

**Novel Fourier Transform Ion Cyclotron Resonance Mass Spectrometric  
Strategies for Peptide Nucleic Acids and Acidic Biomacromolecules:  
Qualitative Characterization and Quantification**

by

**Hangtian Song**

**A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Chemistry)  
in The University of Michigan  
2012**

**Doctoral Committee:**

**Associate Professor Kristina I. Håkansson, Chair  
Professor Zhan Chen  
Professor Mark E. Meyerhoff  
Assistant Professor Nina Lin**

---

© Hangtian Song  
2012

To My Dear Family

## **Acknowledgements**

After five and half years at the University of Michigan, I would like to express my greatest appreciation to those who offered help and support to me.

First and foremost I would like to thank my Ph. D. advisor Dr. Kristina I. Håkansson for her brilliant advice and continuous support. Her excellent personality encouraged me to persist and overcome hurdles in my studies. Her vast knowledge of science provided me a great resource to lay the cornerstone of my thesis work. Her valuable insights inspired me to explore unknown paths. It is my fortune and honor to work with her out of the past years.

I would like to thank the dissertation committee members for their valuable help. Dr. Meyerhoff provided me an excellent opportunity to experience scientific research in his group during my first year at Michigan. He also sets a perfect example as an instructor and I was proud to be one of his teaching assistants for multiple semesters. Dr. Chen offered great suggestions during committee meetings. Dr. Lin collaborated with our group for a long time. She always provided solid help, motivating suggestions and driving force to move the projects forward.

I also would like to thank Prof. Brandon Ruotolo, Prof. Robert Kennedy and Prof. Carol Fierke for constructive advice and helpful discussions.

In the Håkansson lab I always received guidance and help from former and present group members. Former group members include Jiong Yang, Hye Kyong Kweon, Haichuan Liu, Jingjie Mo, Jason Kieltyka, Julie Adamson, Yibing Kong, Noah Gardner, Natasa Kalli, Ashley Brant, Hyun Ju Yoo, Bo Wang, Chris Rath and Wen Zhou. Current members include Katie Hersberger, Di Gao, Ning Wang and Wendi Hale. I sincerely appreciate their kindness and friendship.

I would like to acknowledge my collaborators for their great effort. They are Dr. Marjan Varedi (Lin group, Dept. Chem. Eng.), Dr. Miri Yoon (Marsh group), Dr. Xiaomu Guan (Fierke group), Dr. Lidong Wang (Simeone group, Med. School), Danming Tang (Wang group, MCDB), Yueyang Zhong (Routolo group), Peng Song, Jing Nie and Shuwen Sun (Kennedy group).

My research was funded by an NSF CAREER Award to Dr. Håkansson, a George Ashworth Analytical Chemistry Fellowship, a Murrill Memorial Scholarship/Winter Departmental Fellowship, a Rackham Graduate Student Research Grant, and financial support from Department of Chemistry at the University of Michigan through teaching assistantship.

At last but not least this dissertation could not be accomplished without the endless encouragement and support from my beloved family and Cacey. I would like to deliver special thanks to them. Also I would like to express my gratitude to friends in Ann Arbor, across the United States, China and other countries.

Hangtian Song

November 18, 2011

Ann Arbor, Michigan

## Table of Contents

<b>Dedication</b> .....	<b>ii</b>
<b>Acknowledgements</b> .....	<b>iii</b>
<b>List of Figures</b> .....	<b>xi</b>
<b>List of Schemes</b> .....	<b>xv</b>
<b>List of Tables</b> .....	<b>xvi</b>
<b>List of Abbreviations</b> .....	<b>xvii</b>
<b>Abstract</b> .....	<b>xix</b>
<b>Chapter</b> .....	
<b>1. Introduction</b> .....	<b>1</b>
1.1. Characterization of Peptides/Proteins with Mass Spectrometry.....	1
1.1.1. Bottom-up Approach of Proteomics.....	3
1.1.2. Top-down Approach of Proteomics .....	6
1.2. Characterization of Nucleic Acids with Mass Spectrometry... ..	6
1.3. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.....	8
1.3.1. Basic Concepts.....	8
1.3.2. Factors Affecting FT-ICR Detection Performance and Solutions for Improvement .....	12
1.3.3. Instrumentation .....	15
1.4. Electrospray Ionization (ESI).....	16

1.4.1. Principle and Implementation.....	16
1.4.2. Characteristics of ESI.....	18
1.5. Tandem Mass Spectrometry (MS/MS, MS <sup>n</sup> ).....	19
1.5.1. Precursor Ion Selection.....	20
1.5.2. Tandem Mass Spectrometry in Time or in Space.....	21
1.5.3. Collision Activated Dissociation (CAD) .....	22
1.5.4. Infrared Multiphoton Dissociation (IRMPD) .....	26
1.5.5. Electron Capture Dissociation (ECD) .....	27
1.5.6. Electron Detachment Dissociation (EDD) .....	30
1.6. Dissertation Overview.....	31
1.7. Bibliography.....	33
<b>2. Gas-phase Ion Activation Techniques for Characterization of Peptide Nucleic Acids (PNAs).....</b>	<b>43</b>
2.1. Introduction.....	43
2.2. Experimental Section.....	45
2.2.2. Sample Preparation.....	45
2.2.3. Mass Spectrometry.....	45
2.3. Results and Discussion.....	46
2.3.1. CAD, IRMPD and ECD of PNA CAGCAG and ATGTCAGTCA in Positive Ion Mode.....	46
2.3.2. CAD, IRMPD and EDD of PNA CAGCAG and ATGTCAGTCA in Negative Ion Mode.....	55
2.4. Conclusions.....	58
2.5. Bibliography.....	59

<b>3. Top-down MS/MS of Ribonucleic Acids and Their Complexes.....</b>	<b>62</b>
3.1. Introduction.....	62
3.2. Experimental Section.....	64
3.2.1. Sample Preparation.....	64
3.2.2. Mass Spectrometry and Ion Mobility-Mass Spectrometry .....	65
3.3. Results and Discussion.....	66
3.3.1. Effects of Organic Solvent in Electrospray Solution on 27-mer RNA Gas-phase Structure.....	66
3.3.2. Dependence of EDD and CAD Fragmentation Efficiency on Precursor Ion Charge State.....	70
3.3.3. EDD and CAD of RNA-ligand Complexes.....	74
3.3.4. Correlation between Specific Binding and RNA R1 Charge State.....	77
3.3.5. Correlation between Binding Specificity and RNA-ligand Molar Ratio.....	80
3.4. Conclusions.....	82
3.5. Bibliography.....	84
<b>4. Top-down MS/MS of Proteins in Negative Ion Mode.....</b>	<b>88</b>
4.1. Introduction.....	88
4.2. Experimental Section.....	92
4.2.1. Sample Preparation.....	92
4.2.2. Mass Spectrometry.....	92
4.3. Results and Discussion.....	93
4.3.1. Infrared Multiphoton Dissociation of $\beta$ -casein Anions .....	94
4.3.2. Activated Ion Electron Detachment Dissociation of $\beta$ -casein Anions..	94

4.3.3. Comparison between Negative and Positive Ion Mode.....	96
4.4. Conclusions.....	99
4.5. Bibliography.....	100
<b>5. Characterization of Oligonucleotides by Negative Ion Electron Capture Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry</b>	<b>104</b>
5.1. Introduction.....	104
5.2. Experimental Section.....	107
5.2.1. Sample Preparation.....	107
5.2.2. Mass Spectrometry.....	107
5.3. Results and Discussion.....	108
5.3.1 Electron Capture of RNA Hexamers (rGCAUAC, rA <sub>6</sub> , rC <sub>6</sub> , rG <sub>6</sub> and rU <sub>6</sub> ) .....	108
5.3.2. Electron Capture of RNA 11-mer and 27-mer.....	113
5.3.3 niECD-IRMPD (MS <sup>3</sup> ) and AI-niECD of RNA rGCAUAC Anions.....	115
5.3.4 niECD-IRMPD of RNA rA <sub>6</sub> , rG <sub>6</sub> , rC <sub>6</sub> and rU <sub>6</sub> Anions and DNA dA <sub>6</sub> , dG <sub>6</sub> , dC <sub>6</sub> and dT <sub>6</sub> Anions .....	116
5.4. Conclusions.....	119
5.5. Bibliography.....	119
<b>6. Gas-Phase Proton Transfer Reactions and Segmented-flow Fraction Collection for Improved niECD of Phosphopeptides</b> .....	<b>123</b>
6.1. Introduction.....	123
6.2. Experimental Section.....	126
6.2.1. Sample Preparation.....	126
6.2.2. Mass Spectrometry.....	128

6.3. Results and Discussion.....	129
6.3.1. niECD of Doubly and Singly Deprotonated Phosphopeptide BBC.....	129
6.3.2. PTR-assisted niECD of Doubly Deprotonated Phosphopeptide BBC....	134
6.3.3. Effect of PTR on Sensitivity and Product Ion Yield of niECD.....	137
6.3.4. Coupling Segmented-flow Fraction Collection and niECD.....	141
6.4. Conclusions.....	144
6.5. Bibliography.....	144
<b>7. Conclusions and Future Outlook.....</b>	<b>148</b>
7.1. Summary of Results.....	149
7.2. Prospects for Future Work.....	152
7.2.1. Comparison of Ionization Efficiency for Detecting Acid Peptides in Positive and Negative Ion Mode.....	152
7.2.2. niECD of Proteins and Larger RNAs.....	156
7.3. Bibliography.....	158
<b>Appendix. The Quantitative Investigation of Degree of Phosphorylation of Sic1 Protein.....</b>	<b>160</b>
A.1. Introduction.....	160
A.2. Experimental Section.....	163
A.3. Results and Discussion.....	164
A.3.1. Comparison of Ionization Efficiency of Sic1 Peptides in Positive and Negative Ion Mode .....	164
A.3.2. Validation of Valine-Leucine Substituted Standard Peptides .....	168
A.3.3. Investigation of Phosphorylation Kinetics of Sic1.....	172

A.3.4. Investigation of Distribution of Phosphorylation States on Sic1 wild type and Mutants.....	172
A.4. Conclusions.....	174
A.5. Bibliography.....	175

## List of Figures

### Figures

- Figure 1.1.** Systematic overview of bottom-up and top-down mass spectrometry for dissecting the primary structure of proteins. 3
- Figure 1.2.** (upper panel) Schematic diagram of an FT-ICR detector and the excitation of an ion. (lower panel) (a) The time-domain signal of ubiquitin ions recorded by image current detection. (b) Frequency domain signal from the time-domain signal through fast Fourier transformation. (c) Mass spectrum from calibration of the frequency domain signal. 11
- Figure 1.3.** (a) Simulated transient of  $C_{60}$  signal with asymmetric overloading of the ADC. (b) Simulated mass spectrum with 1 Da spacing artifacts due to intermodulation of the  $C_{60}$  isotope signals. 15
- Figure 1.4.** Configuration of a commercial 7-T FT-ICR mass spectrometer. 16
- Figure 1.5.** Schematic description of ESI. 17
- Figure 2.1.** Structure of the PNA CAGCAG. 44
- Figure 2.2.** Product ion nomenclature used in this Chapter. 47
- Figure 2.3.** CAD (16 V collision cell voltage), IRMPD (7.5 W, 80 ms), and ECD (- 1 V, 7 ms) MS/MS spectra of the PNA CAGCAG. 48
- Figure 2.4.** Backbone cleavages from CAD, IRMPD, and ECD of the PNA CAGCAG in positive ion mode. 48
- Figure 2.5.**  $MS^3$  (CAD of  $M^{3+}$  followed by SORI-CAD of  $[M - H_2O]^{3+}$ ) of the PNA CAGCAG 51
- Figure 2.6.** IRMPD of triply protonated PNA CAGCAG with and without simultaneous double resonance (DR) ejection of  $[M - H_2O]^{3+}$ . 52

<b>Figure 2.7.</b> Abundance changes of precursor and y-type product ions with and without DR in IRMPD of the PNA CAGCAG.	52
<b>Figure 2.8.</b> Backbone cleavages from CAD, IRMPD and ECD o the PNA ATGTCAGTCA in positive ion mode.	53
<b>Figure 2.9.</b> IRMPD spectra of the non-acetylated 10-mer PNA ATGTCAGTCA (upper panel) and its acetylated variant (bottom panel).	54
<b>Figure 2.10.</b> Negative ion mode CAD (19 V collision cell voltage), IRMPD (7.5 W, 180 ms), and EDD (- 18 V, 2 s) MS/MS spectra of the PNA CAGCAG.	55
<b>Figure 2.11.</b> Backbone cleavages from CAD, IRMPD and EDD of the PNA CAGCAG in negative ion mode.	56
<b>Figure 2.12.</b> Backbone cleavages from CAD, IRMPD and EDD of the PNA ATGTCAGTCA in negative ion mode.	58
<b>Figure 3.1.</b> Sequence and secondary structure of the 27-mer RNA R1.	67
<b>Figure 3.2.</b> 3D structure of aminoglycoside-bound sites on tRNA, mRNA, and 70S prokaryotic rRNA. 70S rRNA incorporates 16S rRNA as a subunit.	67
<b>Figure 3.3.</b> Negative ion ESI spectrum of the 27-mer RNA R1.	68
<b>Figure 3.4.</b> IM-MS of RNA R1 dissolved in 1% triethylamine and 50% (v/v) methanol in water.	69
<b>Figure 3.5.</b> CCS comparison of RNA R1 electrosprayed from solutions with different methanol content.	70
<b>Figure 3.6.</b> AI-EDD spectrum of the RNA R1 at a charge state of 8-.	71
<b>Figure 3.7.</b> Backbone cleavages observed in AI-EDD of the RNA R1 at charge states of 5- to 8-.	72
<b>Figure 3.8.</b> Charge dependence of EDD and CAD efficiency.	72
<b>Figure 3.9.</b> <i>d/w</i> -type ions observed in AI-EDD experiments of R1-PM complexes at (a) 6-, (b) 7-, and (c) 8- charge states.	75
<b>Figure 3.10.</b> <i>c/y</i> -type ions observed in CAD experiments of R1-PM complexes at (a) 6-, (b) 7-, and (c) 8- charge states.	76
<b>Figure 3.11.</b> Structures of aminoglycosidic ligands: (a) paromomycin (PM), (b)(Be)kanamycin (BK), (c) apramycin (AP), and (d) ribostamycin (RM).	78

<b>Figure 3.12.</b> ESI-MS of RNA R1 and ligand mixture. The mixture contained 2 $\mu$ M RNA R1 and 0.1 $\mu$ M of PM, BK, AP, and RM.	79
<b>Figure 3.13.</b> ESI-MS of RNA R1 and ligand mixture.	81
<b>Figure 4.1.</b> (a) Negative ion IRMPD of all observed charge states (15- to 26-) of $\beta$ -casein anions generated by electrospray ionization. Inset: partial IRMPD spectrum. Noise peaks are labeled with asterisks. (b) AI-EDD of all observed charge states of $\beta$ -casein anions generated by electrospray ionization. Only <i>a</i> -type fragments are labeled. Spectra are summed over 256 scans. (c) Fragmentation map from (a). (d) Fragmentation map from (b).	93
<b>Figure 4.2.</b> Comparison of $\beta$ -casein fragmentation patterns in negative and positive ion mode.	96
<b>Figure 4.3.</b> Comparison of calmodulin fragmentation patterns between negative ion mode and positive ion mode.	97
<b>Figure 4.4.</b> Comparison of RNase B fragmentation patterns between negative ion mode and positive ion mode.	98
<b>Figure 4.5.</b> Comparison of the numbers of backbone cleavages in negative and positive ion mode top down MS/MS of (a) $\beta$ -casein; (b) calmodulin; and (c) RNase B.	99
<b>Figure 5.1.</b> Electron capture by rGCAUAC.	110
<b>Figure 5.2.</b> niECD of RNA hexamers: (a) rA <sub>6</sub> , (b) rG <sub>6</sub> , (c) rC <sub>6</sub> , (d) rU <sub>6</sub> .	110
<b>Figure 5.3.</b> Electron capture efficiencies for different oligoribonucleotides.	112
<b>Figure 5.4</b> Electron capture efficiencies for RNA hexamers: (a) rA <sub>6</sub> , (b) rG <sub>6</sub> , (c) rC <sub>6</sub> , (d) rU <sub>6</sub> .	112
<b>Figure 5.5</b> Electron capture by UAACUAUGACG in its 2- charge state following irradiation with 4.5 eV electrons for 6 s.	114
<b>Figure 5.6.</b> Electron capture by rGGCGUCACACCUUCGGGUGAAGUCGCC in its 5- charge state following irradiation with 8.4 eV electrons for 3s.	114
<b>Figure 5.7.</b> MS <sup>3</sup> (niECD-IRMPD) of singly deprotonated rGCAUAC.	116
<b>Figure 5.8</b> AI-niECD of singly deprotonated rGCAUAC.	116
<b>Figure 6.1.</b> Overlay of two negative ion mode ESI mass spectra of the phosphopeptide H-FQ-pS-EEQQQTEDELQDK-OH at 0.5 $\mu$ M.	129

<b>Figure 6.2.</b> niECD of the singly (a) and doubly (b) deprotonated phosphopeptide H-FQpSEEQQQTEDELQDK-OH.	131
<b>Figure 6.3.</b> Mass spectrum following PTR between doubly deprotonated peptide anions and protonated fluoranthene.	134
<b>Figure 6.4.</b> PTR-niECD spectrum of the doubly deprotonated phosphopeptide H-FQpSEEQQQTEDELQDK-OH.	135
<b>Figure 6.5.</b> (a) niECD of the singly deprotonated phosphopeptide H-FQ-pS-EEQQQTEDELQDK-OH at 0.1 $\mu$ M. The spectrum is shown at 100% scale. (b) PTR-niECD of the doubly deprotonated form of the same peptide..	137
<b>Figure 6.6.</b> Comparison of fragmentation efficiencies in non-PTR niECD and PTR-niECD at two peptide concentrations.	139
<b>Figure 6.7.</b> Total ion chromatograms of a peptide mixture (NPF, BBC, DAM1) separated by nano-LC.	142
<b>Figure 6.8.</b> niECD spectra of (a) NPF, (b) BBC and (c) DAM1. Peptides were collected by segmented flow fraction collection from nano-LC and infused into a nanoESI-FT-ICR mass spectrometer.	143
<b>Figure 7.1.</b> Distribution of journal articles related to mass spectrometry in ten different scientific areas.	148
<b>Figure 7.2.</b> Comparison of extracted ion chromatograms for the peptide H-TPQKPSQNLVPVTPSTTK-OH in unphosphorylated (0p), singly (1p), and doubly (2p) phosphorylated forms in positive and negative ion mode.	154
<b>Figure 7.3.</b> Systematic diagram of the principle for fine-tuning LC-MS flow rate.	155
<b>Figure 7.4.</b> AI-niECD of ubiquitin with 5 negative charges.	157
<b>Figure A.1.</b> Scheme of Sic1 regulatory activity in the cell cycle of <i>S. cerevisiae</i> .	161
<b>Figure A.2.</b> Comparison of extracted ion chromatograms for selected tryptic peptides from the protein Sic1 in positive and negative ion mode.	166
<b>Figure A.3.</b> Comparison of EICs for H-TPQKPSQNLVPVTPSTTK-OH and its valine-leucine substituted reference peptide H-TPQKPSQNLLPVTPSTTK-OH.	169
<b>Figure A.4.</b> Phosphorylation kinetics of T45 (left) and S76 (right) in a Sic1-2p mutant	170
<b>Figure A.5.</b> Distribution of Sic1 phosphoforms obtained from top down mass spectrometry.	173

## List of Schemes

### Schemes

<b>Scheme 1.1.</b> Nomenclature for peptide/protein fragmentation.	25
<b>Scheme 1.2.</b> Generation of <i>b</i> and <i>y</i> -type ions from CAD of peptides/proteins.	25
<b>Scheme 1.3.</b> Nomenclature for oligonucleotide fragmentation.	26
<b>Scheme 1.4.</b> ( <i>a</i> - base) and <i>w</i> -type ions from CAD of DNA.	26
<b>Scheme 1.5.</b> Fragmentation pathway in ECD.	28
<b>Scheme 1.6.</b> Fragments from peptide ECD: <i>c</i> and <i>z</i> <sup>•</sup> -type ions.	28
<b>Scheme 1.7.</b> Fragmentation route in EDD.	31
<b>Scheme 3.1.</b> Schematic summary of the correlation between AI-EDD efficiency, binding specificity, precursor charge state, and ligand amount in the electrospray solution.	83

## List of Tables

### Tables

<b>Table 2.1.</b> List of assigned peaks in the positive ion mode CAD spectrum of the PNA CAGCAG.	50
<b>Table 2.2.</b> List of assigned peaks in negative ion mode CAD of the PNA CAGCAG.	57
<b>Table 5.1.</b> Summary of fragments identified in niECD-IRMPD spectra of RNA and DNA hexamers A <sub>6</sub> , G <sub>6</sub> , C <sub>6</sub> and U <sub>6</sub> (T <sub>6</sub> ).	118
<b>Table 6.1.</b> Comparison of sequence coverage in the form of <i>c/z</i> -type fragments from PTR-niECD of doubly deprotonated peptides and niECD of the same peptides in their singly charged form.	138

## List of Abbreviations

ADC	Analog-to-digital converter
AGC	Automatic gain control
AI-ECD	Activated-ion electron capture dissociation
AI-EDD	Activated-ion electron detachment dissociation
BIRD	Blackbody infrared radiative dissociation
CAD	Collision activated dissociation
CCS	Collision cross sections
CID	Collision-induced dissociation
CE	Capillary electrophoresis
CI	Chemical ionization
DNA	Deoxyribonucleic acids
DR	Double resonance
ECD	Electron capture dissociation
EDD	Electron detachment dissociation
EID	Electron induced dissociation
EIC	Extracted ion chromatogram
ESI	Electrospray ionization
ETD	Electron transfer dissociation
FAB	Fast atom bombardment
FEL	Free electron laser
FT	Fourier transform
FWHM	Full width at half maximum
GC	Gas chromatography
HDX	Hydrogen-deuterium exchange
HeCAD	High-energy CAD
ICAT	Isotope-coded affinity tags
ICR	Ion cyclotron resonance
IR	Infrared
IM	Ion mobility
IRMPD	Infrared multiphoton dissociation
IVR	Intramolecular vibrational-energy redistribution

LC	Liquid chromatography
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
MS/MS or MS <sup>n</sup>	Tandem mass spectrometry
<i>m/z</i>	Mass-to-charge ratio
Nano ESI	Nano electrospray ionization
NETD	Negative electron transfer dissociation
niECD	Negative ion electron capture dissociation
NMR	Nuclear magnetic resonance
PCR	Polymerase chain reaction
pI	Isoelectric point
PNA	Peptide nucleic acids
ppm	Parts-per-million
ppb	Parts-per-billion
PTM	Posttranslational modification
PTR	Proton transfer reaction
Q	Quadrupole
RNA	Ribonucleic acids
RF	Radio frequency
RP	Reverse phase
SCS	Strong cation exchange
SILAC	Stable isotope labeling with amino acids in cell culture
S/N	Signal to noise
SNP	Single base polymorphisms
SID	Surface induced dissociation
SORI-CAD	Sustained off-resonance irradiation collision activated dissociation
SWIFT	Stored waveform inverse Fourier transform
T	Tesla, a unit of magnetic field
TOF	Time of flight
UniProt	Universal Protein Resource
UHPLC	Ultrahigh-pressure liquid chromatography
UHV	Ultra-high vacuum
UVPD	Ultraviolet photodissociation

## Abstract

Mass spectrometry is widely applied to characterize oligo- and poly-nucleotides, peptides, and proteins. In tandem mass spectrometry (MS/MS), analytes are dissociated in the gas phase and their fragments are analyzed to generate structural information. Among the various MS/MS techniques, few are well suited for acidic biomolecules, particularly nucleic acids and acidic peptides/proteins.

In this dissertation emerging MS/MS methods are applied to new problems and to molecules not previously studied by such techniques. Gas-phase fragmentation of peptide nucleic acid (PNA) cations and anions was examined in collision activated dissociation (CAD), infrared multiphoton dissociation (IRMPD), electron capture dissociation (ECD), and electron detachment dissociation (EDD). Negative-ion CAD and IRMPD provided the most sequence information and were simplest to interpret. Positive-ion mode MS<sup>3</sup> and double resonance experiments revealed mechanistic details on the water loss pathway. N-terminal acetylation blocked this pathway and simplified spectrum interpretation. A 27-mer ribonucleic acid (RNA) and its complexes with aminoglycoside antibiotics were characterized by ion-mobility MS, negative-ion CAD, and EDD. Higher binding specificity and lower gas-phase charge state were correlated with a more compact RNA structure. EDD fragmentation efficiency was higher for higher charge states. Negative-ion IRMPD and EDD were compared to positive-ion

IRMPD and ECD for dissociating intact proteins. For acidic proteins, negative ion mode techniques provided complementary sequence information to positive ion mode. Negative-ion ECD (niECD) was investigated for oligonucleotide characterization. Activated ion-niECD and niECD-IRMPD MS3 provided similar fragmentation patterns as EDD, possibly due to similar gas-phase precursor ion structures. Characteristic “ECD-like” radical fragments also suggested similarities to ECD, which may be due to DNA/RNA zwitterionic structures in negative ion mode.

Proton transfer reaction (PTR) was performed prior to niECD of phosphopeptides. For peptides larger than 1700 Da, PTR reduced the ion accumulation time and increased the detection efficiency of niECD fragments. niECD was also coupled to offline nano-LC by segmented flow fraction collection, which allowed more time for signal averaging, compared with online LC coupling.

Overall this dissertation presents method development for improved MS/MS analysis in systems biology, i.e., the study of organisms as a whole rather than focusing on individual biochemical pathways.

# Chapter 1

## Introduction

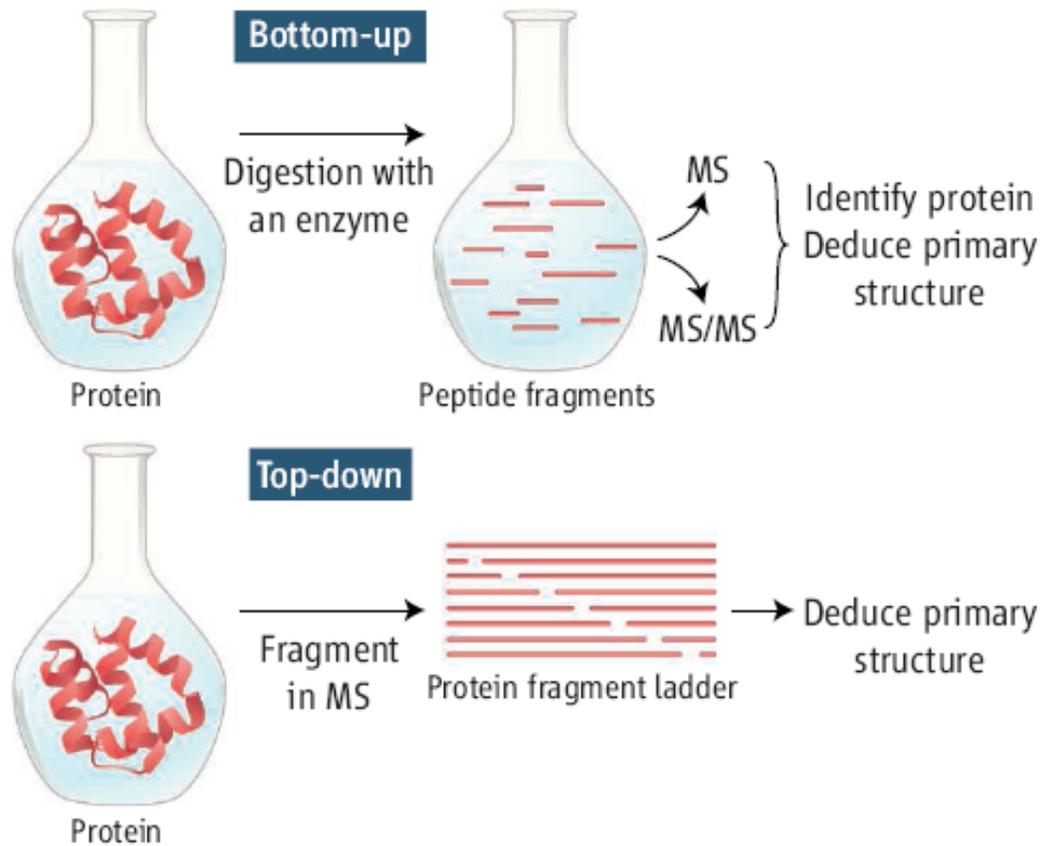
### 1.1 Characterization of Peptides/Proteins with Mass Spectrometry

The proteome,<sup>1,2</sup> in analogy with the genome, is the collection of all proteins present in an organism, cell, or system. Proteomics, an analog to genomics, aim to investigate the whole spectrum of proteins, including their identities, quantities, modifications, interactions, and functions to understand a biological system. The challenges for researchers lie in the qualitative and quantitative complexity of the proteome. As revealed by the Human Genome Project, the human genome encodes 20,000-25,000 protein-coding genes.<sup>3</sup> However, post-translational modifications (PTMs) are commonly observed. According to the Universal Protein Resource (UniProt, a well established online protein database), 434 different PTMs have been reported.<sup>4</sup> Also, protein-protein and protein-ligand interactions dramatically increase the qualitative complexity. It has been estimated by statistical methods that around 650,000 protein-protein interactions are possible in humans.<sup>5</sup> Quantitative investigation is also challenging due to the requirement of high sensitivity and dynamic range. The abundance of a protein is affected by its expression level, PTM level, conditions within the cell or the organism, biological activities and so on. Due to these complexities,

technologies for studying proteomics should be able to detect, identify, and quantify proteins with high sensitivity, specificity, and efficiency. Proteomics combines techniques from biology, biochemistry, chemistry, biophysics, bioinformatics, biomedical engineering, chemical engineering, and other disciplines.

Mass spectrometry is one powerful technology contributing to current proteomics. State-of-the-art mass spectrometers, along with other tools, provide high sensitivity, high specificity, and high through-put.<sup>6,7</sup> Peptides can be detected at low attomole<sup>8,9</sup> or even zeptomole<sup>10</sup> amounts. Mass accuracy at the hundreds of parts-per-billion (ppb) level has been achieved.<sup>11</sup> From coupling with separation techniques, around 1000 proteins can be identified in one hour.<sup>12</sup>

There are two complementary approaches for protein analysis by mass spectrometry: bottom-up and top-down analysis, as shown in Figure 1.1.<sup>13</sup> These two approaches are discussed in the following sections.



**Figure 1.1.** Systematic overview of bottom-up and top-down mass spectrometry for dissecting the primary structure of proteins. Adapted from Chait.<sup>13</sup>

### 1.1.1 Bottom-up Approach of Proteomics

Bottom-up mass spectrometry for proteomics is widely applied for protein identification, quantification, PTM analysis, and structural analysis. The sample is a collection of proteins of interest, ranging from one particular protein to a whole cell lysate or tissue extract. The collection of proteins is digested by proteases such as trypsin to generate a pool of peptides. Then the peptides are analyzed by mass spectrometry. By piecing back the identified peptides, facilitated by bioinformatics tools and appropriate databases, primary structure of proteins can be identified.

Several techniques are closely associated with mass spectrometry in bottom-up proteomics. Various separation techniques, such as liquid chromatography (LC), capillary electrophoresis (CE), and gel electrophoresis, play important roles in the workflow. Their ability to fractionate proteins or peptides according to various physical and chemical properties reduces sample complexity. More proteins/peptides can be identified due to the increased peak capacity of the combined separation-MS workflow.<sup>14</sup> For example, a 12T Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer can offer a peak capacity of ~25,000 for peptides,<sup>15</sup> while ultrahigh-pressure liquid chromatography (UHPLC) offers peak capacity of 1500.<sup>16</sup> It has been estimated that coupling reversed-phase (RP) HPLC to FT-ICR-MS provides an additional two orders of magnitude increase in peak capacity, and further coupling of strong cation exchange (SCX) fractionation provides another order of magnitude increase.<sup>15</sup> Two-dimensional (2-D) gel electrophoresis also has been widely demonstrated to separate proteins from a whole proteome.<sup>1</sup> Similar to 2-D liquid chromatography, the two separation methods should be orthogonal, i.e., based on different properties of proteins, such as size and isoelectric point (pI). 2D electrophoresis is capable of separating thousands of proteins and compatible with mass spectrometry through in-gel digestion.<sup>17</sup> However, compared to LC-MS this approach is labor intensive and low-throughput.

With the development of bioinformatics tools, the bottom-up approach is widely used for identifying proteins. One data set from an LC-MS/MS run in large-scale proteomics can contain tens of thousands of spectra. Fully manual interpretation is not feasible and thus computerized data searching tools such as Mascot<sup>18</sup> and SEQUEST,<sup>19</sup> as well as other specialized software, compare the experimental data to public or custom-

built databases. By balancing between sensitivity and specificity, automated data interpretation can yield acceptable results.<sup>20</sup>

Protein quantification by mass spectrometry is also achievable. Due to the fact that mass spectral ion abundances are not directly correlated to protein concentration, some forms of internal standards or references are required. Absolute and relative quantification can be achieved by labeling techniques such as isotope-coded affinity tags (ICAT),<sup>21</sup> stable isotope labeling with amino acids in cell culture (SILAC),<sup>22</sup> and isobaric tagging,<sup>23</sup> or by label-free techniques such as spectral counting.<sup>24</sup>

Structural analysis by bottom-up approach is available through chemical crosslinking methods.<sup>25</sup> The idea is to covalently link two amino acid residues in spatial proximity.<sup>26</sup> From protease digestion and identification of crosslinked peptides, a map of protein intramolecular interactions or protein-protein interactions can be built.<sup>27</sup> Chemical crosslinkers can be designed to fit particular applications. Crosslinkers with different lengths probe distance between two interacting amino acids whereas crosslinkers with different reactive end groups can probe different types of interacting amino acids. Crosslinkers with chemically enrichable motifs or mass spectrometric signature motifs can enhance the detection of crosslinked peptides.

The bottom-up approach, compared with the top-down approach, offers several advantages: Peptides in general, compared to proteins, are easier to dissolve in water, are more stable, and have higher ionization efficiency. The bottom-up approach is also considered high throughput in large-scale proteomics. However, typically not all peptides from one protein are detected. Also, one peptide can originate from different

proteins. Therefore the correlation from peptides back to proteins can be problematic and requires support from bioinformatics tools.<sup>13</sup>

### **1.1.2 Top-down Approach of Proteomics**

Top-down mass spectrometry for proteomics is applied for protein identification, PTM analysis, structural analysis, protein-protein interaction analysis, and protein-ligand analysis. In the top-down approach, intact proteins are introduced to the gas phase without protease digestion and their intact mass is measured. Tandem mass spectrometric techniques allow protein ions to be fragmented. Based on the fragment masses, protein primary structure can be obtained facilitated by bioinformatics tools and appropriate databases. Electron capture dissociation (ECD)<sup>28</sup> and electron transfer dissociation (ETD)<sup>29</sup> benefit top-down mass spectrometry based on their ability to dissociate the protein backbone more randomly than collision activated dissociation (CAD) and their ability to retain labile PTMs. PTMs can thus be directly located.<sup>30, 31</sup> Protein structure can also be probed by hydrogen/deuterium exchange (HDX)<sup>32-34</sup> or ECD.<sup>35</sup> The top-down approach is able to directly detect and localize PTMs. It is also able to directly characterize protein-protein and protein-ligand complexes. However, relatively low throughput is a major disadvantage, as the sample is required to be relatively pure and abundant.

## **1.2 Characterization of Nucleic Acids with Mass Spectrometry**

The genome is the collection of an organism's hereditary information. This information is encoded in the form of deoxyribonucleic acids (DNAs) and ribonucleic acids (RNAs). Genomics aims to investigate the genome in terms of sequences,

structures, interactions, and functions. Many tools from biology, chemistry, physics, engineering, and informatics have been developed for genomics studies, such as DNA sequencing technology invented by Nobel laureate Frederic Sanger in 1977.<sup>36</sup> Direct analysis of intact oligo- and poly-nucleotides by mass spectrometry has been limited by the ionization techniques. Nucleic acids are highly polar and difficult to volatilize.<sup>37</sup> Also, the glycosidic bonds in nucleotides are labile and readily cleaved if the ionization process is not soft enough.<sup>38</sup> The introduction of “soft” ionization techniques such as matrix-assisted laser desorption/ionization (MALDI)<sup>39,40</sup> and electrospray ionization (ESI)<sup>41</sup> greatly extended the application of mass spectrometry to genomics, due to their abilities to ionize polynucleotides efficiently without fragmentation.

Fast atom bombardment (FAB), before ESI and MALDI, was applied to ionize oligonucleotides, and was demonstrated to generate intact molecular ions along with series of fragment ions.<sup>42,43</sup> However, in practice, FAB is only able to provide molecular ions up to 10-mers, and the overall sensitivity is low.<sup>44</sup>

MALDI was shown to ionize intact oligonucleotides soon after its invention,<sup>45</sup> including successful ionization of large nucleic acids such as tRNAs.<sup>46</sup> Compared with ESI, MALDI is advantageous due to the tolerance of sample interferences such as salt and also its high sensitivity.<sup>38</sup> However, MALDI typically generates singly charged ions and the upper mass limit is thus determined by the  $m/z$  range of the mass spectrometer. Multiple charging offered by ESI brings  $m/z$  ratios of large molecules down to the detection range of most common mass spectrometers. Also, some tandem mass spectrometric techniques, such as electron detachment dissociation (EDD), require multiply charged precursor ions.

