETD and UVPD as a function of precursor ion charge state for (A) ubiquitin and (B) myoglobin. (C) UVPD product ion array for the 10+ charge state of ubiquitin showing the distribution of cleavages throughout the protein. Green and blue areas indicate the observed cleavage for N-terminal and C-terminal product ions, respectively.

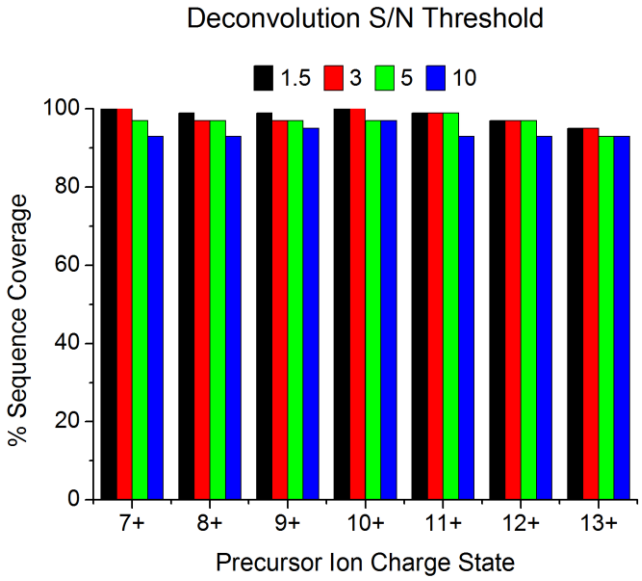

Figure 7.5 Graph of precent sequence coverage obtained for isolated precursor ion charge states of ubiquitin as a function of Xtract deconvolution S/N threshold.

The extensive fragmentation and array of product ion types produced by UVPD enabled *de novo* sequencing of every residue of ubiquitin based on the product ions observed; however, the development of necessary informatics tools and practical implementation of this methodology is not trivial and would require significant effort for larger proteins. UVPD of the 10+ charge state yielded nearly complete sequence coverage from both N-terminal and C-terminal product ions (only 4 and 3 missed cleavages from N-terminal and C-terminal product ions, respectively). Product ions from two or more N-terminal or C-terminal ion series (i.e. *a, b, c* or *x, y, z* ions) were observed for 46 and 63 of the 75 inter-residue bonds sequenced from the N-terminus and C-terminus, respectively (**Figure 7.4C**).

The same experiments were also performed for selected charge states of myoglobin (152 residues, 17 kDa) (**Figures 7.3B** and **7.4B**), and the advantages of UVPD are even more pronounced. As with ubiquitin, HCD and CID yielded greater

sequence coverage for the lower charge states of myoglobin with HCD producing the greatest coverage (66%) for the 16+ charge state. ETD did not exhibit the same preference for higher charge states as noted for ubiquitin, yielding approximately equal sequence coverage (64-71%) for 16+ to 24+ charge states. The lower than expected performance of ETD compared to the results for ubiquitin may be due to the greater number of non-covalent interactions in myoglobin which maintain secondary and tertiary structure and prevent release and detection of many of the product ions. UVPD clearly outperforms HCD, CID and ETD for myoglobin, yielding greater than 93% sequence coverage (by manual interpretation) for each individual charge state interrogated here. There is slightly greater discrepancy in sequence coverage between the manual interpretation method and the use of ProSightPC for myoglobin compared to ubiquitin due to increasing spectral complexity and the reasons discussed above.

Figure 7.6 shows the comparative spectra for CID, HCD, ETD and UVPD of the 20+ charge state of myoglobin. For all charge states analyzed, the CID and HCD spectra were dominated by cleavages N-terminal to proline residues,[30] C-terminal to acidic residues[31] and neutral losses of water and ammonia. Due to the higher energy deposition of HCD compared to CID, the HCD spectra also contained a significant number of low mass internal ions produced by multiple cleavages of the protein backbone. The predominant products in the ETD spectra of myoglobin were charge reduced precursor ions. For the corresponding UVPD spectra, there were surprisingly few neutral losses from the precursor or product ions despite the high energy deposition of 193 nm UVPD (6.4 eV photons). The conversion efficiency (i.e. precursor to product ions) for both ETD and UVPD were rather low for the parameters used in these experiments, but this could be modulated by adjusting the ETD reaction time and laser pulse energy for UVPD. 100% conversion is possible for both UVPD and ETD; however, greater precursor

depletion does not necessarily produce a better outcome for UVPD or ETD due to secondary dissociation of product ions.

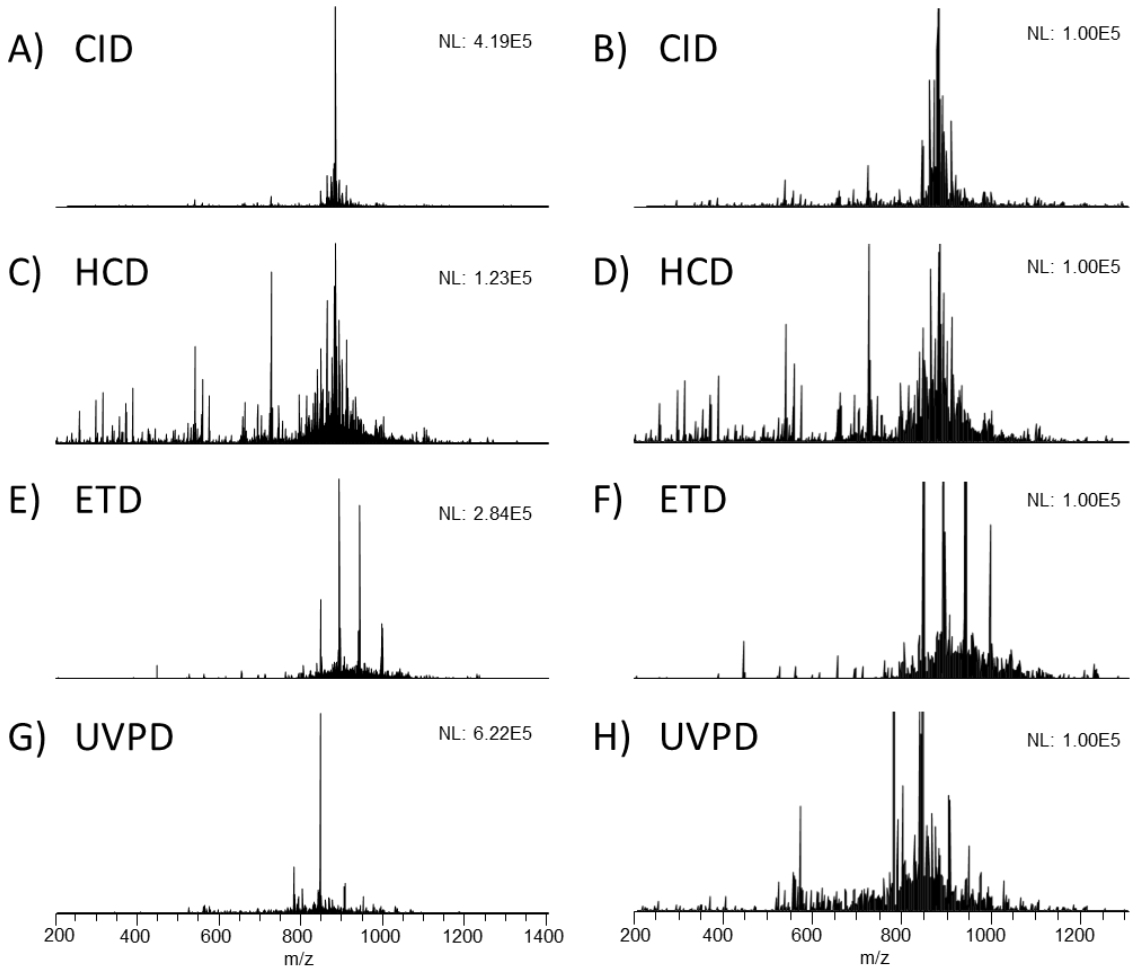

Figure 7.6 Full scale spectra produced by (A) CID, (C) HCD, (E) ETD and (G) UVPD of the 20+ charge state of myoglobin. The same (B) CID, (D) HCD, (F) ETD and (H) UVPD spectra of myoglobin scaled to the same absolute intensity (1.00E5).

The dominant ion series produced by UVPD for all charge states of ubiquitin and myoglobin (**Figure 7.7A, B**) were *a, x, y* and *z* ions with relatively few *b*, *c*, *v*, *w*, and *d* ions observed. The *a, x, y* and *z* ions exhibited significant contributions from both even

and odd electron forms, thus yielding mass shifts of ±1.0078 Da and favoring the $a + 1$, x, $y - 1$, and z species as summarized for myoglobin in **Figure 7.8**. In general, our results are consistent with previous peptide and small protein UVPD studies, suggesting 193 nm UVPD promotes backbone cleavage through pathways occurring prior to intramolecular vibrational re-distribution as well as others on longer time scales subsequent to intramolecular vibrational re-distribution.[22,32,33]

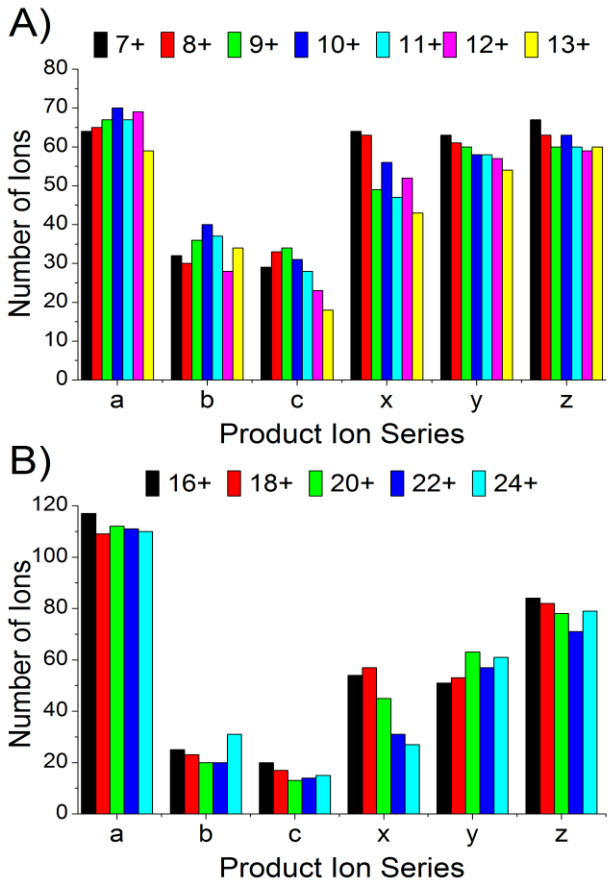

Figure 7.7 Number of each product ion type observed as a function of precursor ion charge state for (A) ubiquitin and (B) myoglobin.

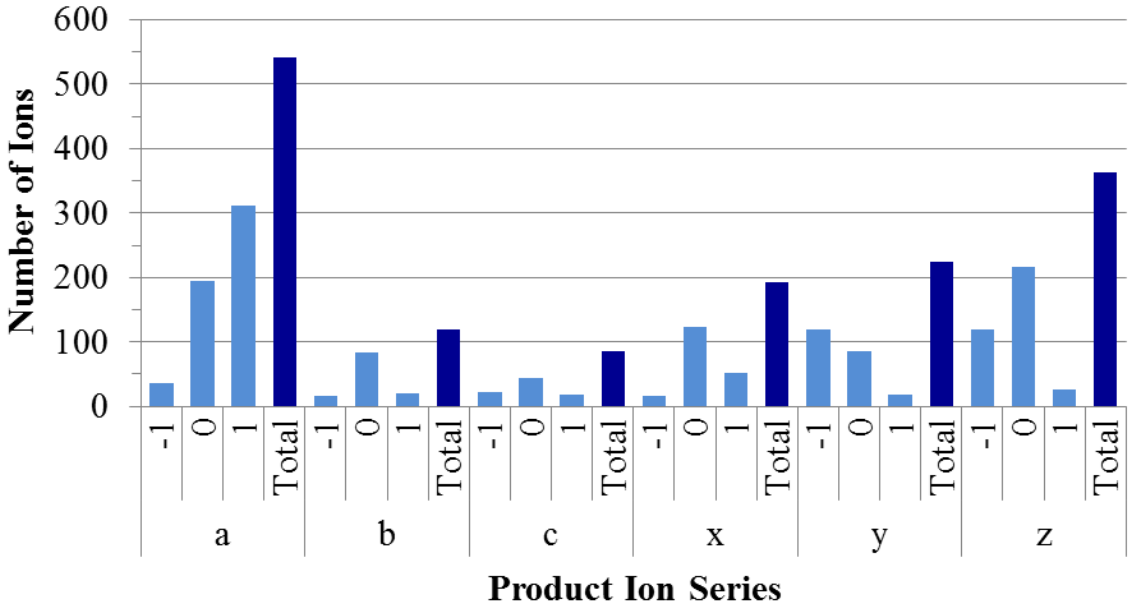

Figure 7.8 Distribution of UVPD product ions for myoglobin. Deconvoluted UVPD spectra were automatically compared to theoretical fragment ion masses using an altered version of ProSightPC created to annotate and aggregate single protein data with and without hydrogen atom shifts. Fragment ions were grouped based on the loss or gain of one hydrogen atom observed collectively for the 16+, 18+, 20+, 22+ and 24+ charge states of myoglobin. Zero represents the even electron species for all product ion series except for *z* ions where zero represents the radical species; -1 and +1 indicate the loss or gain of a hydrogen atom, respectively. Some amount of +/- 1 Da shifts are inevitable in automated interpretation of isotopic distributions in dense data sets like those created by UVPD; the amount of +/- 1 Da species for the *b*- and *c*-type ions reflect this approximate level. The +/- 1 Da shifts are enhanced for the *a, x, y* and *z* ions. Inspection of the analogous trends for other proteins in this study mirrored that found for myoglobin.