ESI is able to ionize large polynucleotides such as tRNA<sup>47</sup> and DNA PCR products<sup>48, 49</sup> in intact form. Isotopic distributions of intact nucleic acid ions can be resolved by mass analyzers with high resolving power such as FT-ICR and intact mass can be determined accurately within ppm range. In addition to intact mass measurements, nucleic acid ions can be fragmented in the gas phase. Vibrational activation such as CAD and infrared multiphoton dissociation (IRMPD),<sup>50, 51</sup> as well as ion-electron reactions such as ECD<sup>52</sup> and EDD<sup>53</sup> have been applied to generate backbone fragments. Sequence information can be retrieved from the fragment ion series. Nucleic acid complexes such as duplexes,<sup>54-57</sup> oligonucleotide-ligand complexes,<sup>58-61</sup> and RNA-protein complexes<sup>62, 63</sup> can also be identified and characterized. Combined with gas-phase H/D exchange, gas-phase structure of nucleic acids can be accessed.<sup>64-66</sup> In addition to the advantage of multiple charging, ESI is considered “softer” than MALDI. In MALDI metastable decay has been observed between the ionization and detection events,<sup>67</sup> whereas non-covalent complexes can be preserved from ESI. Despite these advantages, ESI is less tolerant to salts than MALDI. Nucleic acids are hydrophilic and highly negatively charged, thus having high affinity to metal ions such as sodium and potassium. When samples are not treated with proper desalting procedures, spectra can be complicated by adducts with various numbers of metals, which also lowers sensitivity. This problem may be reduced by precipitation from ammonium acetate and organic additives.<sup>68</sup>

### **1.3 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry**

#### **1.3.1 Basic Concepts**

Early ideas on measuring mass-to-charge (or charge-to-mass) ratios from ion cyclotron resonance frequencies in a magnetic field emerged in the 1940s and 1950s.<sup>69, 70</sup>

The first realization of this concept was the measurement of the proton charge-to-mass ratio by Sommer and Hipple in 1951.<sup>71</sup> FT-ICR-MS was first introduced by Comisarow and Marshall in 1974.<sup>72, 73</sup> Since its invention, FT-ICR-MS has been extensively studied, developed, and applied. The FT-ICR mass analyzer is the most powerful type of mass spectrometer in terms of mass accuracy and resolution. The development of “soft” ionization techniques such as ESI and MALDI, allowed FT-ICR-MS to be extensively applied for analysis of biomacromolecules such as proteins/peptides and nucleic acids.

FT-ICR-MS is based on the cyclotron motion of an ion in a highly homogenous magnetic field, usually provided by superconducting magnets. When an ion with mass  $m$ , charge  $q = ze$  in which  $e$  is the elementary charge and  $z$  is the number of elementary charges, and velocity  $\mathbf{v}$  is placed into a magnetic field  $\mathbf{B}$ , it experiences the Lorentz force with a magnitude of  $F = zevB$  and a direction perpendicular to  $\mathbf{v}$  and  $\mathbf{B}$ , as described in equation 1. This force is a centripetal force and can also be expressed as in equation (2). Combination of equations (1) and (2) leads to an expression for the angular frequency,  $v/r$  or  $\omega$ , of the cyclotron motion driven by the Lorentz force as shown in equation (3). If  $B$  is known, the mass-to-charge ratio can be determined by measuring  $\omega$ . In other words, each ion with a specific  $m/z$  ratio has a characteristic frequency in a given magnetic field.

$$F = ze\mathbf{v} \times \mathbf{B} \quad (1)$$

$$F = m \frac{v^2}{r} \quad (2)$$

$$\omega = B \frac{ze}{m} \quad (3)$$

The key component of an FT-ICR mass spectrometer is the ICR cell. There are several ICR cell geometries, such as cubic traps, cylindrical traps, and hyperbolic traps<sup>74</sup> and also various modified geometries for better performance. Despite this variety, common ICR cells consist of three basic components: trapping electrodes, excitation electrodes, and detection electrodes.

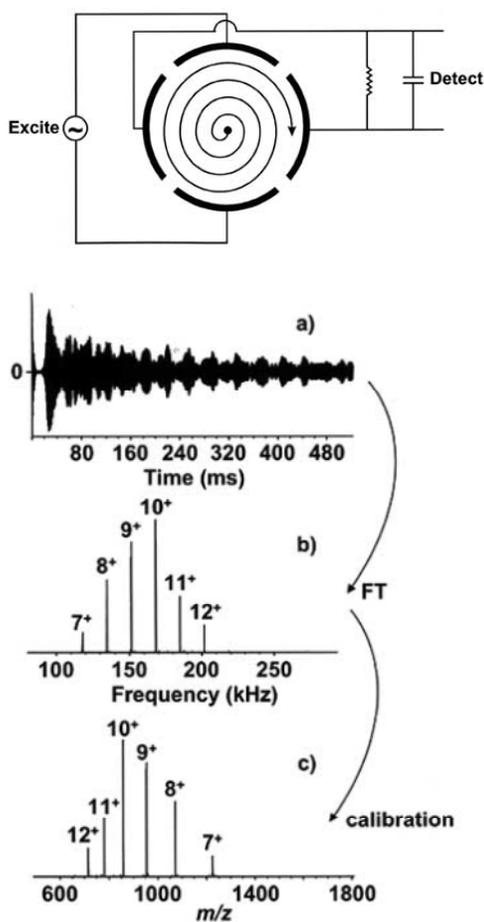
When ions are transferred along the axial direction (z-direction) into the ICR cell, trapping plates with small (~1 V) DC potentials prevent ions from escaping the cell axially. Ideally, ions are oscillating axially in a one-dimensional potential well.<sup>75</sup>

The initial ion radial velocity typically generates cyclotron motion with a radius much smaller (~1 mm) than the radius of the cell (~10 cm). A modulated waveform, or “chirp”, serving two purposes, is applied to the excitation electrodes. The first function of the chirp, which contains the frequencies of all ions within the m/z range of interest, is to cause absorption of energy by ions in resonance with the applied frequencies and thus increase their cyclotron radii. Second, the initial phases of the ion cyclotron motions are random but the excitation waveform, with a specific phase, synchronizes the phases of all ions and forces ions with the same m/z ratio into one compact ion packet.

After ions are excited and their ion cyclotron radii are close to the detection electrodes, the movement of the ion packets induces a local electric field change on the surface of the detection electrodes. The corresponding image current, flowing from one detection electrode to the other is recorded. At least about 100 charges of a given m/z are needed to induce a detectable current.<sup>76</sup> Compared with other detectors in other types of mass spectrometers, such as electron multipliers with a gain of  $10^7$ ,<sup>77</sup> the ICR detection

principle has lower sensitivity. However, ICR detectors are non-destructive and ion packets can be remeasured. The measured image current is digitized and stored as a time-domain signal. Through fast Fourier transform, time-domain signal is converted to frequency-domain signal. Frequencies are then converted to  $m/z$  ratios based on equation (3). The longer the image current is recorded, the better resolution the mass spectrum has.<sup>78</sup>

A schematic view of FT-ICR detection is shown in Figure 1.2.<sup>79</sup>



**Figure 1.2** (Upper panel) Schematic diagram of an FT-ICR detector and the excitation of an ion. (Lower panel) (a) The time-domain signal of ubiquitin ions recorded by image current detection. (b) Frequency domain signal from the time-domain signal through fast Fourier transformation. (c) Mass spectrum from calibration of the frequency domain signal.

The Orbitrap<sup>80, 81</sup> is a mass analyzer based on the Kingdon trap developed in 1923.<sup>82</sup> Similar to the detection principle in FT-ICR, Orbitrap ion detection is also based on image current and Fourier transformation. However, ion motion is confined by a combined quadro-logarithmic electrostatic field, instead of the homogeneous magnetic field in FT-ICR-MS. The geometry of an Orbitrap detector consists of an inner spine-like central electrode and an outer barrel-like electrode. The shape of these electrodes determines the shape of the electric field. Trapped ions move harmonically along the axis and also orbit the inner electrode radially. Mass-to-charge ratio is derived from the frequency of axial motion. Orbitrap mass spectrometers provide higher mass accuracy than other mass spectrometers, except for FT-ICR-MS and high performance time-of-flight (TOF) mass spectrometers.

### **1.3.2 Factors Affecting FT-ICR Detection Performance and Solutions for Improvement**

There are several factors making the performance of FT-ICR detection deviate from the ideal case.

Magnetron motion of ion packets is induced by the combination of the trapping electric field and the magnetic field. The magnetron frequency is one to three orders of magnitude smaller than the cyclotron frequency.<sup>74</sup> The observed image current frequency corresponds to the difference between the ion cyclotron frequency and the magnetron frequency. In practice, to first order approximation, the observed frequency can be expressed as shown in Equation 4.

$$m/z = \frac{A}{f} + \frac{B}{f^2} \quad (4)$$

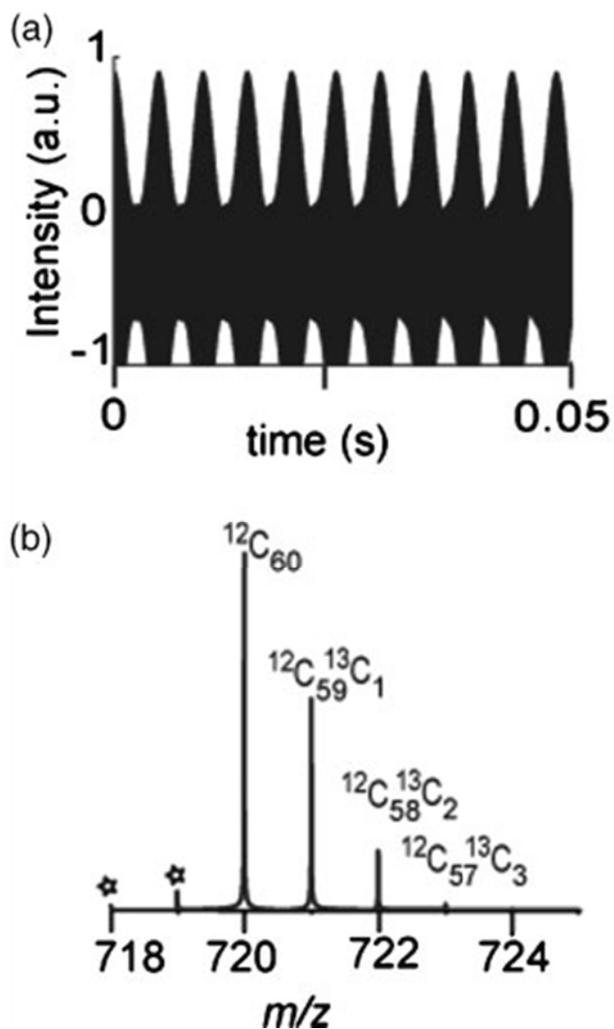
The constants A and B are experimentally determined by fitting observed frequencies for known m/z ratios to Equation 4. The magnitude of the magnetron motion is dependent on the trapping potential. Due to the fact that the electric field induced by the trapping voltages is not uniform in space, the magnitude of magnetron motion is dependent on the ion position. The magnitude variation of the magnetron motion induces peak broadening. Ways to improve the uniformity of the trapping potential include, e.g., addition of rf-shimming wires,<sup>83</sup> addition of compensation electrodes,<sup>84, 85</sup> and reshaping of excitation/detection electrodes.<sup>86</sup> Collisional damping combined with magnetron motion may be detrimental to resolution and calibration. Ideally ICR cells should be operated under absolute vacuum to maximize image current duration. In practice ICR cells are under ultra high vacuum (<10<sup>-9</sup> mBar range). However, due to the high ion velocity, collisions with neutral gas molecules still occur, resulting in partial loss of both radial and axial velocity, and thus magnetron motion at varying magnitude. If the magnetron motion is not constant within an image current transient, mass spectral resolution and calibration will deteriorate.

Space-charge effects are also affecting spectral resolution. Due to Coulomb repulsion, ion packets expand after the excitation stage. Increased ion packet sizes broaden observed frequency distributions, increase mass spectral peak widths, induce peak splitting, and even generate artificial peaks.<sup>87</sup> Ideally, few ions should be allowed into the ICR cell for optimum resolution. In practice, automatic gain control (AGC)<sup>88</sup> can be used to balance resolution and sensitivity.

Harmonic signals are observed as artificial peaks in image current detection. Given that the ion cyclotron frequency at one m/z ratio is  $\omega_c$ , odd harmonic frequencies at

$3\omega_c$ ,  $5\omega_c$ , *etc.* are commonly observed. Odd harmonics mainly arise from the discontinuous image current recorded from the detection electrode pair, due to the gap between the two detection electrodes. Optimizing the span of detection/excitation electrodes helps reduce the odd harmonics, particularly for the third harmonic.<sup>89</sup> Even harmonics are also observed but of less abundance compared to odd harmonics. Even harmonics arise from off-axis ions<sup>74</sup> and also from a mismatch in the two channels of the detection circuit.<sup>90</sup> Even harmonics can be reduced by summing time-domain signals at integer powers of 2, i.e., setting the scan number to 2, 4, 8, *etc.*

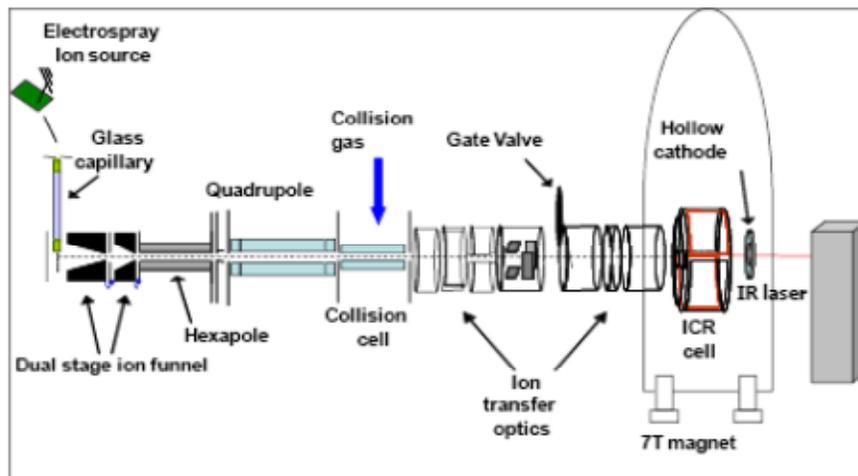
Artificial peaks at roughly 1 Da to the side of an abundant isotope cluster are observed, due to overloading of the analog-to-digital converter (ADC) for converting analog time-domain image current to discrete signal for computer input, as demonstrated in Figure 1.3.<sup>90</sup> Implementation of automatic-gain control, which controls the number of ions in the ICR cell, helps minimize this effect.



**Figure 1.3.** (a) Simulated transient of  $C_{60}$  signal with asymmetric overloading of the ADC. (b) Simulated mass spectrum with 1 Da spacing artifacts due to intermodulation of the  $C_{60}$  isotope signals.

### 1.3.3 Instrumentation

A schematic diagram of a commercial FT-ICR mass spectrometer is shown in Figure 1.4.



**Figure 1.4.** Configuration of a commercial 7-T Q-FT-ICR mass spectrometer.

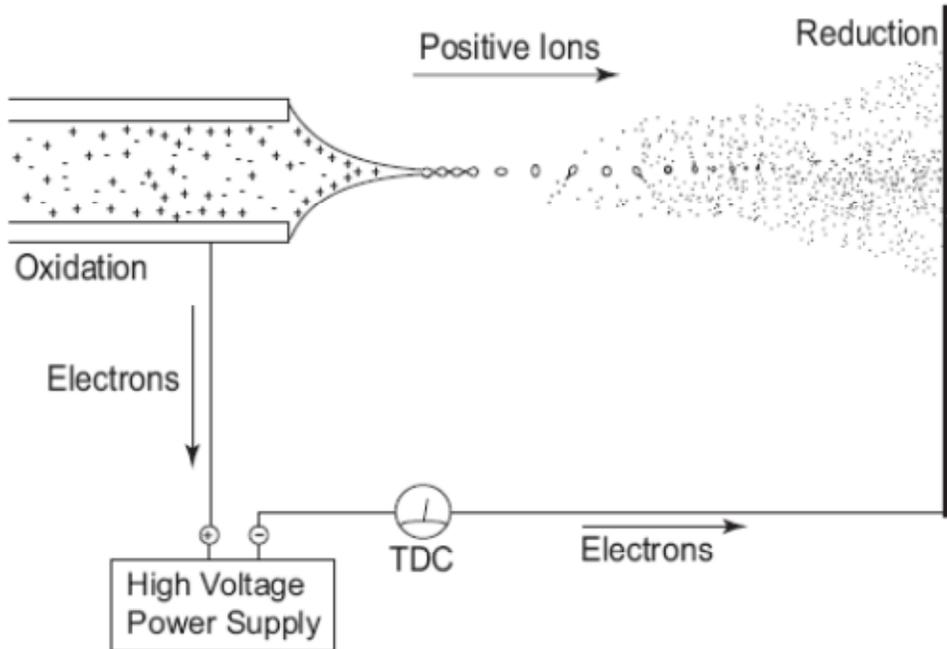
Basic components include an electrospray ionization source, a quadrupole (Q), a hexapole (collision cell), ion optics, and the ICR cell. The ESI source principles are discussed in section 1.2 below. The quadrupole can be used to allow only ions within a certain  $m/z$  window to pass. The hexapole collision cell is used for external accumulation of ions, and for performing gas-phase reactions such as CAD, electron transfer dissociation (ETD), and hydrogen-deuterium exchange. Transfer optics is used to transport ions through the magnetic field gradient into the ICR cell. Inside the ICR cell ions can be subjected to activation such as sustained off-resonance irradiation (SORI) CAD, IRMPD, ECD, and EDD. These tandem mass spectrometric techniques are discussed in detail in the following sections. The ICR cell is also used for ion excitation and detection, as described in the previous section.

## 1.4 Electrospray Ionization (ESI)

### 1.4.1 Principle and Implementation

Mass spectrometry measures mass-to-charge ratio in the gas phase. However, most biomolecules are soluble in aqueous phase and not volatile. Ionization techniques are necessary to bring biomolecules from solution phase to the gas phase, while preserving the native structure as much as possible. ESI<sup>41</sup> and MALDI<sup>40, 91</sup> are typical “soft” ionization methods. Through ESI, large biomolecules and even non-covalent complexes can be brought into the gas phase in intact form.<sup>92</sup>

Figure 1.5 shows a typical ESI apparatus.



**Figure 1.5.** Schematic description of ESI.<sup>93</sup>

During positive ion mode ESI, in which analytes typically carry positive charge in solution, a solution containing analytes is forced through a needle-shaped tip. The solution near the tip is connected to a highly positive voltage, usually several thousand volts. Due to the repulsion between positively charged species and due to the electric

field gradient, bulk solution exiting the emitter overcomes the surface tension, forms a Taylor cone,<sup>94</sup> and breaks into positively charged droplets. Negative charges move towards the inner wall of the emitter and are neutralized. Excessive negative charges (electrons) flow to the power supply circuit. Positively charged droplets fly towards the inlet of the mass spectrometer driven by the electric field and the pressure gradient. With the help of heated drying gas or a heated inlet, droplets desolvate and shrink. When Coulomb repulsion between charges on the surface of a droplet exceeds the surface tension (Rayleigh charge limit<sup>95-97</sup>), the droplet splits into smaller droplets. Eventually solvent completely evaporates and the charge carrying non-volatile species from the original spray solution are collected by the inlet of the mass spectrometer.

The mechanism behind ion generation in ESI is debated.<sup>98,99</sup> There are two main models: the ion evaporation model<sup>100</sup> and the charged residue model.<sup>101</sup> According to the ion evaporation model, charges are distributed primarily on the droplet surfaces and analyte ions on the surface may be ejected from the droplet into the gas phase due to field desorption. The charged residue model states that non-volatile species remain inside the droplets during the repetitive Coulomb fission process. The non-volatile species are then forced to carry charges in the gas phase when solvent is completely evaporated.

If the spray solution contains water, hydrophobic analytes are distributed more on the surface of the droplets due to solubility.<sup>102</sup> Hydrophilic molecules prefer the center of the droplets and have less chance to be ejected out of the droplets, and thus lower ionization efficiency is observed for such species.<sup>103</sup>

#### **1.4.2 Characteristics of ESI**

First, ESI is considered to be a “soft” ionization technique preserving the native structure of analytes. It has been shown to preserve not only covalent bonds but also non-covalent bonds, such as large protein-protein complexes<sup>104</sup> and DNA duplexes.<sup>54</sup> The “soft” nature mainly comes from the fact that most ionization energy is dispersed by solvent evaporation before analyte decomposition occurs<sup>41</sup> and there is no sudden energy change for the analyte molecules during the ESI process.

Second, ESI offers multiple charging. Multiple charging brings the  $m/z$  ratio of large molecules below the upper  $m/z$  limit thus making detection of large molecules possible, for example, whole TMV virus at 39-42 MDa.<sup>105</sup> Multiple charging also enhances the signal abundance in FT-ICR-MS, as the image current is directly proportional to the number of charges on the analyte ion. For tandem MS (MS/MS), multiple charging of precursor ions often improves fragmentation efficiency.<sup>29, 106-108</sup>

### **1.5 Tandem Mass Spectrometry (MS/MS, MS<sup>n</sup>)**

Measuring accurate masses helps identify analytes. However, due to the presence of isomers, in many cases mass alone cannot confirm the identity of an analyte. Also, information regarding the primary structures of analytes is of interest in many cases. In tandem mass spectrometry, ions of a selected  $m/z$  ratio are fragmented in the mass spectrometer. Based on the fragmentation chemistry of a particular MS/MS activation method and the fragments observed, additional structural information can be obtained.

Depending on the application, the desired MS/MS fragmentation pattern varies. In order to obtain sequence information for biomolecules composed of similar building blocks, the appropriate MS/MS technique should generate cleavages between each

building block specifically. For example, for a peptide, the proper MS/MS technique should be able to provide cleavages between each amino acid. In order to probe a certain property of a molecule, MS/MS should be able to provide specific information regarding that property. For example, in order to examine if a peptide contains phosphorylated serine or threonine, the MS/MS technique should generate cleavages directly related to the phosphate.

Among various tandem mass spectrometric activation techniques, there are two common classes: vibrational activation and radical-driven reactions. Vibrational activation includes, e.g., CAD<sup>109</sup> or collision-induced dissociation (CID), IRMPD,<sup>110, 111</sup> blackbody infrared radiative dissociation (BIRD),<sup>112, 113</sup> nozzle-skimmer fragmentation,<sup>114</sup> and multipole storage assisted dissociation (MSAD).<sup>115, 116</sup> Ion-electron reactions include ECD,<sup>28</sup> EDD,<sup>117</sup> hot ECD,<sup>118</sup> electron-induced dissociation (EID),<sup>119, 120</sup> and negative ion electron capture dissociation (niECD).<sup>121</sup> Ion-ion reactions include ETD<sup>29</sup> and negative electron transfer dissociation (NETD).<sup>122</sup> There are many other types of tandem mass spectrometric activation techniques widely studied and applied, e.g., ultraviolet photodissociation (UPVD).<sup>123</sup> CAD, IRMPD, ECD, EDD, and niECD are discussed in the following sections.

### **1.5.1 Precursor Ion Selection**

MS/MS can be performed on ions of various  $m/z$  ratios or, more commonly, one particular  $m/z$  of interest. In the latter case, a precursor ion selection step prior to activation/dissociation is required. Precursor ion selection can be implemented in a variety of ways:

In the FT-ICR cell, precursor ion selection can be achieved by stored waveform inverse Fourier transform (SWIFT).<sup>124, 125</sup> In this method, a frequency domain waveform including frequencies corresponding to all unwanted  $m/z$  ratios is computed. Then this frequency domain waveform is inversely Fourier transformed to a time domain waveform. The stored time domain waveform is applied to one pair of ICR cell electrodes,<sup>126</sup> resulting in excitation of ions with cyclotron frequencies included in the applied waveform, similar to the excitation step preceding ICR detection. However, ions are excited to a greater extent to reach the electrodes and be neutralized. In theory ions with cyclotron frequencies not included in the waveform are not affected.

A quadrupole mass filter can also be used as a mass-selecting device.<sup>127</sup> In mass-resolving mode, an RF and DC combined potential is applied to the two pairs of rod electrodes. This potential defines an  $m/z$  window outside of which ions cannot undergo stable oscillatory motion between the rods, and thus will be ejected radially. Only ions with  $m/z$  ratios within the defined window can migrate through the quadrupole. A quadrupole coupled with an FT-ICR mass analyzer can function as an external mass filter.<sup>128</sup> Similar to a linear quadrupole, a quadrupole ion trap can also be used for external mass selection.<sup>129</sup>

### **1.5.2 Tandem Mass Spectrometry in Time or in Space**

In tandem in time mass spectrometry, precursor ion selection and activation/fragmentation are performed in the same space but sequential in time. Thus the cycle of selection/fragmentation can continue until ion abundance is below the detection limit, meaning that the number of cycles is limited by the starting ion

abundance. In other words,  $MS^n$  is possible where  $n$  can be a fairly large number. However, in practice  $n$  can be  $4^{130, 131}$  or  $5^{132}$  for an ICR cell, and as large as  $10^{133, 134}$  for an ion trap.

In tandem in space mass spectrometry, precursor ion selection and activation/fragmentation are performed sequentially with different mass analyzers. For example, in the configuration shown in Figure 1.4, a hexapole is located between the quadrupole and the ICR cell. This hexapole functions as an external device for gas-phase reactions and collisional activation of mass-selected ion. Fragments from CAD are transferred to the ICR cell for detection.

### **1.5.3 Collision Activated Dissociation (CAD)**

CAD or CID is a widely used technique for fragmenting gaseous ions. The basic process involves inelastic collisions between accelerated ions and neutral molecules such as argon. Depending on the kinetic energy of the accelerated ions, there are two types of CAD: high energy CAD in the kilo eV range and low energy CAD in the 1-100 eV range. High energy CAD, which mainly causes electronic excitation,<sup>135</sup> can be performed in tandem sector instruments and TOF/TOF-type instruments. Low energy CAD can be performed in multipoles, FT-ICR cell, ion trap, and ion sources close to ambient pressure. Low energy CAD excitation is mostly vibrational.<sup>136</sup> In this thesis, the acronym “CAD” refers to low energy CAD. During the low energy CAD activation process, each inelastic collision deposits a small amount of internal energy into the precursor ion. However, the energy from each collision is not sufficient to break a chemical bond but, after multiple collisions, the accumulated energy is high enough to cause bond cleavage. For this

reason, CAD is considered one of the “slow-heating” techniques. During collisional activation, the internal energy is redistributed through intramolecular vibrational-energy redistribution (IVR)<sup>137</sup>. Thus, in general, CAD and other slow-heating techniques tend to cleave the most labile bonds in a molecule.

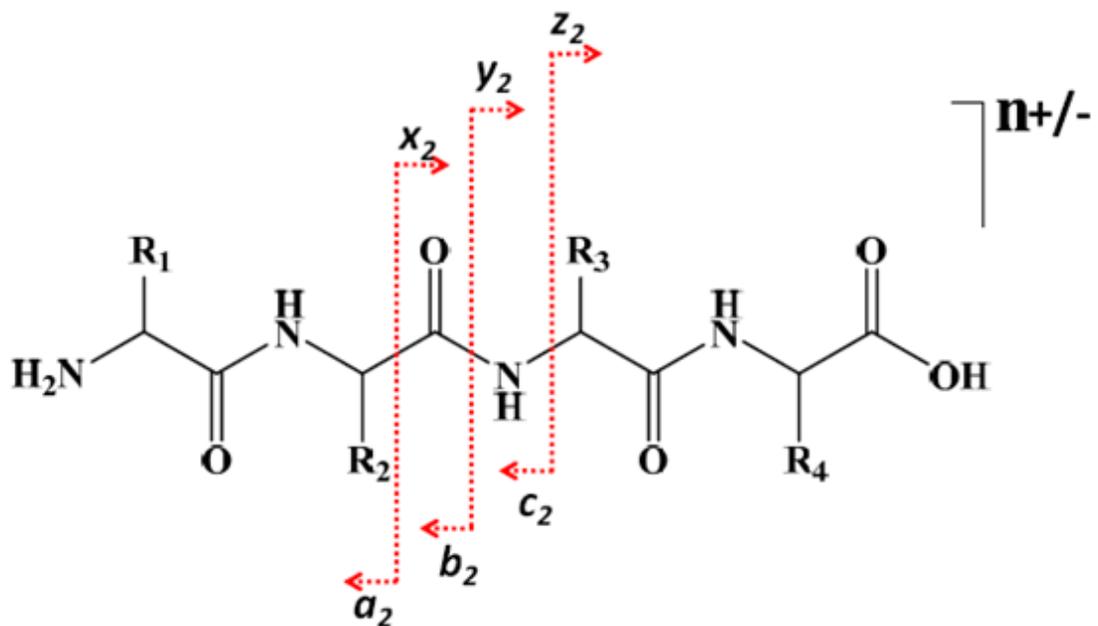
There are several methods to implement CAD:

In beam-type CAD, precursor ions are accelerated by an electric field into a chamber filled with neutral gas. In the configuration in Figure 1.4, the hexapole located in between the quadrupole and the ICR cell is filled with argon to a pressure of  $\sim 10^{-6}$  mBar. Precursor ions selected by the quadrupole are accelerated into the hexapole, collide with argon, and fragment. The radial RF electric field confines fragments inside the hexapole. Triple quadrupole and Q-TOF-type instruments also contain similar geometries for beam-type CAD.

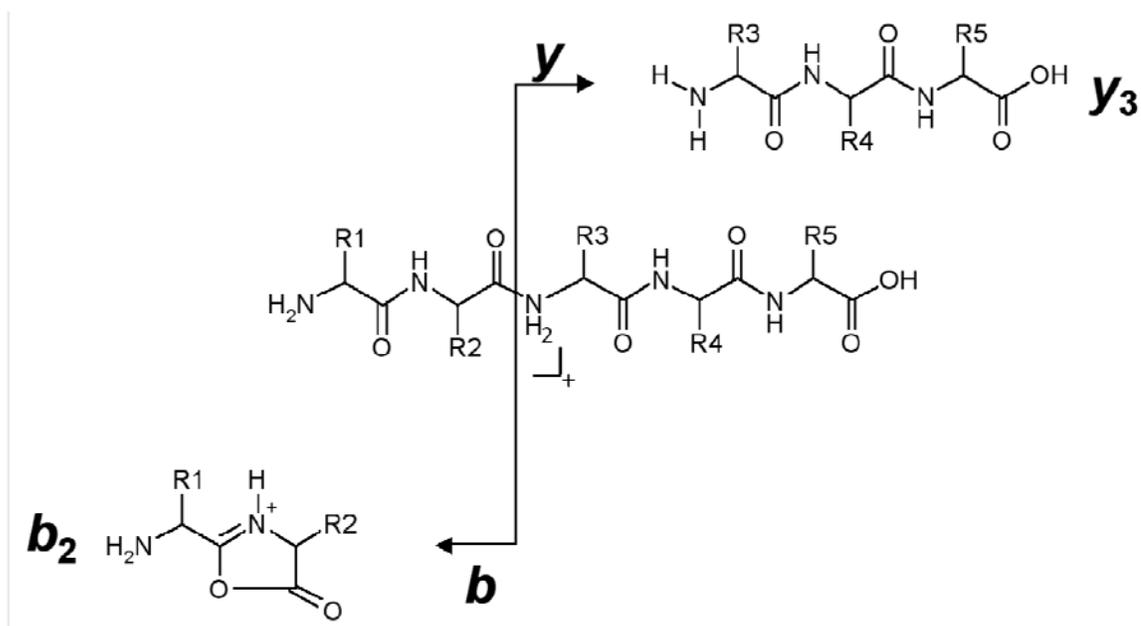
In ion trap-type CAD, precursor ions are accelerated by resonant excitation. In 3D ion trap-type instruments the RF amplitude is increased to axially excite precursor ions. The excited precursor ions collide with the background helium gas and dissociate. One difference between beam-type CAD and ion trap-type CAD is the amount of energy imparted per collision. In beam-type CAD a relatively high amount of energy is deposited per collision with an activation time of  $\sim 10^{-4}$  s. However, in ion trap CAD relatively low energy is deposited per collision and the activation time is  $\sim 10^{-1}$  to  $10^{-3}$  s, i.e., significantly slower than beam-type CAD. Ion trap CAD shows higher favorability for cleavage of the most labile bonds due to the narrower energy distribution at lower collision energy.<sup>138</sup>

SORI-CAD<sup>139</sup> is a CAD implementation in FT-ICR cells. Because the pressure in ICR cells ( $<10^{-9}$  mBar) is significantly lower than that in ion traps ( $\sim 10^{-3}$  mBar) or multipoles ( $\sim 10^{-6}$  mBar), neutral gas is pulsed in to the cell. An excitation waveform with a frequency close ( $\pm 1$  kHz) to the frequency of selected precursor ions is applied to the excitation electrodes. Precursor ions partially absorb excitation energy and collide with neutral gas as their cyclotron radii increase. However, the off-resonant nature of the excitation pulse causes precursor ions to be accelerated and decelerated in a periodic manner, thus generating product ions near the center of the ICR cell.<sup>79</sup> By contrast, on-resonance excitation only involves an acceleration step and energy uptake is therefore limited by the size of the ICR cell. Also, product ions are formed off-axis in resonant CAD, which limits further fragmentation and lowers ICR resolution.<sup>140</sup> Overall, SORI-CAD is similar to ion trap type CAD whereas resonant CAD is more similar to beam-type CAD (with the advantages and disadvantages stated above). One disadvantage of performing CAD inside the ICR cell is that collision gas needs to be leaked in. Thus, after a CAD event, several seconds are needed to restore high vacuum, which reduces the duty cycle. If the high vacuum is not restored, collisional damping will be more significant and resolution is deteriorated.

CAD is widely used for structural analysis of peptides, proteins, and nucleic acids. For unmodified peptides and proteins, CAD primarily provides cleavages at peptide amide bonds to form *b* and *y*-type ions<sup>141</sup> (Schemes 1.1-1.2). Neutral loss of small molecules such as water and ammonia also occurs.

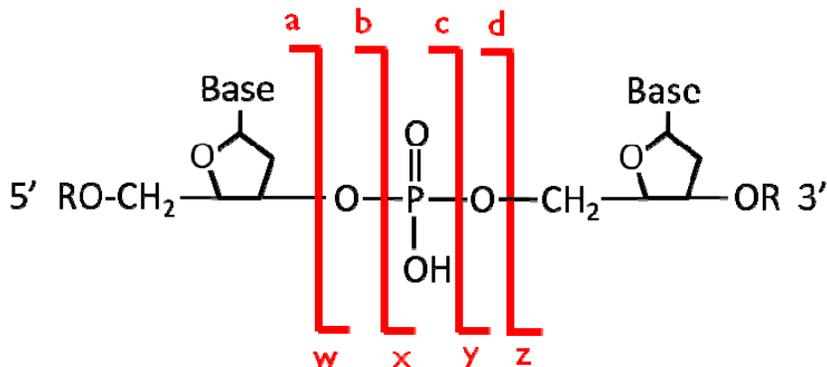


**Scheme 1.1.** Nomenclature for peptide/protein fragmentation. Adapted from Roepstorff *et al.*<sup>141</sup>

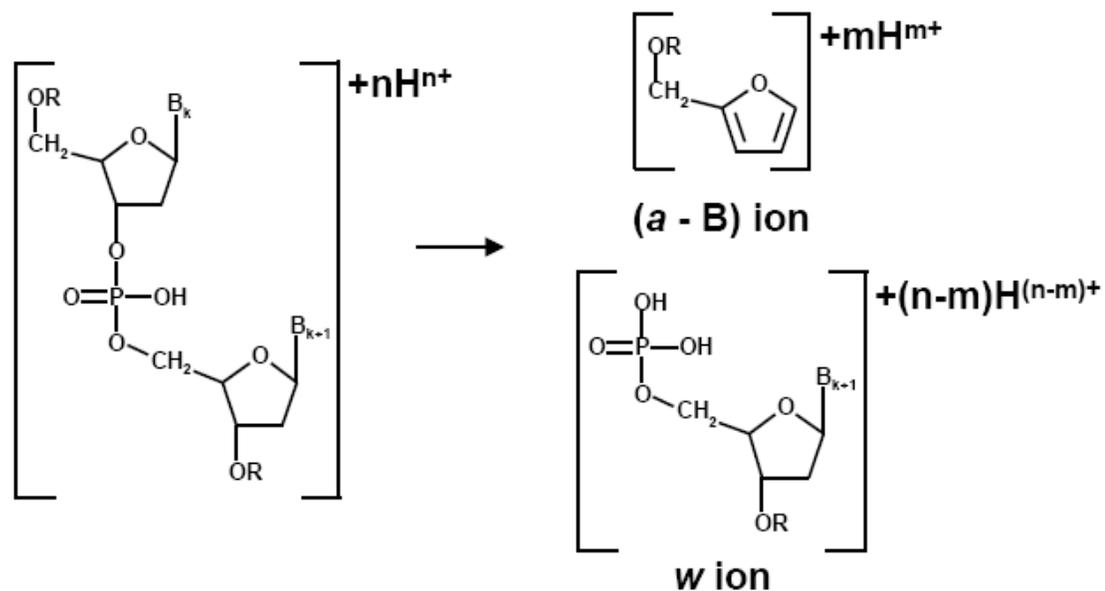


**Scheme 1.2.** Generation of  $b$  and  $y$ -type ions from CAD of peptides/proteins.<sup>79</sup>

For DNA, CAD provides primarily (*a* - base) and *w*-type ions,<sup>142, 143</sup> as shown in Schemes 1.3<sup>143</sup> and 1.4.<sup>79</sup>



**Scheme 1.3.** Nomenclature for oligonucleotide fragmentation. Adapted from McLuckey, *et al.*<sup>143</sup>



**Scheme 1.4.** (*a* - base) and *w*-type ions from CAD of DNA.<sup>79</sup>

#### 1.5.4 Infrared Multiphoton Dissociation (IRMPD)

IRMPD<sup>144</sup> is another type of slow-heating tandem mass spectrometric technique. During IRMPD, precursor ions are trapped in the ICR cell and irradiated with an IR laser

beam. Precursor ions absorb multiple photons to accumulate internal energy until they reach their dissociation threshold. Infrared photons are typically provided by a 10.6  $\mu\text{m}$   $\text{CO}_2$  laser but can also be provided from a free electron laser (FEL)<sup>145</sup> or optical parametric oscillator (OPO)<sup>146, 147</sup> systems. IRMPD has also been implemented in an external hexapole<sup>148</sup> and in quadrupole ion traps.<sup>149</sup> However, the pressure in the external hexapole or ion trap is several orders of magnitude higher than the UHV inside an ICR cell. Under high pressure vibrationally excited ions are cooled through collisions with buffer gas molecules. Thus IRMPD in ion trap devices is not as efficient as in ICR cells.