The excess energy imparted upon absorption of a 193 nm photon (~3 eV greater than typical peptide backbone bond energies) and the combination of fast and slow fragmentation timescales are likely the key to the success of UVPD for the dissociation of larger proteins. UVPD of the 34+ charge state of carbonic anhydrase II (29 kDa, 259 residues) yielded 150 *a*, 38 *b*, 15 *c*, 54 x, 71 *y*, and 64 *z* ions (**Figure 7.9**). These fragments corresponded to cleavage at 225 (87%) of the 258 interresidue sites. Extensive

fragmentation is observed at the termini of the protein as well as the central portion of the sequence. This is the greatest sequence coverage that has been reported for carbonic anhydrase (to the best of our knowledge). HCD and ETD of the 34+ charge state of carbonic anhydrase yielded 30% and 70% sequence coverage, respectively, and are consistent with results recently published also using a Thermo Scientific Orbitrap Elite mass spectrometer.[24] Previously, plasma-ECD was shown to produce cleavages at 197 of the 258 inter-residue bonds of carbonic anhydrase II from spectra acquired under multiple experimental conditions.[34]

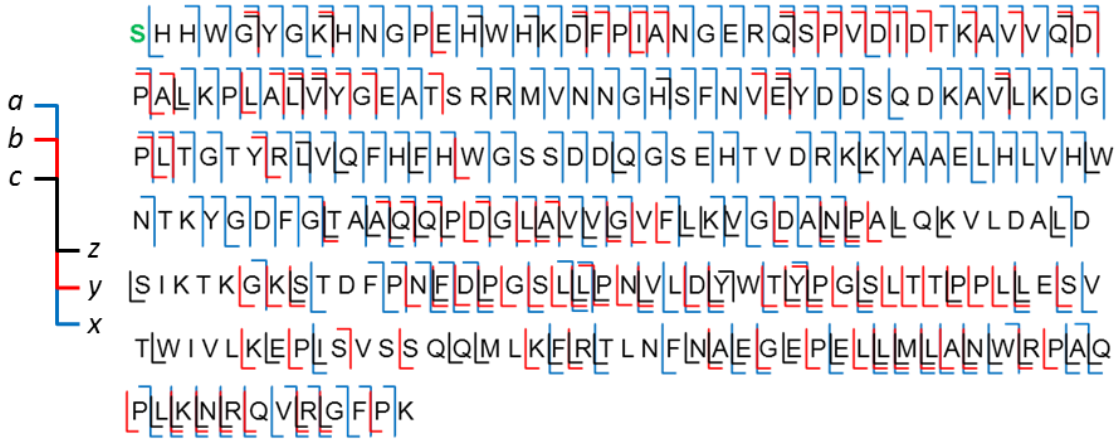

Figure 7.9 Fragment ion map of UVPD product ions from the 34+ charge state of bovine carbonic anhydrase II. Green "S" indicates N-terminal acetylation.

The potential to provide unambiguous determination of SNPs, sequence truncations and the combinatorial nature of PTMs are the major attractors of top-down mass spectrometry, and 193 nm UVPD provides a means for fulfilling these goals. To demonstrate, we studied the oxidation of peptidyl-prolyl *cis/trans* isomerase Pin1. Human Pin1 (UniProt Accession Number: O13526) regulates the signal transduction by conformational changes of phosphorylated Ser/Thr-Pro peptide bonds upon proline isomerization.[35] Due to the prevalence of the Ser/Thr-Pro motif in signaling pathways,

the physiological function of Pin1 has been demonstrated to be critical for the development of pathological processes including cancer and Alzheimer's disease.[36] Pin1 has been shown to be inactivated in neurons of patients with Alzheimer's disease consistent with the accumulation of *cis* form of phosphorylated tau observed at this stage.[37] Pin1 activity is reduced in the early stage of AD through modification by oxidation.[38,39] The identification of oxidation sites of Pin1 requires considerable sequencing detail and has not been accomplished.

Intact mass analysis of Pin1, incubated with hydrogen peroxide for 3 hours (**Figure 7.10** top right), reveals that the primary products are doubly, triply and quadruply oxidized species. 193 nm UVPD of all 22+ charge state species produced 96% sequence coverage for Pin1 and allowed unambiguous determination of dioxidation of C113 (to cysteine sulfinic acid) as the major oxidation product with lower abundance oxidations detected at M5 and M19 and no detectable oxidation present at C57, M130 or M146. Inspection of a crystal structure of Pin1[40] (**Figure 7.10** top left) shows the catalytically-essential C113 is located in a loop at the surface of the C-terminal PPIase domain, whereas C57 is buried at the interface of two β-sheets and an α-helix. This supports the observation of C113 oxidation and no oxidation of C57. Consecutive *a* ions allow unambiguous localization of a single residue mutation of R14A and partial oxidation of M15, while consecutive *x* ions show that oxidation of M146 is not observed in detectable abundance (**Figure 7.10** bottom). A mechanistic rationale for this particular oxidation motif with respect to the progression of Alzheimer's disease would be premature, but this result demonstrates the ability of UVPD to provide such rich fragmentation spectra that the precise characterization of individual proteoforms associated with a certain disease type becomes possible.

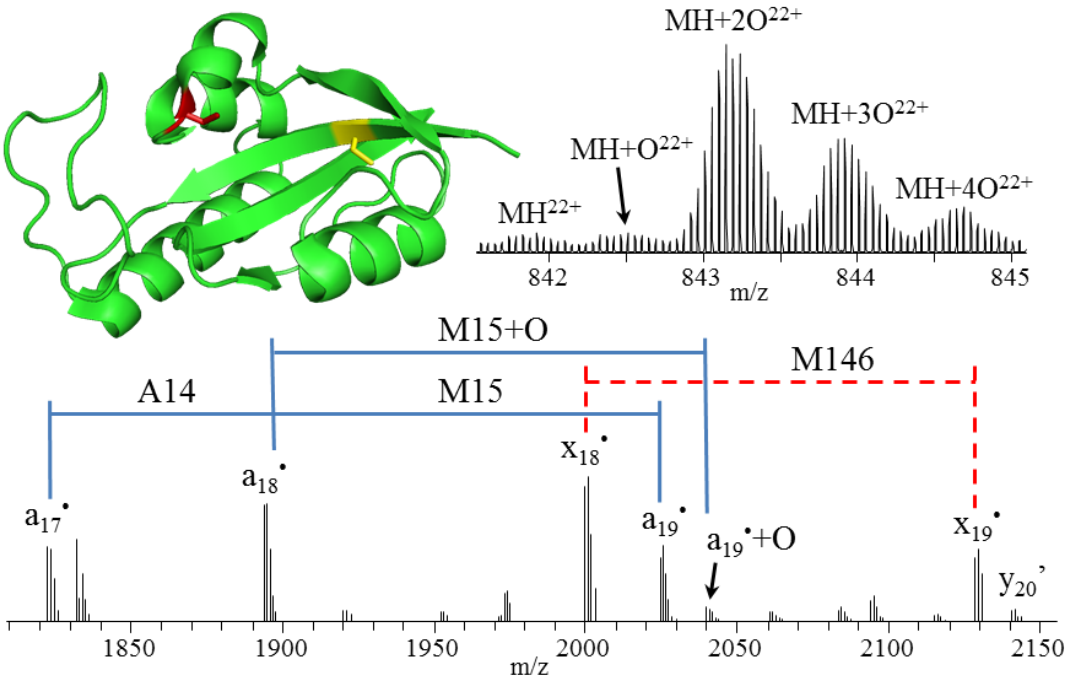

Figure 7.10 Crystal structure (PDB ID 3NTP) of Pin1 (top left) showing the locations of C57 and C113, which are observed in non-oxidized (C57, yellow) and oxidized (C113, red) forms, respectively. Mass spectrum of intact Pin1 (top right) showing doubly and triply oxidized species as the most abundant. Zoomed-in section of a 193 nm UVPD mass spectrum (bottom) of the 22+ charge state species, deconvolved to singly protonated species. Apostrophe (') indicates the neutral loss of water.

Proteins labile PTMs, such as phosphorylation, have proven difficult to analyze with conventional dissociation methods like CID due to abundant neutral losses of the PTM. As a result, protein sequence determination and localization of the PTM is insufficient. One of the major reasons for the success of ECD and ETD for protein analysis is the ability to dissociate protein ions while preserving the labile PTMs. To explore the capablities of UVPD for labile PTM analysis, bovine alpha-casein, which contains eight phosphorylation sites, was analyzed. To our surprise, despite the high energy activation and wide range of dissociation timescales produced by UVPD, minimal

neutral loss of the phosphate was observed from the precursor or product ions. **Figure 7.11A** shows the raw UVPD mass spectrum of the 20+ charge of α-casein. Two very abundant product ions are observed at approximately *m/z* 400. These ions are y_3 and a_{11}, which both correspond to cleavage N-terminal to proline residues (α-casein contains 17 proline residues). The neutral monoisotopic mass spectrum zoomed in on the precursor mass range (**Figure 7.11B**) shows a complete absence of neutral loss of phosphate ions from the precursor. The majority of the ions in this region are the result of side chain losses from the precursor, including a small amount of the phosphorylated serine side chain. The fragment ion map (**Figure 7.11C**) for UVPD of the 20+ charge state of α-casein shows that very good sequence coverage (85%) is achieved and every phosphorylation can be localized to a single residue.