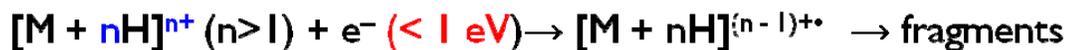
IRMPD has been applied to dissociate proteins/peptides<sup>111</sup> and also oligonucleotides.<sup>51, 52, 150, 151</sup> IRMPD provides similar fragmentation patterns to low energy CAD. One advantage of IRMPD in ICR cells over SORI-CAD is that IRMPD does not require introduction of gas into the ICR cell. Thus, no pumping stage is needed after each fragmentation event. Second, collisional damping of the time-domain signal is not affected in IRMPD so resolution and calibration are not worsened.

IR irradiation inside the ICR cell can also be used as an activation method rather than a dissociation method. IR irradiation preferentially breaks the weakest bonds within a molecule, such as non-covalent interactions.<sup>152, 153</sup>

### **1.5.5 Electron Capture Dissociation (ECD)**

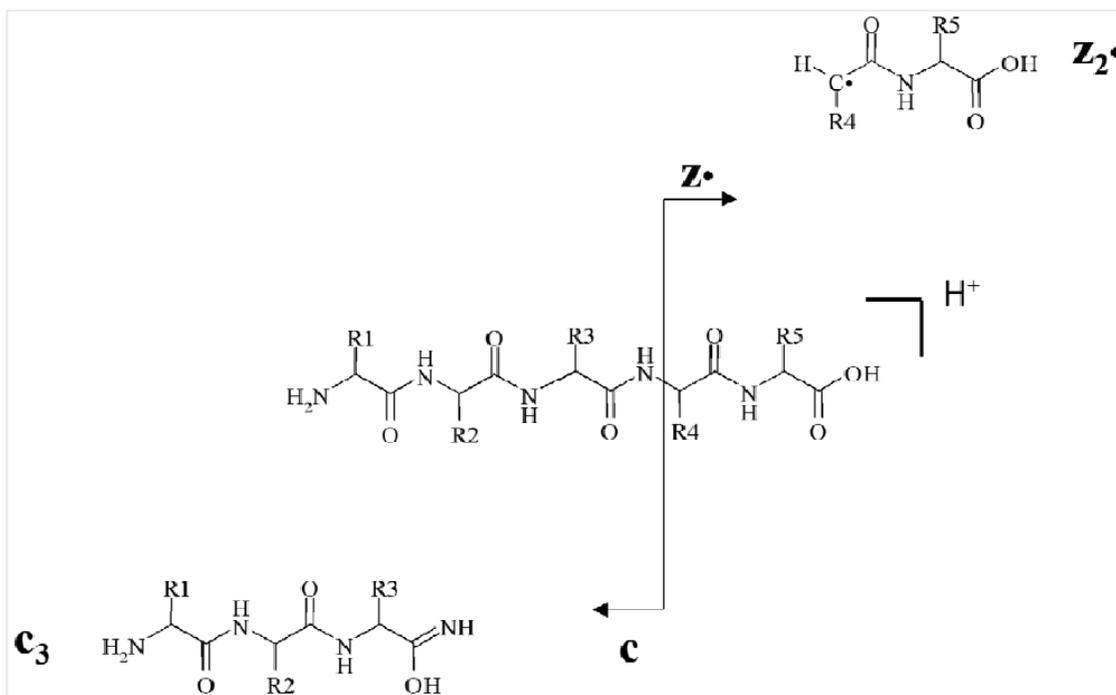
ECD is an MS/MS technique based on gas-phase ion-electron reactions. It was first applied to dissociate peptides/proteins.<sup>28</sup> ECD is performed by irradiating multiply charged precursor cations with low energy (<1 eV) electrons. Precursor cations capture electrons and form meta-stable radical intermediates that dissociate into fragments, as

shown in Scheme 1.5. For larger molecules, IR activation prior to or after electron irradiation is required to improve dissociation efficiency. Typically ECD is performed in ICR cells. Ion traps can be modified to perform ECD.<sup>154, 155</sup> However, such implementation is still challenging because electrons have masses well below the low-mass cut-off of typical ion traps designed to trap molecular analytes.



**Scheme 1.5.** Fragmentation pathway in ECD.

For peptides/proteins, ECD preferentially cleaves N-C<sub>α</sub> bonds to produce *c* and *z*<sup>•</sup>-type ions,<sup>28</sup> as shown in Scheme 1.6.



**Scheme 1.6.** Fragments from peptide ECD: *c* and *z*<sup>•</sup>-type ions.<sup>79</sup>

The mechanism of ECD is still under debate. The “hot hydrogen” model<sup>2, 156, 157</sup> proposes that electron capture occurs at protonated sites (i.e., N-terminus or basic side

chains). The proton is neutralized to hydrogen and migrates to a nearby backbone carbonyl. The weakened backbone N-C<sub>α</sub> bond dissociates to form *c* and z•-type ions. The “amide super base” model<sup>158, 159</sup> proposes that electron capture occurs into the LUMO of the carbonyl double bond. *Ab initio* calculations revealed that electron attachment and subsequent N-C<sub>α</sub> bond cleavage are thermodynamically favored with a stabilizing Coulomb potential.<sup>159</sup> In this model the amide backbone directly captures an electron, facilitated by positive charges in spatial proximity. After electron capture the excited carbonyl double bond is highly basic (“superbase”) and attracts a proton, explaining the hydrogen transfer observed based on the fragment masses.

ECD has been proposed to be a nonergodic process.<sup>2, 28, 157</sup> However, the detailed mechanism is still under debate. It has been stated that, in ECD of cytochrome *c*, dissociation of the radical intermediate generated by electron capture, is faster than unfolding via IVR and faster than hydrogen-deuterium scrambling.<sup>28</sup> Compared with slow-heating activation approaches, ECD provides more extensive backbone cleavages.<sup>160-162</sup>

The radical-driven nature of ECD also provides complementary backbone bond cleavages for peptides/proteins compared to slow-heating techniques. For example, ECD frequently retains labile post-translational modifications (PTMs) such as phosphorylation,<sup>31, 163</sup> sulfonation,<sup>164</sup> and glycosylation<sup>165-167</sup>. By contrast, PTMs are preferentially lost in slow-heating methods such as CAD and IRMPD. This characteristic of ECD allows direct localization of PTM sites within peptides/proteins. ECD has also been applied for characterizing the structure of intact proteins<sup>168</sup> and protein complexes.<sup>169</sup> ECD effectively cleaves backbone amine bonds but does not disrupt non-

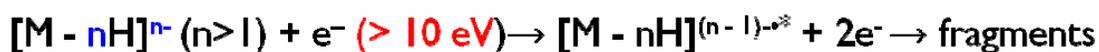
covalent bonds. Gas-phase protein structure contains intramolecular non-covalent interactions such as charge-charge interactions and hydrogen bonding. Thus, fragments from ECD may still be connected with non-covalent bonds such that the total mass does not change. To observe ECD of proteins, additional activation, e.g., IR irradiation, is often required. Consequently, ECD of proteins without additional activation may result in low fragmentation efficiency, which could be improved by additional activation such as IR irradiation or collisional activation.<sup>170-172</sup>

Nucleic acids contain an acidic phosphate backbone and have an overall pI around 2. Unless the electrospray solution is highly acidic, the phosphate backbone tends to be negatively charged in the gas phase. Thus detection of nucleic acids is more efficient in negative ion mode. However, nucleobases are basic and can be protonated both in solution and in the gas phase. ECD of oligonucleotides has been demonstrated to yield primarily *a/z* and *d/w*-type ions, as well as *c/x*-type ions.<sup>52, 173</sup> Due to the sequence symmetry of the investigated oligonucleotides, *a*- and *z*-ions have the same mass and thus cannot be differentiated in the mass spectra. Similar ambiguity is also present in *b/y*, *c/x* and *d/w*-type ions.

Recently, negative ion electron capture dissociation (niECD) has been demonstrated to fragment peptides in the gas phase in a manner analogous to ECD in positive ion mode,<sup>121</sup> however, niECD operates in negative ion mode. In niECD, peptide anions capture 3.5-6.5 eV electrons to form charge-increased radicals that fragment to *c/z*-type ions.

### **1.5.6 Electron Detachment Dissociation (EDD)**

ECD requires multiply charged cations as precursor ions. EDD is a negative ion mode ion-electron reaction in which multiply charged anions are irradiated with electrons of energy higher than 10 eV. This higher electron energy (compared to 3.5-6.5 eV in niECD) ejects an electron from the precursor anions to form an electron-deficient odd-electron intermediate different from the electron-rich intermediate formed in niECD. This radical intermediate further dissociates, as shown in Scheme 1.7. EDD has been implemented in FT-ICR and quadrupole ion trap instruments.<sup>174</sup> For peptides, the major product ions in EDD are *a*• and *x*-type ions, as well as abundant CO<sub>2</sub> loss.<sup>117</sup> EDD has shown the ability to retain PTMs such as phosphorylation<sup>175</sup> and sulfonation.<sup>117</sup>



**Scheme 1.7.** Fragmentation route in EDD.

EDD has also been demonstrated to efficiently dissociate oligonucleotides.<sup>53, 176</sup> For oligodeoxyribonucleotides, EDD primarily provides *a*, *c*, *d* and *w*-type ions whereas, for oligoribonucleotides, primarily *d* and *w*-type ions are observed. As a radical-driven process, EDD provides complementary sequence information compared to slow-heating techniques such as CAD and IRMPD. In contrast to CAD, nucleobase loss is not favored in EDD, as observed after carefully minimizing collisional activation during the experiment.<sup>177</sup>

## 1.6 Dissertation Overview

This dissertation research is on mass spectrometric method development for detection and characterization of biomacromolecules, with a particular focus on acidic peptides, proteins, and nucleic acids. This chapter provided a broad overview of basic

concepts related to ESI-FT-ICR-MS and its applications. The following chapters are grouped to answer three general questions from methodology development in mass spectrometry.

Chapters 2-5 demonstrate how to extend existing mass spectrometric techniques to molecules not previously studied by these techniques. In Chapter 2, various MS/MS techniques are evaluated for sequencing peptide nucleic acids (PNAs). In Chapter 3, various MS/MS techniques are applied to large RNAs and RNA-ligand complexes for obtaining sequence information and ligand binding sites. In Chapter 4, negative ion mode MS/MS techniques are applied to acidic proteins. Sequence information obtained from negative ion mode is compared to positive ion mode analysis. In Chapter 5, a recently developed MS/MS technique is evaluated for sequencing oligonucleotides. Chapter 6 demonstrates how to improve new mass spectrometric techniques. In this Chapter, niECD, an MS/MS technique recently developed in our laboratory, is modified to improve fragmentation efficiency and practical utility.

Appendix A demonstrates how to apply mass spectrometric techniques to solve practical problems from biology through collaborative efforts. Through collaboration with Dr. Xiaoxia (Nina) Lin and Dr. Marjan Varedi, workflows for bottom-up and top-down approaches are established to understand the regulatory activity of the yeast protein Sic1.

The Chapters in this dissertation are prepared in the format of journal articles that will be submitted for publication.

## 1.7 Bibliography

- (1) Wilkins, M. R.; Pasquali, C.; Appel, R. D.; Ou, K.; Golaz, O.; Sanchez, J.-C.; Yan, J. X.; Gooley, A. A.; Hughes, G.; Humphery-Smith, I.; Williams, K. L.; Hochstrasser, D. F. *Nat Biotech* **1996**, *14*, 61-65.
- (2) 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.
- (3) Human Genome Sequencing, C. *Nature* **2004**, *431*, 931-945.
- (4) Swiss Institute of Bioinformatics (SIB); Geneva, S.; European Bioinformatics Institute (EBI); Hinxton, U. K.; Protein Information Resource (PIR); Washington DC, U., 27-Jul-2011 ed.
- (5) Stumpf, M. P. H.; Thorne, T.; de Silva, E.; Stewart, R.; An, H. J.; Lappe, M.; Wiuf, C. *Proc. Natl. Acad. Sci.* **2008**, *105*, 6959-6964.
- (6) Pandey, A.; Mann, M. *Nature* **2000**, *405*, 837-846.
- (7) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198-207.
- (8) Valaskovic, G. A.; Kelleher, N. L.; Little, D. P.; Aaserud, D. J.; McLafferty, F. W. *Anal. Chem.* **1995**, *67*, 3802-3805.
- (9) Valaskovic, G. A.; Kelleher, N. L.; McLafferty, F. W. *Science* **1996**, *273*, 1199-1202.
- (10) Andren, P.; Emmett, M.; Caprioli, R. *J. Am. Soc. Mass. Spectrom.* **1994**, *5*, 867-869.
- (11) Schaub, T. M.; Hendrickson, C. L.; Horning, S.; Quinn, J. P.; Senko, M. W.; Marshall, A. G. *Anal. Chem.* **2008**, *80*, 3985-3990.
- (12) Shen, Y.; Smith, R. D.; Unger, K. K.; Kumar, D.; Lubda, D. *Anal. Chem.* **2005**, *77*, 6692-6701.
- (13) Chait, B. T. *Science* **2006**, *314*, 65-66.
- (14) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. *Nat Biotech* **1999**, *17*, 676-682.
- (15) Frahm, J. L.; Howard, B. E.; Heber, S.; Muddiman, D. C. *J. Mass Spectrom.* **2006**, *41*, 281-288.

- (16) Jorgenson, J. W. *Annual Review of Analytical Chemistry* **2010**, *3*, 129-150.
- (17) Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Rochon, Y.; Aebersold, R. *Proc. Natl. Acad. Sci.* **2000**, *97*, 9390-9395.
- (18) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* **1999**, *20*, 3551-3567.
- (19) Eng, J.; McCormack, A.; Yates, J. *J. Am. Soc. Mass. Spectrom.* **1994**, *5*, 976-989.
- (20) Elias, J. E.; Gygi, S. P. *Nat Meth* **2007**, *4*, 207-214.
- (21) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994-999.
- (22) Ong, S.-E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. *Mol. Cell. Proteomics* **2002**, *1*, 376-386.
- (23) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlett-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. *Mol. Cell. Proteomics* **2004**, *3*, 1154-1169.
- (24) Washburn, M. P.; Wolters, D.; Yates, J. R. *Nat Biotech* **2001**, *19*, 242-247.
- (25) Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. *Proc. Natl. Acad. Sci.* **2000**, *97*, 5802-5806.
- (26) Sinz, A. *Mass Spectrom. Rev.* **2006**, *25*, 663-682.
- (27) Sinz, A. *J. Mass Spectrom.* **2003**, *38*, 1225-1237.
- (28) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
- (29) 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.
- (30) Pesavento, J. J.; Kim, Y. B.; Taylor, G. K.; Kelleher, N. L. *J. Am. Chem. Soc.* **2004**, *126*, 3386-3387.
- (31) Shi, S. D. H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, *73*, 19-22.
- (32) McLafferty, F. W.; Guan, Z. Q.; Haupts, U.; Wood, T. D.; Kelleher, N. L. *J. Am. Chem. Soc.* **1998**, *120*, 4732-4740.
- (33) Powell, K. D.; Ghaemmaghami, S.; Wang, M. Z.; Ma, L.; Oas, T. G.; Fitzgerald, M. C. *J. Am. Chem. Soc.* **2002**, *124*, 10256-10257.

- (34) Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Int. J. Mass spectrom.* **1999**, *187*, 565-575.
- (35) Breuker, K.; Oh, H. B.; Horn, D. M.; Cerda, B. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124*, 6407-6420.
- (36) Sanger, F.; Nicklen, S.; Coulson, A. R. *Proc. Natl. Acad. Sci.* **1977**, *74*, 5463-5467.
- (37) Pomerantz, S. C.; McCloskey, J. A.; James, A. M. In *Methods Enzymol.*; Academic Press, 1990; Vol. Volume 193, pp 796-824.
- (38) Jacobson, K. B.; Arlinghaus, H. F.; Buchanan, M. V.; Chen, C. H.; Glish, G. L.; Hettich, R. L.; McLuckey, S. A. *Genet. Anal. - Biomol. Eng.* **1991**, *8*, 223-229.
- (39) Karas, M.; Bachmann, D.; Hillenkamp, F. *Anal. Chem.* **1985**, *57*, 2935-2939.
- (40) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-153.
- (41) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.
- (42) Williams, D. H.; Bradley, C.; Bojesen, G.; Santikarn, S.; Taylor, L. C. E. *J. Am. Chem. Soc.* **1981**, *103*, 5700-5704.
- (43) Grotjahn, L.; Frank, R.; Blocker, H. *Nucleic Acids Res.* **1982**, *10*, 4671-4678.
- (44) Grotjahn, L.; Blöcker, H.; Frank, R. *Biol. Mass Spectrom.* **1985**, *12*, 514-524.
- (45) Karas, M.; Bahr, U. *TrAC, Trends Anal. Chem.* **1990**, *9*, 321-325.
- (46) Nordhoff, E.; Ingendoh, A.; Cramer, R.; Overberg, A.; Stahl, B.; Karas, M.; Hillenkamp, F.; Crain, P. F.; Chait, B. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 771-776.
- (47) Smith, R. D.; Loo, J. A.; Edmonds, C. G.; Barinaga, C. J.; Udseth, H. R. *Anal. Chem.* **1990**, *62*, 882-899.
- (48) Muddiman, D. C.; Null, A. P.; Hannis, J. C. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1201-1204.
- (49) Schultz, J. C.; Hack, C. A.; Benner, W. H. *J. Am. Soc. Mass. Spectrom.* **1998**, *9*, 305-313.
- (50) Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1994**, *116*, 4893-4897.

- (51) Little, D. P.; Aaserud, D. J.; Valaskovic, G. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **1996**, *118*, 9352-9359.
- (52) Hakansson, K.; Hudgins, R. R.; Marshall, A. G.; O'Hair, R. A. J. *J. Am. Soc. Mass. Spectrom.* **2003**, *14*, 23-41.
- (53) Yang, J.; Mo, J. J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, *77*, 1876-1882.
- (54) Light-Wahl, K. J.; Springer, D. L.; Winger, B. E.; Edmonds, C. G.; Camp, D. G.; Thrall, B. D.; Smith, R. D. *J. Am. Chem. Soc.* **1993**, *115*, 803-804.
- (55) Doktycz, M. J.; Habibi-Goudarzi, S.; McLuckey, S. A. *Anal. Chem.* **1994**, *66*, 3416-3422.
- (56) Aaserud, D. J.; Kelleher, N. L.; Little, D. P.; McLafferty, F. W. *J. Am. Soc. Mass. Spectrom.* **1996**, *7*, 1266-1269.
- (57) Schnier, P. D.; Klassen, J. S.; Strittmatter, E. E.; Williams, E. R. *J. Am. Chem. Soc.* **1998**, *120*, 9605-9613.
- (58) Gale, D. C.; Goodlett, D. R.; Light-Wahl, K. J.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 6027-6028.
- (59) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 10129-10133.
- (60) Hofstadler, S. A.; Sannes-Lowery, K. A.; Crooke, S. T.; Ecker, D. J.; Sasmor, H.; Manalili, S.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 3436-3440.
- (61) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.
- (62) Cheng, X.; Harms, A. C.; Goudreau, P. N.; Terwilliger, T. C.; Smith, R. D. *Proc. Natl. Acad. Sci.* **1996**, *93*, 7022-7027.
- (63) Hagan, N.; Fabris, D. *Biochemistry* **2003**, *42*, 10736-10745.
- (64) Robinson, J. M.; Greig, M. J.; Griffey, R. H.; Mohan, V.; Laude, D. A. *Anal. Chem.* **1998**, *70*, 3566-3571.
- (65) Freitas, M. A.; Shi, S. D. H.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Chem. Soc.* **1998**, *120*, 10187-10193.
- (66) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *J. Mass Spectrom.* **2000**, *35*, 62-70.
- (67) Stemmler, E. A.; Buchanan, M. V.; Hurst, G. B.; Hettich, R. L. *Anal. Chem.* **1995**, *67*, 2924-2930.

- (68) Limbach, P. A.; Crain, P. F.; McCloskey, J. A. *J. Am. Soc. Mass. Spectrom.* **1995**, *6*, 27-39.
- (69) Hipple, J. A.; Sommer, H.; Thomas, H. A. *Physical Review* **1949**, *76*, 1877-1878.
- (70) Sommer, H.; Thomas, H. A.; Hipple, J. A. *Physical Review* **1950**, *80*, 487-487.
- (71) Sommer, H.; Thomas, H. A.; Hipple, J. A. *Physical Review* **1951**, *82*, 697-702.
- (72) Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *25*, 282-283.
- (73) Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *26*, 489-490.
- (74) Marshall, A. G.; Grosshans, P. B. *Anal. Chem.* **1991**, *63*, A215-A229.
- (75) Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325-1337.
- (76) Grosshans, P. B.; Marshall, A. G. *Int. J. Mass Spectrom. Ion Processes* **1990**, *100*, 347-379.
- (77) Allen, J. S. *Rev. Sci. Instrum.* **1947**, *18*, 739-749.
- (78) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.
- (79) Hakansson, K.; Cooper, H. J.; Hudgins, R. R.; Nilsson, C. L. *Curr. Org. Chem.* **2003**, *7*, 1503-1525.
- (80) Makarov, A. *Anal. Chem.* **2000**, *72*, 1156-1162.
- (81) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. *J. Mass Spectrom.* **2005**, *40*, 430-443.
- (82) Kingdon, K. H. *Physical Review* **1923**, *21*, 408.
- (83) Wang, M. D.; Marshall, A. G. *Anal. Chem.* **1990**, *62*, 515-520.
- (84) Brustkern, A. M.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 1281-1285.
- (85) Bruce, J. E.; Anderson, G. A.; Lin, C. Y.; Gorshkov, M.; Rockwood, A. L.; Smith, R. D. *J. Mass Spectrom.* **2000**, *35*, 85-94.
- (86) Nikolaev, E. N.; Boldin, I. A.; Jertz, R.; Baykut, G. *J. Am. Soc. Mass. Spectrom.* **2011**, *22*, 1125-1133.
- (87) Hendrickson, C. L.; Beu, S. C.; Laude, D. A. *J. Am. Soc. Mass. Spectrom.* **1993**, *4*, 909-916.

- (88) Belov, M. E.; Zhang, R.; Strittmatter, E. F.; Prior, D. C.; Tang, K.; Smith, R. D. *Anal. Chem.* **2003**, *75*, 4195-4205.
- (89) Hendrickson, C. L.; Beu, S. C.; Blakney, G. T.; Kaiser, N. K.; McIntosh, D. G.; Quinn, J. P.; Marshall, A. G. Series, *Proceedings of the 57th American Society for Mass Spectrometry Conference on Mass Spectrometry and Allied Topics*, Philadelphia, PA, May 31–June 4, 2009 2009.
- (90) Mathur, R.; O'Connor, P. B. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 523-529.
- (91) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.
- (92) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.
- (93) Kebarle, P.; Verkerk, U. H. *Mass Spectrom. Rev.* **2009**, *28*, 898-917.
- (94) Taylor, G. I. *Proc. Roy. Soc.* **1964**, 383.
- (95) Rayleigh, L. *Philosophical Magazine* **1882**, *14*, 184.
- (96) Taflin, D. C.; Ward, T. L.; Davis, E. J. *Langmuir* **1989**, *5*, 376-384.
- (97) Li, K.-Y.; Tu, H.; Ray, A. K. *Langmuir* **2005**, *21*, 3786-3794.
- (98) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Mass Spectrom. Rev.* **1990**, *9*, 37-70.
- (99) Smith, R. D.; Loo, J. A.; Loo, R. R. O.; Busman, M.; Udseth, H. R. *Mass Spectrom. Rev.* **1991**, *10*, 359-451.
- (100) Iribarne, J. V.; Thomson, B. A. *On the evaporation of small ions from charged droplets*; AIP, 1976.
- (101) Malcolm, D.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *Molecular Beams of Macroions*; AIP, 1968.
- (102) Fenn, J. J. *Am. Soc. Mass. Spectrom.* **1993**, *4*, 524-535.
- (103) Cech, N. B.; Enke, C. G. *Anal. Chem.* **2000**, *72*, 2717-2723.
- (104) van Duijn, E.; Bakkes, P. J.; Heeren, R. M. A.; van den Heuvel, R. H. H.; van Heerikhuizen, H.; van der Vies, S. M.; Heck, A. J. R. *Nat Meth* **2005**, *2*, 371-376.
- (105) Bothner, B.; Siuzdak, G. *ChemBioChem* **2004**, *5*, 258-260.
- (106) Jockusch, R. A.; Schnier, P. D.; Price, W. D.; Strittmatter, E. F.; Demirev, P. A.; Williams, E. R. *Anal. Chem.* **1997**, *69*, 1119-1126.
- (107) Kalli, A.; Hakansson, K. *J. Proteome. Res.* **2008**, *7*, 2834-2844.

- (108) Iavarone, A. T.; Paech, K.; Williams, E. R. *Anal. Chem.* **2004**, *76*, 2231-2238.
- (109) Hayes, R. N.; Gross, M. L. *Methods Enzymol.* **1990**, *193*, 237-263.
- (110) Woodin, R. L.; Bomse, D. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1978**, *100*, 3248-3250.
- (111) Little, D. P.; Speir, J. P.; Senko, M. W.; Oconnor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.
- (112) Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859-866.
- (113) Dunbar, R. C.; McMahon, T. B. *Science* **1998**, *279*, 194-197.
- (114) Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. *Anal. Chim. Acta* **1990**, *241*, 167-173.
- (115) Sannes-Lowery, K.; Griffey, R. H.; Kruppa, G. H.; Speir, J. P.; Hofstadler, S. A. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1957-1961.
- (116) Hakansson, K.; Axelsson, J.; Palmblad, M.; Hakansson, P. *J. Am. Soc. Mass. Spectrom.* **2000**, *11*, 210-217.
- (117) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.
- (118) Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. *Chem. Phys. Lett.* **2002**, *356*, 201-206.
- (119) Cody, R. B.; Freiser, B. S. *Anal. Chem.* **1979**, *51*, 547-551.
- (120) Budnik, B. A.; Haselmann, K. F.; Elkin, Y. N.; Gorbach, V. I.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 5994-6001.
- (121) Yoo, H. J.; Wang, N.; Zhuang, S.; Song, H.; Hakansson, K. *J. Am. Chem. Soc.* **2011**, *133*, 16790-16793.
- (122) Coon, J. J.; Shabanowitz, J.; Hunt, D. F.; Syka, J. E. P. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 880-882.
- (123) Bowers, W. D.; Delbert, S. S.; Hunter, R. L.; McIver, R. T. *J. Am. Chem. Soc.* **1984**, *106*, 7288-7289.
- (124) Wang, T. C. L.; Ricca, T. L.; Marshall, A. G. *Anal. Chem.* **1986**, *58*, 2935-2938.
- (125) Guan, S. H.; Marshall, A. G. *Int. J. Mass spectrom.* **1996**, *157*, 5-37.
- (126) Chen, L.; Marshall, A. G. *Int. J. Mass Spectrom. Ion Processes* **1987**, *79*, 115-125.

- (127) Paul, W.; Steinwedel, H. *Zeitschrift Naturforschung Teil A* **1953**, *8*, 448.
- (128) Belov, M. E.; Nikolaev, E. N.; Anderson, G. A.; Udseth, H. R.; Conrads, T. P.; Veenstra, T. D.; Masselon, C. D.; Gorshkov, M. V.; Smith, R. D. *Anal. Chem.* **2001**, *73*, 253-261.
- (129) Syka, J. E. P.; Marto, J. A.; Bai, D. L.; Horning, S.; Senko, M. W.; Schwartz, J. C.; Ueberheide, B.; Garcia, B.; Busby, S.; Muratore, T.; Shabanowitz, J.; Hunt, D. F. *J. Proteome. Res.* **2004**, *3*, 621-626.
- (130) Huang, Y.; Pasa-Tolic, L.; Guan, S.; Marshall, A. G. *Anal. Chem.* **1994**, *66*, 4385-4389.
- (131) Zhang, J.; Schubothe, K.; Li, B.; Russell, S.; Lebrilla, C. B. *Anal. Chem.* **2004**, *77*, 208-214.
- (132) Wu, Q. *Anal. Chem.* **1998**, *70*, 865-872.
- (133) Ngoka, L. C. M.; Gross, M. L. *J. Am. Soc. Mass. Spectrom.* **1999**, *10*, 732-746.
- (134) Zucker, S.; Lee, S.; Webber, N.; Valentine, S.; Reilly, J.; Clemmer, D. *J. Am. Soc. Mass. Spectrom.* **2011**, *22*, 1477-1485.
- (135) Hitoshi, Y.; Pham, D.; Jean, D. *Energetics of the Collision-Induced Dissociations  $C_2H_2 \rightarrow C_2H + H$  and  $C_2H_2 \rightarrow H + C_2H$* ; AIP, 1969.
- (136) Schwartz, R. N.; Slawsky, Z. I.; Herzfeld, K. F. *Calculation of Vibrational Relaxation Times in Gases*; AIP, 1952.
- (137) Stannard, P. R.; Gelbart, W. M. *J. Phys. Chem.* **1981**, *85*, 3592-3599.
- (138) Wells, J. M.; McLuckey, S. A. In *Biol. Mass Spectrom.*; Elsevier Academic Press Inc: San Diego, 2005; Vol. 402, pp 148-185.
- (139) Gauthier, J. W.; Trautman, T. R.; Jacobson, D. B. *Anal. Chim. Acta* **1991**, *246*, 211-225.
- (140) Guan, S. H.; Marshall, A. G.; Wahl, M. C. *Anal. Chem.* **1994**, *66*, 1363-1367.
- (141) Roepstorff, P.; Fohlman, J. *Biol. Mass Spectrom.* **1984**, *11*, 601-601.
- (142) McLuckey, S. A.; Habibigoudarzi, S. *J. Am. Chem. Soc.* **1993**, *115*, 12085-12095.
- (143) McLuckey, S. A.; Van Berker, G. J.; Glish, G. L. *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 60-70.
- (144) Basov, N. G. *Jetp Letters* **1971**, *14*, 165.

- (145) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. *Rev. Sci. Instrum.* **2005**, *76*.
- (146) Oh, H.; Breuker, K.; Sze, S. K.; Ge, Y.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15863-15868.
- (147) Bush, M. F.; O'Brien, J. T.; Prell, J. S.; Saykally, R. J.; Williams, E. R. *J. Am. Chem. Soc.* **2007**, *129*, 1612-1622.
- (148) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 2067-2070.
- (149) Payne, A. H.; Glish, G. L. *Anal. Chem.* **2001**, *73*, 3542-3548.
- (150) Little, D. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **1995**, *117*, 6783-6784.
- (151) Hofstadler, S. A.; Griffey, R. H.; Pasa-Tolic, L.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1400-1404.
- (152) Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256-3262.
- (153) Mo, J. J.; Hakansson, K. *Anal. Bioanal. Chem.* **2006**, *386*, 675-681.
- (154) Baba, T.; Hashimoto, Y.; Hasegawa, H.; Hirabayashi, A.; Waki, I. *Anal. Chem.* **2004**, *76*, 4263-4266.
- (155) Silivra, O. A.; Kjeldsen, F.; Ivonin, I. A.; Zubarev, R. A. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 22-27.
- (156) Zubarev, R. A. *Mass Spectrom. Rev.* **2003**, *22*, 57-77.
- (157) Zubarev, R. A.; Haselmann, K. F.; Budnik, B.; Kjeldsen, F.; Jensen, F. *Eur. J. Mass Spectrom.* **2002**, *8*, 337-349.
- (158) Syrstad, E. A.; Turecek, F. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 208-224.
- (159) Sobczyk, M.; Anusiewicz, I.; Berdys-Kochanska, J.; Sawicka, A.; Skurski, P.; Simons, J. *J. Phys. Chem. A* **2004**, *109*, 250-258.
- (160) Axelsson, J.; Palmblad, M.; Hakansson, K.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 474-477.
- (161) Kruger, N. A.; Zubarev, R. A.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass spectrom.* **1999**, *187*, 787-793.

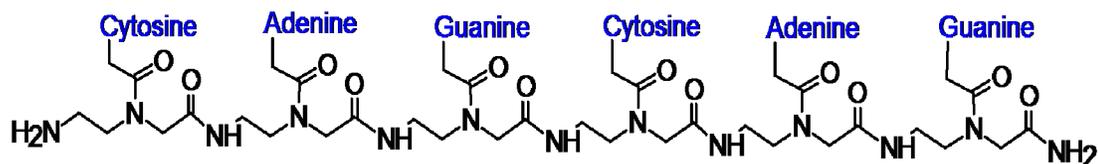
- (162) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563-573.
- (163) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793-1800.
- (164) Kelleher, R. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250-4253.
- (165) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431-4436.
- (166) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.
- (167) Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnsen, A. H.; Zubarev, R. A. *Anal. Chem.* **2001**, *73*, 2998-3005.
- (168) Horn, D. M.; Breuker, K.; Frank, A. J.; McLafferty, F. W. *J. Am. Chem. Soc.* **2001**, *123*, 9792-9799.
- (169) Loo, J. A.; Thanabal, V.; Mei, H. Y. *Mass Spectrometry in Biology & Medicine* **2000**, 73-90.
- (170) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 10313-10317.
- (171) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.
- (172) Haselmann, K. F.; Jorgensen, T. J. D.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 2260-2265.
- (173) Schultz, K. N.; Hakansson, K. *Int. J. Mass spectrom.* **2004**, *234*, 123-130.
- (174) Kjeldsen, F.; Silivra, O. A.; Ivonin, I. A.; Haselmann, K. F.; Gorshkov, M.; Zubarev, R. A. *Chem. Eur. J.* **2005**, *11*, 1803-1812.
- (175) Kweon, H. K.; Hakansson, K. *J. Proteome. Res.* **2008**, *7*, 749-755.
- (176) Yang, J.; Hakansson, K. *J. Am. Soc. Mass. Spectrom.* **2006**, *17*, 1369-1375.
- (177) Taucher, M.; Rieder, U.; Breuker, K. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 278-285.

## Chapter 2

### Gas-phase Ion Activation Techniques for Characterization of Peptide Nucleic Acids (PNAs)

#### 2.1 Introduction

Peptide nucleic acids (PNAs) are nucleic acid analogs in which the backbone is replaced by a synthetic peptide-like backbone, as shown in Figure 2.1. They were designed to bind nucleic acids in a sequence-specific manner, through Watson-Crick bonding rules.<sup>1-3</sup> In PNAs, 2-aminoethyl-glycine units replace the labile phosphate-sugar backbone units in regular DNA/RNA. The length and flexibility of the PNA backbone offer the ability of PNAs to bind single-strand and duplex DNA/RNA.<sup>4,5</sup> The uncharged nature of the PNA backbone offers higher stability than non-PNA complexes, due to reduction in electrostatic repulsion. The PNA backbone is resistant to proteases and nucleases because it is neither peptide- nor nucleic acid-like. The design of the PNA backbone provides higher thermal stability and chemical resistance of PNA containing complexes compared with regular DNA or RNA complexes,<sup>6</sup> and also results in unique ionic properties.



**Figure 2.1.** Structure of the PNA CAGCAG.

PNAs have been investigated and applied in many areas. First they were developed as drug candidates. Through inhibition of DNA replication,<sup>7</sup> mRNA transcription,<sup>8,9</sup> and protein translation<sup>10,11</sup> PNAs have shown promise in antigene and antisense applications.<sup>12</sup> However, delivery of PNAs through the cell membrane is one major challenge preventing use of PNAs as drugs.<sup>13,14</sup> Second PNAs have been adapted in biotechnologies such as biosensors for detection of single nucleotide polymorphisms (SNP)<sup>15,16</sup> and mutations,<sup>17</sup> as well as for enhancing PCR efficiency<sup>18,19</sup> and telomere analysis.<sup>20</sup>

Sequencing of PNA can be performed by tandem mass spectrometric techniques such as collision activated dissociation (CAD), infrared multiphoton dissociation (IRMPD),<sup>21,22</sup> and electron capture dissociation (ECD).<sup>23</sup> In positive ion mode, the nucleobases and terminal primary amines are most likely to be protonated. Water loss, methylene carbonyl linker loss, nucleobase loss, and backbone cleavages are observed in CAD<sup>24-28</sup> and similar fragmentation patterns are seen in IRMPD.<sup>27,28</sup> In ECD, water loss is limited compared to CAD and IRMPD. However, the overall fragmentation efficiency is lower in ECD.<sup>26-28</sup> Despite the lack of acidity of the PNA backbone PNA-DNA or PNA-RNA complexes show improved ionization efficiency in negative ion mode compared to positive ion mode, due to the acidic DNA or RNA backbone. PNA itself

also ionizes well in negative ion mode.<sup>25</sup> Backbone cleavages and neutral loss fragments are observed in negative ion CAD with backbone cleavages being more dominant than in positive ion mode.<sup>25</sup> Even less water loss was observed in CAD of PNAs acetylated at the N-terminus and similar patterns were observed in IRMPD.<sup>28</sup>

In this Chapter, non-acetylated PNA fragmentation in both positive and negative ion mode by CAD, IRMPD, ECD (positive ion mode only) and electron detachment dissociation (EDD, negative ion mode only) is examined. Also, double resonance and acetylation are employed to investigate the mechanism of water loss.

## **2.2 Experimental Section**

### **2.2.1 Sample Preparation**

PNAs including CAGCAG and ATGTCAGTCA were purchased from Biosynthesis Inc. (Lewisville, TX). All reagents were used without further purification. In acetylation reactions, 2  $\mu\text{L}$  of PNA at 100  $\mu\text{M}$  was mixed with 100 mL of 1:3 v/v acetic anhydride (Sigma-Aldrich)/methanol (Fisher). The reaction mixture was kept at room temperature for 24 hours and then evaporated to dryness. PNAs and acetylated PNAs were diluted into 1:1 v/v isopropanol (Fisher, Fair Lawn, NJ) /water (Fisher) with 10 mM ammonium acetate (Fisher) to 5  $\mu\text{M}$  for negative ion mode, and into 1:1 v/v methanol (Fisher)/water with 2% v/v acetic acid (Sigma-Aldrich) to 1  $\mu\text{M}$  for positive ion mode analysis.