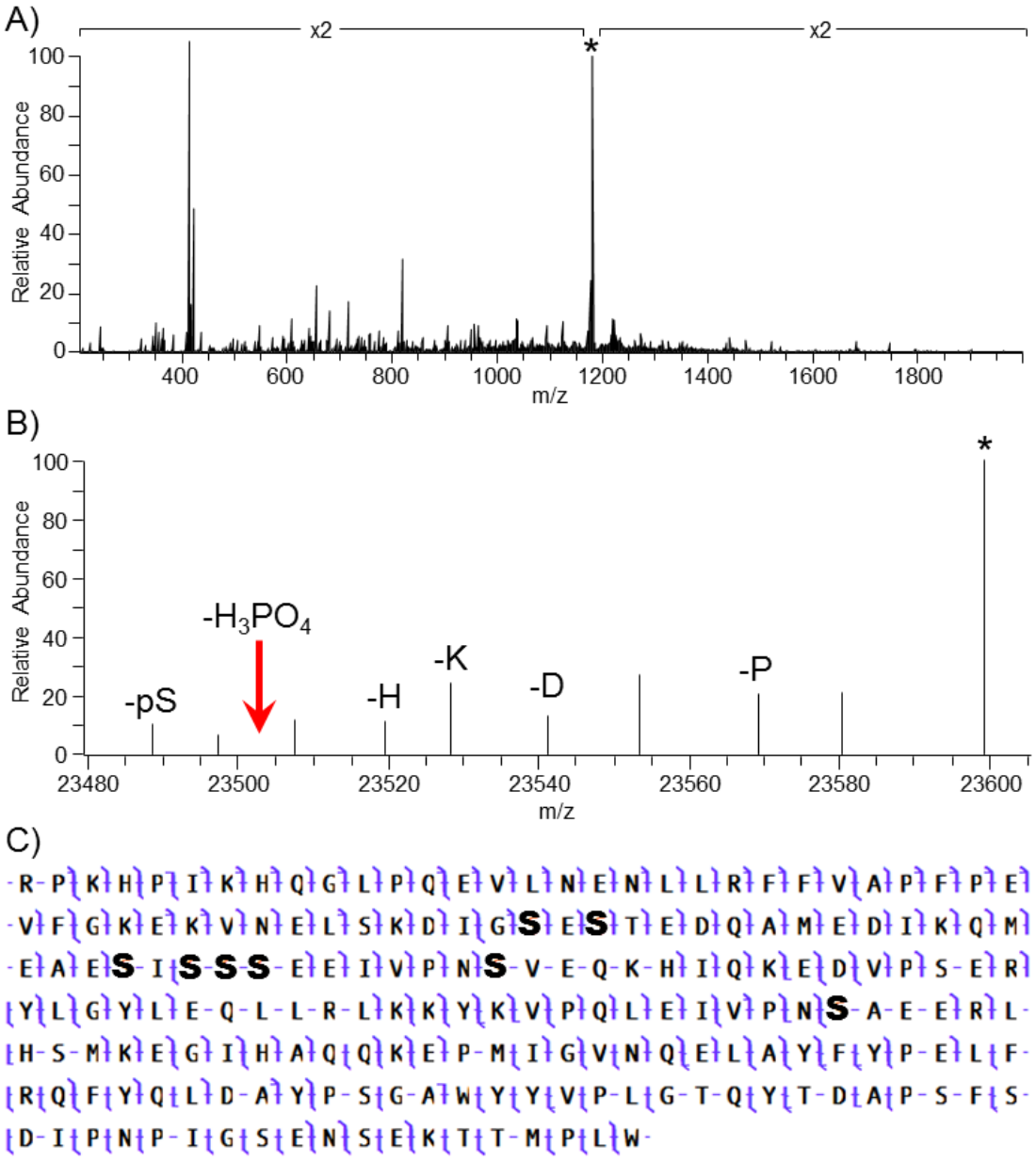

Figure 7.11 Raw UVPD spectrum of the 20+ charge state of α-casein (A), neutral monoisotopic mass spectrum zoomed in on the precursor ion regions (B) and the UVPD fragment ion map (C). Asterisks indicate precursor ions and bold 'S' indicates a phosphorylated serine residue.

Seven GELFrEE fractions of less than 30 kDa proteins from a yeast whole cell lysate were analyzed in triplicate by LC-MS/UVPD. Observed protein molecular weights

ranged from approximately 4 kDa to 20 kDa and there was a large overlap in the proteins observed between fractions. Laser pulse energy was set to 4 mJ/pulse for the lowest molecular weight fraction and incremented down for increasing molecular weight fractions. A resolving power of 60,000 at *m/z* 400 was chosen to minimized acquisition time while still maintaining adequate performance for these relatively low molecular weight proteins. In total, 177 unique accession numbers and 201 unique proteoforms were identified.

The raw UVPD spectrum for Zeocin resistance protein 1 (accession # Q08245) and fragment ion map are shown if **Figure 7.12A** and illustrates the ability to acquire high quality data on an LC-timescale. High S/N product ions are observed throughout the protein as is evidenced by the fragment ion map (**Figure 7.12 A**). 64 a-ions, 29 x-ions, 16 y-ions and 40 z-ions were identified by ProSightPC which yielded 75% sequence coverage and the ability to unambiguously localize the site of acetylation and phosphorylation to the N-terminus and Ser39, respectively. Equally high performance is achieved for 60S ribosomal protein L27-A (accession # P0C2H6) as shown in **Figure 7.12B**. An extensive series of N- and C-terminal product ions are observed for this 15 kDa protein.

To access the quality of fractionation yielded by the GELFrEE method, the percent overlap between fractions in terms of protein accession number identifications is shown in **Figure 7.13**. All fractions showed at least 80% overlap in the protein accession numbers identified. In total 177 unique protein accession numbers were identified; however, none of the fractions yielded more than 14 unique accession numbers that were not identified in other fractions. More refinement of the GELFrEE method would yield significantly less overlap and improve the overall result for LC-MS/UVPD.

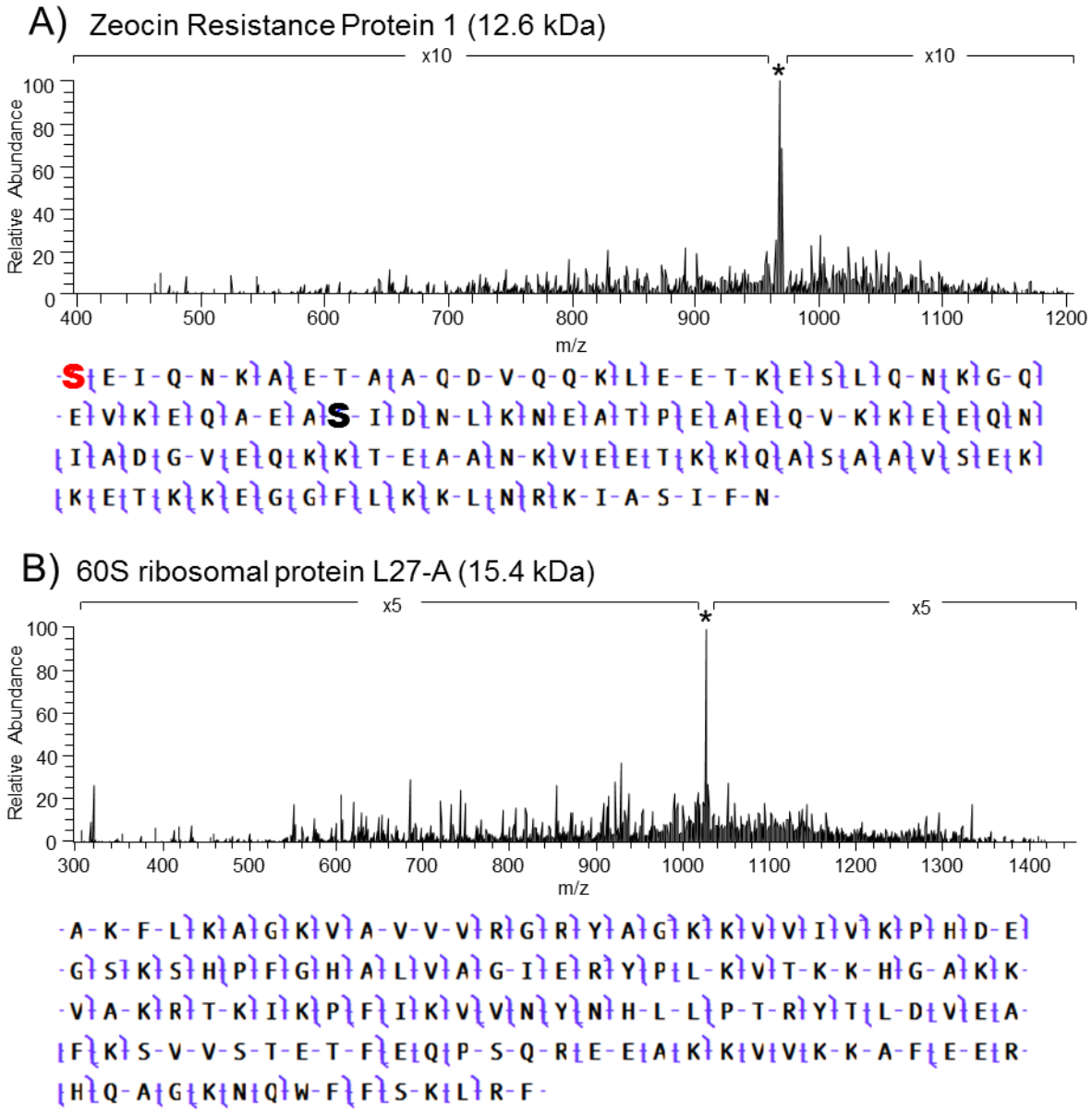

Figure 7.12 Example LC-MS/UVPD spectra and the corresponding fragment ion maps for Zeocin Resistance Proein 1 (A) and 60S ribosomal protein L27-A (B). UVPD spectra were collected using 6 microscans at a resolving power of 60,000 at *m/z* 400. Asterisks indicate precursor ions and bold 'S' indicates a phosphorylated serine residue.

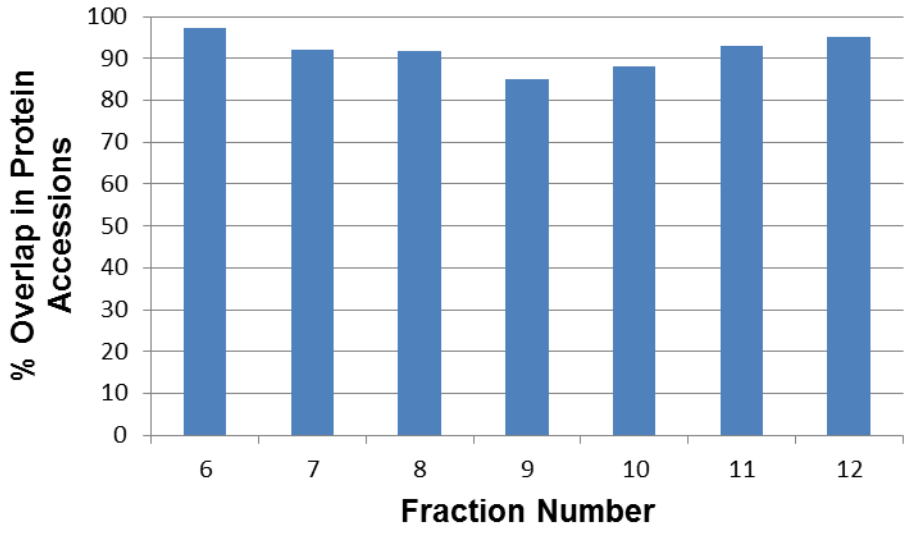

Figure 7.13 Graph of the percent overlap in protein accession number identifications in each GELFrEE fraction.

7.5 Conclusions

Overall, 193 nm UVPD yields major performance gains for the characterization of intact proteins compared to any existing collision- or electron-based dissociation methods. A single 5 ns laser pulse results in extensive sequence coverage and the ability to identify and localize labile PTMs in exquisite detail, demonstrating the significant potential of UVPD to drive an expansion of top-down proteomics. Analysis of low molecular weight proteins from a fractionated yeast lysate illustrates the compatibility of UVPD with LC-timescale.

7.6 References

(1) Duncan, M. W.; Aebersold, R.; Caprioli, R. M. *Nat Biotech* **2010**, *28*, 659–664.

(2) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 806–812.

(3) McLafferty, F. W.; Breuker, K.; Mi Jin; Xuemei Han; Infusini, G.; Honghai Jiang; Xianglei Kong; Begley, T. P. *FEBS Journal* **2007**, *274*, 6256–6268.

(4) Zhou, H.; Ning, Z.; E. Starr, A.; Abu-Farha, M.; Figeys, D. *Anal. Chem.* **2011**, *84*, 720–734.

(5) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809–2815.

(6) Raspopov, S. A.; El-Faramawy, A.; Thomson, B. A.; Siu, K. W. M. *Anal. Chem.* **2006**, *78*, 4572–4577.

(7) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.

(8) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528 –9533.

(9) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778–4784.

(10) Håkansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256–3262.

(11) Ledvina, A. R.; Beauchene, N. A.; McAlister, G. C.; Syka, J. E. P.; Schwartz, J. C.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J. *Anal. Chem.* **2010**, *82*, 10068–10074.

(12) Loo, J. A.; Edmonds, C. G.; Smith, R. D. *Anal. Chem.* **1991**, *63*, 2488–2499.

(13) Han, X.; Jin, M.; Breuker, K.; McLafferty, F. W. *Science* **2006**, *314*, 109 –112.

(14) Capriotti, A. L.; Cavaliere, C.; Foglia, P.; Samperi, R.; Laganà, A. *J. Chromatogr. A* **2011**, *1218*, 8760–8776.

(15) Bowers, W. D.; Delbert, S. S.; Hunter, R. L.; McIver, R. T. *J. Am. Chem. Soc.* **1984**, *106*, 7288–7289.

(16) Hunt, D. F.; Shabanowitz, J.; Yates, J. R. *J. Chem. Soc., Chem. Commun.* **1987**, *0*, 548–550.

(17) Thompson, M. S.; Cui, W.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1439–1452.

(18) Williams, E. R.; Furlong, J. J. P.; McLafferty, F. W. *Journal of the American Society for Mass Spectrometry* **1990**, *1*, 288–294.