### **2.2.2 Mass Spectrometry**

All mass spectra were acquired on a Bruker Apex-Q 7T ESI-FT-ICR mass spectrometer (Bruker Daltonics, Billerica, MA). Samples were infused via an Apollo II

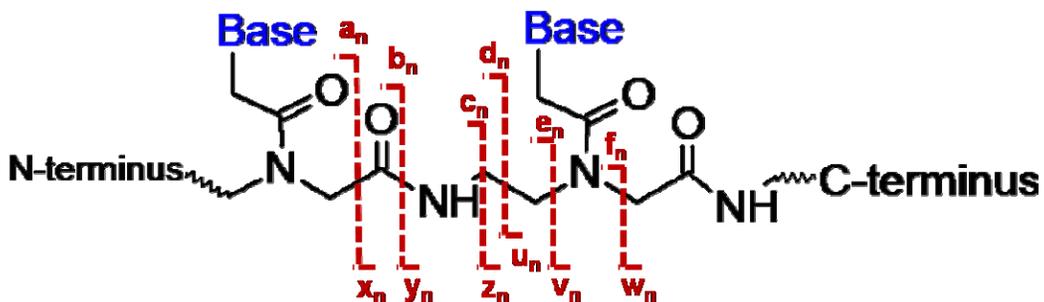
electrospray source (Bruker Daltonics, Billerica, MA) at 50  $\mu\text{L}/\text{h}$ . Precursor ions for MS/MS were selected by the quadrupole and accumulated externally in the collision cell (hexapole). CAD experiments were conducted by acceleration of precursor ions into the collision cell filled with argon at  $6 \times 10^{-6}$  mBar. IRMPD experiments were conducted by irradiating precursor ions in the ICR cell by a laser beam from a  $\text{CO}_2$  laser at 12.5 W for 20-100 ms. ECD and EDD experiments were conducted by irradiating precursor ions in the ICR cell by an electron beam from an indirectly heated cathode. For ECD experiments, the cathode bias voltage was set to -1 V and a lens located between the cathode and the ICR cell was set to 1 V. For EDD experiments, the cathode bias voltage was between -18 and -25 V and the lens voltage was kept 1 V lower than the cathode bias voltage, i. e. -19 and -26 V. For double resonance IRMPD experiments, a waveform containing the frequency corresponding to the  $m/z$  ratio of interest was applied to the excitation electrodes while the laser beam was irradiating the precursor ions. For  $\text{MS}^3$  experiments, after the first MS/MS event, product ions of interest were isolated in the ICR cell by a correlated harmonic excitation field (CHEF)<sup>29</sup> waveform containing all frequencies corresponding to ions to be ejected, then subjected to IRMPD. Mass spectra were averaged over 64 to 256 scans. Data files were processed by MIDAS<sup>30</sup> software.

## 2.3 Results and Discussion

### 2.3.1 CAD, IRMPD and ECD of the PNAs CAGCAG and ATGTCAGTCA in Positive Ion Mode

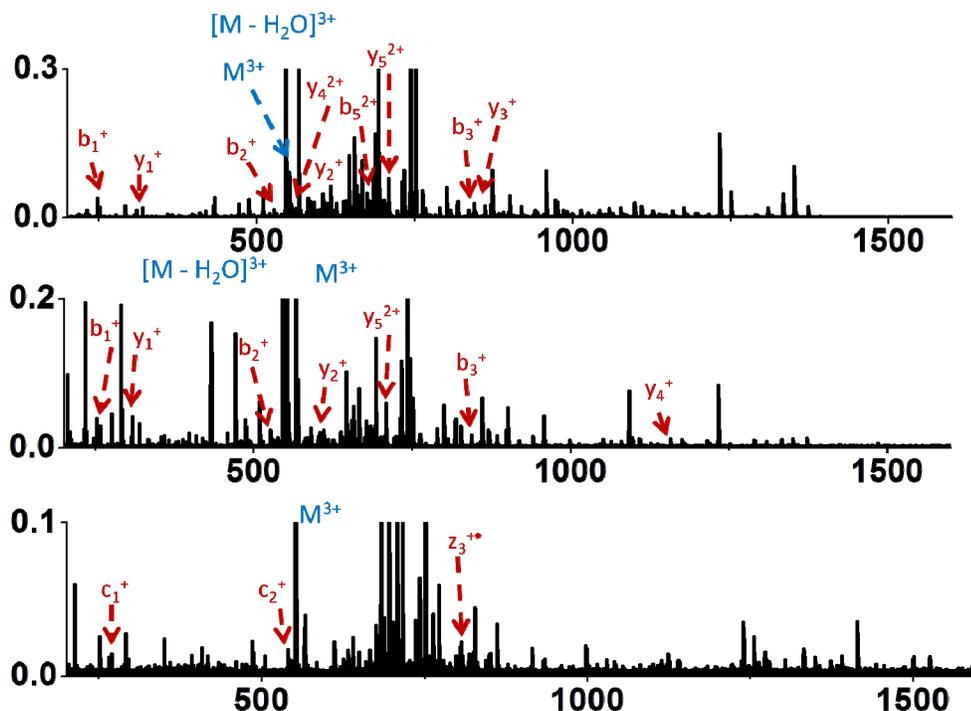
The structure of the PNA 5'-CAGCAG-3' is shown in Figure 2.1. Different nomenclatures for PNA fragmentation were proposed by Takao *et al.*,<sup>31</sup> Muddiman *et al.*,<sup>24</sup> and Zubarev *et al.*<sup>26</sup> In this Chapter the nomenclature from Zubarev *et al.*, which is based on commonly accepted peptide notation is adopted.<sup>32</sup> In the Zubarev nomenclature

*a*, *b*, *c*, *d*, *x*, *y*, *z* and *w*-type ions correspond to *b*, *c*, *d*, *e*, *v*, *w*, *x*, and *y*-type ions in the Muddiman nomenclature. The Zubarev nomenclature is shown in Figure 2.2.

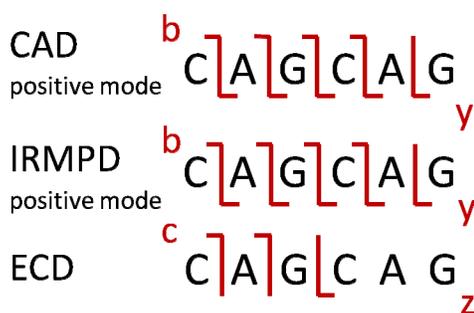


**Figure 2.2.** Product ion nomenclature used in this Chapter. Adapted from Zubarev *et al.*<sup>26</sup>

CAD, IRMPD and ECD were performed on triply protonated precursor cations,  $[M + 3H]^{3+}$ . Mass spectra are shown in Figure 2.3. Backbone cleavages are mapped onto the PNA sequence, as shown in Figure 2.4. Other assigned peaks in the CAD spectrum, not labeled in Figure 2.3, are listed in Table 2.1.



**Figure 2.3.** CAD (16 V collision cell voltage), IRMPD (7.5 W, 80 ms), and ECD (- 1 V, 7 ms) MS/MS spectra of the PNA CAGCAG. Only *b*, *c*, *y* and *z*-type ions without neutral loss are labeled.



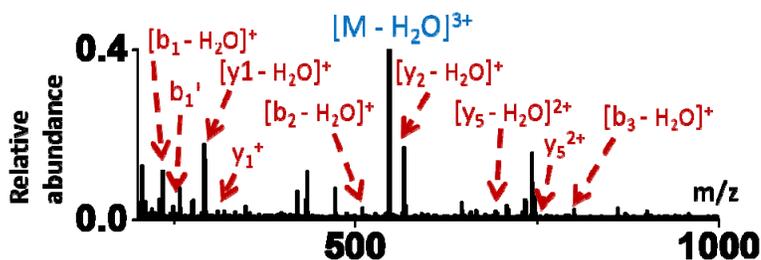
**Figure 2.4.** Backbone cleavages from CAD, IRMPD, and ECD of the PNA CAGCAG in positive ion mode. Product ion labels include combined backbone/neutral loss cleavages. For example, “b” labels include *b*, (*b* - H<sub>2</sub>O), and (*b* - base) ions.

Observed m/z	Calculated m/z	Assignment	Abundance (%)	Error (ppm)
234.1002	234.0991	$(b_1 - \text{H}_2\text{O})^+$	0.53	4.60
291.1331	291.1312	$(y_1 - \text{H}_2\text{O})^+$	2.38	6.38
433.2067	433.2054	$(y_2 - \text{G})^+$	3.90	2.88
433.2067	433.2054	$(c_2 - \text{C})^+$	3.90	2.88
471.1855	471.1847	$(w_2 - \text{A} - \text{H}_2\text{O})^+$	2.60	1.79
509.2125	509.2116	$(b_2 - \text{H}_2\text{O})^+$	5.26	1.80
554.7301	554.7307	$(y_4 - \text{H}_2\text{O})^{2+}$	2.56	-0.93
566.2443	566.2443	$(y_2 - \text{H}_2\text{O})^+$	31.27	0.06
604.7623	604.7571	$(c_5 - \text{H}_2\text{O} - \text{A})^{2+}$	4.70	8.55
674.7901	674.7864	$[\text{M} - \text{A} - \text{G} - \text{H}_2\text{O}]^{2+}$	4.63	5.47
684.2792	684.2778	$(z_5 - \text{H}_2\text{O})^{2+}$	0.38	2.03
686.7954	686.7920	$[\text{M} - \text{H}_2\text{O} - \text{C} - \text{G}]^{2+}$	16.59	4.90
692.2848	692.2875	$(y_5 - \text{H}_2\text{O})^{2+}$	34.74	-3.86
733.3094	733.3084	$[\text{M} - \text{G} - 2\text{H}_2\text{O}]^{2+}$	9.20	1.35
742.3141	742.3134	$[\text{M} - \text{G} - \text{H}_2\text{O}]^{2+}$	100.00	0.90
746.2928	746.2977	$(w_3 - \text{C} - \text{H}_2\text{O})^+$	5.34	-6.63
751.3198	751.3190	$[\text{M} - \text{G}]^{2+}$	48.66	1.10
762.3107	762.3167	$[\text{M} - \text{C} - \text{H}_2\text{O}]^{2+}$	5.25	-7.94
800.3142	800.3195	$(b_3 - \text{H}_2\text{O})^+$	6.25	-6.66
817.3428	817.3461	$(y_3 - \text{H}_2\text{O})^+$	2.54	-3.96
957.4049	957.4046	$(y_4 - \text{G} - \text{H}_2\text{O})^+$	8.97	0.28

957.4049	957.4046	$(c_4 - C - H_2O)^+$	8.97	0.28
1193.4903	1193.4955	$(b_5 - G)^+$	0.79	-4.37
1232.5105	1232.5177	$(c_5 - C - H_2O)^+$	17.73	-5.82
1232.5105	1232.5177	$(y_5 - G - H_2O)^+$	17.73	-5.82
1250.5179	1250.5283	$(c_5 - C)^+$	5.22	-8.28
1250.5179	1250.5283	$(y_5 - C)^+$	5.22	-8.28
1350.5807	1350.5807	$[M - 2G]^+$	10.08	0.00

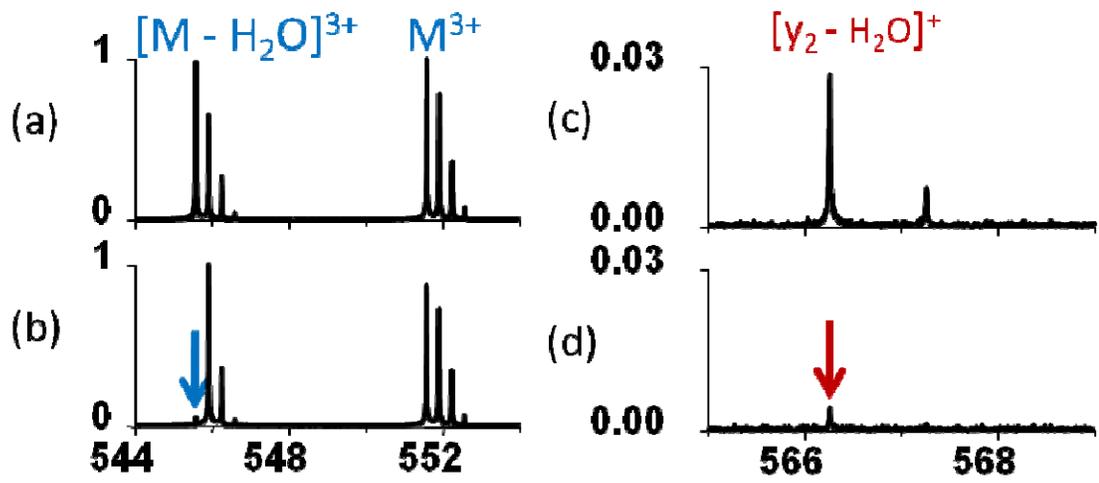
**Table 2.1.** List of assigned peaks in the positive ion mode CAD spectrum of the PNA CAGCAG. These assignments are not included in Figure 2.3.

For the PNA CAGCAG, CAD and IRMPD provide complete sequence coverage. However, abundant neutral losses are observed, that complicate spectrum interpretation. Neutral losses include water loss and nucleobase loss. Previous experimental and theoretical studies have proposed PNA fragmentation mechanisms.<sup>25</sup> As for water loss, two mechanisms corresponding to N-terminal<sup>25</sup> and C-terminal<sup>33</sup> water losses have been proposed. Proposed C-terminal water loss is charge independent whereas the proposed N-terminal water loss is dependent on the protonation sites. An MS<sup>3</sup> experiment was performed in which  $[M - H_2O]^{3+}$  generated from CAD in the external collision cell was fragmented by SORI-CAD in the ICR cell. The corresponding spectrum is shown in Figure 2.5.

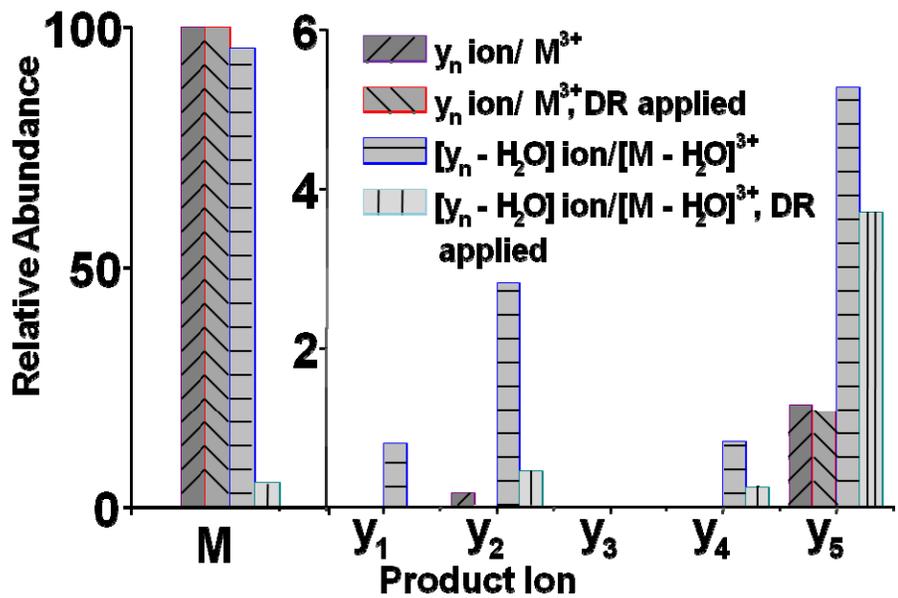


**Figure 2.5.** MS<sup>3</sup> (CAD of M<sup>3+</sup> followed by SORI-CAD of [M - H<sub>2</sub>O]<sup>3+</sup>) of the PNA CAGCAG.

In the MS<sup>3</sup> spectrum,  $b_1$  and ( $b_1 - \text{H}_2\text{O}$ ),  $y_1$  and ( $y_1 - \text{H}_2\text{O}$ ), as well as  $y_5$  and ( $y_5 - \text{H}_2\text{O}$ ) ions are observed, indicating that water loss occurs from both the N- and C-termini. This result is consistent with Ballistreri *et al.*<sup>33</sup> and Flora *et al.*<sup>25</sup> In order to further explore the mechanism of water loss, double resonance experiments were performed in which [M - H<sub>2</sub>O]<sup>3+</sup> was continuously ejected from the ICR cell while IRMPD was being performed. The results of the experiment are shown in Figure 2.6.



**Figure 2.6.** IRMPD of the triply protonated PNA CAGCAG with and without simultaneous double resonance (DR) ejection of  $[M - H_2O]^{3+}$ . (a) Zoomed spectrum without DR, showing  $[M - H_2O]^{3+}$  and  $M^{3+}$ . (b) Zoomed spectrum with DR. (c) Zoomed spectrum without DR, showing  $[y_2 - H_2O]^+$ . (d) Zoomed spectrum with DR.

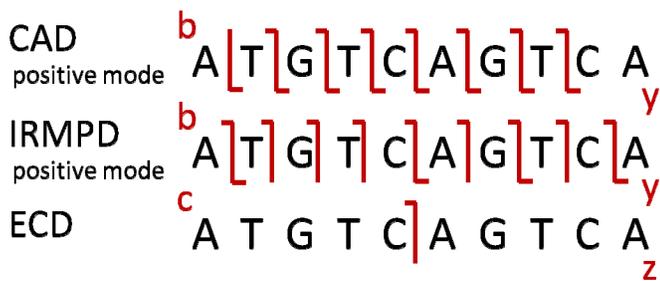


**Figure 2.7.** Abundance changes of precursor and y-type product ions with and without DR in IRMPD of the PNA CAGCAG.

Figure 2.7 illustrates the abundance of precursor ions and  $y$ -type fragment ions in IRMPD experiments with and without DR. Simultaneous ejection of  $[M - H_2O]^{3+}$  dramatically affected the formation of  $(y_1 - H_2O)$ ,  $(y_2 - H_2O)$ , and  $(y_4 - H_2O)$  as well as decreased the formation of  $(y_5 - H_2O)$ . This result indicates that water loss either primarily occurs from the C-terminus, or that one of the  $y$ -ion formation pathways involves water loss.

In the ECD spectrum of CAGCAG, only  $c_1$ ,  $c_2$ , and  $z_3$  fragments of low abundance are observed. Water loss, base loss, and unassignable peaks constitute the major product ions. This result appears consistent with previous observations by Olsen *et al.*<sup>26</sup>

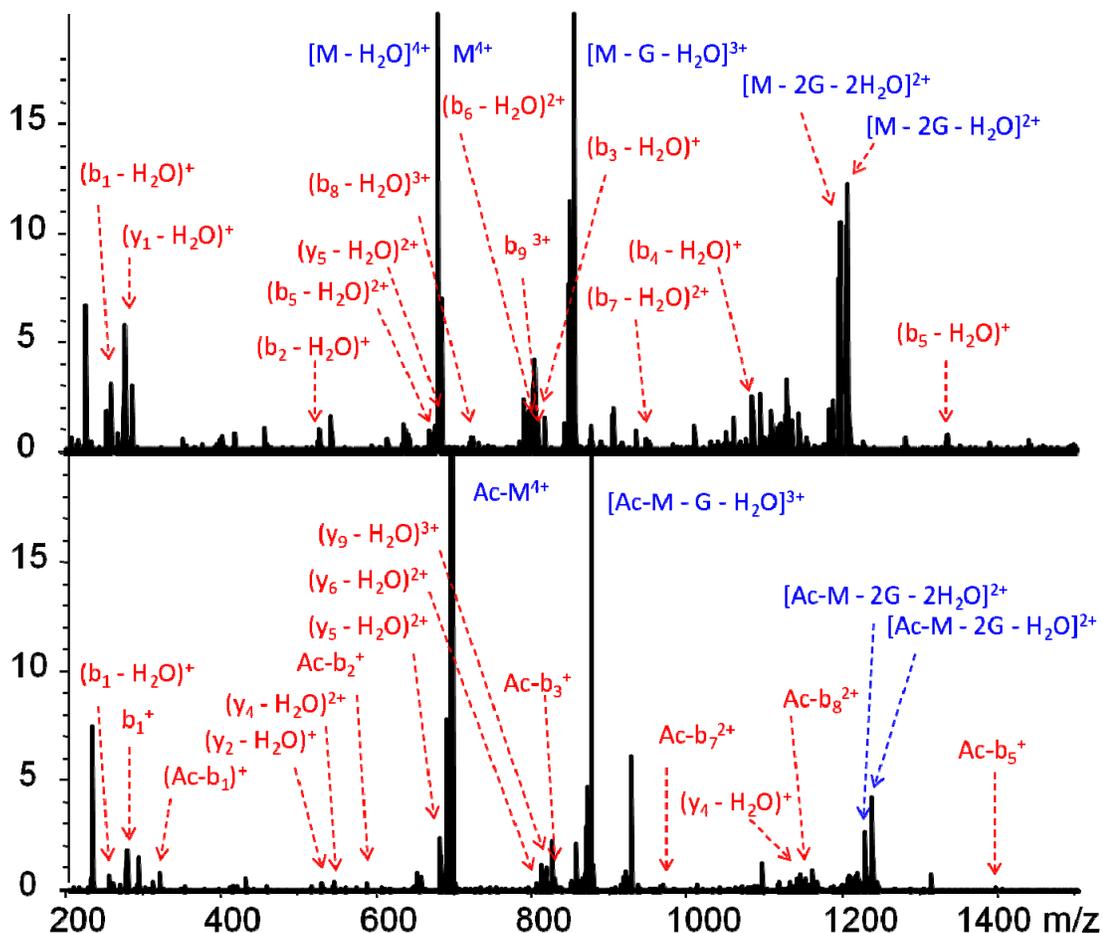
CAD, IRMPD and ECD of the quadruply protonated PNA ATGTCAGTCA were also performed. Backbone fragments are shown in Figure 2.8.



**Figure 2.8.** Backbone cleavages from CAD, IRMPD and ECD of the PNA ATGTCAGTCA in positive ion mode. Product ion labels include combined backbone/neutral loss cleavages. For example, “b” labels include  $b$ ,  $(b - H_2O)$ , and  $(b - \text{base})$  ions.

From CAD and IRMPD spectra, full sequence coverage was obtained. However, abundant water loss, base loss, and unassignable peaks dominate the spectra. In order to

reduce water loss, acetylation of the N-terminus was performed. IRMPD spectra of the 10-mer PNA ATGTCAGTCA and its acetylated variant are shown in Figure 2.9.



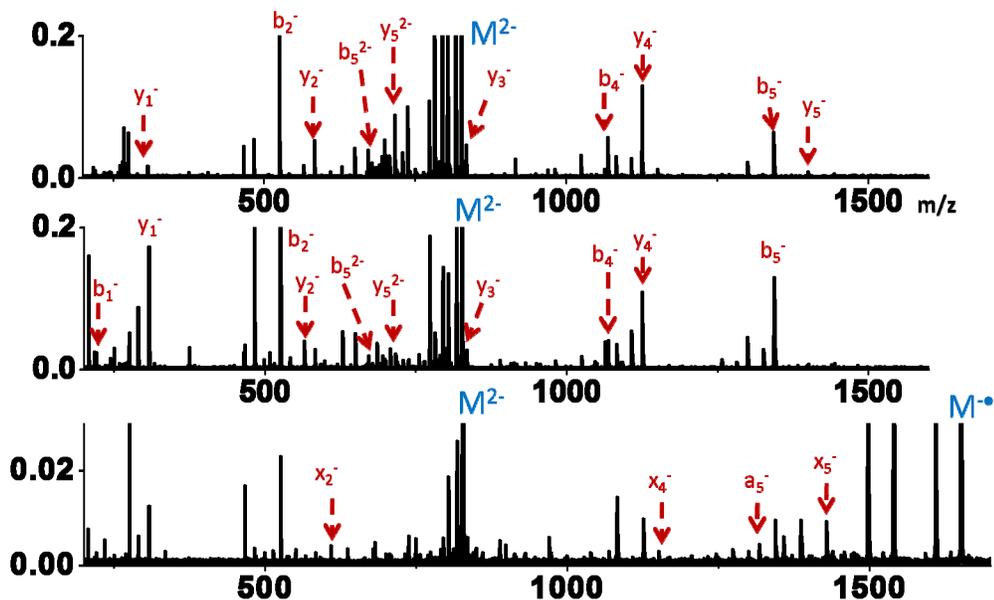
**Figure 2.9.** IRMPD spectra of the non-acetylated 10-mer PNA ATGTCAGTCA (upper panel) and its acetylated variant (bottom panel).

From Figure 2.9 it can be seen that, before acetylation,  $b_1$  through  $b_8$  ions are observed with water loss. By contrast, after acetylation,  $b_1$ ,  $b_2$ ,  $b_3$ ,  $b_5$ ,  $b_7$ , and  $b_8$  ions are observed without water loss. However, all  $y$  ions are observed with water loss following acetylation. This result indicates that acetylation blocks one of the water loss pathways – the one corresponding to N-terminal water loss. The IRMPD spectrum of the acetylated 10-mer PNA is easier to interpret than the one for the non-acetylated form. This result is consistent with previous results obtained in negative ion mode.<sup>34</sup>

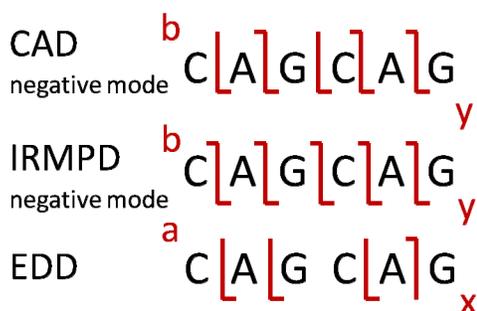
ECD of the PNA ATGTCAGTCA provided only one backbone cleavage. Peaks from neutral loss and unassignable peaks dominate the spectrum. Similar fragmentation patterns were previously observed for 12-mer and 13-mer PNAs.<sup>28</sup>

### 2.3.2 CAD, IRMPD and EDD of the PNAs CAGCAG and ATGTCAGTCA in Negative Ion Mode

CAD, IRMPD and EDD were performed on the triply deprotonated anionic PNA CAGCAG,  $[M - 3H]^{3-}$ , as shown in Figure 2.10. Backbone cleavages were mapped to the sequence as shown in Figure 2.11. Other assigned peaks in the CAD spectrum, not labeled in Figure 2.10, are listed in Table 2.2.



**Figure 2.10.** Negative ion mode CAD (19 V collision cell voltage), IRMPD (7.5 W, 180 ms), and EDD (- 18 V, 2 s) MS/MS spectra of the PNA CAGCAG. Only *a*, *b*, *x*, and *y*-type ions without neutral loss are labeled.



**Figure 2.11.** Backbone cleavages from CAD, IRMPD and EDD of the PNA CAGCAG in negative ion mode. Product ion labels include combined backbone/neutral loss cleavages. For example, “b” labels include  $b$ ,  $(b - \text{H}_2\text{O})$ , and  $(b - \text{base})$  ions.

Observed m/z	Calculated m/z	Species	Abundance (%)	Error (ppm)
482.2062	482.2018	$(a_2 - \text{NH}_3)^-$	5.82	9.10
627.7596	627.7549	$(e_5 - \text{C} - \text{H}_2\text{O})^{2-}$	1.49	7.37
649.2636	649.2578	$(a_5 - \text{NH}_3)^{2-}$	4.54	8.92
690.2767	690.2718	$(y_5 - \text{H}_2\text{O})^{2-}$	1.89	7.13
706.7993	706.7951	$[\text{M} - 2\text{C} - \text{H}_2\text{O}]^{2-}$	3.27	5.87
715.8053	715.8004	$[\text{M} - 2\text{C}]^{2-}$	9.46	6.80
740.302	740.298	$[\text{M} - \text{H}_2\text{O} - \text{G}]^{2-}$	0.97	5.40
749.3121	749.3041	$[\text{M} - \text{G}]^{2-}$	1.20	10.65
790.3391	790.3357	$a_3^-$	2.35	4.28
794.3249	794.3198	$[\text{M} - \text{CONH} - \text{H}_2\text{O}]^{2-}$	36.39	6.42
803.3298	803.3251	$[\text{M} - \text{CONH}]^{2-}$	100.45	5.92
815.8263	815.8235	$[\text{M} - \text{H}_2\text{O}]^{2-}$	30.25	3.34
816.3228	816.3155	$[\text{M} - \text{NH}_3]^{2-}$	38.24	8.88
981.408	981.4057	$(e_4 - \text{C} - \text{H}_2\text{O})^-$	1.18	2.26

1024.414	1024.41	$(a_4 - \text{NH}_3)^-$	3.22	3.27
1081.443	1081.432	$(d_4 - \text{NH}_3)^-$	3.27	9.83
1106.43	1106.438	$(y_4 - \text{H}_2\text{O})^-$	2.79	-7.27
1150.429	1150.429	$x_4^-$	1.26	-0.26

**Table 2.2.** List of assigned peaks in negative ion mode CAD of the PNA CAGCAG. These peaks are not included in Figure 2.10.

From CAD and IRMPD spectra, full sequence coverage was obtained, respectively. Fragmentation patterns are similar between CAD and IRMPD. Neutral loss including water loss, ammonia loss, base loss, and -CONH loss are observed for both the precursor ion and fragment ions. However, neutral loss peaks are more abundant for the precursor ion compared to the fragment ions. In terms of  $b/y$  ion series in the CAD spectrum,  $b_2$ ,  $b_4$ ,  $b_5$ , and  $y_1 - y_3$  ions are observed only in the form without water loss. By contrast,  $y_4$  and  $y_5$  are observed both with and without water loss. Compared to CAD in positive mode, CAD in negative mode provided more intact  $b/y$  ions without extensive neutral losses, thus generating a spectrum less complicated to interpret.

From EDD, four out of 10 possible backbone cleavages are observed from  $a/x$  ion series. However, peaks associated with neutral loss and unassignable peaks are still dominant.

CAD, IRMPD and EDD were also performed on the quadruply deprotonated PNA ATGTCAGTCA. Backbone fragments are shown in Figure 2.12.



**Figure 2.12.** Backbone cleavages from CAD, IRMPD, and EDD of the PNA ATGTCAGTCA in negative ion mode. Product ion labels include combined backbone/neutral loss cleavages. For example, “b” labels include  $b$ ,  $(b - H_2O)$ , and  $(b - \text{base})$  ions.

From CAD and IRMPD spectra, full sequence coverage was obtained, respectively. Fragmentation patterns are similar between CAD and IRMPD. Similar to CAD and IRMPD of CAGCAG, neutral loss, including water loss, ammonia loss, base loss and -CONH loss is observed for both the precursor ion and fragment ions. In CAD,  $b_2$ ,  $b_3$ ,  $b_4$ ,  $b_7$ ,  $b_8$ ,  $y_5$ ,  $y_6$ , and  $y_9$  ions are observed only in intact form without water loss whereas  $b_5$ ,  $b_6$ ,  $b_9$ , and  $y_3$  ions are observed both with and without water loss.  $y_4$  and  $y_7$  ions are observed only in association with water loss. Intact  $b/y$  ions were more abundant compared to their water loss variants. This fragmentation pattern is consistent with CAD of the PNA CAGCAG in negative mode. In EDD, no backbone cleavages are observed and neutral loss and unassignable peaks are dominant.

## 2.4 Conclusions

In this Chapter, the MS/MS fragmentation patterns of a 6-mer and a 10-mer PNA in positive and negative ion mode were examined. The MS/MS techniques studied include CAD, IRMPD, ECD, and EDD.

For CAD in positive ion mode, full sequence coverage was obtained. However, neutral loss, including water and nucleobase loss from both precursor and product ions,

dominated the spectra. An MS<sup>3</sup> experiment revealed that water loss occurs from both the N- and C-terminus. A double resonance experiment revealed that water loss occurs from the C-terminus, or that one of the y-ion formation pathways involves water loss. IRMPD provided similar fragmentation patterns as CAD. Acetylation on the N-terminus of the 10-mer PNA blocked the water loss pathway involving the N-terminus. The IRMPD spectrum of the acetylated 10-mer PNA was easier to interpret compared to the spectrum of the non-acetylated variant due to reduced water loss from the fragments. ECD of 6-mer and 10-mer PNAs provided less backbone cleavages than CAD and IRMPD and also yielded abundant neutral loss peaks.

For CAD in negative ion mode, full sequence coverage was also obtained. Neutral loss including water loss, ammonia loss, nucleobase loss, and -CONH loss was observed for precursor and fragment ions. However, compared to positive ion mode, ions with neutral loss were of less abundance. Thus CAD spectra in negative mode were easier to interpret compared with positive ion mode. This result is consistent with previously reported results.<sup>28, 34</sup> EDD of the 6-mer and 10-mer PNAs provided less backbone cleavages than CAD and IRMPD and also abundant neutral loss peaks.

Overall, operating CAD and IRMPD in the negative ion mode, as well as acetylating the N-terminus, reduce the abundance of neutral loss peaks and simplify spectrum interpretation.

## 2.5 Bibliography

- (1) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497-1500.

- (2) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897.
- (3) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566-568.
- (4) Ray, A.; Norden, B. *The FASEB Journal* **2000**, *14*, 1041-1060.
- (5) Peter E, N. *Curr. Opin. Biotechnol.* **2001**, *12*, 16-20.
- (6) Demidov, V. V.; Yavnilovich, M. V.; Belotserkovskii, B. P.; Frank-Kamenetskii, M. D.; Nielsen, P. E. *Proc. Natl. Acad. Sci.* **1995**, *92*, 2637-2641.
- (7) Taylor, R. W.; Chinnery, P. F.; Turnbull, D. M.; Lightowlers, R. N. *Nat. Genet.* **1997**, *15*, 212-215.
- (8) Boffa, L. C.; Morris, P. L.; Carpaneto, E. M.; Louissaint, M.; Allfrey, V. G. *J. Biol. Chem.* **1996**, *271*, 13228-13233.
- (9) Good, L.; Nielsen, P. E. *Nat Biotech* **1998**, *16*, 355-358.
- (10) Knudsen, H.; Nielsen, P. E. *Nucleic Acids Res.* **1996**, *24*, 494-500.
- (11) Good, L.; Nielsen, P. E. *Proc. Natl. Acad. Sci.* **1998**, *95*, 2073-2076.
- (12) Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, C. F.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Bruckenstein, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E. *Science* **1992**, *258*, 1481-1485.
- (13) Pooga, M.; Soomets, U.; Hallbrink, M.; Valkna, A.; Saar, K.; Rezaei, K.; Kahl, U.; Hao, J.-X.; Xu, X.-J.; Wiesenfeld-Hallin, Z.; Hokfelt, T.; Bartfai, T.; Langel, U. *Nat Biotech* **1998**, *16*, 857-861.
- (14) Richard, J. P.; Melikov, K.; Vives, E.; Ramos, C.; Verbeure, B.; Gait, M. J.; Chernomordik, L. V.; Lebleu, B. *J. Biol. Chem.* **2003**, *278*, 585-590.
- (15) Ross, P. L.; Lee, K.; Belgrader, P. *Anal. Chem.* **1997**, *69*, 4197-4202.
- (16) Komiyama, M.; Ye, S.; Liang; Yamamoto, Y.; Tomita, T.; Zhou, J.-M.; Aburatani, H. *J. Am. Chem. Soc.* **2003**, *125*, 3758-3762.
- (17) Orum, H.; Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O.; Stanley, C. *Nucleic Acids Res.* **1993**, *21*, 5332-5336.
- (18) Demers, D. B.; Curry, E. T.; Egholm, M.; Sozer, A. C. *Nucleic Acids Res.* **1995**, *23*, 3050-3055.
- (19) Dose, C.; Ficht, S.; Seitz, O. *Angew. Chem. Int. Ed.* **2006**, *45*, 5369-5373.

- (20) Lansdorp, P. M.; Verwoerd, N. P.; van de Rijke, F. M.; Dragowska, V.; Little, M.-T. r. s.; Dirks, R. W.; Raap, A. K.; Tanke, H. J. *Hum. Mol. Genet.* **1996**, *5*, 685-691.
- (21) Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1994**, *116*, 4893-4897.
- (22) Little, D. P.; Aaserud, D. J.; Valaskovic, G. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **1996**, *118*, 9352-9359.
- (23) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
- (24) Flora, J. W.; Muddiman, D. C. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 759-762.
- (25) Flora, J. W.; Shillady, D. D.; Muddiman, D. C. *J. Am. Soc. Mass. Spectrom.* **2000**, *11*, 615-625.
- (26) Olsen, J. V.; Haselmann, K. F.; Nielsen, M. L.; Budnik, B. A.; Nielsen, P. E.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 969-974.
- (27) Han, X.; Morrison, H.; Kobaslija, M.; McQuade, D. T.; McLafferty, F. W. Series, Peptide Nucleic Acid Characterization by Electrospray FT-ICR Mass Spectrometry, *The 52nd ASMS Conference on Mass Spectrometry and Allied Topics*, Nashville, TN, May 23 - 27 2004.
- (28) Ziehe, M.; Grossmann, T. N.; Seitz, O.; Linscheid, M. W. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1132-1138.
- (29) de Koning, L. J.; Nibbering, N. M. M.; van Orden, S. L.; Laukien, F. H. *Int. J. Mass Spectrom. Ion Processes* **1997**, *165-166*, 209-219.
- (30) Senko, M. W.; Canterbury, J. D.; Guan, S. H.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1839-1844.
- (31) Takao, T.; Fukuda, H.; Coull, J.; Shimonishi, Y. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 925-928.
- (32) Roepstorff, P.; Fohlman, J. *Biol. Mass Spectrom.* **1984**, *11*, 601-601.
- (33) Ballistreri, A.; Garozzo, D.; Maravigna, P.; Montaudo, G.; Giuffrida, M. *J. Polym. Sci., Part A: Polym. Chem.* **1987**, *25*, 1049-1063.
- (34) Flora, J. W.; Muddiman, D. C. *Anal. Chem.* **2001**, *73*, 3305-3311.

## Chapter 3

### Top-down MS/MS of Ribonucleic Acids and Their Complexes

#### 3.1 Introduction

Characterization of biomolecules by mass spectrometry (MS) and tandem mass spectrometry (MS/MS) has been gaining popularity since the emergence of electrospray ionization (ESI),<sup>1</sup> a “soft” ionization technique able to introduce labile molecules into the gas phase. The “top-down” approach of mass spectrometry, i.e., analysis of large biomacromolecules in their intact form, has been primarily applied for characterization of proteins.<sup>2,3</sup> However, top-down analysis has also been successfully applied to nucleic acids. For example, intact masses of transfer ribonucleic acids (tRNAs)<sup>4</sup> and polymerase chain reaction (PCR) products<sup>5,6</sup> have been obtained from mass spectrometry. In top-down MS/MS, intact nucleic acid ions are dissociated in the gas phase. Based on the fragments observed, sequence or structural information can be obtained. Top-down MS/MS has been applied to tRNAs<sup>7</sup> small interfering RNAs (siRNAs),<sup>8</sup> RNA-protein complexes,<sup>9,10</sup> RNA-ligand complexes,<sup>11-14</sup> and HIV-1 SL1 RNA<sup>15</sup>.

”Slow heating” MS/MS techniques, including collision activated dissociation (CAD)<sup>16,17</sup> and infrared multiphoton dissociation (IRMPD)<sup>18,19</sup> are the most common for

analysis of RNAs. In CAD, precursor ions are heated by gaseous inelastic collisions between accelerated ions and neutral molecules such as argon. During the activation process, each inelastic collision deposits a small amount of internal energy to the precursor ion. The energy from each collision is not sufficient to dissociate a molecular bond. However, energy is accumulated from multiple collisions to overcome the dissociation threshold. For RNAs, CAD provides primarily *c* and *y*-type backbone product ions, as well as (*a* - B)/*w*-type ions and product ions from neutral loss of small molecules.<sup>8, 20-22</sup> CAD has been applied to both small and large (up to 108 nt) DNAs.<sup>23, 24</sup> Recently it was applied to dissociate RNAs with sizes up to 61 nt.<sup>7, 22, 25</sup> IRMPD is another type of slow-heating technique, that activates ions by depositing energy from IR photon absorption. IRMPD typically provides comparable backbone fragmentation patterns to CAD.<sup>26</sup> However, oligonucleotides are more labile in IRMPD compared to proteins and peptides due to strong absorption of 10.6  $\mu\text{m}$  photons from resonance with the P-O stretching frequency of the oligonucleotide phosphodiester backbone.<sup>18, 27</sup> When implemented in ion trap instruments, IRMPD provides advantages compared with CAD because ion loss due to collisional scattering is minimized, and also no rf voltage is needed for precursor ion excitation, thus eliminating the instability (i.e., 1/3 rule) for low *m/z* product ions.<sup>28</sup>

Electron capture dissociation (ECD)<sup>29-31</sup> is a reaction between multiply positively charged precursor ions and low energy electrons. During this reaction, a cation captures one electron to form a meta-stable radical intermediate that further dissociates into fragments. ECD has been demonstrated to dissociate oligonucleotides.<sup>32, 33</sup> However, due to the acidity of the phosphate backbone, nucleic acids are detected more sensitively

in negative ion mode compared with positive ion mode. Thus ECD has been limited to the analysis of short oligonucleotides rather than polynucleotides.

Electron detachment dissociation (EDD)<sup>34</sup> is a reaction between multiply negatively charged precursor ions and electrons of relatively high energy. During this reaction, an anion is bombarded with electrons to lose electrons and forms a meta-stable radical intermediate that further dissociates into fragments. EDD has been demonstrated to dissociate oligonucleotides<sup>35-37</sup> and, more recently, large RNAs (up to 61 nt)<sup>25</sup>. EDD was also attempted on several DNA complexes, however, only electron detachment was observed.<sup>38</sup> Major products ions from EDD of nucleic acids are *d*- and *w*-type ions. For DNAs it has been shown that base loss is a secondary fragmentation pathway while backbone cleavage is primary.<sup>39,40</sup> In other work, it was observed that EDD is able to preserve non-covalent bonding in the gas phase.<sup>39,41</sup>

Here an effort is made to characterize a 27-mer RNA and its small molecule complexes by MS/MS techniques including EDD and negative ion CAD, as well as ion mobility mass spectrometry (IM-MS). The 27-mer RNA R1 is representative of the A-site structure of 16S rRNA in prokaryotic cells and is known to form complexes with aminoglycoside antibiotics.<sup>11,12</sup>

## **3.2 Experimental Section**

### **3.2.1 Sample Preparation**

The 27-mer RNA R1 (5'-GGCGUCACACCUUCGGGUGAAGUCGCC-3') was purchased from Integrated DNA Technologies, Inc. (Coralville, IA) and desalted by ethanol precipitation (protocol modified from Limbach et al.<sup>42</sup>). 1 nmole of RNA was

mixed with 20  $\mu\text{L}$  of 7.5 M  $\text{NH}_4\text{OAc}$  and 80  $\mu\text{L}$  of cold ethanol. This mixture was vortexed, stored at - 80  $^\circ\text{C}$  for 3 hours, and centrifuged for 30 minutes. The supernatant was discarded. 200  $\mu\text{L}$  of 90% cold ethanol was added to the precipitate and stored at - 80  $^\circ\text{C}$  for 2 h. The resulting mixture was then centrifuged at 12,000 g for 30 minutes and the supernatant was discarded. The precipitate was dissolved in electrospray solution for electrospray ionization. Aminoglycosidic antibiotics, including paromomycin (PM), ribostamycin (RM), (Be)kanamycin (BK), and apramycin (AP) were purchased from Sigma-Aldrich (St. Louis, MO).

### **3.2.2 Mass Spectrometry and Ion Mobility-Mass Spectrometry**

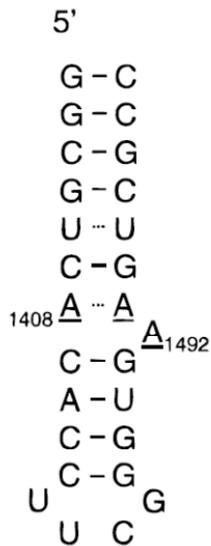
Mass spectra were obtained with a 7-T quadrupole-FT-ICR mass spectrometer (Apex-Q, Bruker Daltonics, Billerica, MA). The desalted RNA was dissolved in 1% triethylamine (TEA) and 10%, 30%, or 50% methanol in water (v/v). Solutions were infused via an Apollo II electrospray ion source at a flow rate of 70  $\mu\text{L}/\text{hour}$ . Ions with mass to charge ratios of interest were selected by an external quadrupole, accumulated in a hexapole, and transported to the ICR cell for dissociation and detection. Beam-type CAD experiments were performed by accelerating selected precursor ions into the hexapole filled with argon at  $6 \times 10^{-6}$  mbar. EDD experiments were performed in the ICR cell by irradiating precursor ions with an electron beam generated by an indirectly heated hollow dispenser cathode for 1-2 s. The cathode bias voltage was adjusted between - 20 and - 30 V. A lens located between the cathode and the ICR cell was set to around 1 V higher than the cathode bias voltage, i. e. between - 19 V and - 29 V. For activated ion (AI) EDD experiments, precursors ions in the ICR cell were irradiated with an infrared laser beam from a 25-W, 10.6- $\mu\text{m}$ ,  $\text{CO}_2$  laser (Synrad, Mukilteo, WA) at 5%

power for 20-50 ms, then irradiated with the electron beam for 1-2 s as described above. Electrons were extracted from an indirectly heated ring cathode heated with a current of 1.8 A. Mass spectra were processed by DataAnalysis software (Bruker Daltonics) and manually interpreted. Ion mobility mass spectra were obtained with a Q-IM-TOF instrument (Synapt G2, Waters, Milford, MA, USA). Solutions were infused via a nanoESI source with in-house generated fused silica capillary tips. Data sets were processed by MassLynx software (Waters).

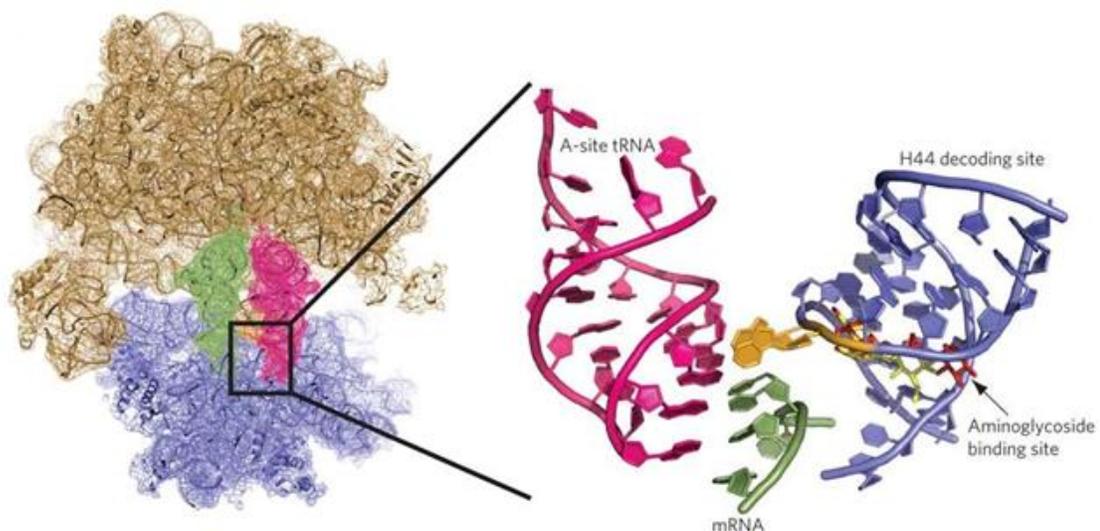
### **3.3 Results and Discussion**

#### **3.3.1 Effects of Organic Solvent in Electrospray Solution on 27-mer RNA Gas-phase Structure**

The 27-mer RNA R1 5'-GGCGUCACACCUUCGGGUGAAGUCGCC-3' has been proposed to form a hairpin structure, as shown in Figure 3.1.<sup>43</sup> It has been revealed by CAD of an RNA R1 analog that A1408, A1492, and A1493 are binding sites for aminoglycoside antibiotic drugs.<sup>11,43</sup> The R1 analog consisted of 2'-O-methylribonucleotides except for the potential ligand binding sites, which consisted of deoxynucleotides with the same nucleobases as RNA R1. X-ray crystallography and cryo-electron microscopy have been employed to determine the 3-D structure of aminoglycoside-bound 16S rRNA, as shown in Figure 3.2.<sup>44-46</sup> One aminoglycoside ligand binds into the groove containing A1408, A1492, and A1493 of 16S rRNA.<sup>44-46</sup>



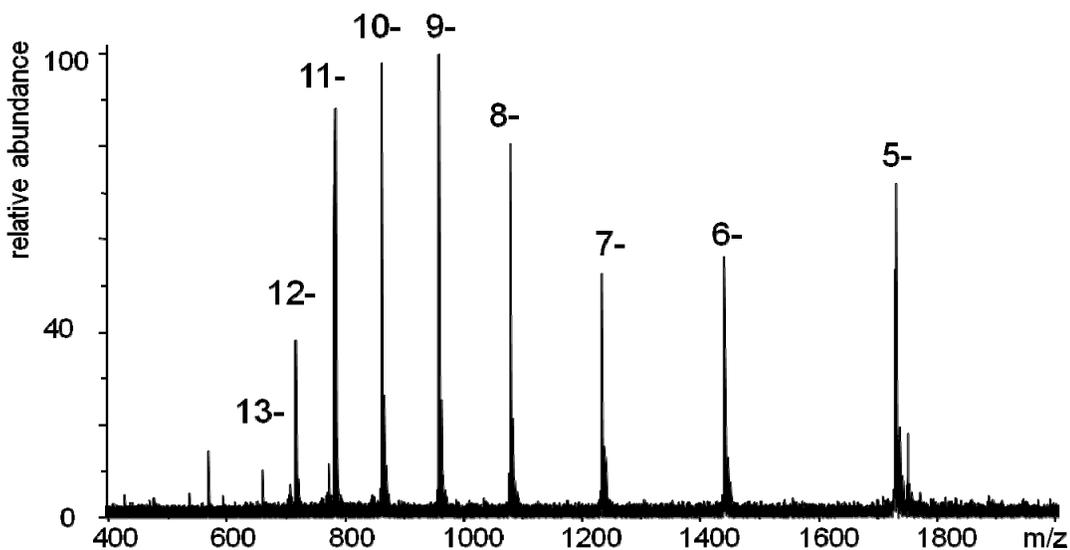
**Figure 3.1.** Sequence and secondary structure of the 27-mer RNA R1. The numbering is relative to 16S ribosomal RNA. Adapted from Griffey *et al.*<sup>43</sup>



**Figure 3.2.** 3D structure of aminoglycoside-bound sites on tRNA, mRNA, and 70S prokaryotic rRNA. 70S rRNA incorporates 16S rRNA as a subunit. Figure adapted from [44].

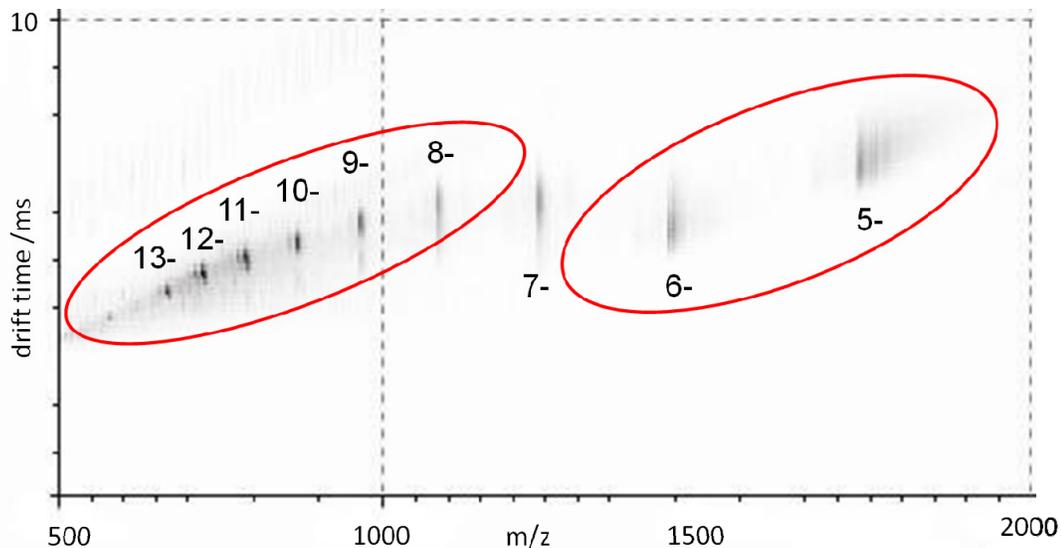
R1 was electrosprayed from a solution containing 1% triethylamine and 50% methanol in water. The electrospray solution was adapted from Taucher *et al.*,<sup>25</sup> who demonstrated that RNAs up to 61 nt could be dissociated by CAD and EDD. We observed charge states ranging from 5- to 13-, as shown in Figure 3.3. The average

charge density is  $\sim 0.3$  per nucleotide (nt), which is lower than  $\sim 0.4$  for a 34-mer riboswitch aptamer RNA as reported previously.<sup>22</sup> The latter 34-mer RNA has been proposed to contain less base-pairing than the present 27-mer RNA, and thus should have a less ordered structure.<sup>47</sup>



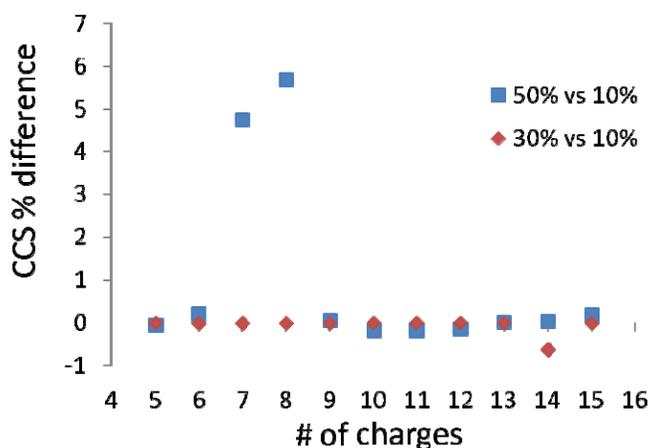
**Figure 3.3.** Negative ion ESI spectrum of the 27-mer RNA R1.

There are two charge state distributions, or two “envelopes”, in the ESI spectrum of RNA R1, which is an indication that more than one gas-phase conformation may exist simultaneously. An ion mobility mass spectrum of RNA R1 shows two distinct groups of conformations (Figure 3.4).



**Figure 3.4.** IM-MS of RNA R1 dissolved in 1% triethylamine and 50% (v/v) methanol in water.

Collision cross sections (CCSs), or sizes of RNA R1 anions with higher charge states (8- to 13-) are larger than lower charge states (6- to 5-), demonstrating that anions with more charges are more unfolded and anions with less charges are more folded. The CCS of the 7- charge state lies between the two groups, indicating that 7- anions are partially unfolded. Within the group of higher charge states, CCSs at higher charge states are slightly larger than for lower charge states. When the RNA R1 was electrosprayed from 25 mM ammonium acetate in 50% (v/v) methanol in water, similar overall CCSs were observed. The percentage of organic solvent in the electrospray solution may impact the gas phase structure. The differences of observed CCSs of RNA R1 anions electrosprayed from 10%, 30%, and 50% methanol in water are summarized in Figure 3.5.



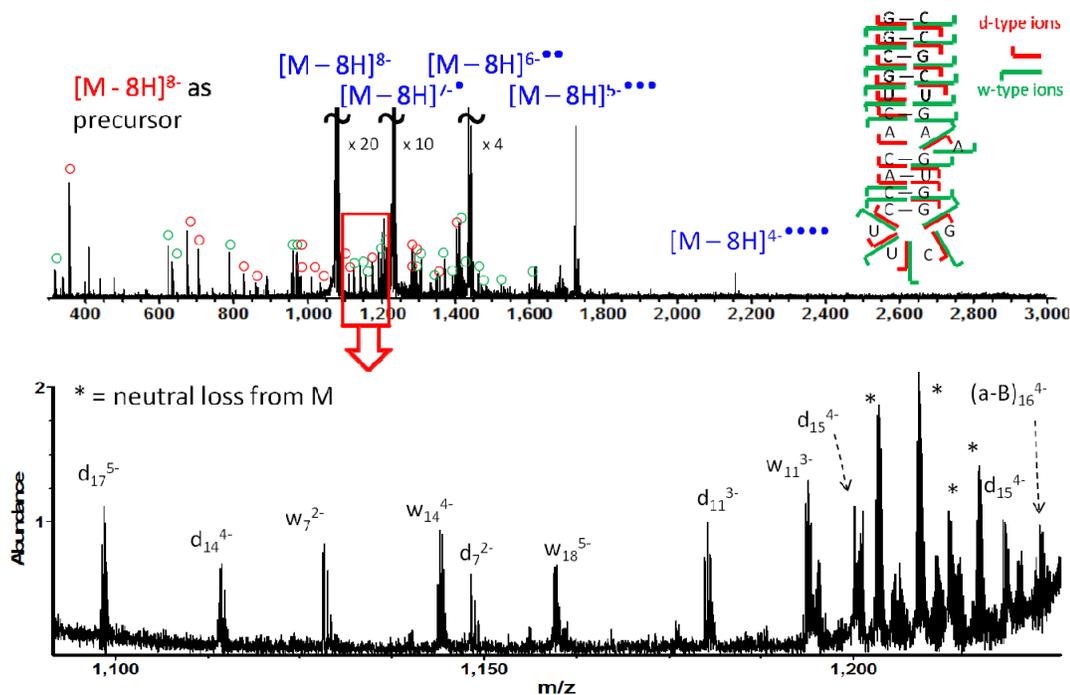
**Figure 3.5.** CCS comparison of RNA R1 electrospayed from solutions with different methanol content.

CCSs of RNA R1 at the 7- and 8- charge states electrospayed from 50% methanol are larger than from 10% methanol while, at other charge states, CCSs are similar. This result indicates that 50% methanol content aids the unfolding of RNA R1 at the 7- and 8- charge states while it does not affect gas-phase structures of RNA R1 at other charge states. Based on these data, it can be proposed that RNA gas-phase structure is strongly dependent on charge state. Also, lower methanol content does not aid the preservation of RNA folding at certain charge states. RNA R1 electrospayed from 30% methanol content share similar sizes as from 10% methanol content for all observed charge states. Up to 10% organic solvent was used in previous RNA-ligand binding specificity studies.<sup>11, 43</sup>

### 3.3.2 Dependence of EDD and CAD Fragmentation Efficiency on Precursor Ion Charge State

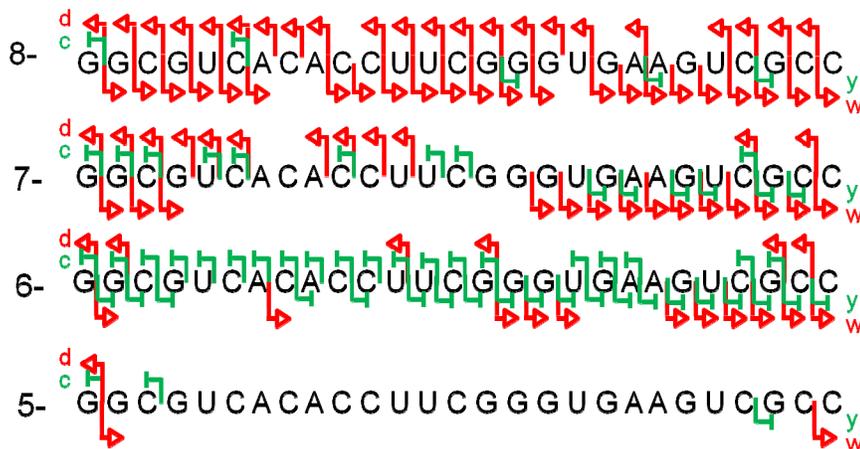
EDD and CAD were performed on RNA R1 at charge states from 5- to 8-, electrospayed from 50% methanol (v/v) in water with 1% trimethylamine. Infrared laser irradiation was applied prior to EDD to activate the precursor ions. Absence of IR

activation resulted in limited EDD backbone cleavages. An activated ion (AI)-EDD spectrum of  $[M - 8H]^{8-}$  is shown in Figure 3.6.



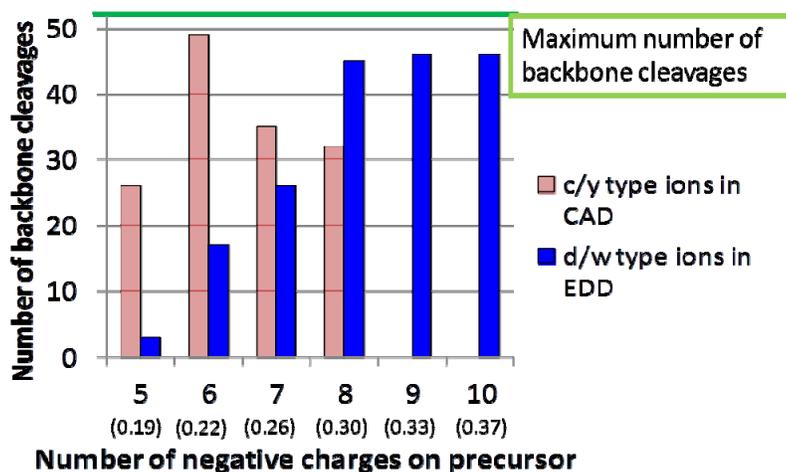
**Figure 3.6.** AI-EDD spectrum of the RNA R1 at a charge state of 8-. IR activation prior to EDD was performed by irradiating the  $[M - 8H]^{8-}$  ion with a 10.6  $\mu\text{m}$  laser beam at 0.125 W for 150 ms. EDD was performed by irradiating precursor ions with electrons for 0.4 s. Electrons were extracted at a voltage of -25 V and filtered by a lens voltage set at -24 V. Upper panel: *d* and *w*-type ions are labeled as circles.

Backbone *d*- and *w*-type ions observed in AI-EDD of different charge states are mapped to the RNA R1 sequence, as shown in Figure 3.7.



**Figure 3.7.** Backbone cleavages observed in AI-EDD of the RNA R1 at charge states of 5- to 8-.

AI-EDD of the RNA R1 at the 10- charge state provided 46 out of 52 possible backbone cleavages. However, AI-EDD of the 5- charge state only provided three backbone cleavages in the form of *d/w*-type ions. Figure 3.8 illustrates the comparison of the numbers of *d/w*-type ions observed in AI-EDD spectra from different precursor charge states, as well as the numbers of *c/y*-type ions observed in CAD spectra.



**Figure 3.8.** Charge dependence of EDD and CAD efficiency. Charge density (charge/nt) is noted under each number of negative charges.

As the charge state of the precursor ions increased from 5- to 8-, AI-EDD provided more backbone cleavages in the form of *d/w*-type ions. The number of *d/w*-type ions does not significantly change between the 8-, 9- and 10- precursor ion charge states. A dependence of fragmentation efficiency on precursor charge state has been observed in other ion-electron reactions such as ECD of peptides<sup>48</sup> and ETD of peptides.<sup>49</sup> The effect of precursor charge state on the number of observed backbone fragments may be due to two reasons: One reason is the fact that, in FT-ICR detection, signal from image current is proportional to the number of charges. Product ions from precursor ions with more charges may carry more charges and be detected with higher sensitivity. The other reason is the correlation between gas-phase size and number of charges. Coulomb repulsion between charges on the RNA may facilitate unfolding of the molecule in the gas phase, similar to proteins.<sup>50</sup> A larger cross-section increases the chance of overlap between precursor anions and the electron cloud during the EDD event. Also a more unfolded gas-phase structure contains less non-covalent bonding, which may hinder dissociation of complementary fragments after backbone cleavage occurs. Figure 3.4 shows that RNA anions are more expanded as the charge state increased from 6- to 8-. The size expansion from 8- (0.30 charges/nt) to higher charge states is not as significant as from 6- to 8-. The trend of size expansion with charge increase is consistent with the trend of increased observation of *d/w*-type ions. Size expansion in the gas phase along with charge increase was also observed for three DNAs (28-mer, 40-mer, and 55-mer).<sup>51</sup> In the latter study the gas-phase size expanded dramatically with the charge increase when the charge density was lower than 0.30 charges/nt. By contrast the size expansion was not as significant when the charge density was higher than 0.30 charges/nt.

Interestingly, in this study of RNA R1, 0.3 charges/nt is also the turning point in the trend line of size expansion and the trend line of the number of observed *d/w*-type ions.

Taucher *et al.* also compared EDD fragmentation efficiencies of 22-mer RNA anions at two different charge states.<sup>25</sup> Both charge states had higher charge densities than 0.30 charges/nt (0.45 and 0.59, respectively) and their fragmentation efficiencies were similar. Our EDD and IM-MS data, combined with previously published results from EDD and CCS measurements, support the existence of a correlation between EDD fragmentation efficiency, gas-phase anion size, and precursor charge state.

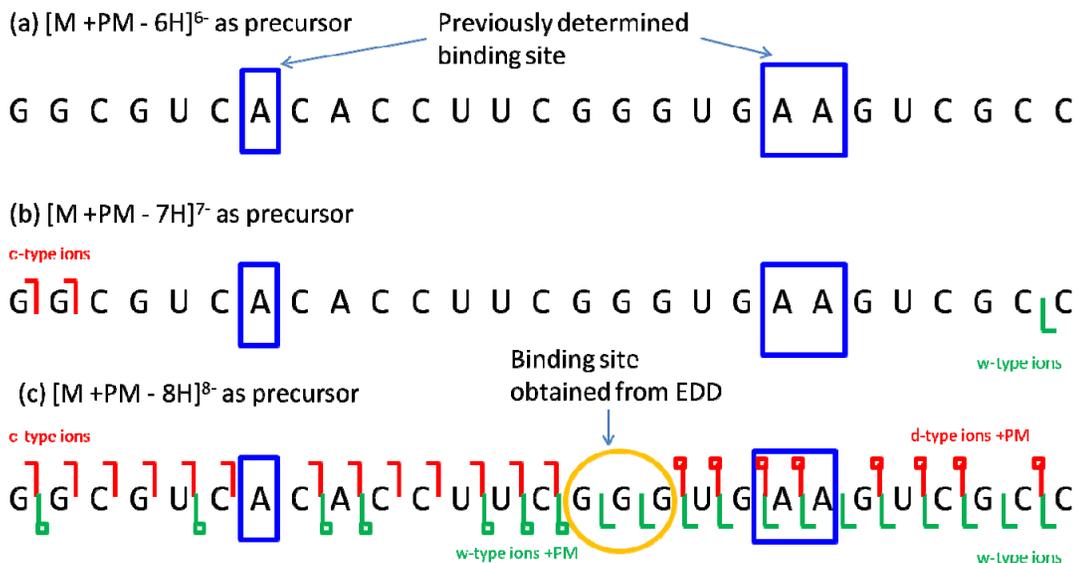
In the case of CAD, as the charge state of the precursor ions increased from 5- to 8-, the number of backbone cleavages in the form of *c/y*-type ions did not increase. As a vibrational activation technique, CAD disrupts non-covalent bonding within the gas-phase structure of an oligonucleotide.<sup>52</sup> Thus, the number of observed backbone cleavages did not significantly decrease as the precursor charge state decreased.

### **3.3.3 EDD and CAD of RNA-ligand Complexes**

Paramomycin (PM) is known to bind specifically to 16S RNA at A1408, A1492 and A1493, as revealed by CAD of the RNA R1 analog complexed with PM.<sup>11,43</sup> This CAD approach required synthesis of the RNA analog. Here, the feasibility of using EDD to detect ligand binding sites directly from the RNA-ligand complex, without the synthesis of RNA analogs, is evaluated. The R1-PM complex was chosen as a model system for this investigation because the PM binding site was already confirmed.

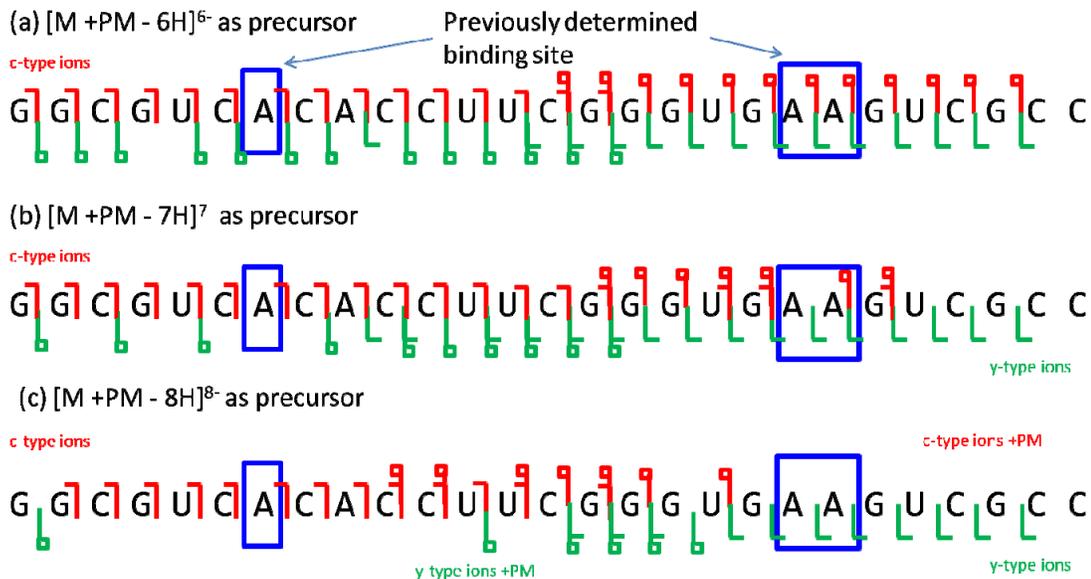
RNA R1 was mixed with PM at a 1:1 molar ratio and electrosprayed from 50% methanol (v/v) in water with 1% trimethylamine. AI-EDD experiments of R1-PM

complexes with 6 to 8 negative charges were performed. Observed fragments are mapped to the backbone, as shown in Figure 3.9.



**Figure 3.9.** *d/w*-type ions observed in AI-EDD experiments of R1-PM complexes at (a) 6-, (b) 7-, and (c) 8- charge states. PM binding sites derived from *d/w* ion series are circled and PM binding sites obtained previously<sup>11</sup> are shown in the boxes.

Compared to the AI-EDD experiments of free RNA R1, AI-EDD of R1-PM complexes at 6- and 7- charge states yielded less backbone fragments and PM binding sites could not be determined from *d/w* ion series. The low fragmentation efficiency may be due to the noncovalent interactions introduced from ligand binding, as observed previously.<sup>14, 53</sup> Here we observed that AI-EDD of R1-PM complexes at the 8- charge state yielded almost complete sequence coverage. From *d/w*-type ions the PM binding site was directly localized, as shown in Figure 3.9. However, the obtained binding site was not consistent with the previously reported site.<sup>11</sup> CAD of the RNA R1 at the 6-, 7-, and 8- charge states yielded almost complete sequence coverage, as shown in Figure 3.10. However, the PM binding sites determined from CAD were also not consistent with the previously reported site.<sup>11</sup>



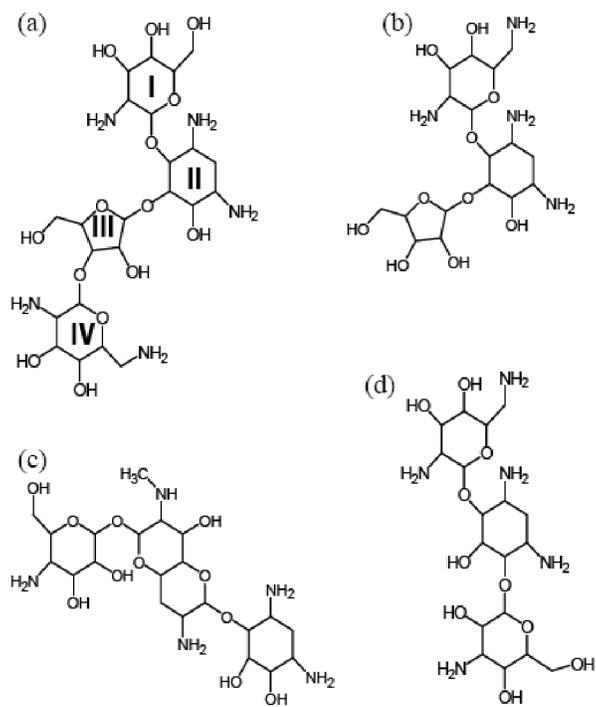
**Figure 3.10.** *c/y*-type ions observed in CAD experiments of R1-PM complexes at (a) 6-, (b) 7-, and (c) 8- charge states. PM binding sites obtained previously<sup>11</sup> are shown in the boxes.

From both AI-EDD and CAD of R1-PM complexes the observed binding sites were not consistent with results from CAD of the RNA R1 analog, X-ray crystallography, or cryo-electron microscopy.<sup>44-46</sup> This result is an indication that the complexes observed in the mass spectra may not be due to specific binding. The binding sites obtained from AI-EDD and CAD experiments may thus be from non-specific binding. Further experiments, such as a systematic investigation of solvent conditions, to confirm this hypothesis are needed.

Another observation was that, in CAD experiments of the R1-PM complex, *c/y* fragments containing the ligand were observed, indicating that ligand loss was energetically less favored than backbone dissociation. Such strong non-covalent interactions have also been observed in DNA/RNA-metal complexes and DNA/RNA-small molecule complexes.<sup>53</sup>

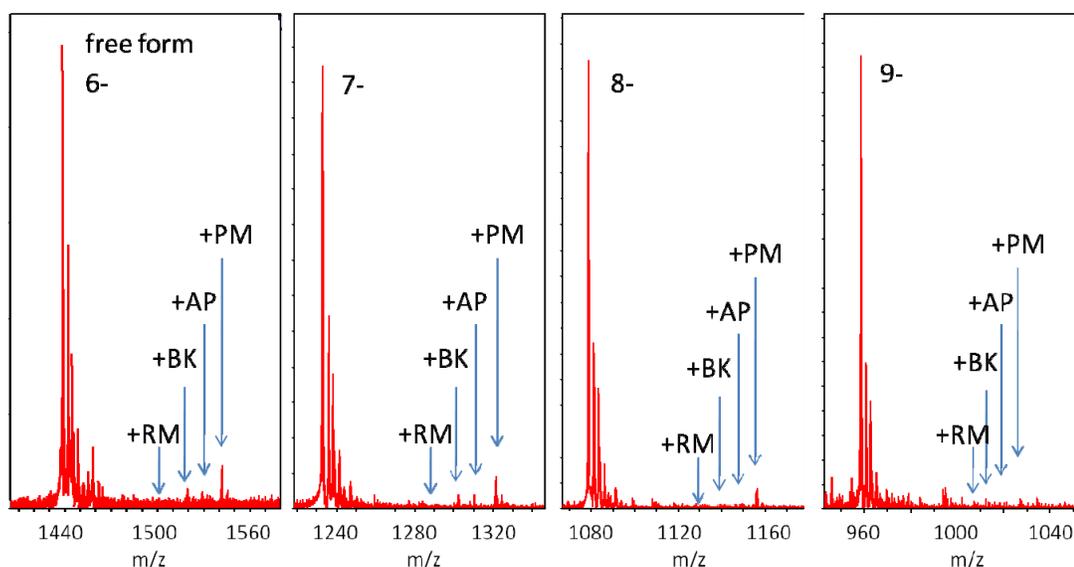
### 3.3.4 Correlation between Specific Binding and the RNA R1 Charge State

As shown above, AI-EDD was only able to dissociate RNA R1 at the 8- charge state and higher. Also, the binding site obtained from AI-EDD experiments was not consistent with the previously reported site, indicating that the R1-PM complex with 8 negative charges may be a result of nonspecific binding. It is reasonable to propose a correlation between specific binding and RNA R1 charge state. A competitive binding experiment was performed to investigate this correlation. RNA R1 has been demonstrated to form complexes with four aminoglycosidic ligands, including paromomycin (PM), (Be)kanamycin (BK), apramycin (AP), and ribostamycin (RM), as shown in Figure 3.11. Binding affinities were measured by ESI-MS to be 0.11  $\mu\text{M}$ , 2  $\mu\text{M}$ , 2  $\mu\text{M}$ , and 16  $\mu\text{M}$ , respectively.<sup>11</sup>



**Figure 3.11.** Structures of aminoglycosidic ligands: (a) paromomycin (PM), (b) (Be)kanamycin (BK), (c) apramycin (AP), and (d) ribostamycin (RM).