(19) Reilly, J. P. *Mass Spectrom. Rev.* **2009**, *28*, 425–447.

(20) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9*, 4205–4214.

(21) Guan, Z.; Kelleher, N. L.; O'Connor, P. B.; Aaserud, D. J.; Little, D. P.; McLafferty, F. W. *Int. J. Mass Spectrom.* **1996**, *157–158*, 357–364.

(22) Kim, T. Y.; Thompson, M. S.; Reilly, J. P. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 1657–1665.

(23) Moon, J. H.; Shin, Y. S.; Cha, H. J.; Kim, M. S. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 359–368.

(24) Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Müller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M.; Dunyach, J.-J.; Cox, J.; Horning, S.; Mann, M.; Makarov, A. *Mol. Cell. Proteomics* **2012**, *11*, 10.1074/mcp.O111.013698.

(25) Jez, J. M.; Ferrer, J.-L.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Biochemistry* **2000**, *39*, 890–902.

(26) Zhang, Y.; Daum, S.; Wildemann, D.; Zhou, X. Z.; Verdecia, M. A.; Bowman, M. E.; Lücke, C.; Hunter, T.; Lu, K. P.; Fischer, G. *ACS Chem. Biol.* **2007**, *2*, 320–328.

(27) Sultana, R.; Boyd-Kimball, D.; Poon, H. F.; Cai, J.; Pierce, W. M.; Klein, J. B.; Markesbery, W. R.; Zhou, X. Z.; Lu, K. P.; Butterfield, D. A. *Neurobiol. Aging* **2006**, *27*, 918–925.

(28) Vasicek, L. A.; Ledvina, A. R.; Shaw, J. B.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 1105–1108.

(29) Senko, M. W.; Beu, S. C.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 52–56.

(30) Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1996**, *31*, 1185–1187.

(31) Gu, C.; Tsaprailis, G.; Breci, L.; Wysocki, V. H. *Anal. Chem.* **2000**, *72*, 5804–5813.

(32) Morgan, J. W.; Russell, D. H. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 721–729.

(33) Yoon, S. H.; Chung, Y. J.; Kim, M. S. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 645–655.

(34) Sze, S. K.; Ge, Y.; Oh; McLafferty, F. W. *Anal. Chem.* **2003**, *75*, 1599–1603.

(35) Ping Lu, K.; Hanes, S. D.; Hunter, T. *Nature* **1996**, *380*, 544–547.

(36) Lu, K. P.; Zhou, X. Z. *Nat Rev Mol Cell Biol* **2007**, *8*, 904–916.

(37) Nakamura, K.; Greenwood, A.; Binder, L.; Bigio, E. H.; Denial, S.; Nicholson, L.; Zhou, X. Z.; Lu, K. P. *Cell* **2012**, *149*, 232–244.

(38) Butterfield, D. A.; Poon, H. F.; St Clair, D.; Keller, J. N.; Pierce, W. M.; Klein, J. B.; Markesbery, W. R. *Neurobiol. Dis.* **2006**, *22*, 223–232.

(39) Sultana, R.; Butterfield, D. A. *J Alzheimers Dis.* **2013**, *33*, S243–S251.

(40) Xu, G. G.; Zhang, Y.; Mercedes-Camacho, A. Y.; Etzkorn, F. A. *Biochemistry* **2011**, *50*, 9545–9550.

Chapter 8

Extending the Isotopically Resolved Mass Range of Orbitrap Mass Spectrometers

8.1 Overview

The routine analysis of large biomolecules (greater than 30 kDa) has been a challenge for Orbitrap mass spectrometers due to the relatively high kinetic energy of ions entering and within the Orbitrap mass analyzer. This characteristic results in rapid signal decay for large biomolecules due to energetic collisions with background gas molecules. Here, we report a method to significantly enhance the analysis of large biomolecules in an Orbitrap mass spectrometer. The combination of reduced C-trap and HCD cell bath gas pressure, using helium as the bath gas and trapping ions in the HCD prior to mass analysis, greatly increased sensitivity and reduced signal decay for large protein ions. As a result, isotopic resolution of monoclonal IgG was achieved, and we have established a new high mass record for which accurate mass measurement and isotopic resolution has been achieved.

8.2 Introduction

Accurate mass determination of large biomolecules (e.g. proteins of molecular weight greater than 100 kDa) by mass spectrometry has long been exclusively within the realm of performance achievable by Fourier transform mass spectrometry (FTMS), more specifically that of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS).[1,2] In 1997, Kelleher et al. reported unit mass resolution of 112 kDa chondroitinase with a resolving power of 170,000 and 3 Da (~25 ppm) mass accuracy.[3]

Not until 2009 was the next significant advance reported based on the measurement of 115 kDa cardiac myosin binding protein C with 4 ppm mass accuracy.[4] More recently, Marshall and coworkers utilized a substantially redesigned 9.4T FT-ICR mass spectrometer and phase correction to acquire a unit mass baseline resolved spectrum of a 147.7 kDa monoclonal antibody with an average resolving power of 530,000 at *m/z* 2,600 in absorption mode.[5]

The Orbitrap mass analyzer, a relatively new member of the FTMS family, was first described in 2000[6] and commercialized in the form of a hybrid linear ion trap/Orbitrap mass spectrometer in 2005.[7] The Orbitrap mass spectrometer has rapidly gained popularity for a wide variety of applications due to its ultrahigh resolving power, high mass accuracy, short acquisition times and robust operation.[8–15] Development of the compact high-field Orbitrap mass analyzer[16] in addition to advanced signal processing methods[17] provided up to a further 3.6 fold improvement in resolving power (or speed),[15] placing Orbitrap mass spectrometers well within the realm of performance achievable with high field FT-ICR mass spectrometers. This was exemplified in a recent study by Denisov *et al*. which demonstrated the capability of the high-field Orbitrap mass analyzer to achieve resolving power in excess of 1,000,000 using a 3 s detection time.[18]

Despite the exceptionally high performance and *m/z* independent trapping conditions (i.e. theoretically unlimited mass range) afforded by Orbitrap mass analyzers, routine analysis of proteins typically has been limited to small proteins up to about 30 kDa.[7,19] This was attributed to the inherently higher kinetic energies of ions (1.5 to 3.6 keV/charge) in Orbitrap mass spectrometers compared to FT-ICR systems, resulting in more rapid signal decay due to higher center-of-mass collision energies.[16,20] Makarov and Denisov reported the major sources of signal decay for protein ions in Orbitrap mass spectrometers to be fragmenting collisions upon the ejection of ions from the C-trap en

route to the Orbitrap mass analyzer as well as scattering collisions with residual gas within the mass analyzer.[20] In the same study, isotopic resolution of bovine serum albumin (BSA, 66.4 kDa) and bovine apo-transferrin (78 kDa) using a LTQ Orbitrap XL mass spectrometer (a standard Orbitrap mass analyzer) was achieved by blocking the flow of nitrogen to the HCD cell and performing a 12 hr bake-out at 180 °C instead of the standard 110 °C to obtain a lower ultimate pressure in the analyzer vacuum compartment.[20] Similar results were also achieved for bovine apo-transferrin using an Orbitrap Elite mass spectrometer (a high-field Orbitrap mass analyzer) and 3.04 s transient acquisitions.[18] Heck and coworker have recently demonstrated the ability to analyze large intact proteins and noncovalent protein complexes using native mass spectrometry and Orbitrap mass analysis.[21,22] To facilitate trapping of high *m/z* ions (10,000-12,000 *m/z* for the species analyzed) produced by native electrospray conditions, ions were trapped in the HCD cell prior to mass analysis and xenon was used in place of nitrogen as C-trap and HCD cells bath gas.[21] These modifications enabled very sensitive analysis of protein complexes up to 800 kDa (bacterial chaperonin GroEL) using short transient acquisitions (64 ms).

Here, we report a simple yet effective method to significantly extend the practical upper mass limit for which isotopic resolution can be achieved using Orbitrap mass spectrometers, and we establish a new high mass record for which isotopic resolution and accurate mass determination has been achieved. This work builds upon the conclusions reported by Makarov and Denisov.[20] Replacing the C-trap and HCD cell nitrogen bath gas with helium and operating at reduced pressure significantly decreases signal decay as a result of fewer collisions and lower center-of-mass collision energies upon ion ejection from the C-trap and within the Orbitrap mass analyzer. Also, trapping ions in the HCD cell prior to Orbitrap mass analysis provides greater sensitivity via greater trapping

efficiency of higher *m/z* ions due to the longer path length available for collisional cooling compared to the relatively small dimensions of the C-trap.[23] Combined, the above modifications allowed isotopic resolution and accurate mass determination of the glycoforms of an intact monoclonal IgG1 antibody of mass up to 148,706.3391 Da within 3.1 ppm (0.5 Da). Three second transient acquisitions yielded an average resolving power of 330,000 for the 53+ charge state of IgG1 at *m/z* 2800.

8.3 Experimental

8.3.1 Materials

Monoclonal IgG1 (product number: 186006552) was purchased from Waters (Milford, MA). Horse myoglobin, bovine carbonic anhydrase and bovine serum albumin (BSA) were purchased from Sigma Aldrich (St. Louis, MO). All other chemicals were purchased from Fisher Scientific (Pittsburgh, PA).

8.3.2 Sample Preparation

Horse myoglobin, bovine carbonic anhydrase and BSA were used without further purification. Monoclonal IgG1 was cleaned-up using centrifugal 30 kDa molecular weight cutoff filters (Millipore, Billerica, MA) by washing three times with 10 mM ammonium acetate and three times with LC-MS grade water. Protein samples were prepared at 10 μM in 50:50 (v/v) water:acetonitrile with 0.1% formic acid.

8.3.4 Mass Spectrometry

All experiments were performed using an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). The instrument firmware was modified to allow trapping of ions in the HCD cell and the acquisition of 3 s transients. Helium or nitrogen was used as the HCD cell and C-trap bath gases. Automatic gain control (AGC) targets of 1E6 and 1E5 were used for FT and linear ion trap acquisitions, respectively. The bath gas pressure was controlled using an electronic regulator controlled through the instrument software, and the pressure was measured via a flange-mounted pirani gauge connected to the HCD cell housing by a piece of 6 mm tubing. Ions were transferred to the HCD cell with normalized collision energy of 0.3-1% (1 eV). The standard automated calibration procedures were used for mass calibration, AGC and advanced signal processing. No additional manual tuning was needed to optimize the Orbitrap mass analyzer potentials.

Myoglobin and BSA were analyzed by direct infusion at a flow rate of 3 μL/min. Monoclonal IgG was infused by flow injection using the divert/inject valve on the front of the mass spectrometer. IgG was injected into a 20 μL loop then back-flushed at flow rate of 200 - 300 nL/min to a nanospray source by a nanoflow liquid chromatography pump (Eksigent, Framingham, MA). Uncoated fused silica emitters with 30 μm tips (New Objective, Woburn, MA) were used with a typical spray voltage of 1.5 kV. In-source CID of 70 – 80V was critical for desolvation and removal of adducts from IgG. Decharged and decharged monoisotopic spectra were produced using Xtract (Thermo Fisher Scientific) with a signal-to-noise ratio of 10.

8.4 Results and Discussion

Achieving high mass accuracy and isotopic resolution of large biomolecules required optimization of the nature and pressure of the bath gas used in the C-trap and HCD cell. Helium was chosen as an alternative to nitrogen due to its lower mass. It was hypothesized that the reduced center-of-mass collision energy afford by helium would yield a reduction in fragmenting collisions as ions exit the C-trap and scattering of ions in the Orbitrap mass analyzer. To compensate for changes in the trapping efficiency in the C-trap, the pressure of nitrogen or helium used was optimized to ensure approximately equal ion populations reached the Orbitrap mass analyzer when using either gas. This was accomplished by monitoring the amplitude of the first/only beat in short transients (48 ms, R=15,000 at *m/z* 400). Due to limited access to transients and no access to raw transient data, ion populations were judged based on intensity of the first isotopic beat relative to signal spikes caused by detector saturation upon ion injection (pulse/ramp of high voltages) into Orbitrap mass analyzer. The amplitude of these spikes in constant between acquisitions, and thus 100% relative intensity is constant between acquisitions. The first 10 ms of a transient acquired for isolated 25+ carbonic anhydrase ions is shown in **Figure 8.1**. The signal spikes caused by ion injection can be seen in the first ~500 μs of the transient. Gas pressure in the C-trap and HCD cell was adjusted to give approximately the same intensity for the first beat. In the case of carbonic anhydrase, the relative intensity was ~80% of the detector saturation. **Figure 8.1B** and **C** are 48 ms transients acquire for isolated carbonic anhydrase 25+ ions using 3 mTorr of nitrogen or 5mTorr of helium bath gas, respectively. Based on the nearly identical appearance of these transients, it was assumed that approximately equal ion populations were reaching the Orbitrap mass analyzer when using nitrogen or helium bath gases. This was confirmed prior to data acquisition for myoglobin and BSA as well. For the smaller

proteins analyzed in this study (myoglobin, carbonic anhydrase and BSA), the optimal pressure in the HCD cell was found to be 3 mTorr and 5 mTorr for nitrogen and helium, respectively. Due the greater conductance of helium into the mass analyzer vacuum chamber and the lower pumping rate for helium compared to nitrogen by turbomolecular pumps, the pressure in the UHV chamber containing the Orbitrap mass analyzer was ~$2x10^{-10}$ Torr with helium (corrected pressure reading) compared to $3x10^{-11}$ Torr with nitrogen. This point emphasizes the significant contribution of fragmenting collisions as ions exit the C-trap (vide infra) to signal decay.