Among the four ligands, the binding affinity order is  $PM > BK = AP > RM$ . For a competitive binding experiment, a mixture containing R1, PM, BK, AP, and RM with the ligands at equal concentration was electrosprayed. If all ligands bind to R1 specifically, the abundances of R1-ligand peaks in the mass spectrum should follow the order of binding affinity. Figure 3.12 shows ESI-MS of a mixture containing  $2 \mu\text{M}$  RNA R1 and  $0.1 \mu\text{M}$  of PM, BK, AP, and RM. According to previously measured binding affinities, signal abundances of R1-PM, R1-BK, R1-AP, and R1-RM complexes at the same charge state should be at a ratio of 1:0.52:0.52:0.12.



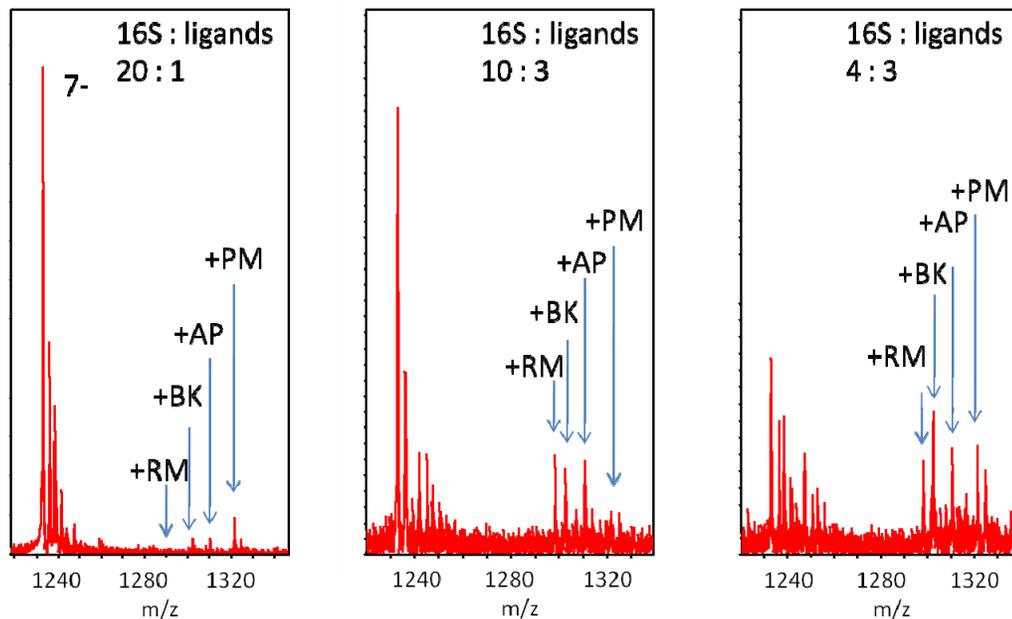
**Figure 3.12.** ESI-MS of RNA R1 and ligand mixture. The mixture contained 2  $\mu\text{M}$  RNA R1 and 0.1  $\mu\text{M}$  of PM, BK, AP, and RM. The mixture was electrosprayed in 1% TEA and 50% methanol in water (v/v).

From Figure 3.12, the abundance ratio of R1-ligand complexes (with PM, BK, AP, and RM respectively) at the 6- charge state was 1:0.48:0.43:0.20, whereas the abundance ratio of R1-ligand complexes at the 7- charge state was 1:0.37:0.48 with the R1-RM complex not observed. In the case of 8- and 9- charge states, only the R1-PM complex was observed. As for charge states 10- and higher, no R1 complex was observed at all. In general, binding affinities between ligands and RNA R1 at charge states 6- and 7- were close to previously reported values. Thus complexes observed at 6- and 7- charge states from ESI-MS of the solution with an RNA-ligand molar ratio of 20:1 are suggestive of specific binding. Binding between ligands and R1 at charge states 8- and higher was weaker than for the 6- and 7- charge state. This correlation between specific binding and charge state supports the observed difference in gas-phase structure of RNA R1 at different charge states (Figure 3.4). As probed by IM-MS, RNA R1 at the 5- and 6- charge states is more folded than at the 8- and higher charge states. Combining the

results from the competitive binding experiment and IM-MS, it is likely that RNA R1 at the 5- and 6- charge states is folded in its native form, and R1 at 8- and higher charge states is unfolded. Little RNA-ligand formation at 8- and higher charge states also challenges the sensitivity of AI-EDD experiments. In practice, AI-EDD of R1-PM complexes at the 8- charge state under the experimental conditions in Figure 3.12 yielded no detectable product ions, due to the low abundance of precursor ions.

### **3.3.5 Correlation between Binding Specificity and RNA-Ligand Molar Ratio**

As shown above, R1-ligand complexation could be specific at charge states 6- and 7-, when R1 and ligands were prepared in solution at a ratio of 20:1. However, CAD of R1-PM complexes at the 6-, 7- and 8- charge states, when R1 and ligand were prepared in solution at a ratio of 1:1, revealed that PM binding was mostly nonspecific. Another competitive binding experiment was performed to investigate the correlation between binding specificity and RNA ligand molar ratio in the sample solution. Mixtures containing RNA R1, PM, BK, AP, and RM at different ratios were subjected to ESI-MS, as shown in Figure 3.13.



**Figure 3.13.** ESI-MS of RNA R1 and ligand mixture. The mixture contained 2  $\mu\text{M}$  RNA R1 and equal molar amounts of PM, BK, AP, and RM at (left panel) 0.1  $\mu\text{M}$ , (middle panel) 0.6  $\mu\text{M}$ , and (right panel) 1.5  $\mu\text{M}$ . The mixture was electrosprayed in 1% TEA and 50% methanol in water (v/v).

From Figure 3.14, the observed abundance ratio of R1-ligand complexes at the 7- charge state was 1:0.37:0.48 (the R1-RM complex was not observed) when R1 and ligands were mixed at a ratio of 20:1. When R1 and ligands were mixed at a ratio of 10:3, the observed abundance ratio of R1-ligand complexes was 1:1.85:1.74:1.79 and when R1 and ligands were mixed at a 4:3 ratio, the observed abundance ratio of complexes was 1:1.03:1.07:0.96. Only an R1/ligand molar ratio of 20:1 yielded observed ratios close to the predicted value based on previously reported binding constants. This observation indicates that, as RNA ligand molar ratio decreases, nonspecific binding increases. The observed correlation between binding specificity and RNA-ligand molar ratio may support the observation that binding sites obtained from CAD of R1-PM complexes at the 6-, 7-, and 8- charge states were not the same as the specific binding sites previously reported.

### 3.4 Conclusions

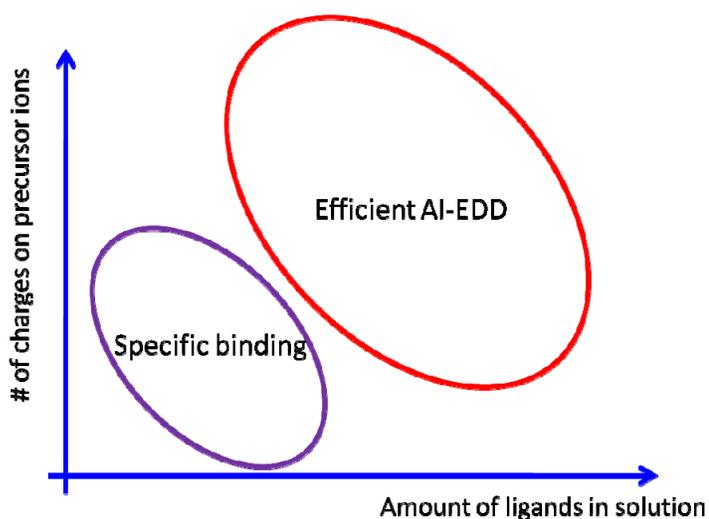
In this Chapter we first investigated the correlation between gas-phase structures of RNA R1 and its charge state. IM-MS showed unfolding of R1 as the number of charges increases. The transition between folded and more unfolded forms occurred between the 7- and 8- charge states. When RNA R1 was electrosprayed from 50% methanol, R1 at the 6- charge state was observed to be folded, the 7- charge state was partially unfolded, and the 8- charge state appeared fully unfolded. This trend was further confirmed by competitive binding experiments. When RNA R1 was electrosprayed from less (30% or 10%) methanol content, R1 at the 8- charge state was observed folded and the 9- charge state appeared to be fully unfolded, from IM-MS.

AI-EDD and CAD were applied to dissociate RNA R1 at different charge states. Sequence coverage from AI-EDD increased as the charge state of precursor ions increased. Full sequence coverage was only obtained when the number of precursor ion charges was higher than 8. Increased number of charges could contribute to the unfolding of RNA gas-phase structure and also the improvement of image current detection in FT-ICR-MS.

AI-EDD and CAD were also applied to dissociate R1 complexed with the aminoglycosidic antibiotic drug paramomycin (PM). The complex was electrosprayed from a solution containing RNA and ligand at a molar ratio of 1:1. AI-EDD of the R1-PM complex at the 6- and 7- charge states provided limited sequence coverage and the binding site could not be determined from *d/w* ion series. AI-EDD of R1-PM at the 8- charge state provided sufficient sequence coverage to determine the binding site.

However, the determined binding site was not consistent with the specific binding site reported previously. CAD of R1-PM at the 6-, 7-, and 8- charge states provided sufficient sequence coverage to locate the binding sites. Unfortunately the determined binding sites were also not consistent with specific binding.

Further competitive experiments revealed that specific binding was dependent on both charge state and R1-ligand molar ratio in solution. General trends were: first, more charges correspond to less specific binding and, second, the more ligand in solution the less specific binding. However, AI-EDD requires both relatively high precursor ion charge state and abundant precursor ions while, in practice, both requirements decrease the binding specificity, as illustrated in Scheme 3.1. So far we have not found an appropriate experimental condition to obtain both good AI-EDD efficiency and good binding specificity. However, further fine-tuning of experimental conditions may help finding a “sweet spot” to achieve both.



**Scheme 3.1.** Schematic summary of the correlation between AI-EDD efficiency, binding specificity, precursor charge state, and ligand amount in the electrospray solution.

Overall, given the correlation between AI-EDD efficiency, binding specificity, precursor charge state, ligand amount in electrospray solution, and the organic solvent content, care should be taken when ESI-MS and AI-EDD are applied for the investigation of RNA-ligand complexes to avoid nonspecific binding.

### 3.5 Bibliography

- (1) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.
- (2) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.
- (3) Chait, B. T. *Science* **2006**, *314*, 65-66.
- (4) Smith, R. D.; Loo, J. A.; Edmonds, C. G.; Barinaga, C. J.; Udseth, H. R. *Anal. Chem.* **1990**, *62*, 882-899.
- (5) Muddiman, D. C.; Null, A. P.; Hannis, J. C. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1201-1204.
- (6) Schultz, J. C.; Hack, C. A.; Benner, W. H. *J. Am. Soc. Mass. Spectrom.* **1998**, *9*, 305-313.
- (7) Huang, T. Y.; Liu, J. A.; McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 890-898.
- (8) Huang, T. Y.; Liu, J.; Liang, X. R.; Hodges, B. D. M.; McLuckey, S. A. *Anal. Chem.* **2008**, *80*, 8501-8508.
- (9) Cheng, X.; Harms, A. C.; Goudreau, P. N.; Terwilliger, T. C.; Smith, R. D. *Proc. Natl. Acad. Sci.* **1996**, *93*, 7022-7027.
- (10) Hagan, N.; Fabris, D. *Biochemistry* **2003**, *42*, 10736-10745.
- (11) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 10129-10133.
- (12) Hofstadler, S. A.; Sannes-Lowery, K. A.; Crooke, S. T.; Ecker, D. J.; Sasmor, H.; Manalili, S.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 3436-3440.
- (13) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

- (14) Keller, K. M.; Breeden, M. M.; Zhang, J.; Ellington, A. D.; Brodbelt, J. S. *J. Mass Spectrom.* **2005**, *40*, 1327-1337.
- (15) Turner, K. B.; Hagan, N. A.; Fabris, D. *J. Mol. Biol.* **2007**, *369*, 812-828.
- (16) McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 599-614.
- (17) Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1994**, *116*, 4893-4897.
- (18) Little, D. P.; Speir, J. P.; Senko, M. W.; Oconnor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.
- (19) Sannes-Lowery, K. A.; Hofstadler, S. A. *J. Am. Soc. Mass. Spectrom.* **2003**, *14*, 825-833.
- (20) Tromp, J. M.; Schürch, S. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 2348-2354.
- (21) Huang, T. Y.; Kharlamova, A.; Liu, J.; McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 1832-1840.
- (22) Taucher, M.; Rieder, U.; Breuker, K. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 278-285.
- (23) McLuckey, S. A.; Vanberkel, G. J.; Glish, G. L. *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 60-70.
- (24) Little, D. P.; Aaserud, D. J.; Valaskovic, G. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **1996**, *118*, 9352-9359.
- (25) Taucher, M.; Breuker, K. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 918-929.
- (26) Keller, K. M.; Brodbelt, J. S. *Anal. Biochem.* **2004**, *326*, 200-210.
- (27) Hofstadler, S. A.; Griffey, R. H.; Pasa-Tolic, L.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1400-1404.
- (28) Crowe, M. C.; Brodbelt, J. S. *J. Am. Soc. Mass. Spectrom.* **2004**, *15*, 1581-1592.
- (29) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
- (30) Zubarev, R. A. *Mass Spectrom. Rev.* **2003**, *22*, 57-77.
- (31) Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom. Rev.* **2005**, *24*, 201-222.
- (32) Hakansson, K.; Hudgins, R. R.; Marshall, A. G.; O'Hair, R. A. J. *J. Am. Soc. Mass. Spectrom.* **2003**, *14*, 23-41.

- (33) Schultz, K. N.; Hakansson, K. *Int. J. Mass spectrom.* **2004**, *234*, 123-130.
- (34) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.
- (35) Yang, J.; Mo, J. J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, *77*, 1876-1882.
- (36) Yang, J.; Hakansson, K. *J. Am. Soc. Mass. Spectrom.* **2006**, *17*, 1369-1375.
- (37) Yang, J.; Hakansson, K. *Int. J. Mass spectrom.* **2008**, *276*, 144-148.
- (38) Xu, Y.; Afonso, C.; Wen, R.; Tabet, J.-C. *J. Mass Spectrom.* **2008**, *43*, 1531-1544.
- (39) Kinet, C.; Gabelica, V. r.; Balbeur, D. e.; De Pauw, E. *Int. J. Mass spectrom.* **2009**, *283*, 206-213.
- (40) Yang, J.; Hakansson, K. *Eur. J. Mass Spectrom.* **2009**, *15*, 293-304.
- (41) Mo, J. J.; Hakansson, K. *Anal. Bioanal. Chem.* **2006**, *386*, 675-681.
- (42) Limbach, P. A.; Crain, P. F.; McCloskey, J. A. *J. Am. Soc. Mass. Spectrom.* **1995**, *6*, 27-39.
- (43) Griffey, R. H.; Greig, M. J.; An, H.; Sasmor, H.; Manalili, S. *J. Am. Chem. Soc.* **1998**, *121*, 474-475.
- (44) Selmer, M.; Dunham, C. M.; Murphy, F. V.; Weixlbaumer, A.; Petry, S.; Kelley, A. C.; Weir, J. R.; Ramakrishnan, V. *Science* **2006**, *313*, 1935-1942.
- (45) Borovinskaya, M. A.; Pai, R. D.; Zhang, W.; Schuwirth, B. S.; Holton, J. M.; Hirokawa, G.; Kaji, H.; Kaji, A.; Cate, J. H. D. *Nat Struct Mol Biol* **2007**, *14*, 727-732.
- (46) Feldman, M. B.; Terry, D. S.; Altman, R. B.; Blanchard, S. C. *Nat Chem Biol* **2010**, *6*, 54-62.
- (47) Rieder, U.; Lang, K.; Kreutz, C.; Polacek, N.; Micura, R. *ChemBioChem* **2009**, *10*, 1141-1144.
- (48) Kalli, A.; Hakansson, K. *J. Proteome. Res.* **2008**, *7*, 2834-2844.
- (49) Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. *Mol. Cell. Proteomics* **2007**, *6*, 1942-1951.
- (50) Chowdhury, S. K.; Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1990**, *112*, 9012-9013.

- (51) Moradian, A.; Scalf, M.; Westphall, M. S.; Smith, L. M.; Douglas, D. J. *Int. J. Mass spectrom.* **2002**, *219*, 161-170.
- (52) Schnier, P. D.; Klassen, J. S.; Strittmatter, E. E.; Williams, E. R. *J. Am. Chem. Soc.* **1998**, *120*, 9605-9613.
- (53) Keller, K. M.; Zhang, J.; Oehlers, L.; Brodbelt, J. S. *J. Mass Spectrom.* **2005**, *40*, 1362-1371.

## Chapter 4

### Top-down MS/MS of Proteins in Negative Ion Mode

#### 4.1. Introduction

The development of modern mass spectrometers together with the rapid advances in genome sequencing have enabled the field of proteomics. Two complementary approaches, i.e., “bottom-up”<sup>1</sup> and “top-down”<sup>2, 3</sup> analysis have evolved for proteomic research. The more widely utilized bottom-up approach targets peptides with masses less than a few kDa. Such peptide mixtures are either naturally present or generated by protease digestion. Top-down proteomics, on the other hand, involves introduction of undigested proteins into the mass spectrometer. Masses of whole proteins are measured and further dissociation of protein ions (tandem mass spectrometry or MS/MS) provides structural information. Several MS/MS strategies are employed in top-down proteomics, including collision activated dissociation (CAD),<sup>4-9</sup> infrared multiphoton dissociation (IRMPD),<sup>10</sup> blackbody infrared radiative dissociation (BIRD),<sup>11, 12</sup> electron capture dissociation (ECD),<sup>13-16</sup> and electron transfer dissociation (ETD).<sup>17</sup> The three former techniques are referred to as “slow-heating” methods<sup>18</sup> because activation of precursor ions is achieved by accumulation of energy from multiple collisions with collision gas

molecules or from absorption of multiple infrared photons. Due to intramolecular vibrational energy redistribution (IVR),<sup>19</sup> vibrational energy gained from collisions or from photon absorption is redistributed to thermodynamically favored fragmentation pathways. Thus, facile losses of labile post-translational modifications (PTMs) frequently dominate MS/MS spectra of peptides.<sup>20</sup> For example, for peptides phosphorylated on serine or threonine, abundant loss of phosphate and/or phosphoric acid is observed.<sup>21, 22</sup> There are examples showing that labile PTM loss is less frequent when intact proteins (>9 kDa) are fragmented.<sup>23, 24</sup> This difference between bottom-up and top-down analysis may be due to gas-phase protein higher-order structures that stabilize PTMs and thus favor fragmentation at backbone amide bonds rather than PTM loss.<sup>25</sup> In contrast to vibrational activation of even-electron protein ions, ECD involves gas-phase radical ion chemistry<sup>13, 26</sup> following capture of low-energy electrons (<2 eV) by multiply protonated peptide and protein ions. Backbone N-C<sub>α</sub> bonds are cleaved to produce *c* and *z*<sup>•</sup>-type ions without loss of PTMs.<sup>27-29</sup> Analogous fragmentation behavior is observed in electron transfer dissociation (ETD), which also involves radical intermediates but from ion-ion reactions.<sup>30</sup>

Phosphorylation is one of the most common PTMs and plays critical roles in various biological processes such as cell signaling.<sup>31</sup> Top-down proteomics provides benefits in the characterization of phosphoproteins, including determination of phosphorylation sites, and possible quantification of site-specific phosphorylation.<sup>32, 33</sup> ECD and ETD, in particular, provide extensive backbone cleavages without phosphate or phosphoric acid loss.<sup>28, 30, 34</sup>

To date, the vast majority of top-down approaches involves positive ion mode in which proteins are typically electrosprayed in acidified solvent. Protein N-termina and basic amino acid residues are protonated or become protonated during the electrospray process. However, for phosphorylated proteins, the high acidity of phosphate affects the overall protonation following electrospray ionization. Furthermore, salt bridges may form between negatively charged phosphates and protonated groups if favorable steric configurations exist.<sup>35, 36</sup> Because ECD at conventional low electron energies does not disrupt non-covalent interactions,<sup>37-39</sup> backbone cleavages near phosphorylated sites may not be observed. Lack of such cleavages is unfortunate because they are frequently required to precisely assign phosphorylation sites. Higher electron energy can improve ECD efficiency, presumably by disrupting noncovalent bonds and unfolding peptide/protein gas-phase structure.<sup>36</sup> Alternatively, an infrared laser pulse or collisions with neutral gas before or after electron irradiation may unfold the gas-phase protein structure and consequently increase ECD backbone cleavage coverage, an approach termed “activated ion” ECD (AI-ECD).<sup>40, 41</sup> In “plasma” ECD,<sup>42</sup> protonated protein cations are activated by gas collisions. In addition, electrons are decelerated and forced to travel in the same direction as the precursor ions to maximize the cross-section of the ion-electron reaction. However, even after AI-ECD or plasma ECD, backbone cleavages near phosphorylated sites are scarcer than in non-phosphorylated protein regions.<sup>42</sup>

Negative ion mode, in which proteins are typically electrosprayed in basic solvent, promotes deprotonation of protein C-termina, acidic PTMs, acidic amino acid residues, and, to a certain extent, backbone amides. In regions with multiple acidic PTMs or acidic amino acid residues, electrostatic forces may contribute to unfolding of the local

noncovalent tertiary structure, thus potentially enhancing the observation of backbone cleavages from such regions compared to positive ion mode. Formation of salt bridges may also be circumvented because protonation is suppressed by the basic electrospray solvent. In bottom-up approaches, CAD and IRMPD have been demonstrated to generate useful structural information for deprotonated peptides.<sup>43-47</sup> However, vibrational activation of peptide anions typically results in abundant internal fragments, neutral losses, and side chain losses, in addition to the backbone cleavages that dominate in positive ion mode.<sup>45, 47</sup> Thus, MS/MS spectra are more complex and less useful structural information is generated in negative ion mode. Electron detachment dissociation (EDD),<sup>48</sup> operates in negative ion mode and involves higher electron energy (>10 eV) than ECD. EDD has been shown to yield mainly *a*<sup>•</sup> and *x*-type product ions for peptides and proteins,<sup>48-50</sup> mainly *d* and *w*-type ions for nucleic acids,<sup>51-54</sup> and extensive glycosidic and cross-ring fragments for oligosaccharides.<sup>55, 56</sup> However, one drawback of EDD of polypeptides is its low fragmentation efficiency, compared to other fragmentation methods.<sup>57</sup>

Top-down analysis of large biomolecules in negative ion mode has only recently begun to be explored. Taucher and Breuker reported ~90% sequence coverage for an 18.2 kDa ribonucleic acid in negative ion CAD<sup>52</sup> and McLuckey and co-workers demonstrated ~60% sequence coverage for a 25 kDa intact tRNA.<sup>58</sup> Very recently, Breuker and co-workers explored EDD for top-down protein analysis and reported ~19% sequence coverage for an acidic 147-residue protein.<sup>50</sup> However, to our knowledge, negative ion top-down analysis of proteins with acidic modifications, including phosphorylation and glycosylation has not previously been explored. Here we employ

negative ion IRMPD and AI-EDD for top-down analysis of three proteins, including multiply phosphorylated  $\beta$ -casein, the acidic protein calmodulin, and glycosylated ribonuclease B.

## **4.2. Experimental Section**

### **4.2.1. Sample Preparation**

For negative ion mode electrospray ionization (ESI), protein solvent consisted of 0.2% v/v piperidine (Sigma-Aldrich, St. Louis, MO) in 1:1 v/v HPLC grade water (Fisher Scientific, Fair Lawn, NJ) /HPLC grade methanol (Fisher Scientific, Fair Lawn, NJ). For positive ion mode ESI, the solvent was 0.5% v/v formic acid (Acros Organics, Morris Plains, NJ) in 1:1 v/v water/methanol.  $\beta$ -casein from bovine milk, calmodulin from bovine testes, and RNase B from bovine pancreas (Sigma-Aldrich) were dissolved separately into electrospray solution and infused via an Apollo II electrospray ion source at 60  $\mu$ L/h.

### **4.2.2. Mass Spectrometry**

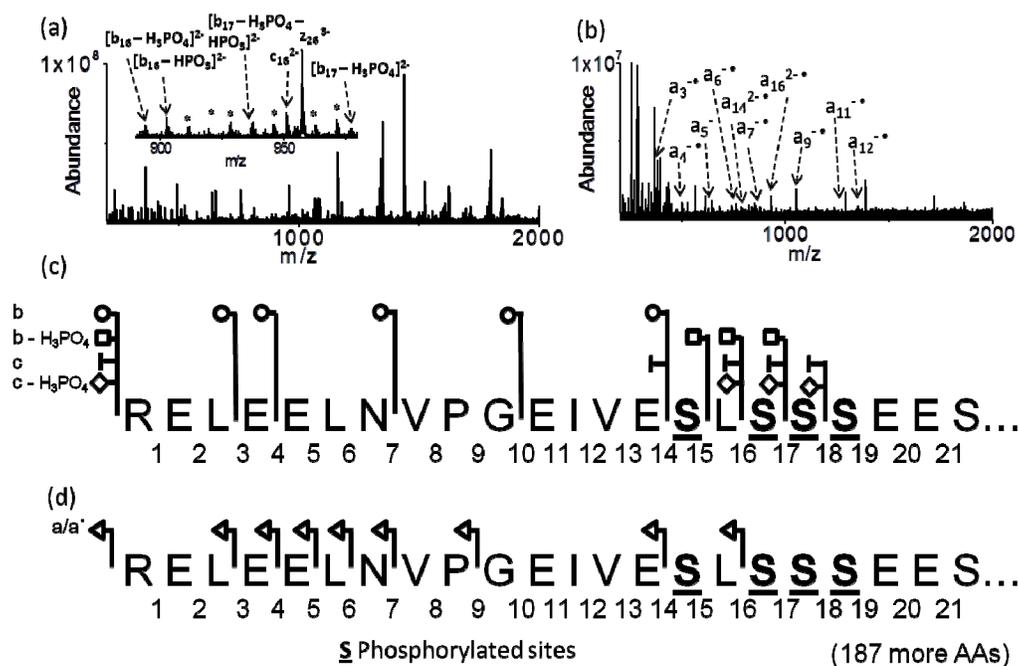
Mass spectra were obtained with a 7-T quadrupole-FT-ICR mass spectrometer (Bruker Daltonics, Billerica, MA). IRMPD experiments were performed with a vertically mounted 25 W, 10.6 $\mu$ m, CO<sub>2</sub> laser (Synrad, Mukilteo, WA). Photon irradiation was performed at 50% laser power with two different pulse sequences, either one 200-300 ms pulse or multiple 15 ms pulses with a 200 ms delay between each pulse. IRMPD peak assignments are based on the combination of spectra from two pulse sequences. AI-EDD experiments were performed in the ICR cell by irradiating protein anions with a 20 ms IR laser pulse (50% power) followed by 20-30 eV electron irradiation for 1 second. Electrons were generated by a hollow cathode. AI-ECD experiments were performed by

irradiating protein cations with a 20 ms IR laser pulse (50% power) followed by 0.01 eV electrons for 0.5-1 s.

Raw spectra were acquired with Bruker XMass software and summed over 256 or 512 scans. Spectra were processed by DataAnalysis 4.0 software (Bruker Daltonics) and peaks were assigned by a combination of manual assignment and an in-house written Microsoft Excel macro.

### 4.3. Results and Discussion

#### 4.3.1. Infrared Multiphoton Dissociation of $\beta$ -casein Anions



**Figure 4.1.** (a) Negative ion IRMPD of all observed charge states (15- to 26-) of  $\beta$ -casein anions generated by electrospray ionization. Inset: partial IRMPD spectrum. Noise peaks are labeled with asterisks. (b) AI-EDD of all observed charge states of  $\beta$ -casein anions generated by electrospray ionization. Only  $a$ -type fragments are labeled. Spectra are summed over 256 scans. (c) Fragmentation map from (a). (d) Fragmentation map from (b). Only the first 22 amino acids are shown.

Figure 4.1a shows negative ion IRMPD of  $\beta$ -casein anions. Intact  $\beta$ -casein was electrosprayed at 0.1 mg/mL from a solution containing 1:1 v/v water/acetonitrile with

0.25% piperidine and all charge states were irradiated with a 180 ms laser pulse (7.5 W). IRMPD of the individual charge states 15- and 22- yielded virtually identical spectra but with lower signal to noise ratio than IRMPD of all charge states. N-terminal *b* and *c*-type ions dominate the IRMPD spectrum, as summarized in Figure 1c, but extensive internal fragments were also observed. The overall sequence coverage was 7% (15 backbone cleavages out of 208). However, extensive fragmentation was observed within the region Glu-14 to Glu-21, which contains seven acidic or phosphorylated side chains out of eight amino acids. Glu-14 to Glu-21 has a pI of 1.6 as predicted by Scansite Molecular Weight and Isoelectric Point Calculator (<http://scansite.mit.edu/cgi-bin/calcp1>). Partial phosphate loss is observed for both *b*-type and *c*-type ions. Phosphorylation sites at Ser-15, Ser-17, Ser-18, and Ser-19 are clearly identified by *c*-type fragments and a combination of *b* and *b* - H<sub>3</sub>PO<sub>4</sub>-type ions also provides identification of phosphorylation sites at Ser-15, Ser-17, and Ser-18.

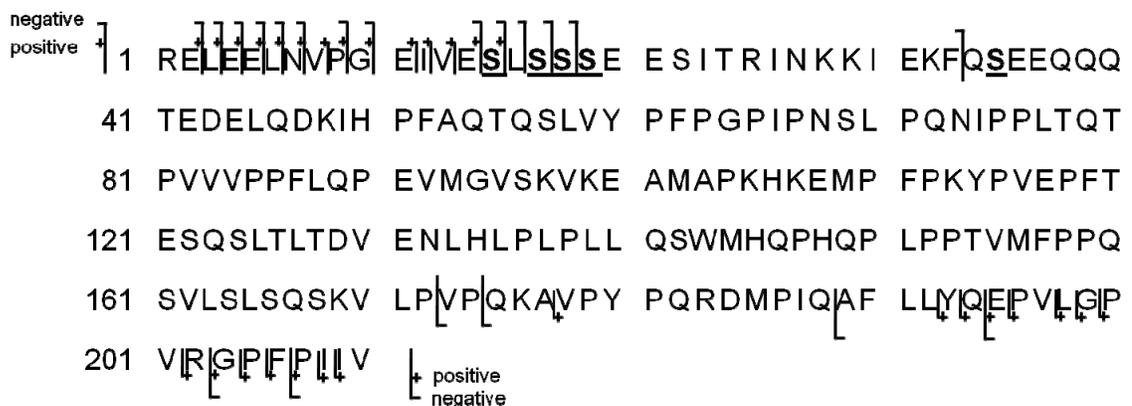
#### **4.3.2. Activated Ion Electron Detachment Dissociation of $\beta$ -casein Anions**

AI-EDD of  $\beta$ -casein anions is shown in Figure 1b. All observed charge states were irradiated with a 90 ms IR laser pulse (7.5 W), followed by a beam of  $\sim 19$  eV electrons for 1s. AI-EDD of  $\beta$ -casein yields predominantly *a*-type radical ions corresponding to backbone cleavage close to the N-terminus, as shown in Figure 1d. The EDD fragmentation mechanism for peptides has been investigated by Kjeldsen, *et al.* and Anusiewicz, *et al.* by computational approaches.<sup>59, 60</sup> In their work it was proposed that negative charges are located on backbone nitrogen atoms and that electron ejection occurs from such deprotonated backbone nitrogens. These authors showed that formation of *a/x*<sup>•</sup> product ions is both thermodynamically and kinetically less favored than

formation of  $a\bullet/x$  product ions. In a recent investigation of protein EDD  $a/x\bullet$  product ions were not reported<sup>50</sup> and such product ions are also dominant in the work presented here. However, even-electron ions  $a_4$ ,  $a_9$  and radical  $x_2\bullet$ ,  $x_3\bullet$  ions were also observed in our EDD spectra. One hypothesis is that even-electron  $a$ -type ions are from hydrogen transfer events, similar to those observed in ECD<sup>61, 62</sup>. Another hypothesis is that, even though formation of  $a/x\bullet$  product ions is less favored, it is still statistically possible, noting that electron irradiation time in our experiments was 1 s, rather than the 100-300 ms used by Ganisl *et al.*<sup>50</sup> Neutral CO<sub>2</sub> loss, which was reported as a major fragmentation pathway by Ganisl *et al.*, was not observed in our experiments, even for product ions with terminal glutamic acid residues. However, the product ion abundance in our experiments was rather low and thus fragments corresponding to CO<sub>2</sub> loss may be below our detection limit. Ganisl *et al.*, also reported that  $a\bullet/x$  product ion formation was facilitated by neighboring basic residues. However, the fragments observed in our study did not result from cleavage near lysine or arginine. Compared to negative ion IRMPD, AI-EDD provides less sequence coverage, ~5% (10/208 backbone cleavages observed). However, the total number of peaks in the AI-EDD spectrum corresponds to 28% of the number of peaks (with an S/N ratio >2) in the IRMPD spectrum. 13% of the peaks in the AI-EDD spectrum are sequence informative compared to 5% in IRMPD. Thus, AI-EDD results are more facile to interpret. The simplicity of EDD spectra and the specificity of  $a$ -type backbone cleavage appear advantageous compared to IRMPD although EDD fragmentation efficiency is low.

#### **4.3.3. Comparison between Negative and Positive Ion Mode**

Positive ion IRMPD and AI-ECD were applied to  $\beta$ -casein for comparison with the negative ion results. Positive ion IRMPD cleaved 21 backbone amide bonds, with extensive internal fragments, water losses, and phosphoric acid losses. Under our instrumental conditions, AI-ECD cleaved 26 backbone amine bonds. The combination of IRMPD and AI-ECD provided 29 backbone cleavages. Previously reported AI-ECD and plasma ECD of intact  $\beta$ -casein provided 87 backbone cleavages.<sup>34</sup> In negative ion mode, the combination of AI-EDD and IRMPD yielded 20 backbone cleavages. As shown in Figure 4.2, AI-EDD and negative ion IRMPD provided nine unique backbone cleavages compared to AI-ECD and positive ion IRMPD. In particular, negative ion IRMPD yielded extensive fragmentation within the highly acidic region Glu-14 to Glu-21. Four phosphorylated sites were directly identified by backbone cleavages in this region. The latter analysis was not feasible in positive ion mode with our instrument, nor from previously published AI-ECD.<sup>34</sup> Plasma ECD data<sup>42</sup> located all phosphorylation sites; however, it did not provide backbone cleavages between each phosphorylated residue.



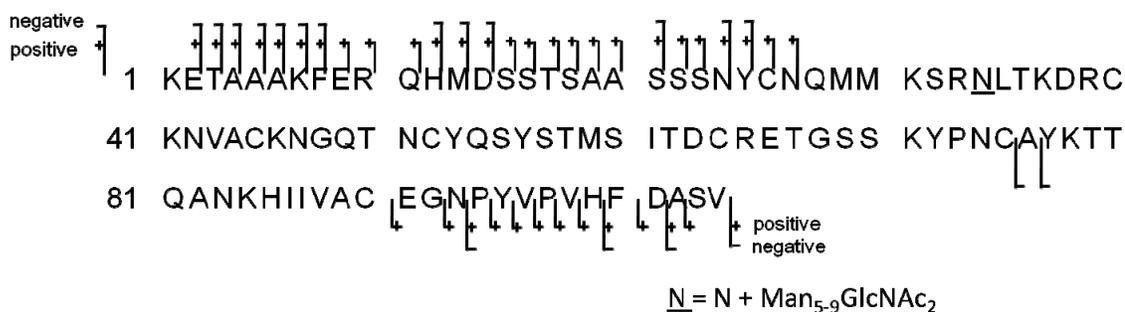
**Figure 4.2.** Comparison of  $\beta$ -casein fragmentation patterns in negative and positive ion mode. All backbone cleavages from AI-EDD and negative ion IRMPD are labeled ‘negative’ and all backbone cleavages from AI-ECD and positive ion IRMPD are labeled ‘positive’.

Calmodulin (pI = 4.0) and glycosylated ribonuclease B (RNaseB, pI = 8.6) were also investigated in both positive and negative ion top down MS/MS. For calmodulin, AI-ECD and positive ion IRMPD provided 40 backbone cleavages out of 148 possible peptide bonds. AI-EDD and negative ion IRMPD yielded an additional six backbone cleavages, as shown in Figure 4.3.



**Figure 4.3.** Comparison of calmodulin fragmentation patterns in negative and positive ion mode. All backbone cleavages from AI-EDD and IRMPD in negative ion mode are labeled ‘negative’ and all backbone cleavages from AI-ECD and IRMPD in positive ion mode are labeled ‘positive’.

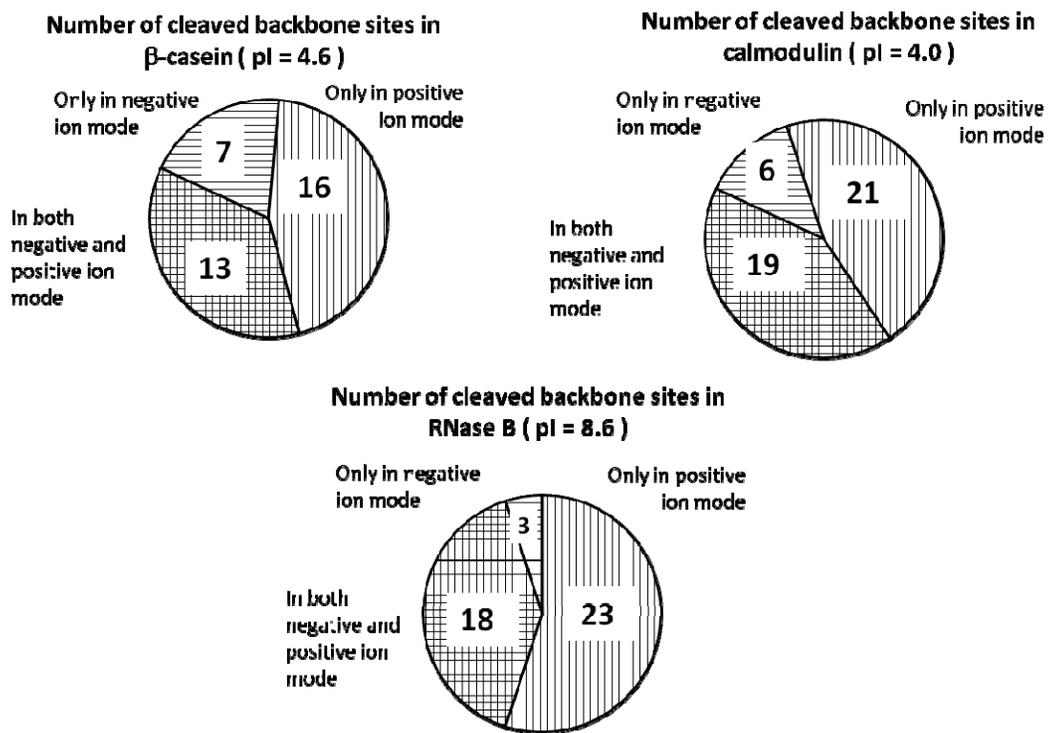
For RNase B, AI-ECD and positive ion IRMPD provided 41 backbone cleavages out of 124 possible peptide bonds whereas AI-EDD and negative ion IRMPD yielded an additional two backbone cleavages, as shown in Figure 4.4.