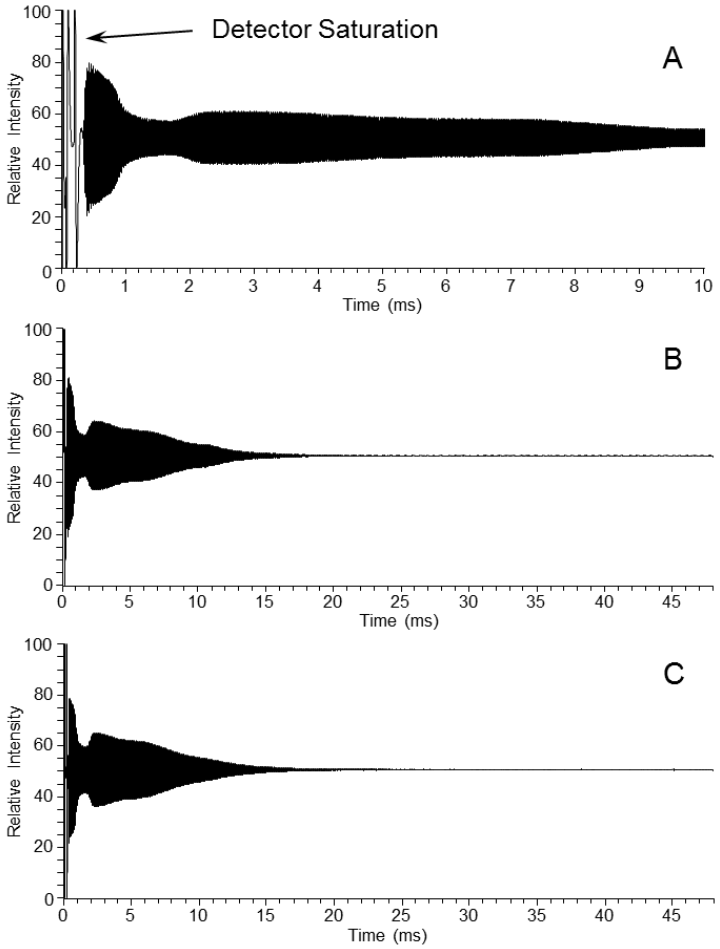

Figure 8.1 First 10 ms of a transient for the 25+ charge state of carbonic anhydrase (A) showing showing detector saturation in the first ~500 μs of the transient. 48 ms transients (R=15,000 at *m/z* 400) acquired for the 25+ charge state of carbonic anhydrase using nitrogen (B) or helium (C) as the C-trap and HCD cell bath gas.

Spectra were acquired for the isolated 20+ charge state of myoglobin (m/z 854, 16.9 kDa) and 25+ charge state of carbonic anhydrase (m/z 1162, 29 kDa) using transient acquisition times from 96 to 3072 ms with helium or nitrogen bath gas. The signal-to-noise ratio (S/N) of the most abundant isotopic peak using helium or nitrogen as well as the resolution observed are graphed as a function of transient acquisition time in **Figure 8.2A** and **8.2B** for myoglobin and carbonic anhydrase, respectively. For myoglobin,

helium yielded a minimum of 40% increase in S/N at the shortest transient acquisition utilized (96 ms) and up to 72% increase in S/N at 1536 ms transient acquisition for myoglobin. The observed resolution for myoglobin was nearly identical for helium and nitrogen. For carbonic anhydrase, helium yielded a 15% increase in S/N at the shortest transient acquisition (192 ms), 65% increase at 768 ms (longest commercial transient) and 140% increase in S/N at the longest transient acquisition time (3072 ms).

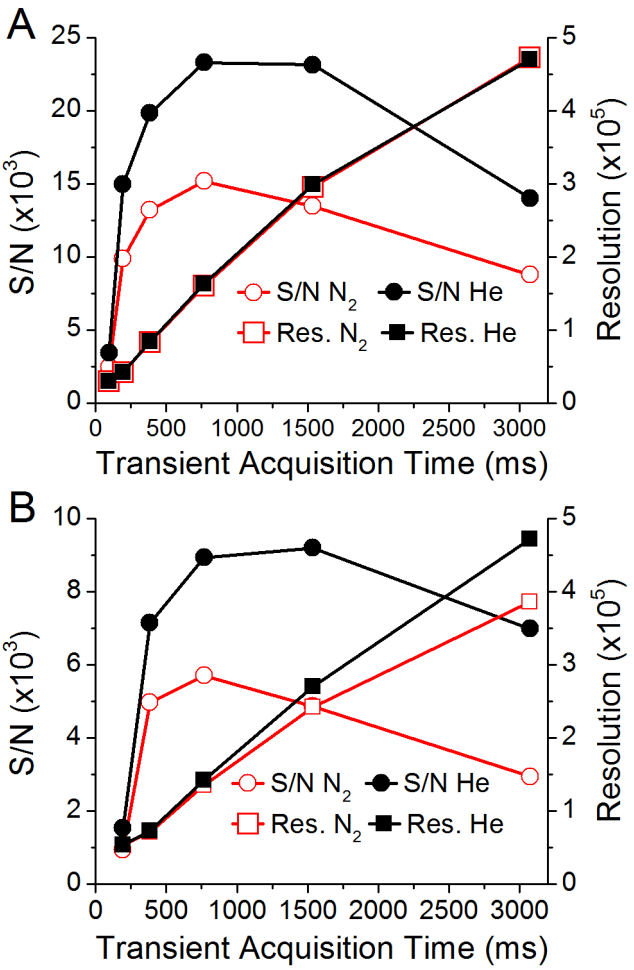

Figure 8.2 Signal to noise ratio and resolution observe as a function of transient acquisition time for the 20+ charge state of myoglobin (A) and the 25+ charge state of carbonic anhydrase (B).

The use of helium also yielded a significant increase in resolution and the linearity of the increase in resolution as a function of transient acquisition time for carbonic anhydrase. The change in resolution observed for carbonic anhydrase is a result of substantially reduced transient decay. Examples of 3072 ms transients acquired using helium or nitrogen for 20+ myoglobin and 25+ carbonic anhydrase are shown in **Figure 8.3** and **8.4**, respectively. A small but noticeable change in amplitude of the isotopic beat pattern for myoglobin in the transient using helium results in increased S/N; however, the isotopic beat pattern for both bath gases persists for the full transient acquisition. On the other hand, the isotopic beat pattern of 25+ carbonic anhydrase in the transient acquired using nitrogen had completely decayed by approximately 1.5 s, whereas the isotopic beat pattern persists for the entire 3 s transient when using helium. This yielded an increase in resolution of approximately 90,000 (20%) for the 3 s acquisition.

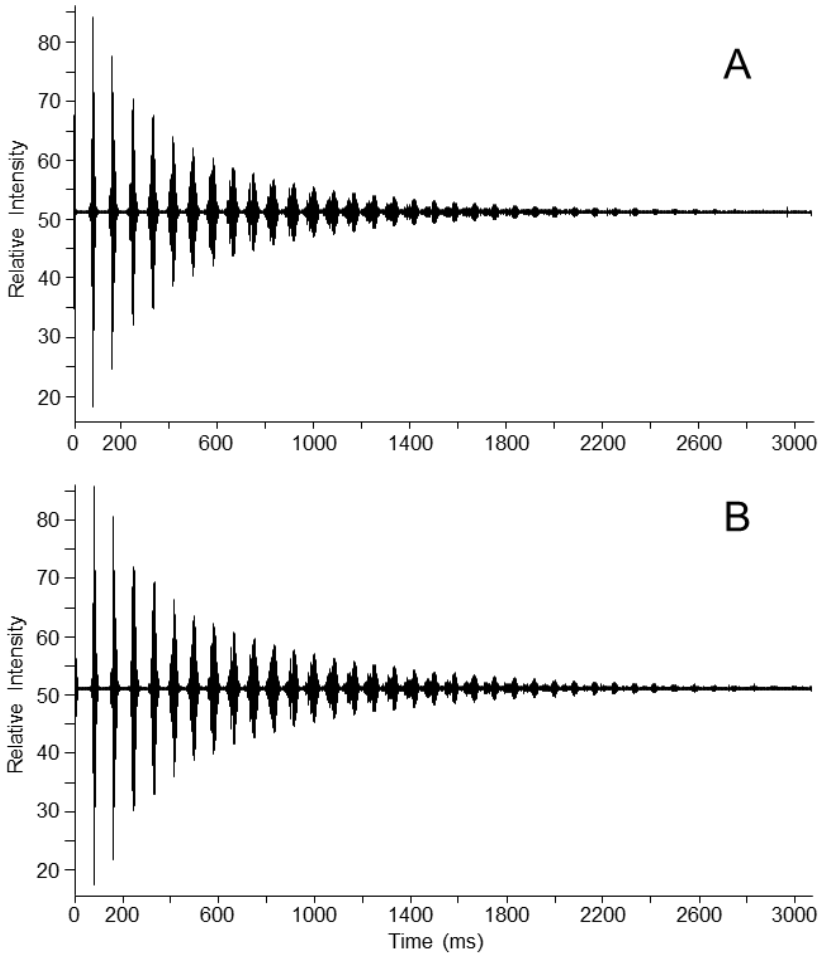

Figure 8.3 3072 ms transient acquired for the 20+ charge state of myoglobin using nitrogen (A) or helium (B) as the C-trap and HCD cell bath gas.

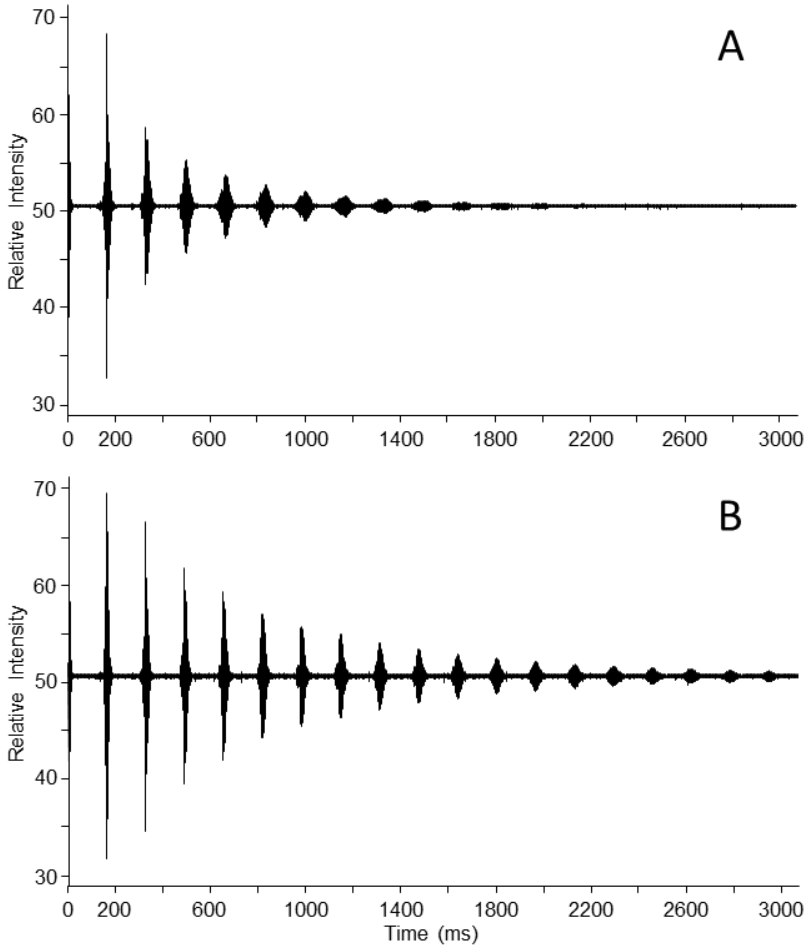

Figure 8.4 3072 ms transient acquired for the 25+ charge state of carbonic anhydrase using nitrogen (A) or helium (B) as the C-trap and HCD cell bath gas.

Spectra of BSA acquired using 768 ms transients (R = 240,000 at *m/z* 400, maximum commercial resolution setting) and nitrogen or helium bath gas are shown in **Figure 8.5.** Broadband spectra of the charge state distribution of BSA acquired using nitrogen and helium are shown **Figure 8.5A** and **8.5B**, respectively. To achieve similar S/N for the charge state distributions, 500 averages were needed using nitrogen compared to 100 averages when using helium bath gas. The S/N of the center isotope of the 49+ charge state isotopic distribution was 57 with nitrogen and 65 with helium (S/N of 137 with 500 averages and helium), and expansions of these isotopic distributions are shown

in **Figure 8.5C** and **8.5D**. **Figure 8.5E** and **8.5F** display isotopic distributions of the isolated 49+ charge state of BSA acquired using nitrogen and helium, respectively, and only 5 averages. The use of nitrogen bath gas does not yield an interpretable isotopic distribution; however, the spectrum acquired with helium bath gas exhibits a fairly symmetric isotopic distribution, in which the center isotope is observed at S/N 251. This result is very promising for top-down data-dependent LC-MS/MS acquisitions which utilize a low resolution broad mass range MS^1 scan and data-dependent selected ion monitoring (SIM) scans for charge state and accurate mass determination.[15]

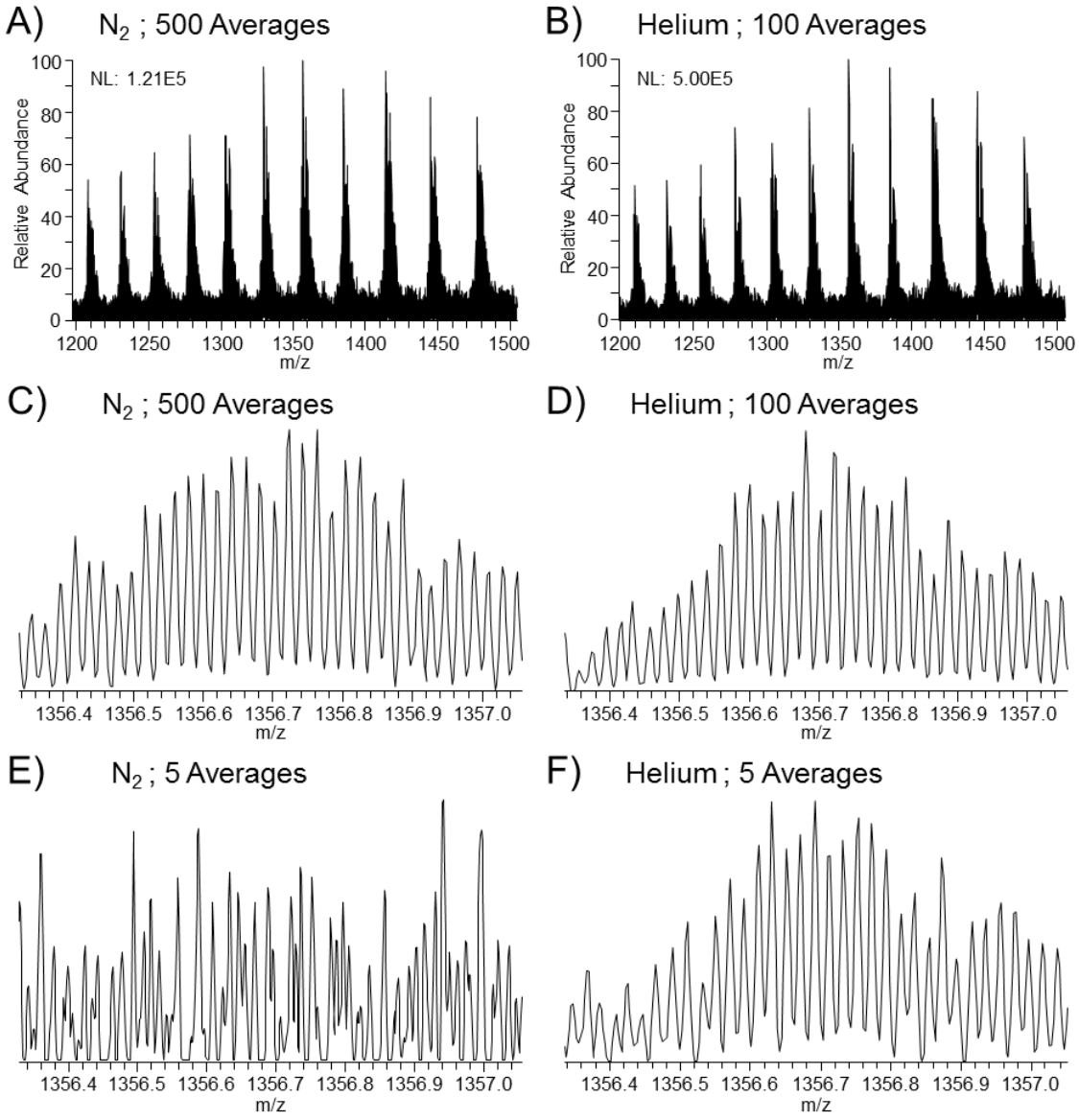

Figure 8.5 High resolution MS^1 spectra of BSA (A, B), expansion of the isotopic distributions of the 49+ charge state (C, D) and isotopic distributions of the isolated 49+ charge state acquired using nitrogen (A, C, E) or helium (B, D, F) bath gas in the C-trap and HCD cell.

Upon inspection of transients acquired for BSA using nitrogen (**Figure 8.6A**) or helium (**Figure 8.6B**) bath gas, the source of the substantially increased performance is immediately evident. The isotopic beat pattern for the transient acquired using nitrogen

contains two isotopic beats compared to four isotopic beats with helium. These transients clearly illustrate the reduced signal decay rate resulting from less energetic center-of-mass collisions afforded by the use of helium bath gas in the C-trap. The exaggerated difference in the decay rate of the transient isotopic beat patterns between nitrogen or helium bath gas for the analysis BSA (66.4 kDa) compared to myoglobin (16.9 kDa) and carbonic anhydrase (29 kDa) is a result of the significantly larger collision cross-section of BSA and demonstrates the significant benefit of using helium for the analysis of large proteins. The larger collision cross-section of BSA results in a greater number of collisions upon ejection from the C-trap and in the Orbitrap mass analyzer. The greater number of both fragmenting and scattering collisions yields more rapid decay of the transient isotopic beat pattern characteristic of the analysis of large proteins. Despite the nearly 10-fold higher pressure in the Orbitrap mass analyzer vacuum chamber when using helium, there was a significant reduction in transient decay rate and increased S/N observed for BSA. This result emphasized the significance of fragmenting collisions upon ejection of ions from the C-trap. Lower energy collisions with helium produced less fragmentation despite the higher pressure of helium present in the C-trap and thus greater number of collisions.

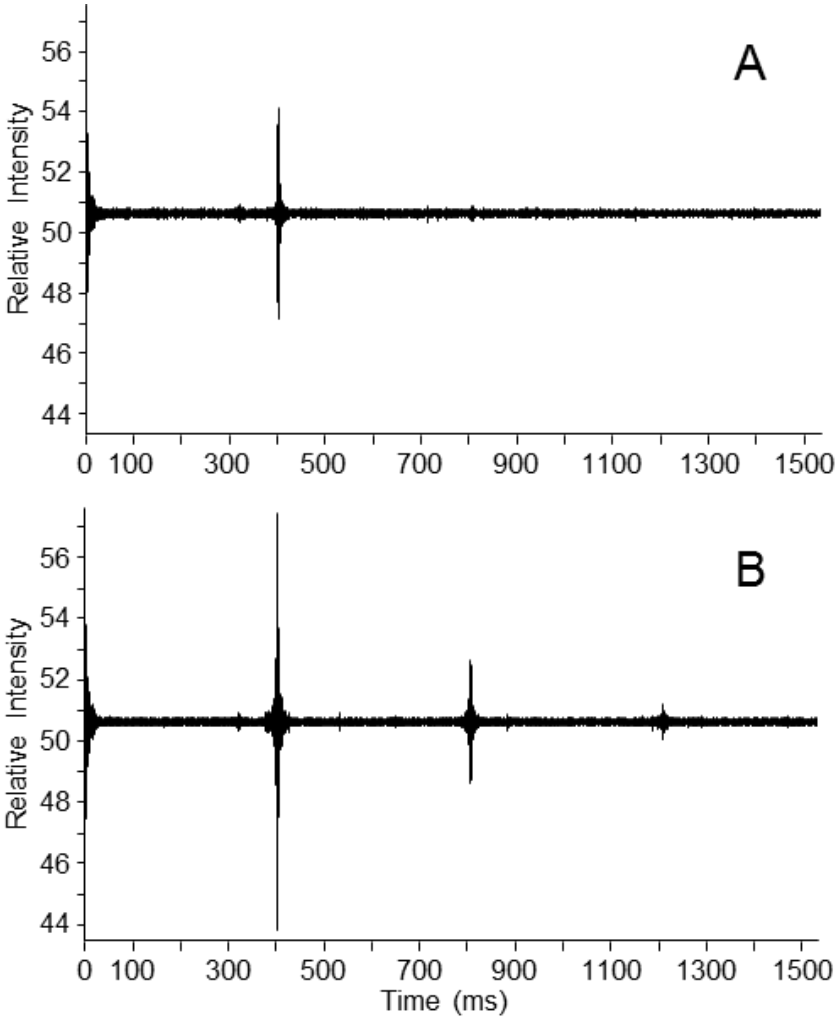

Figure 8.6 1.5 s transient acquisitions of isolated ions of the 49+ charge state of BSA using (A) nitrogen or (B) helium.

The composite effects of trapping ions in the HCD cell prior to Orbitrap mass analysis, helium C-trap and HCD cell bath gas (7 mTorr in HCD cell), and 3 s transient acquisitions allowed baseline isotopic resolution of intact monoclonal IgG1. A low resolution linear ion trap mass spectrum of the charge state distribution of IgG1 is shown in **Figure 8.7A**. Ions of *m/z* 2800 ± 150 *m/z* were isolated in the linear ion trap and transferred to the HCD cell prior to high resolution mass analysis in the Orbitrap mass

analyzer using 3 s transient acquisitions (**Figure 8.7B**). An expansion of the 53+ charge state isotopic distribution of the most abundant glycoform is shown in **Figure 8.7C**. Baseline resolution is achieved with an average resolving power of 330,000 at *m/z* 2800. The decharged monoisotopic mass spectrum and decharged mass spectrum produced by deconvolution of the spectrum in **Figure 8.7B** using the Xtract algorithm (Thermo Fisher Scientific) are shown in **Figure 8.8A** and **8.8B**, respectively. A number of salt adducts (sodium and phosphate) are evident in addition to the major glycoforms of the antibody. This particular monoclonal IgG1 has 17 disulfide bonds, fixed pyroglutamic acid from Gln at the N-terminus of each heavy chain, and one N-linked glycosylation on each heavy chain. The elemental composition of the antibody without glycosylations is $C_{6472}H_{9940}N_{1698}O_{2008}S_{62}$ corresponding to an average molecular weight of 145,329.7 Da. The experimentally determine monoisotopic masses (**Figure 8.8A**) of the G0F/G0F, G0F/G1F, G1F/G1F and G1F/G2F glycoforms were within 15 ppm (2.2 Da), 0.95 ppm (0.12 Da), 8.4 ppm (1.2 Da) and 4.9 ppm (0.73 Da) of the theoretical masses, respectively. The lower accuracy of the G0F/G0F and G1F/G1F glycoform monoisotopic masses is due to deconvolution of less ideally shaped isotopic distributions as seen in the decharged spectrum (**Figure 8.8B)**. The central isotope of the isotopic distributions of the G0F/G0F, G0F/G1F, G1F/G1 and G1F/G2F glycoforms are within 2.3, 2.5, 2.6 and 3.1 ppm, respectively. More extensive desalting would afford greater accuracy in the determination of monoisotopic masses as a result of greater S/N and more ideally shaped isotopic distributions.