**Figure 4.4.** Comparison of RNase B fragmentation patterns in negative and positive ion mode. All backbone cleavages from AI-EDD and IRMPD in negative ion mode are labeled ‘negative’ and all backbone cleavages from AI-ECD and IRMPD in positive ion mode are labeled ‘positive’.

A summary of observed negative and positive ion mode fragmentation for all three proteins is shown in Figure 4.5. For  $\beta$ -casein and calmodulin, which are both acidic proteins, negative ion dissociation, particularly IRMPD, provides more complementary sequence information than for the basic protein RNase B. AI-EDD of  $\beta$ -casein, calmodulin and RNase B yielded similar numbers of backbone  $a\bullet/x$  ions. In all three cases backbone cleavages were mainly observed close to the protein termini. Kjeldsen *et al.* reported that EDD backbone cleavages in peptides are favored near acidic residues.<sup>60</sup> By contrast, such a preference was not observed in EDD of the proteins ubiquitin and melittin.<sup>50</sup> In our work, the largest  $a\bullet$  ion observed from  $\beta$ -casein was  $a\bullet_{16}$ . For this protein, 11 of the first 16 (i.e., 69%) possible backbone cleavages, counting from the N-terminus, represent cleavage N- or C-terminal to acidic residues (D, E and pS). Among the observed eight backbone cleavages in this region, five (i.e., 63%) represent backbone cleavage close to acidic residues. For calmodulin, 62% of the first 13 possible backbone cleavages are close to acidic residues and, among the observed nine backbone cleavages, 56% are close to acidic residues. For RNase B, 50% of the first 16 possible backbone cleavages are close to acidic residues and, among the observed nine backbone cleavages,

33% are close to acidic residues. Thus, for these proteins a cleavage preference near acidic residues was not apparent. However, a general analysis was hampered by the low sequence coverage. As for the previously mentioned proposed favored cleavage near basic residues<sup>50</sup> all three proteins investigated here contain similar total numbers of K and R residues: 15, 14 and 14, respectively. The total numbers of  $a\bullet/x$ -type backbone bond cleavages were similar as well (10, 9, and 10, respectively). Further, the N-terminal regions of these three protein sequences contain few K and R residues. Thus, a clear correlation between  $a\bullet/x$  cleavage and the presence of basic residues is difficult to derive.



**Figure 4.5.** Comparison of the numbers of backbone cleavages in negative and positive ion mode top down MS/MS of (a)  $\beta$ -casein; (b) calmodulin; and (c) RNase B.

#### 4.4. Conclusions

Top-down protein analysis in negative ion mode was investigated by employing two dissociation techniques on three intact proteins. In general, negative ion IRMPD and

AI-EDD provided complementary sequence information compared to positive ion IRMPD and AI-ECD. However, less overall sequence coverage was seen in negative ion mode. More complementary fragments were observed in negative ion mode for the acidic proteins  $\beta$ -casein and calmodulin than for the basic protein RNase B. A possible explanation for this observation may be that salt-bridge formation, which can reduce the observation of backbone fragments, is favored in positive ion mode for acidic proteins. In particular, IRMPD of  $\beta$ -casein, with multiple phosphorylated sites close to each other, yielded backbone cleavages between the first four phosphorylated amino acids from the N-terminus. Such cleavages, which are essential to precisely locate phosphorylation sites, were not observed in positive ion IRMPD nor AI-ECD but were partially observed in previous plasma ECD.<sup>42</sup>

#### 4.5. Bibliography

- (1) Chait, B. T. *Science* **2006**, *314*, 65-66.
- (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.; Fridriksson, E. K.; Horn, D. M.; Lewis, M. A.; Zubarev, R. A. *Science* **1999**, *284*, 1289-1290.
- (4) Loo, J. A.; Edmonds, C. G.; Smith, R. D. *Science* **1990**, *248*, 201-204.
- (5) Smith, R. D.; Loo, J. A.; Barinaga, C. J.; Edmonds, C. G.; Udseth, H. R. *J. Am. Soc. Mass. Spectrom.* **1990**, *1*, 53-65.
- (6) Stephenson, J. L.; McLuckey, S. A. *Anal. Chem.* **1998**, *70*, 3533-3544.
- (7) Stephenson, J. L.; Cargile, B. J.; McLuckey, S. A. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2040-2048.
- (8) Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Anal. Chem.* **2001**, *73*, 3274-3281.

- (9) Loo, J. A.; Quinn, J. P.; Ryu, S. I.; Henry, K. D.; Senko, M. W.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 286-289.
- (10) Little, D. P.; Speir, J. P.; Senko, M. W.; Oconnor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.
- (11) Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859-866.
- (12) Ge, Y.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass spectrom.* **2001**, *210*, 203-214.
- (13) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
- (14) 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.
- (15) McLafferty, F. W.; Horn, D. M.; Breuker, K.; Ge, Y.; Lewis, M. A.; Cerda, B.; Zubarev, R. A.; Carpenter, B. K. *J. Am. Soc. Mass. Spectrom.* **2001**, *12*, 245-249.
- (16) Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J. H.; Begley, T. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124*, 672-678.
- (17) Coon, J. J.; Ueberheide, B.; Syka, J. E. P.; Dryhurst, D. D.; Ausio, J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 9463-9468.
- (18) McLuckey, S. A.; Goeringer, D. E. *J. Mass Spectrom.* **1997**, *32*, 461-474.
- (19) Stannard, P. R.; Gelbart, W. M. *J. Phys. Chem.* **1981**, *85*, 3592-3599.
- (20) McLachlin, D. T.; Chait, B. T. *Curr. Opin. Chem. Biol.* **2001**, *5*, 591-602.
- (21) Carr, S. A.; Huddleston, M. J.; Annan, R. S. *Anal. Biochem.* **1996**, *239*, 180-192.
- (22) Annan, R. S.; Carr, S. A. *Anal. Chem.* **1996**, *68*, 3413-3421.
- (23) Meng, F. Y.; Cargile, B. J.; Miller, L. M.; Forbes, A. J.; Johnson, J. R.; Kelleher, N. L. *Nat. Biotechnol.* **2001**, *19*, 952-957.
- (24) Reid, G. E.; Stephenson, J. L.; McLuckey, S. A. *Anal. Chem.* **2002**, *74*, 577-583.
- (25) Siuti, N.; Kelleher, N. L. *Nat. Methods* **2007**, *4*, 817-821.
- (26) Kruger, N. A.; Zubarev, R. A.; Carpenter, B. K.; Kelleher, N. L.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass spectrom.* **1999**, *182*, 1-5.
- (27) Kelleher, R. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250-4253.

- (28) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793-1800.
- (29) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.
- (30) 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.
- (31) Hunter, T. *Cell* **2000**, *100*, 113-127.
- (32) Mann, M.; Ong, S. E.; Gronborg, M.; Steen, H.; Jensen, O. N.; Pandey, A. *Trends Biotechnol.* **2002**, *20*, 261-268.
- (33) Breuker, K.; Jin, M.; Han, X. M.; Jiang, H. H.; McLafferty, F. W. *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 1045-1053.
- (34) Shi, S. D. H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, *73*, 19-22.
- (35) Woods, A. S.; Ferre, S. *J. Proteome. Res.* **2005**, *4*, 1397-1402.
- (36) Creese, A. J.; Cooper, H. F. *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 1263-1274.
- (37) Haselmann, K. F.; Jorgensen, T. J. D.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 2260-2265.
- (38) Jackson, S. N.; Dutta, S.; Woods, A. S. *J. Am. Soc. Mass. Spectrom.* **2009**, *20*, 176-179.
- (39) Xie, Y. M.; Zhang, J.; Yin, S.; Loo, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 14432-14433.
- (40) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.
- (41) Chalmers, M. J.; Hakansson, K.; Johnson, R.; Smith, R.; Shen, J. W.; Emmett, M. R.; Marshall, A. G. *Proteomics* **2004**, *4*, 970-981.
- (42) Sze, S. K.; Ge, Y.; Oh, H. B.; McLafferty, F. W. *Anal. Chem.* **2003**, *75*, 1599-1603.
- (43) Steinborner, S. T.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1243-1247.
- (44) Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Mass Spectrom. Rev.* **2002**, *21*, 87-107.
- (45) Ewing, N. P.; Cassady, C. J. *J. Am. Soc. Mass. Spectrom.* **2001**, *12*, 105-116.
- (46) Kalli, A.; Hakansson, K. *Int. J. Mass spectrom.* **2007**, *263*, 71-81.

- (47) Flora, J. W.; Muddiman, D. C. *Anal. Chem.* **2001**, *73*, 3305-3311.
- (48) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.
- (49) Haselmann, K. F.; Budnik, B. A.; Kjeldsen, F.; Nielsen, M. L.; Olsen, J. V.; Zubarev, R. A. *Eur. J. Mass Spectrom.* **2002**, *8*, 117-121.
- (50) Ganisl, B.; Valovka, T.; Hartl, M.; Taucher, M.; Bister, K.; Breuker, K. *Chem. Eur. J.* **2011**, *17*, 4460-4469.
- (51) Yang, J.; Mo, J. J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, *77*, 1876-1882.
- (52) Taucher, M.; Breuker, K. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 918-929.
- (53) Yang, J.; Hakansson, K. *J. Am. Soc. Mass. Spectrom.* **2006**, *17*, 1369-1375.
- (54) Mo, J. J.; Hakansson, K. *Anal. Bioanal. Chem.* **2006**, *386*, 675-681.
- (55) Wolff, J. J.; Amster, I. J.; Chi, L. L.; Linhardt, R. J. *J. Am. Soc. Mass. Spectrom.* **2007**, *18*, 234-244.
- (56) Adamson, J. T.; Hakansson, K. *J. Am. Soc. Mass. Spectrom.* **2007**, *18*, 2162-2172.
- (57) Kweon, H. K.; Hakansson, K. *J. Proteome. Res.* **2008**, *7*, 749-755.
- (58) Huang, T. Y.; Liu, J. A.; McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 890-898.
- (59) Anusiewicz, I.; Jasionowski, M.; Skurski, P.; Simons, J. *J. Phys. Chem. A* **2005**, *109*, 11332-11337.
- (60) Kjeldsen, F.; Silivra, O. A.; Ivonin, I. A.; Haselmann, K. F.; Gorshkov, M.; Zubarev, R. A. *Chem. Eur. J.* **2005**, *11*, 1803-1812.
- (61) O'Connor, P. B.; Lin, C.; Cournoyer, J. J.; Pittman, J. L.; Belyayev, M.; Budnik, B. A. *J. Am. Soc. Mass. Spectrom.* **2006**, *17*, 576-585.
- (62) Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. *J. Am. Soc. Mass. Spectrom.* **2007**, *18*, 113-120.

## Chapter 5

# Characterization of Oligonucleotides by Negative Ion Electron Capture Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

### 5.1 Introduction

Mass spectrometry (MS) and tandem mass spectrometry (MS/MS) have been well-developed and applied to the determination of nucleic acid sequence.<sup>1-3</sup> Commonly employed MS/MS strategies are based on vibrational activation, or “slow-heating”, such as collision activated dissociation (CAD),<sup>4</sup> infrared multiphoton dissociation (IRMPD),<sup>5</sup> and blackbody infrared radiative dissociation (BIRD).<sup>6</sup> Signature (*a* - B) and *w*-type ions (McLucky nomenclature<sup>7</sup>) are observed for the sequencing of oligodeoxyribonucleotides (DNA), whereas, for oligoribonucleotides (RNA), sequence-specific *c* and *y*-type ions are observed. Despite the wide-spread use of “slow-heating” techniques, one major disadvantage is extensive neutral loss, such as water and nucleobase loss. These losses spread available signal over multiple peaks, which complicates spectrum interpretation and lowers sensitivity.

Gas-phase ion-electron reactions such as electron capture dissociation (ECD)<sup>8</sup> in positive ion mode and electron detachment dissociation (EDD)<sup>9</sup> in negative ion mode have been developed as alternatives to vibrational activation techniques. In ECD, multiply charged cations are irradiated with low energy electrons (< 1 eV) and capture

electrons to form radical intermediates that further dissociate into fragments. ECD has been applied to, e.g., peptides,<sup>8, 10</sup> proteins,<sup>8, 11</sup> oligonucleotides,<sup>12, 13</sup> polymers,<sup>14</sup> carbohydrates,<sup>15, 16</sup> antibiotics,<sup>17</sup> and peptide nucleic acids.<sup>18</sup> ECD can retain labile modifications such as peptide phosphorylation<sup>19, 20</sup> and glycosylation<sup>21, 22</sup> that are preferentially lost in slow-heating MS/MS activation methods. ECD of multiply positively charged DNAs has been demonstrated to generate sequence-specific *w/d* and *a•/z•*-type ions<sup>12</sup> with significantly reduced neutral loss compared with CAD and IRMPD. The lower degree of neutral loss simplifies spectrum interpretation.<sup>2</sup> Initial application of ECD to DNA analysis was performed with an FT-ICR mass spectrometer equipped with a directly heated filament. This configuration resulted in limited product ion abundance. Improved fragmentation efficiency was obtained via IR-activated MS<sup>3</sup> of the radical intermediate, similar to the case of phospho<sup>23</sup> and glycopeptides.<sup>24</sup> An alternative approach is IR or collisional activation of the isolated precursor ion prior to electron irradiation.<sup>25, 26</sup> More significant improvement of fragmentation efficiency was achieved by switching to an indirectly heated ring cathode with an increased cross-section, thus generating improved overlap between the electron cloud and cations in the ICR cell.<sup>13</sup> Nevertheless, due to the acidity of the phosphate backbone, oligonucleotides are detected with lower sensitivity in positive ion mode, compared with negative ion mode. EDD, which operates in negative ion mode, benefits from this higher ionization efficiency. In EDD, multiply charged anions are irradiated with energetic electrons (typically between 15 and 30 eV), resulting in ejection of electrons to form radical intermediates that further dissociate.<sup>27</sup> EDD has been applied to peptides,<sup>9</sup> carbohydrates,<sup>28-30</sup> and nucleic acids.<sup>31-33</sup> EDD of multiply charged DNA oligomers has been demonstrated to generate similar

types of fragments as ECD<sup>31</sup> whereas EDD of multiply charged RNA oligomers (6- to 10-mers) has been shown to produce *c*-, *y*-, *d*- and *w*-type ions.<sup>34</sup> EDD of larger RNAs (34- to 61-mers) provided evidence that *c*- and *y*-type ions may result from vibrational activation, such as collisions during the ion transfer process, or charge state-dependent structural relaxation.<sup>33</sup> Both ECD and EDD require precursor ions to be multiply charged and are thus not applicable to singly charged ions.

Recently, our group has demonstrated that electrons within a certain energy range can be captured by singly charged peptide anions.<sup>35</sup> Moreover, ECD-like fragments were observed following the electron capture event. Fragments from investigated phosphorylated peptides retained phosphate during the dissociation process. This new dissociation technique was named negative ion electron capture dissociation (niECD).<sup>35</sup> Several doubly deprotonated peptides were also fragmented by niECD. Gas-phase zwitterionic structures were proposed to be important for successful niECD of peptides. This hypothesis was supported by the observation of increased or decreased niECD efficiency upon alteration of the peptide N-terminus by adding a fixed charge tag (to promote zwitterion formation) or acetylation (to reduce zwitterion formation). Also, it has been shown that gas-phase zwitterionic structures are favored for phosphopeptides<sup>36</sup>,<sup>37</sup> which undergo efficient niECD. Oligonucleotides contain acidic phosphate backbones and basic nucleobases, a composition that may favor gas-phase zwitterion formation. Molecular modeling of mononucleotides suggested that protonation of nucleobase nitrogens is more favored than the non-bridging phosphate oxygens, except for T/U.<sup>38</sup> It has been demonstrated that zwitterion formation is possibly favored when oligonucleotides are electrosprayed from ammonium-based buffers.<sup>39</sup> Further, several

proposed mechanisms for (*a* - *B*)/*w*-type ion formation from vibrational activation involve a zwitterion intermediate.<sup>40-42</sup> One proposed mechanism for *d/w*-type ion formation in EDD of RNA also involves a zwitterionic radical intermediate.<sup>33</sup> If oligonucleotides are zwitterionic in the gas-phase they may undergo efficient niECD, similar to phosphopeptides. Here we explore the application of niECD to deprotonated oligonucleotides.

## **5.2 Experimental Section**

### **5.2.1 Sample Preparation**

All oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). 6-mer and 11-mer DNA and RNA were dissolved in electrospray solution (1% triethylamine in 50% (v/v) H<sub>2</sub>O, 25% methanol and 25% isopropyl alcohol) and electrosprayed without further desalting. The 27-mer RNA 16S R1 (5'-GGCGUCACACCUUCGGGUGAAGUCGCC-3') was desalted by ethanol precipitation (protocol modified from Limbach et al.<sup>43</sup>). 1 nmole of RNA was mixed with 20 μL of 7.5 M NH<sub>4</sub>OAc and 80 μL of cold ethanol. The mixture was vortexed, stored at -80 °C for 3 hours, centrifuged for 30 minutes and the supernatant was discarded. 200 μL of 90% cold ethanol was added to the precipitate and stored at - 80 °C for 2 hours. The mixture was centrifuged at 12,000 g for 30 minutes and the supernatant was discarded. The precipitate was dissolved in electrospray solution (10 mM NH<sub>4</sub>OAc in 50% (v/v) H<sub>2</sub>O, 50% isopropyl alcohol) for electrospray ionization.

### **5.2.2 Mass Spectrometry**

All mass spectra were acquired with a 7-T quadrupole-FT-ICR mass spectrometer (Apex-Q, Bruker Daltonics, Billerica, MA). The oligonucleotide solutions were infused

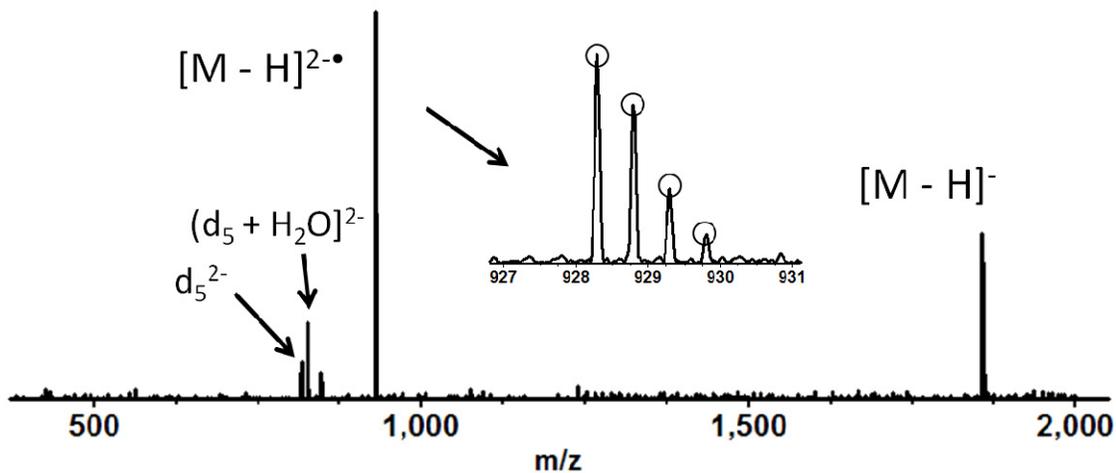
via an Apollo II electrospray ion source at a flow rate of 70  $\mu\text{L}/\text{hour}$ . Ions with mass to charge ratios of interest were selected by the external quadrupole, accumulated in a hexapole, and transported to the ICR cell for further activation and dissociation. niECD experiments were performed in the ICR cell by irradiating precursor ions for 3-5 seconds with an electron beam generated by an indirectly heated hollow dispenser cathode. Electron energy and number were controlled by adjusting the voltages on the cathode and the focusing lens located between the cathode and the ICR cell. The cathode bias voltage was in the range - 6 to - 10 V and the lens voltage was in the range - 4.5 to - 8.5 V. Electron energy distributions were measured by monitoring the current impinging on the focusing lens located on the opposite side of the ICR cell compared with the cathode while changing the voltage of the cathode and the focusing lens. For activated ion (AI) niECD experiments, precursors ions in the ICR cell were irradiated with an infrared laser beam from a 25-W, 10.6- $\mu\text{m}$ ,  $\text{CO}_2$  laser (Synrad, Mukilteo, WA) at 50% power for 20-50 ms, prior to irradiation with the electron beam. For niECD-IRMPD experiments, a correlated harmonic excitation field (CHEF)<sup>44</sup> waveform was applied after the niECD event to eject all ions within the detection range except the charge-increased product ions, then the isolated charge-increased ions were irradiated with an infrared laser for 2-5 s at 12.5 W. Mass spectra were acquired with XMASS (version 6.1, Bruker Daltonics) in broadband mode from  $m/z$  200 to 3000 and analyzed by DataAnalysis (version 4.0, Bruker Daltonics). Product ions were assigned within 10 ppm mass accuracy.

## **5.3 Result and Discussion**

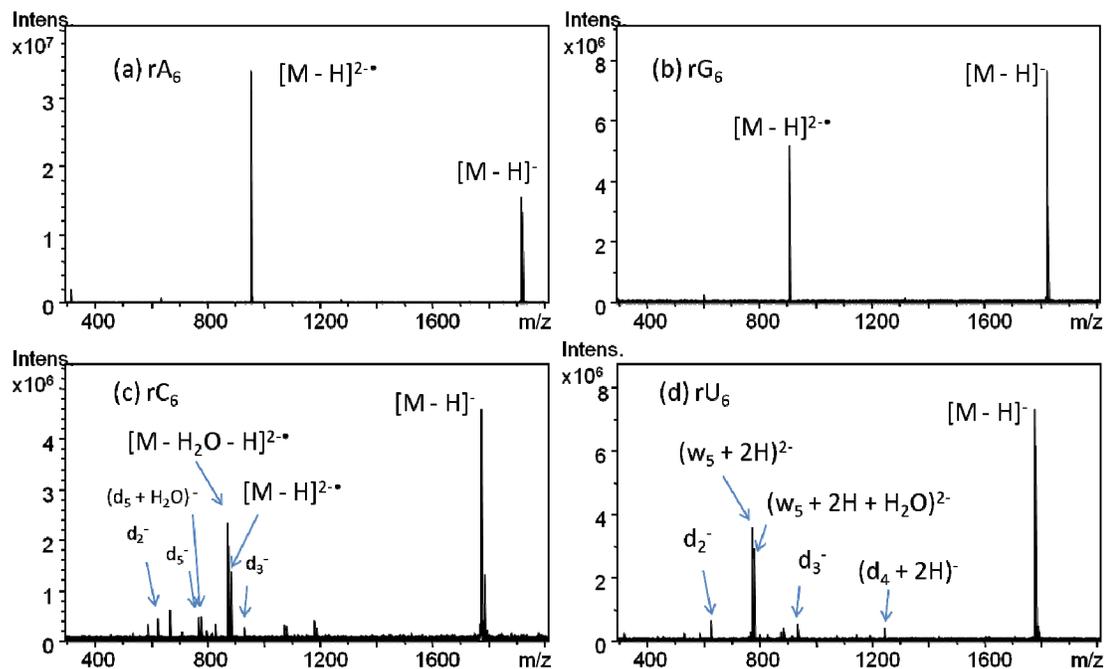
### **5.3.1 Electron Capture of RNA Hexamers (rGCAUAC, rA<sub>6</sub>, rC<sub>6</sub>, rG<sub>6</sub> and rU<sub>6</sub>)**

When the singly deprotonated oligoribonucleotide 5'-GCAUAC-3' (denoted as [M - H]<sup>-</sup>) was irradiated with 4.2 eV electrons for 5 seconds, a dominant product ion peak with an m/z ratio corresponding to [M - H]<sup>2•-</sup> was observed, as shown in Figure 5.1. The isotopic distribution of this peak matches with the calculated distribution for this radical species. The high abundance of this peak indicates that it is not an artifact at twice the ICR frequency of the precursor ions. Also, this isotopic cluster could be further isolated in the ICR cell, as shown in the following section. Electron detachment from oligonucleotide anions has been reported through ion-electron,<sup>31</sup> ion-photon,<sup>45</sup> and ion-ion<sup>46-49</sup> reactions. However, to our knowledge, the result presented here is the first observation of electron capture by oligonucleotide anions. rA<sub>6</sub>, rC<sub>6</sub>, rG<sub>6</sub>, and rU<sub>6</sub> were also subjected to irradiation with electrons of the same energy, as shown in Figure 5.2. In the cases of rA<sub>6</sub> and rG<sub>6</sub>, [M - H]<sup>2•-</sup> was the most abundant product ion whereas, for rC<sub>6</sub>, [M - H<sub>2</sub>O - H]<sup>2•-</sup> and [M - H]<sup>2•-</sup> were the most abundant product ions. For rU<sub>6</sub>, (w<sub>5</sub> + 2H)<sup>2-</sup> and (w<sub>5</sub> + 2H + H<sub>2</sub>O)<sup>2-</sup> were the most abundant product ions but [M - H<sub>2</sub>O - H]<sup>2•-</sup> and [M - H]<sup>2•-</sup> were also observed in lower abundance. (w<sub>5</sub> + 2H)<sup>2-</sup> and (w<sub>5</sub> + 2H + H<sub>2</sub>O)<sup>2-</sup> could also be assigned as [M - 2U - 2H]<sup>2-</sup> and [M - 2U + H<sub>2</sub>O - 2H]<sup>2-</sup>. In ECD of rC<sub>6</sub> cations, water loss was not observed.<sup>34</sup> Water loss in niECD of rC<sub>6</sub> anions may result from structural unfolding due to the charge state increase. Another observation was that backbone fragmentation was limited in niECD of all hexamer oligonucleotides. Similar behavior has been reported in ECD of doubly protonated DNA: under conditions for which most peptides were efficiently dissociated, the dGCATAC dication only captured electrons to form charge-reduced radical species with limited further dissociation.<sup>12, 13</sup>

Further activation, prolonged electron irradiation time, or more electrons injected into the ICR cell were necessary to improve ECD fragmentation efficiency.<sup>24, 50, 51</sup>

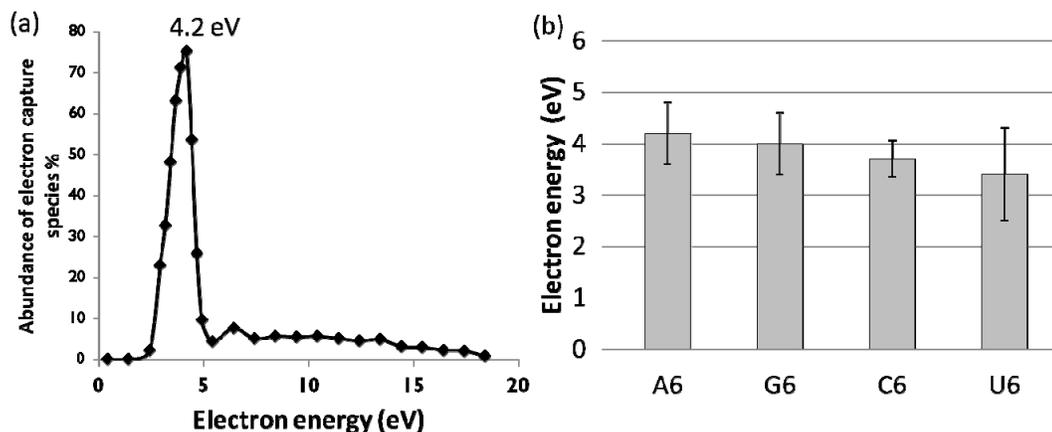


**Figure 5.1.** Electron capture by rGCAUAC. Singly deprotonated precursor anions were irradiated with 4.3 eV electrons for 5 seconds. In the inset, circles illustrate the calculated isotopic distribution of  $[M - H]^{2-\bullet}$  anions.

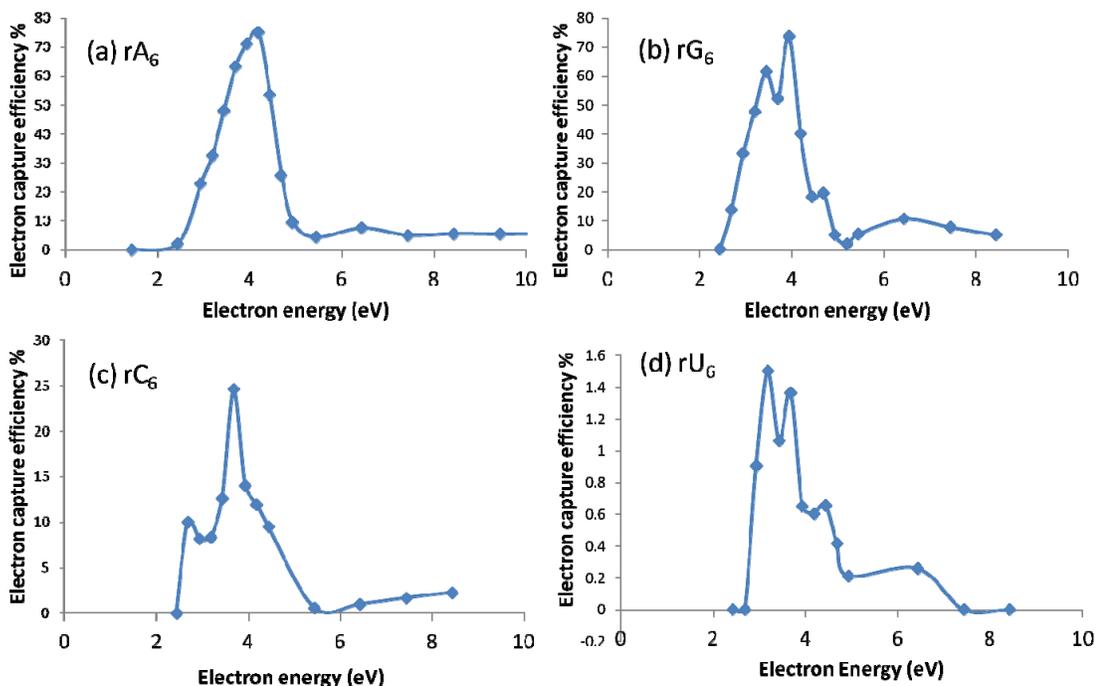


**Figure 5.2.** niECD of RNA hexamers: (a) rA<sub>6</sub>, (b) rG<sub>6</sub>, (c) rC<sub>6</sub>, (d) rU<sub>6</sub>. Singly deprotonated precursor anions were irradiated with 4.3 eV electrons for 3 seconds.

In order to understand the energy dependence of the anion electron capture process, the yield of charge-increased species from electron irradiation at different energies was investigated. This yield was calculated as the ratio between the abundance of charge-increased species (divided by charge state) and the precursor ion abundance prior to electron irradiation. For  $rA_6$  and  $rG_6$ , the charge-increased species was  $[M - H]^{2-}$ . For  $rC_6$  and  $rU_6$ , the charge-increased species was the sum of  $[M - H]^{2-}$  and  $[M - H_2O - H]^{2-}$ . Electron capture efficiency for  $rA_6$  is shown in Figure 5.3 (a) and electron capture efficiencies for  $rA_6$ ,  $rG_6$ ,  $rC_6$ , and  $rU_6$  are shown in Figure 5.4. For  $rA_6$ , maximum efficiency was achieved at 4.2 eV with a full width at half maximum (FWHM) of 1.2 eV. For the utilized hollow cathode implemented on the instrument, the FWHM of the electron energy distribution is 0.6 eV at 4.2 eV. Considering this uncertainty in the electron energy, the accepted energy range for  $rA_6$  anion electron capture may be less than  $4.2 \pm 1.2$  eV. For  $rG_6$ ,  $rC_6$ , and  $rU_6$ , maximum electron capture efficiencies were achieved at 4, 3.7, and 3.4 eV, respectively. A comparison of these maximum efficiencies is shown in Figure 5.3 (b). All maximum efficiencies are around 4 eV and the differences are within the FWHM, indicating that electron capture efficiency does not have a significant nucleobase dependence.



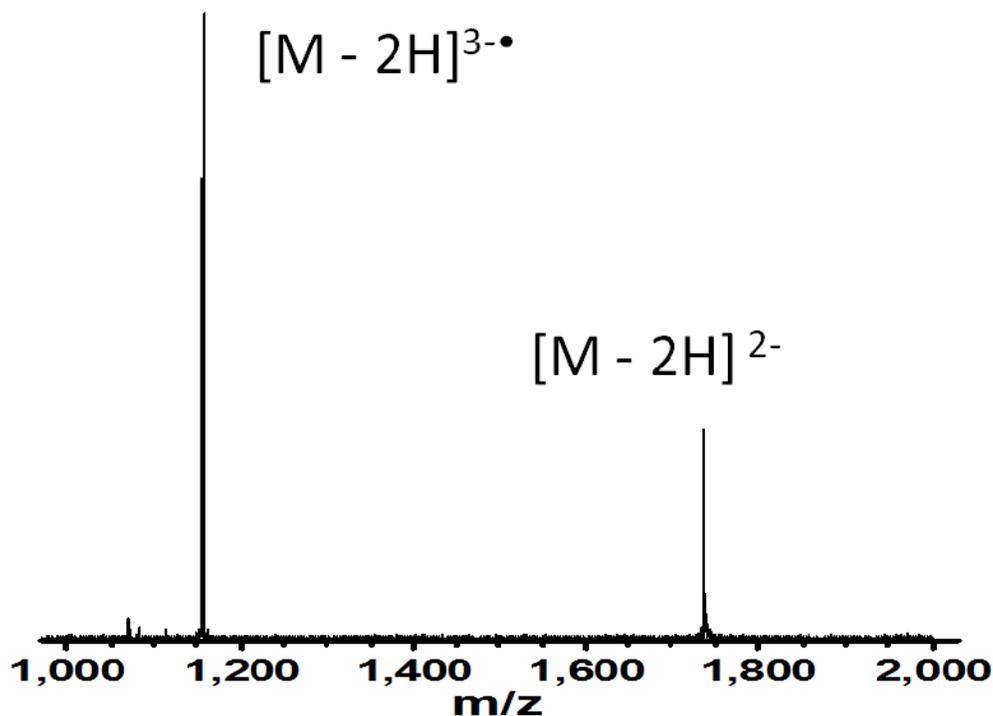
**Figure 5.3.** Electron capture efficiencies for different oligoribonucleotides. (a) Yield of charge-increased species from electron capture by  $rA_6$ . Maximum electron capture efficiency was observed at 4.2 eV. (b) Comparison of electron energies for which maximum electron capture efficiency was observed for  $rA_6$ ,  $rG_6$ ,  $rC_6$ , and  $rU_6$ . Error bars correspond to half of the FWHM of, e.g., the peak in (a).



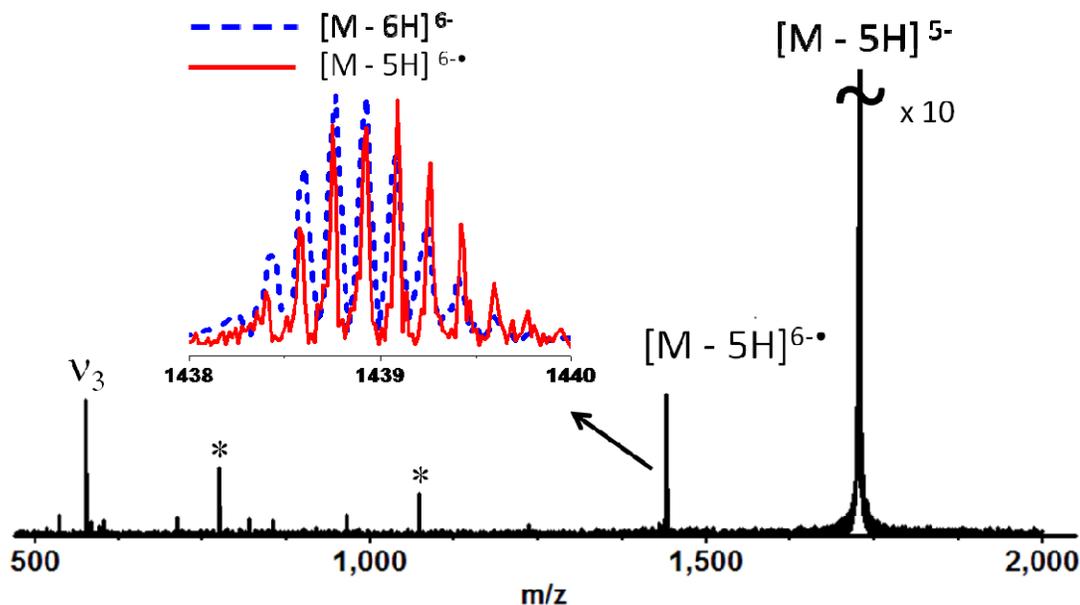
**Figure 5.4.** Electron capture efficiencies for RNA hexamers: (a)  $rA_6$ , (b)  $rG_6$ , (c)  $rC_6$ , (d)  $rU_6$ . Electron capture efficiency was calculated as the ratio between the abundance of charge-increased species (divided by charge state) and the precursor ion abundance prior to electron irradiation. For  $rA_6$  and  $rG_6$ , the charge-increased species was  $[M - H]^{2-\bullet}$ . For  $rC_6$  and  $rU_6$ , the charge-increased species was the sum of  $[M - H]^{2-\bullet}$  and  $[M - H_2O - H]^{2-\bullet}$ .