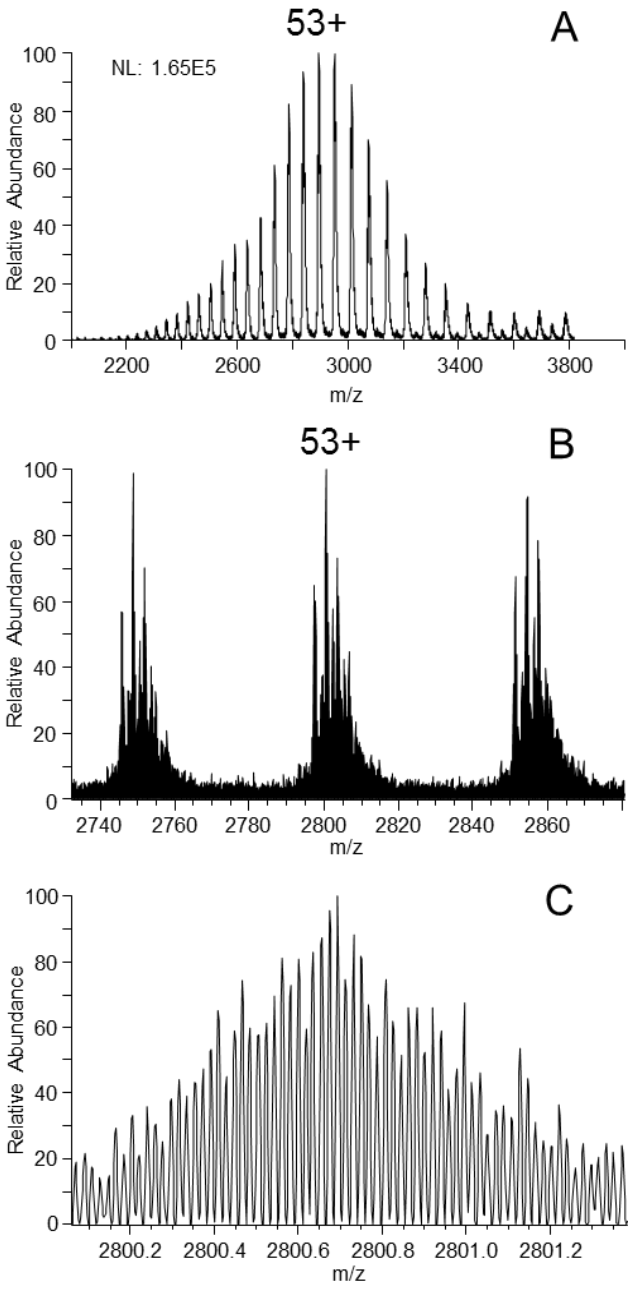

Figure 8.7 Low resolution linear ion trap spectrum of the charge state envelop of intact monoclonal IgG1 (A), portion of a high resolution FTMS spectrum produced using an isolation width of 300 *m/z* centered at *m/z* 2800 (B) and the isotopic distribution of the G0F/G1F glycoform in the 53+ charge state (C).

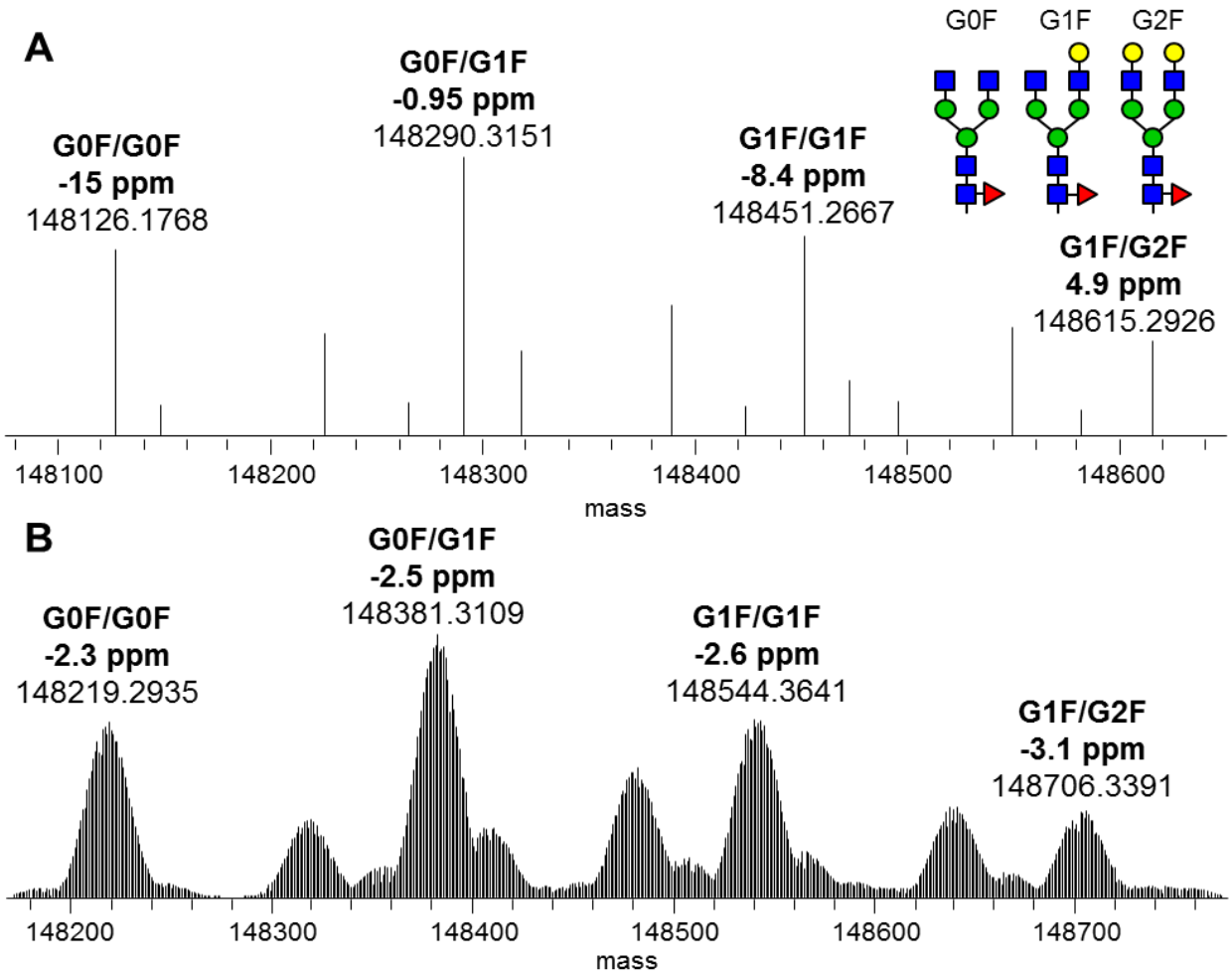

Figure 8.8 Decharged monoisotopic mass spectrum (A) and decharged mass spectrum (B) of monoclonal IgG1.

8.5 Conclusions

To summarize, we have developed a simple yet effective method for significantly improving the performance of Orbitrap mass spectrometers for the analysis of large biomolecules. The use of helium instead of nitrogen as the C-trap and HCD cell bath gas and trapping ions in the HCD cell prior to high resolution mass analysis significantly reduced the signal decay rate for large protein ions and allowed the acquisition of longer

transients. As a result, the intact masses of the major glycoforms of a monoclonal IgG1 antibody were isotopically resolved and accurately determined. We have established a new high mass record for which accurate mass and isotopic resolution has been achieved (148,706.3391 Da ± 3.1 ppm). With sufficient sample purification and homogeneity, isotopic resolution and accurate mass determination of biomolecules well in excess of 150 kDa should be readily achievable with Orbitrap mass spectrometers. One downside of using helium as the HCD cell bath gas is that HCD is rendered completely ineffective. Ions that are accelerated into the HCD cell are not efficiently fragmented or trapped. However, Orbitrap mass analysis of ion trap CID and ETD products is completely unaffected.

8.6 References

(1) Marshall, A. G.; Hendrickson, C. L. *Annu. Rev. Anal. Chem.* **2008**, *1*, 579–599.

(2) Scigelova, M.; Hornshaw, M.; Giannakopulos, A.; Makarov, A. *Mol. Cell. Proteomics* **2011**, *10*, 10.1074/mcp.M111.009431.

(3) Kelleher, N. L.; Senko, M. W.; Siegel, M. M.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 380–383.

(4) Ge, Y.; Rybakova, I. N.; Xu, Q.; Moss, R. L. *P. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 12658–12663.

(5) Valeja, S. G.; Kaiser, N. K.; Xian, F.; Hendrickson, C. L.; Rouse, J. C.; Marshall, A. G. *Analytical chemistry* **2011**, *83*, 8391–8395.

(6) Makarov, A. *Anal. Chem.* **2000**, *72*, 1156–1162.

(7) Makarov, A.; Denisov, E.; Kholomeev, A.; Balschun, W.; Lange, O.; Strupat, K.; Horning, S. *Anal. Chem.* **2006**, *78*, 2113–2120.

(8) Olsen, J. V.; de Godoy, L. M. F.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. *Mol. Cell. Proteomics* **2005**, *4*, 2010 –2021.

(9) Makarov, A.; Denisov, E.; Lange, O.; Horning, S. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 977–982.

(10) Perry, R. H.; Cooks, R. G.; Noll, R. J. *Mass Spectrom. Rev.* **2008**, *27*, 661–699.

(11) Scigelova, M.; Makarov, A. *Proteomics* **2006**, *6*, 16–21.

(12) Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. *Nat. Methods* **2007**, *4*, 709–712.

(13) McAlister, G. C.; Phanstiel, D.; Good, D. M.; Berggren, W. T.; Coon, J. J. *Anal. Chem.* **2007**, *79*, 3525–3534.

(14) Michalski, A.; Damoc, E.; Hauschild, J.-P.; Lange, O.; Wieghaus, A.; Makarov, A.; Nagaraj, N.; Cox, J.; Mann, M.; Horning, S. *Mol. Cell. Proteomics* **2011**, *10*, 10.1074/mcp.M111.011015.

(15) Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Müller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M.; Dunyach, J.-J.; Cox, J.; Horning, S.; Mann, M.; Makarov, A. *Mol. Cell. Proteomics* **2012**, *11*, 10.1074/mcp.O111.013698.

(16) Makarov, A.; Denisov, E.; Lange, O. *Journal of the American Society for Mass Spectrometry* **2009**, *20*, 1391–1396.

(17) Lange, O.; Damoc, E.; Wieghaus, A.; Makarov, A. In *Proceedings of the 59th ASMS Conference on Mass Spectrometry and Allied Topics*; Denver, CO, 2011.

(18) Denisov, E.; Damoc, E.; Lange, O.; Makarov, A. *Int. J. Mass Spectrom.* **2012**, *325–327*, 80–85.

(19) Hardman, M.; Makarov, A. A. *Anal. Chem.* **2003**, *75*, 1699–1705.

(20) Makarov, A.; Denisov, E. *Journal of the American Society for Mass Spectrometry* **2009**, *20*, 1486–1495.

(21) Rose, R. J.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A. J. R. *Nat Meth* **2012**, *9*, 1084–1086.

(22) Rosati, S.; Rose, R. J.; Thompson, N. J.; van Duijn, E.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A. J. R. *Angew. Chem. Int. Ed.* **2012**, *51*, 12992–12996.

(23) Denisov, E.; Damoc, E.; Zeller, M.; Müller, M.; Kuipers, H.; Griep-Raming, J.; Makarov, A.; Nolting, D.; Mohring, T. In *Proceedings of the 60th ASMS Conference on Mass Spectrometry and Allied Topics*; Vancouver, BC, Canada, 2012.

Chapter 9

Conclusions

The rapid growth of mass spectrometry based proteomics in the past decade is largely due to significant improvements in experimental strategy, mass spectrometric instrumentation, and tandem mass spectrometry methods. Nevertheless, due the immense complexity of proteomic analyses, there is still significant room for improvement in all of these areas. Current mass spectrometry based proteomics hinges on the ability to produce meaningful fragmentation patterns for the constituent enzymatically generated peptides (bottom-up approach) or intact proteins (top-down approach) of complex proteome mixtures. Therefore, the continued development of tandem mass spectrometry techniques that produce an extensive and meaningful array of diagnostic product ions is critical to the continued success and advancement of mass spectrometry based proteomics.

In chapter 3, the performance metrics for UVPD and NETD were compared for peptide anion characterization. UVPD allowed the collection of a greater number of MS/MS spectra and activation of a wider variety of peptides during a typical a LC gradient; however, both UVPD and NETD provided additional complementary information to standard peptide cation analyses performed with CID.

In chapter 4, it was shown that the use of additional vibrational excitation afforded significantly increased sequence coverage and product ion abundances for peptide anion characterization by NETD. Supplemental infrared photoactivation outpaced supplemental collisional activation in LC-MS/MS analyses due to complications likely related to space charge effects during the NETD ion/ion reaction.

In chapter 5, it was found that the lower ionization energy of deprotonated tyrosine phenol compared to the carboxylates of acidic residues and the C-terminus

yielded preferential cleavage N-terminal to deprotonated tyrosine residues during NETD and UVPD. Iodo-tyrosine derivatives were utilized to confirm that selective and enhanced *c* and *z* ion formation proceeds through a tyrosyl radical, and high pH LC-MS/MS experiments demonstrated that this cleavage pathway occurs frequently during proteomic experiments.

A method for rapid identification of protein digests which utilized TM-DESI and 1933 nm UVPD was developed in chapter 6. UVPD at 193 nm provided high confidence peptide identification due to the extensive array of sequence ion types produced.

The results in chapter 7 demonstrated the exceptional performance of 193 nm UVPD for intact protein analysis. Overall, 193 nm UVPD yielded major performance gains for the characterization of intact proteins compared to existing collision- or electron-based dissociation methods. The use of a single 5 ns laser pulse yielded extensive sequence coverage and allowed precise localization of labile PTMs. Compatibility of UVPD with liquid chromatography timescale was also established through the analysis of fractionated, less than 30 kDa proteins from a yeast whole cell lysate.

A simple yet effective method for significantly improving the performance of Orbitrap mass spectrometers for the analysis of large biomolecules was developed in chapter 8. The use of helium instead of nitrogen as the C-trap and HCD cell bath gas and trapping ions in the HCD cell prior to high resolution mass analysis significantly reduced the signal decay rate for large protein ions. This allowed isotopic resolution and accurate mass determination of an intact monoclonal antibody. A new high mass record was established for which accurate mass and isotopic resolution (148,706.3391 Da ± 3.1 ppm) was achieved.

The transformative work described in chapters 7 and 8 has laid a solid foundation for rapid expansion and adoption of the top-down proteomics approach. The ultrafast and extensive fragmentation afforded by UVPD of intact proteins will allow detailed characterization of protein PTMs and determination of the combinatorial nature of protein modifications, and thus shed light onto the regulatory mechanisms of cellular pathways. When coupled with the new found sensitivity and overall performance enhancement of Orbitrap mass spectrometers for large intact proteins as described in chapter 8, UVPD has the potential to provide unparalleled performance for simple top-down mass spectrometry experiments involving pure isolated proteins as well as in high throughput top-down analyses of complex proteome samples.

Intact mass analysis of monoclonal antibodies (mAbs), as demonstrated in chapter 8, as well as MS/MS of intact antibodies are of great interest to pharmaceutical and biotechnology research due to their potential use for treatment of a wide variety of human illnesses and disorders. Precise structural characterization of therapeutic mAbs is a critical step in the developmental stages of these therapeutics. 193 nm UVPD of an intact monoclonal antibody (**Figure 9.1**) alone is not sufficient for complete characterization of the primary structure and post-translational modifications. This is due to the extensive network of disulfide bonds that maintain the tertiary and quaternary structure of the mAb. Disulfide bonds are not efficiently cleaved by 193 nm photons, and thus very little information is generally obtained for the regions of the primary structure that are enclosed by disulfide bonds. This is evident by the location of observed backbone cleavages in **Figure 9.1.**

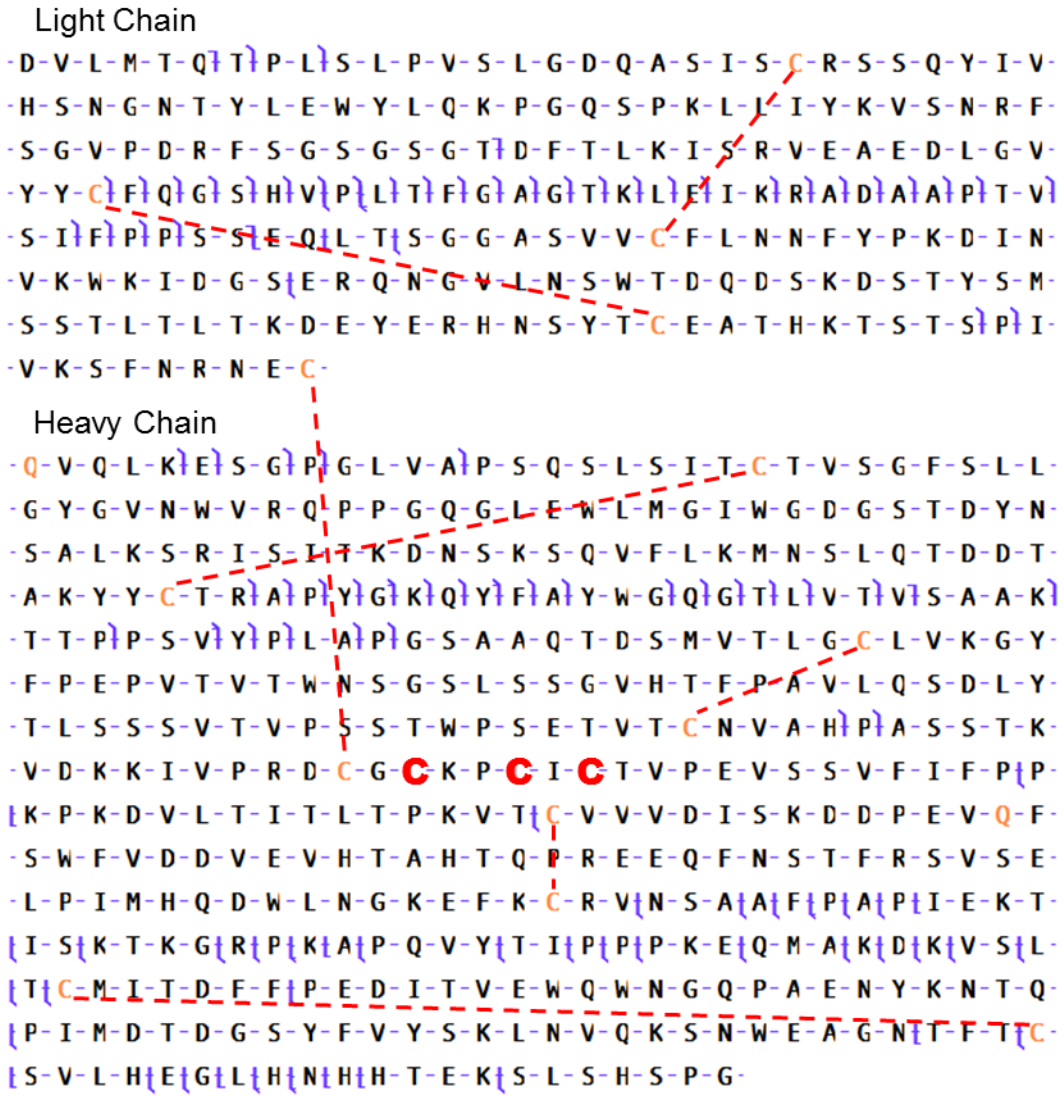

Figure 9.1 UVPD fragment ion map for the light and heavy chains of an intact monoclonal antibody. Dashed lines indicate disulfide bonds and bold 'C' indicates cysteine residues involved in inter-heavy chain disulfide bonds.

One way around this limitation would be to use multiple wavelengths of photons during an experiment. 248 nm and 266 nm photons have been shown to preferentially cleave disulfide bonds in peptides and proteins.[1] An experiment could be performed during which 193 nm photons used to induce backbone fragmentation and 248 nm or 266

nm photons are used to cleave the disulfide bonds and release the disulfide bound fragment ions. These two parts of the experiment could be performed sequentially or simultaneously. Co-alignment of two lasers could be potentially troublesome due to inefficient transmission/reflection of multiple wavelengths of UV light by most optics. Integration of the light source into the ion trapping device where photodissociation is performed is of great interest. Ultraviolet LEDs and lamps offer a less expensive alternative to lasers and could be mounted within the vacuum system of the mass spectrometer. This would allow mounting of the light source in closer proximity to the trapped ions and minimize the impact lower output light sources.

Another area of future interest would be histone analysis. Histones are of great biological interest due to the diverse regulatory roles this group of proteins plays. Histones are involved in the packing of DNA into nuclei and play a critical role in regulation of the transcription through modulation of DNA structure.[2] The regulatory mechanisms of these functions involve extensive post-translational modification of the histone proteins. 193 nm UVPD could be used to greatly enhance our knowledge of the combinatorial nature of the modifications and their phenotypic effects. The extensive fragmentation afforded by UVPD allows precise localization of post-translational modifications throughout the proteins as is illustrated in **Figure 9.2** for human histone H4 and H3.2. Future studies will require extensive chromatography to allow interrogation of individual proteoforms and precise determination of the combinatorial nature of histone modifications.[3–5] the ultrafast activation time scale of UVPD is expected to be a major advantage with regards to compatibility with liquid chromatography time scales.

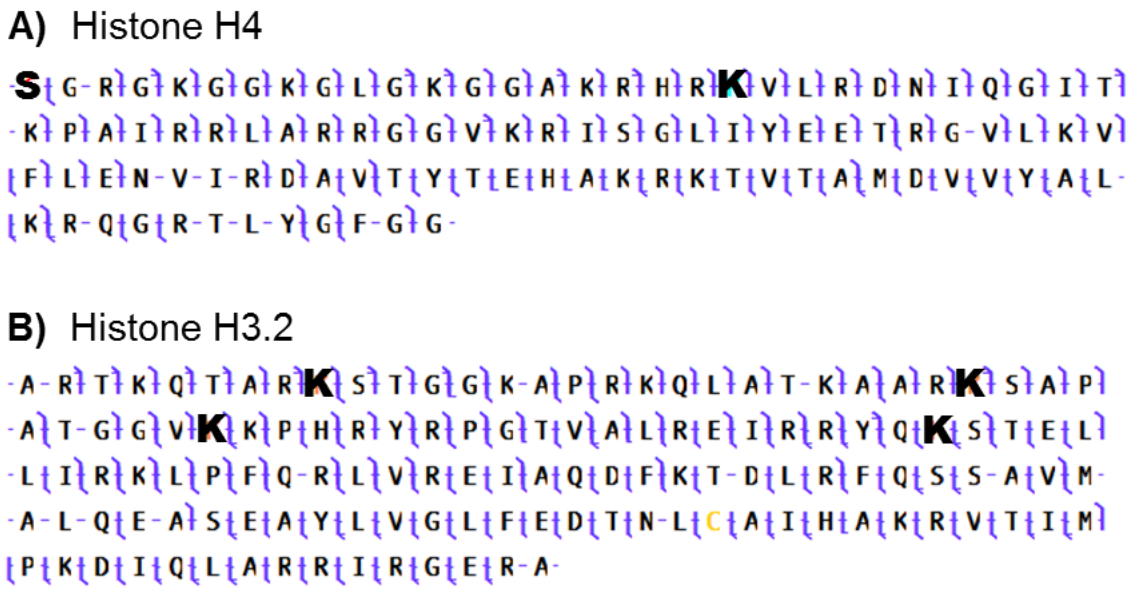

Figure 9.2 UVPD fragment ion maps of Human histone H4 (A) and histone H3.2 (B). For H4, the bold 'S' indicates N-terminal acetylation and the bold 'K' indicates dimethylation. For H3.2, bold 'K' indicates methylation.

References

(1) Agarwal, A.; Diedrich, J. K.; Julian, R. R. *Anal. Chem.* **2011**, *83*, 6455–6458.

(2) Fischle, W.; Wang, Y.; Allis, C. D. *Curr. Opin. Cell Biol.* **2003**, *15*, 172–183.

(3) Garcia, B. A.; Barber, C. M.; Hake, S. B.; Ptak, C.; Turner, F. B.; Busby, S. A.; Shabanowitz, J.; Moran, R. G.; Allis, C. D.; Hunt, D. F. *Biochemistry (Mosc.)* **2005**, *44*, 13202–13213.

(4) Garcia, B. A.; Pesavento, J. J.; Mizzen, C. A.; Kelleher, N. L. *Nat. Methods* **2007**, *4*, 487–489.

(5) Young, N. L.; DiMaggio, P. A.; Plazas-Mayorca, M. D.; Baliban, R. C.; Floudas, C. A.; Garcia, B. A. *Mol. Cell. Proteomics* **2009**, *8*, 2266–2284.