### 5.3.2 Electron capture by RNA 11- and 27-mer

When the doubly deprotonated RNA rUAACUAUCACG ( $[M - H]^-$ ) was irradiated with 4.2 eV electrons,  $[M - 2H]^{3-\bullet}$  was observed as the dominant product ion with limited occurrence of other product ions, as shown in Figure 5.5. Irradiation of the five times deprotonated 27-mer RNA rGGCGUCACACCUUCGGGUGAAGUCGCC ( $[M - 5H]^{5-}$ ) with 8.4 eV electrons resulted in observation of  $[M - 5H]^{6-\bullet}$  as the dominant product ions, as shown in Figure 5.6. The experimentally observed isotopic distribution of  $[M - 5H]^{6-\bullet}$  was compared with the isotopic distribution of  $[M - 6H]^{6-}$  from electrospray ionization on the same instrument, as shown in the inset of Figure 5.6. This comparison lends further support that the product ion is a radical from electron capture rather than an even-electron species. For the 27-mer RNA, higher electron energy was required for electron capture to occur, compared with the 11-mer and the hexamers, possibly due to the higher Coulomb repulsion.



**Figure 5.5.** Electron capture by UAACUAUGACG in its 2- charge state following irradiation with 4.5 eV electrons for 6 s.

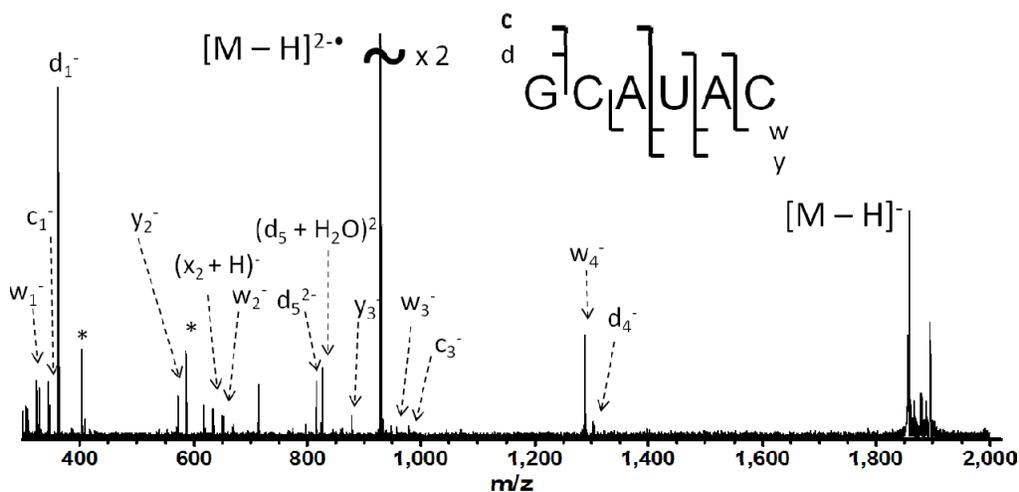


**Figure 5.6.** Electron capture by rGGCGUCACACCUUCGGGUGAAGUCGCC in its 5- charge state following irradiation with 8.4 eV electrons for 3s. The inset shows the isotopic distribution of the observed  $[M - 5H]^{6-\bullet}$  overlaid with  $[M - 6H]^{6-}$  from electro spray ionization on the same instrument.

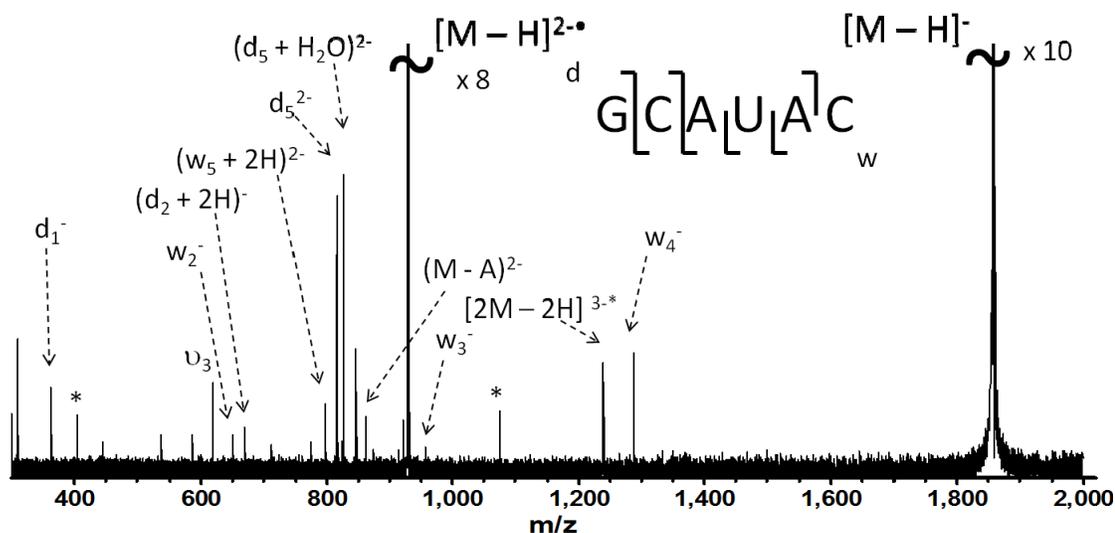
### 5.3.3 niECD-IRMPD ( $MS^3$ ) and AI-niECD of rGCAUAC Anions

Electron capture by singly deprotonated rGCAUAC (Fig. 5.1) yielded limited backbone fragmentation. As shown in Figure 5.7, further activation of  $[M - H]^{2-\bullet}$  by infrared laser irradiation promoted more fragmentation. For IRMPD  $MS^3$ , doubly charged radicals produced from electron capture were activated by IR laser irradiation at 12.5 W for 10 seconds. *d*-, *w*-, *c*-, and *y*-type ions were most common among the observed product ions. Complete *d/w*-type ion series, except for  $d_2$  and  $w_5$ , were detected. Four out of 10 possible *c/y*-type ions were observed. Compared with the EDD spectrum of doubly deprotonated rGCAUAC, niECD-IRMPD of the singly deprotonated form of the same RNA provides the same number of *d/w*-type ions but less *c/y*-type ions, which

may simplify spectrum interpretation. As an alternative to niECD-IRMPD, IR activation can be applied before electron irradiation (denoted as activated ion niECD or AI-niECD). In AI-niECD of rGCAUAC, precursor ions were irradiated with IR photons at 12.5 W laser power for 40 ms and then with 3.5 eV electrons for 3 s.  $d_1$ ,  $d_2$ ,  $d_5$ ,  $w_2$ ,  $w_3$ ,  $w_4$ , and  $w_5$  ions were identified, as shown in Figure 5.8. The optimum electron energy of 3.5 eV in AI-niECD is lower than the 4.5 eV used for the electron capture experiment, possibly due to the lower stability of precursor ions after IR activation. A minor precursor ion corresponding to loss of one adenine base was also observed from IR activation. Compared with niECD-IRMPD, AI-niECD offers less total activation time and dominant  $d/w$ -type product ions, however the fragmentation efficiency in AI-niECD was lower.



**Figure 5.7.** MS<sup>3</sup> (niECD-IRMPD) of singly deprotonated rGCAUAC. 4.5 eV electrons were used for 4 s. Electron capture species were activated by IR laser irradiation at 12.5 W for 4 seconds.



**Figure 5.8.** AI-niECD of singly deprotonated rGCAUAC. IR activation was at 12.5 W laser power for 40 ms. In the following niECD event 3.5 eV electrons were used for 3 s.

### 5.3.4 niECD-IRMPD of rA<sub>6</sub>, rG<sub>6</sub>, rC<sub>6</sub>, and rU<sub>6</sub> Anions and dA<sub>6</sub>, dG<sub>6</sub>, dC<sub>6</sub>, and dT<sub>6</sub> Anions

Similar to the RNA rGCAUAC, the RNAs and DNAs A<sub>6</sub>, G<sub>6</sub>, C<sub>6</sub>, and U<sub>6</sub> (T<sub>6</sub>) were subjected to niECD-IRMPD (MS<sup>3</sup>). The electron capture step was performed by irradiating singly deprotonated precursor ions with 3.5-4.2 eV electrons for 3-5 s. IRMPD (MS<sup>3</sup>) was performed by irradiating [M - H]<sup>2-•</sup> with an IR laser at 12.5 W for 4-10 s. Identified fragments are summarized in Table 5.1, except for nucleobase and water loss from precursor ions. For the RNAs investigated, *d/w*-type ions were more abundant than *c/x*- and *b/y*-type ions. *a/z*-type ions were not detected. Nucleobase and water loss from backbone fragments were not observed either. For the DNAs investigated, *d/w*-type ions were more abundant than *c/x*- and *a/z*-type ions. *b/y*-type ions were not detected. Nucleobase loss was observed for some of the *c/x*- and *a/z*-type ions. Compared with EDD,<sup>31, 34</sup> niECD-IRMPD yields similar types of ions and similar fragmentation efficiency, despite the fact that the intermediate species in EDD is [M - 2H]<sup>•</sup> rather than

$[M - H]^{2-}$ . This similarity may be due to similar initial gas-phase structures of the precursor ions as both techniques utilize anionic precursor ions. Compared with ECD of oligodeoxynucleotides,<sup>12, 13, 31</sup> niECD-IRMPD yields complete *d/w*-type ion series, as well as less abundant *a/z*- and *c/y*- type ions. This product ion pattern simplifies spectrum interpretation. In the case of dC<sub>6</sub>, niECD-IRMPD provided  $(a_3/z_3 + H)^+$  and  $(a_5/z_5 + H)^+$  ions, which have previously been observed<sup>12</sup> in ECD of dC<sub>6</sub>. niECD-IRMPD also provided an  $(a_2 - H)^+$  ion, which has been previously observed in EDD of dC<sub>6</sub>.<sup>31</sup> In the case of dT<sub>6</sub>, niECD-IRMPD provided “ECD-like”  $(a_2 + H)^+$ ,  $(a_4 + H)^+$  and  $(a_5 + H)^+$  ions. However, due to the low proton affinity of thymine nucleobases, ECD of dT<sub>6</sub> has not been reported and thus a direct comparison between niECD-IRMPD and ECD of dT<sub>6</sub> is not possible. niECD-IRMPD of dT<sub>6</sub> also provided  $(a_2 - H)^+$  and  $(a_4 - H)^+$  ions, similar to EDD of dT<sub>6</sub>, which provided  $(a_3 - H)^+$ ,  $(a_4 - H)^+$ , and  $(a_5 - H)^+$  product ions. “ECD-like” and “EDD-like” fragments from niECD-IRMPD indicate that niECD-IRMPD may share similarities with both EDD and ECD. Zwitterionic forms of protonated and deprotonated DNA have been proposed.<sup>13, 39</sup> In this hypothesis, a negative charge on a positively charged oligonucleotide is likely to be stabilized by a nearby positive charge (“salt-bridge” effect). Likewise, a positive charge on a negatively charged oligonucleotide is likely to be stabilized by a nearby negative charge. Thus niECD and ECD may have similar mechanisms.

**Table 5.1. Summary of fragments identified in niECD-IRMPD spectra of RNA and DNA hexamers A<sub>6</sub>, G<sub>6</sub>, C<sub>6</sub> and U<sub>6</sub> (T<sub>6</sub>).**

RNA sequence	Fragments identified in niECD-IRMPD			DNA sequence	Fragments identified in niECD-IRMPD		
rA <sub>6</sub>	<i>b</i> <sub>1</sub> / <i>y</i> <sub>1</sub>	<i>c</i> <sub>1</sub> / <i>x</i> <sub>1</sub>	<i>d</i> <sub>1</sub> / <i>w</i> <sub>1</sub>	dA <sub>6</sub>	( <i>a</i> <sub>2</sub> - H) <sup>•</sup>	<i>c</i> <sub>1</sub> / <i>x</i> <sub>1</sub>	<i>d</i> <sub>1</sub> / <i>w</i> <sub>1</sub>
	<i>b</i> <sub>3</sub> / <i>y</i> <sub>3</sub>	<i>c</i> <sub>2</sub> / <i>x</i> <sub>2</sub>	<i>d</i> <sub>2</sub> / <i>w</i> <sub>2</sub>		( <i>a</i> <sub>3</sub> - B)	( <i>c</i> <sub>1</sub> - B)	<i>d</i> <sub>2</sub> / <i>w</i> <sub>2</sub>
		<i>c</i> <sub>3</sub> / <i>x</i> <sub>3</sub>	<i>d</i> <sub>3</sub> / <i>w</i> <sub>3</sub>		( <i>a</i> <sub>4</sub> - B)	<i>c</i> <sub>2</sub> / <i>x</i> <sub>2</sub>	<i>d</i> <sub>3</sub> / <i>w</i> <sub>3</sub>
			<i>d</i> <sub>4</sub> / <i>w</i> <sub>4</sub>			( <i>c</i> <sub>2</sub> - B)	<i>d</i> <sub>4</sub> / <i>w</i> <sub>4</sub>
						( <i>c</i> <sub>3</sub> - B)	<i>d</i> <sub>5</sub> / <i>w</i> <sub>5</sub>
						( <i>c</i> <sub>4</sub> - B)	
rG <sub>6</sub>	<i>c</i> <sub>1</sub> / <i>x</i> <sub>1</sub>	<i>d</i> <sub>1</sub> / <i>w</i> <sub>1</sub>	<i>d</i> <sub>2</sub> / <i>w</i> <sub>2</sub>	dG <sub>6</sub>	( <i>a</i> <sub>2</sub> - B)	<i>c</i> <sub>1</sub> / <i>x</i> <sub>1</sub>	<i>d</i> <sub>1</sub> / <i>w</i> <sub>1</sub>
			<i>d</i> <sub>3</sub> / <i>w</i> <sub>3</sub>			<i>c</i> <sub>2</sub> / <i>x</i> <sub>2</sub>	<i>d</i> <sub>2</sub> / <i>w</i> <sub>2</sub>
							<i>d</i> <sub>3</sub> / <i>w</i> <sub>3</sub>
							<i>d</i> <sub>4</sub> / <i>w</i> <sub>4</sub>
							<i>d</i> <sub>5</sub> / <i>w</i> <sub>5</sub>
rC <sub>6</sub>	<i>c</i> <sub>1</sub> / <i>x</i> <sub>1</sub>	<i>d</i> <sub>1</sub> / <i>w</i> <sub>1</sub>	<i>d</i> <sub>2</sub> / <i>w</i> <sub>2</sub>	dC <sub>6</sub>	( <i>a</i> <sub>2</sub> - H) <sup>•</sup>	<i>c</i> <sub>1</sub> / <i>x</i> <sub>1</sub>	<i>d</i> <sub>1</sub> / <i>w</i> <sub>1</sub>
	<i>c</i> <sub>2</sub> / <i>x</i> <sub>2</sub>	<i>d</i> <sub>2</sub> / <i>w</i> <sub>2</sub>	<i>d</i> <sub>3</sub> / <i>w</i> <sub>3</sub>		( <i>a</i> <sub>3</sub> - B)		<i>d</i> <sub>2</sub> / <i>w</i> <sub>2</sub>
			<i>d</i> <sub>4</sub> / <i>w</i> <sub>4</sub>		( <i>a</i> <sub>4</sub> - B)		<i>d</i> <sub>3</sub> / <i>w</i> <sub>3</sub>
					( <i>a</i> <sub>3</sub> + H) <sup>•</sup>		<i>d</i> <sub>4</sub> / <i>w</i> <sub>4</sub>
					( <i>a</i> <sub>5</sub> + H) <sup>•</sup>		<i>d</i> <sub>5</sub> / <i>w</i> <sub>5</sub>
rU <sub>6</sub>		<i>d</i> <sub>2</sub> / <i>w</i> <sub>2</sub>	<i>d</i> <sub>3</sub> / <i>w</i> <sub>3</sub>	dT <sub>6</sub>	( <i>a</i> <sub>2</sub> + H) <sup>•</sup>	<i>c</i> <sub>1</sub> / <i>x</i> <sub>1</sub>	<i>d</i> <sub>1</sub> / <i>w</i> <sub>1</sub>
		<i>d</i> <sub>3</sub> / <i>w</i> <sub>3</sub>	<i>d</i> <sub>4</sub> / <i>w</i> <sub>4</sub>		( <i>a</i> <sub>2</sub> - H) <sup>•</sup>	<i>c</i> <sub>2</sub> / <i>x</i> <sub>2</sub>	<i>d</i> <sub>2</sub> / <i>w</i> <sub>2</sub>
					( <i>a</i> <sub>4</sub> + H) <sup>•</sup>		<i>d</i> <sub>3</sub> / <i>w</i> <sub>3</sub>
					( <i>a</i> <sub>4</sub> - H) <sup>•</sup>		<i>d</i> <sub>4</sub> / <i>w</i> <sub>4</sub>
					( <i>a</i> <sub>5</sub> + H) <sup>•</sup>		<i>d</i> <sub>5</sub> / <i>w</i> <sub>5</sub>

Note: Due to the symmetry of investigated RNAs and DNAs, ions labeled *d* could also be *w* ions, *a* could also be *z*, *b* could also be *y*, and *c* could also be *x*.

## 5.4 Conclusions

Here we presented electron capture by singly deprotonated RNA/DNA and multiply deprotonated RNA. IR activation before or after the electron capture event, denoted as AI-niECD and niECD-IRMPD, respectively, provided further fragmentation. Fragments from niECD-IRMPD of the RNA/DNA hexamers A<sub>6</sub>, G<sub>6</sub>, C<sub>6</sub> and U<sub>6</sub>/T<sub>6</sub> were compared with EDD of the same but doubly deprotonated RNA/DNA and with ECD of

the same but doubly protonated RNA/DNA. In general, niECD-IRMPD provided similar fragmentation patterns and fragmentation efficiencies as EDD. However, characteristic “ECD-like” radical fragments also suggested similarities to ECD. The latter similarities may be due to DNA/RNA zwitterionic structures in negative ion mode.

## 5.5 Bibliography

- (1) Murray, K. K. *J. Mass Spectrom.* **1996**, *31*, 1203-1215.
- (2) Hofstadler, S. A.; Sannes-Lowery, K. A.; Hannis, J. C. *Mass Spectrom. Rev.* **2005**, *24*, 265-285.
- (3) Wu, J.; McLuckey, S. A. *Int. J. Mass spectrom.* **2004**, *237*, 197-241.
- (4) McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 599-614.
- (5) Little, D. P.; Speir, J. P.; Senko, M. W.; Oconnor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.
- (6) Aaserud, D. J.; Guan, Z. Q.; Little, D. P.; McLafferty, F. W. *Int. J. Mass spectrom.* **1997**, *167*, 705-712.
- (7) McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 60-70.
- (8) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
- (9) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.
- (10) Axelsson, J.; Palmblad, M.; Hakansson, K.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 474-477.
- (11) Kruger, N. A.; Zubarev, R. A.; Carpenter, B. K.; Kelleher, N. L.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass spectrom.* **1999**, *182*, 1-5.
- (12) Hakansson, K.; Hudgins, R. R.; Marshall, A. G.; O'Hair, R. A. J. *J. Am. Soc. Mass. Spectrom.* **2003**, *14*, 23-41.
- (13) Schultz, K. N.; Hakansson, K. *Int. J. Mass spectrom.* **2004**, *234*, 123-130.

- (14) Cerda, B. A.; Horn, D. M.; Breuker, K.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124*, 9287-9291.
- (15) Budnik, B. A.; Haselmann, K. F.; Elkin, Y. N.; Gorbach, V. I.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 5994-6001.
- (16) Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2007**, *79*, 2901-2910.
- (17) Kleinnijenhuis, A. J.; Duursma, M. C.; Breukink, E.; Heeren, R. M. A.; Heck, A. J. R. *Anal. Chem.* **2003**, *75*, 3219-3225.
- (18) Olsen, J. V.; Haselmann, K. F.; Nielsen, M. L.; Budnik, B. A.; Nielsen, P. E.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 969-974.
- (19) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793-1800.
- (20) Shi, S. D. H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, *73*, 19-22.
- (21) Hakansson, 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) Chalmers, M. J.; Hakansson, K.; Johnson, R.; Smith, R.; Shen, J. W.; Emmett, M. R.; Marshall, A. G. *Proteomics* **2004**, *4*, 970-981.
- (24) Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256-3262.
- (25) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.
- (26) Sze, S. K.; Ge, Y.; Oh, H.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 1774-1779.
- (27) Anusiewicz, I.; Jasionowski, M.; Skurski, P.; Simons, J. *J. Phys. Chem. A* **2005**, *109*, 11332-11337.
- (28) Adamson, J. T.; Hakansson, K. *J. Am. Soc. Mass. Spectrom.* **2007**, *18*, 2162-2172.
- (29) Wolff, J. J.; Chi, L. L.; Linhardt, R. J.; Amster, I. J. *Anal. Chem.* **2007**, *79*, 2015-2022.
- (30) Chi, L. L.; Wolff, J. J.; Laremore, T. N.; Restaino, O. F.; Xie, J.; Schiraldi, C.; Toida, T.; Amster, I. J.; Linhardt, R. J. *J. Am. Chem. Soc.* **2008**, *130*, 2617-2625.

- (31) Yang, J.; Mo, J. J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, *77*, 1876-1882.
- (32) Mo, J. J.; Hakansson, K. *Anal. Bioanal. Chem.* **2006**, *386*, 675-681.
- (33) Taucher, M.; Breuker, K. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 918-929.
- (34) Yang, J.; Hakansson, K. *J. Am. Soc. Mass. Spectrom.* **2006**, *17*, 1369-1375.
- (35) Yoo, H. J.; Wang, N.; Zhuang, S.; Song, H.; Hakansson, K. *J. Am. Chem. Soc.* **2011**, *133*, 16790-16793.
- (36) Creese, A. J.; Cooper, H. F. *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 1263-1274.
- (37) Woods, A. S. *J. Proteome. Res.* **2004**, *3*, 478-484.
- (38) Green-Church, K.; Limbach, P. *J. Am. Soc. Mass. Spectrom.* **2000**, *11*, 24-32.
- (39) Guo, X. H.; Bruist, M. F.; Davis, D. L.; Bentzley, C. M. *Nucleic Acids Res.* **2005**, *33*, 3659-3666.
- (40) Nordhoff, E.; Karas, M.; Cramer, R.; Hahner, S.; Hillenkamp, F.; Kirpekar, F.; Lezius, A.; Muth, J.; Meier, C.; Engels, J. W. *J. Mass Spectrom.* **1995**, *30*, 99-112.
- (41) Wang, Z.; Wan, K. X.; Ramanathan, R.; Taylor, J. S.; Gross, M. L. *J. Am. Soc. Mass. Spectrom.* **1998**, *9*, 683-691.
- (42) Wan, K. X.; Gross, M. L. *J. Am. Soc. Mass. Spectrom.* **2001**, *12*, 580-589.
- (43) Limbach, P. A.; Crain, P. F.; McCloskey, J. A. *J. Am. Soc. Mass. Spectrom.* **1995**, *6*, 27-39.
- (44) de Koning, L. J.; Nibbering, N. M. M.; van Orden, S. L.; Laukien, F. H. *Int. J. Mass Spectrom. Ion Processes* **1997**, *165-166*, 209-219.
- (45) Gabelica, V.; Tabarin, T.; Antoine, R.; Rosu, F.; Compagnon, I.; Broyer, M.; De Pauw, E.; Dugourd, P. *Anal. Chem.* **2006**, *78*, 6564-6572.
- (46) Herron, W. J.; Goeringer, D. E.; McLuckey, S. A. *J. Am. Chem. Soc.* **1995**, *117*, 11555-11562.
- (47) McLuckey, S. A.; Stephenson, J. L.; Ohair, R. A. J. *J. Am. Soc. Mass. Spectrom.* **1997**, *8*, 148-154.
- (48) Wu, J.; McLuckey, S. A. *Int. J. Mass spectrom.* **2003**, *228*, 577-597.
- (49) Barlow, C. K.; Hodges, B. D. M.; Xia, Y.; O'Hair, R. A. J.; McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 281-293.

- (50) Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnsen, A. H.; Zubarev, R. A. *Anal. Chem.* **2001**, *73*, 2998-3005.
- (51) Tsybin, Y. O.; Hakansson, P.; Budnik, B. A.; Haselmann, K. F.; Kjeldsen, F.; Gorshkov, M.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1849-1854.

## Chapter 6

### Gas-Phase Proton Transfer Reactions and Segmented-flow Fraction Collection for Improved niECD of Phosphopeptides

#### 6.1 Introduction

Tandem mass spectrometry (MS/MS) has been well-developed and applied for peptide sequencing. Vibrational activation techniques such as collision activated dissociation (CAD)<sup>1</sup> are most commonly employed. However, labile posttranslational modifications (PTMs) such as phosphorylation and sulfonation are frequently preferentially lost and localization of such PTMs can therefore be challenging.<sup>2-4</sup> On the other hand, gas-phase ion-electron and ion-ion reactions such as electron capture dissociation (ECD)<sup>5</sup> and electron transfer dissociation (ETD)<sup>6</sup> are gaining popularity as alternatives to vibrational activation methods. In ECD multiply charged cations react with electrons of low energy (< 1 eV) to form charge-reduced radical intermediates that further dissociate into fragments. In ETD multiply charged cations react with radical anions and receive electrons to form similar charge-reduced radical intermediates as in ECD. One significant reason for the rising interest in ECD/ETD is their ability to retain labile PTMs.<sup>6-8</sup>

Both ECD and ETD require multiply charged cationic precursor ions. Generation of such precursor ions can be challenging for peptides with acidic PTMs such as phosphorylation and sulfonation. In negative ion mode phosphorylated and sulfonated peptides can be detected with higher sensitivity.<sup>2,9-12</sup> CAD can be performed in negative ion mode, however, neutral loss and internal fragmentation are usually abundant, which complicate spectrum interpretation.<sup>13</sup> EDD<sup>14</sup> and NETD<sup>15,16</sup> operate in negative ion mode but, in both techniques, backbone fragmentation efficiency is lower compared to ECD and ETD. Also, both EDD and NETD require multiply charged anionic precursor ions. Phosphate, water, and CO<sub>2</sub> loss are observed in EDD and NETD spectra of phosphorylated peptides.<sup>16,17</sup> Thus a more informative MS/MS technique in negative ion mode is desired.

Negative ion electron capture dissociation (niECD) is a recently developed MS/MS technique for dissociating peptides<sup>18</sup> and oligonucleotides (see Chapter 5 in this thesis) in negative ion mode. In niECD negatively charged precursor ions capture electrons with energy around 3.5-6.5 eV to form charge-increased radical intermediates. For peptides, niECD yields *c/z*-type ions with phosphate retention, analogous to ECD in positive ion mode. Also, in niECD the increased charge improves sensitivity of FT-ICR detection because induced image current is proportional to the number of charges.<sup>19</sup>

Despite the advantages of niECD, one current drawback is the relatively long irradiation time. In recently published work the reaction time for peptide niECD was 10-20 seconds.<sup>18</sup> By contrast, ECD and ETD reaction times are typically in the sub-second range. The long reaction time is likely due to Coulomb repulsion between precursor anions and electrons, thus lowering the cross section for electron capture. Precursor ion

accumulation time also contributes to the total acquisition time. Fragmentation efficiency in niECD is lower for doubly-charged precursor ions, compared with singly-charged precursors, presumably due to increased Coulomb repulsion. In electrospray ionization (ESI) of peptides, higher masses tend to carry more charges. Thus, for peptides with higher masses, less precursor ions at lower charge states are available for efficient niECD. Also, for peptides electrosprayed from a lower concentration, higher charge states are more abundant due to the higher relative concentration of charge carrying molecules in the solution. Thus, for peptides with higher masses or at lower concentration, precursor accumulation time can be a significant contributor to the total acquisition time.

Such relatively long acquisition time can be a limiting factor for coupling niECD online with liquid chromatography (LC). In chromatography optimum separation and high peak capacity are pursued to improve detection sensitivity and the number of detected analytes. However, a narrower chromatographic peak allows less time for MS detection. Two possible routes can be envisioned to avoid this synchronization problem. One approach is to decrease the niECD acquisition time and the other is to perform offline LC with fraction collection that can preserve LC separation.

Reduction of the niECD acquisition time, particularly for peptides with higher masses or at low concentration, may be achieved by increasing the abundance of lower charge state precursor ions via a proton transfer reaction (PTR).<sup>20, 21</sup> PTRs involve gas-phase reactions between precursor ions and reagent ions of opposite polarity in order to manipulate the charge state of the precursor ions. PTRs have shown utility for reducing the number of precursor ion charges,<sup>22</sup> charge-state purification,<sup>23</sup> charge inversion,<sup>24</sup> and metal addition.<sup>25</sup>

Nano-LC, based on sub-millimeter i.d. capillary columns, offers high chromatographic resolution and high detection sensitivity compared to conventional LC.<sup>26,27</sup> Online coupling of nano-LC with mass spectrometry is widely applied because mass spectrometry offers high sensitivity, high speed, and the capability of analyzing complex mixtures.<sup>28,29</sup> However, fraction collection and offline ESI-MS may be desired because samples can be stored, transferred, and analyzed with more flexibility. Microliter scale fractions can be collected in well-plates.<sup>30</sup> However, smaller fraction volumes are needed to preserve nano-LC separation. Fraction collection via compartmentalization with an immiscible phase has been implemented in capillary electrophoresis<sup>31</sup> and nano-LC.<sup>32</sup> In this technique, nano-LC eluent is collected into a tubing or a microfabricated channel by forming plugs separated by an immiscible oil from a T junction. Collected droplets can be infused for ESI-MS with preservation of LC separation and at variable flow rates for optimum analysis of each plug.<sup>33,34</sup>

Here we investigate how PTR prior to niECD affects the acquisition time and the sensitivity of niECD for phosphorylated and sulfonated peptides with higher masses or in low concentration. Also, we coupled fraction collection by segmented flow to niECD.

## **6.2 Experimental Section**

### **6.2.1 Sample Preparation**

The peptides H-SFVLNPTNIGMpSKSSQGHVTK-OH (DAM1), H-FQpSEEQQTDELQDK-OH (BBC), H-TRDIYETDpYYRIK-OH (KDI1), H-TRDIpYETDpYpYRIK-OH (KDI3), H-DLDVPIPGRFDRRVVpSVAAE-OH (PRSIIS), and H-DFEEIPEEsYLQ-OH (Hirudin-1, 55-65) were purchased from AnaSpec, Inc.

(Fremont, CA). H-RRREEEpSEEEAA-OH and H-DRVYIHPFHL-OH (Angiotensin I) were purchased from Sigma-Aldrich (St. Louis, MO) and H-DsYMGWMDF-NH<sub>2</sub> (CCKS) was purchased from Advanced Chemtech (Louisville, NY). H-TSTEPQpYQPGENL-NH<sub>2</sub> was from Millipore (Billerica, MA) and H-KRSpYEEHIP-OH (NPF) was purchased from GenScript (Piscataway, NJ). Bovine milk  $\alpha$ -casein was from Sigma-Aldrich (St. Louis, MO). All chemicals were used without further purification. Trypsin (Promega, Madison, WI) digestion of  $\alpha$ -casein was performed for 12 h at 37 °C at an enzyme/substrate ratio of 1:50. H-VPQLEIVPNpSAEER-OH ( $\alpha$ -casein 121-134), H-DIGpSEpSTEDQAMEDIK-OH ( $\alpha$ -casein 58-73) and H-TVDMEpSTEVFTK-OH ( $\alpha$ -casein 153-164) were subjected to MS/MS. Peptides and protein digest were dissolved in 1:1 (v/v) water/acetonitrile with 0.1% formic acid. Acidic electrospray solution was used to mimic the typical mobile phase in RPLC. Samples were infused at 80  $\mu$ L/h via an Apollo II (Bruker Daltonics, Billerica, MA) electrospray ion source.

A mixture of BBC, NPF, and DAM1 (0.2 pmoles of each) was separated on an in-house-made capillary C18 column at 300 nL/min provided by an ultra performance liquid chromatography (UPLC) pump (NanoAcquity, Waters, Milford, MA). The 75  $\mu$ m i.d. x 10 cm column was packed with Waters dC18 3 $\mu$ m particles. Mobile phases were, A: 10 mM ammonium formate, 0.15% formic acid, B: 0.15% formic acid in methanol. The solvent gradient was 0 to 70 % B in 22 min. The nano-LC eluent entered a T junction into which perfluorodecalin was pumped at 150 nL/min, following a procedure developed by Li *et al.*<sup>34</sup> This configuration generated 30 nL fractions separated by ~15 nL perfluorodecalin. This segmented flow was collected into a 150  $\mu$ m i.d. high purity

perfluoroalkoxy plus (HPFA+) tubing (Upchurch Scientific, Oak Harbor, OR). The tubing with collected fractions was connected to a Pt-coated fused silica ESI emitter with a 5  $\mu\text{m}$  i.d. tip (New Objective, Woburn, MA). Droplets from the segmented flow were infused at 20-100 nL/min via an Apollo II electrospray ion source modified for nano-ESI. The emitter tip was at a 45° angle with respect to the inlet orifice of the mass spectrometer. The tip position was manually adjusted to optimize the electrospray process. A voltage of ~1300 V was applied to the tip.

### 6.2.2 Mass Spectrometry

PTR mass spectra were acquired on a 7T Solarix ESI-Q-FT-ICR mass spectrometer (Bruker Daltonics) equipped with an ETD module. PTR in the external hexapole was performed by first accumulating precursor anions for 0.5 s to 5 s, followed by accumulation of protonated fluoranthene cations (203 Da) for 0.5 s to 1 s. Singly protonated fluoranthene cations were generated in an external chemical ionization (CI) source with methane as carrier gas. niECD was performed in the ICR cell by irradiating reacted precursor anions with electrons from an indirectly heated hollow cathode filtered by a ring electrode. The cathode bias voltage was - 3.5 to - 4.5 V and a lens located between the cathode and the ICR cell was at - 2.5 to - 3.5 V. The irradiation time was 3 to 5 s. Spectra were processed by DataAnalysis 4.0 software (Bruker Daltonics) and peaks were assigned within 10 ppm error after internal calibration. Precursor ions and the charge-increased radical species were used as internal calibrants.

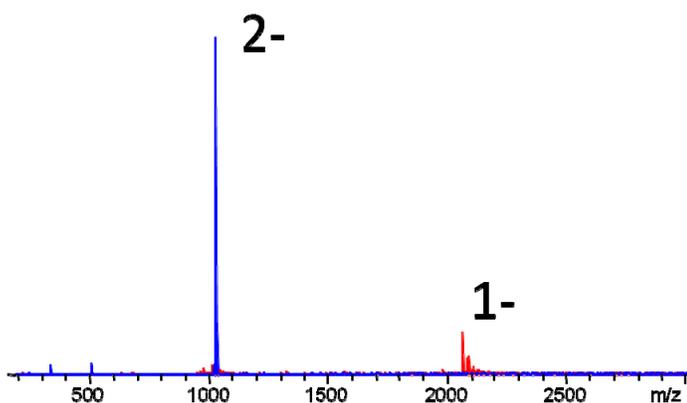
Online nanoESI-MS was performed on an LCQ DECA XP<sup>+</sup> (Thermo-Finnigan, San Jose, CA) ion trap mass spectrometer. niECD mass spectra of fraction-collected

peptides were acquired on a 7T Apex ESI-Q-FT-ICR mass spectrometer (Bruker Daltonics). Precursor anions were accumulated in the external hexapole for 10 s. niECD was performed with the same method as above. The cathode bias voltage was - 6 V and the lens voltage was - 4.8 V with an irradiation time of 4 s.

## 6.3 Results and Discussion

### 6.3.1 niECD of Doubly and Singly Deprotonated Phosphopeptide BBC

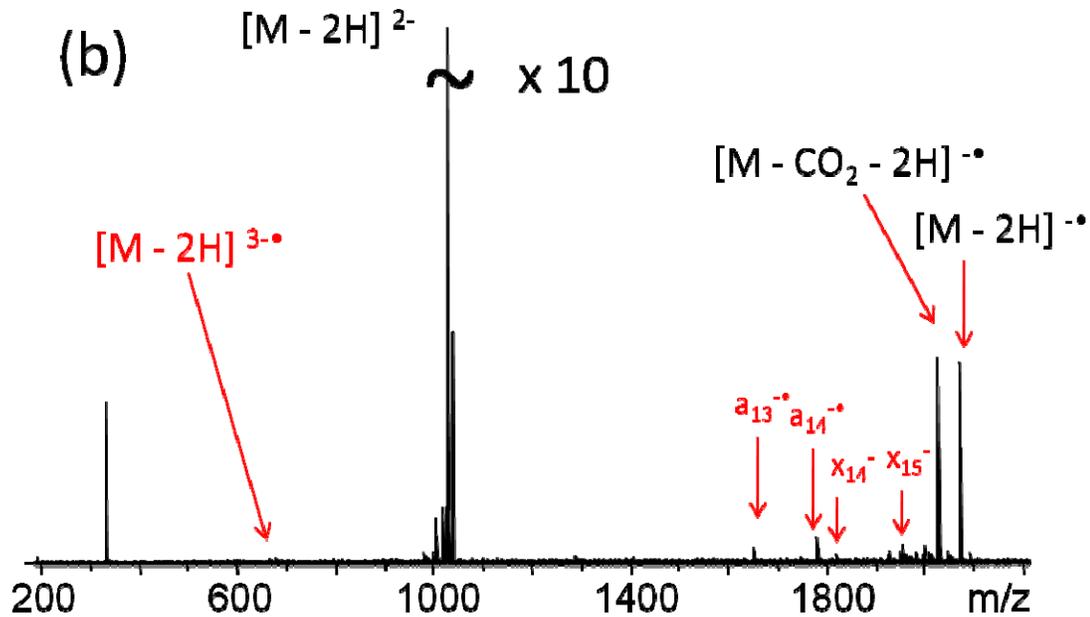
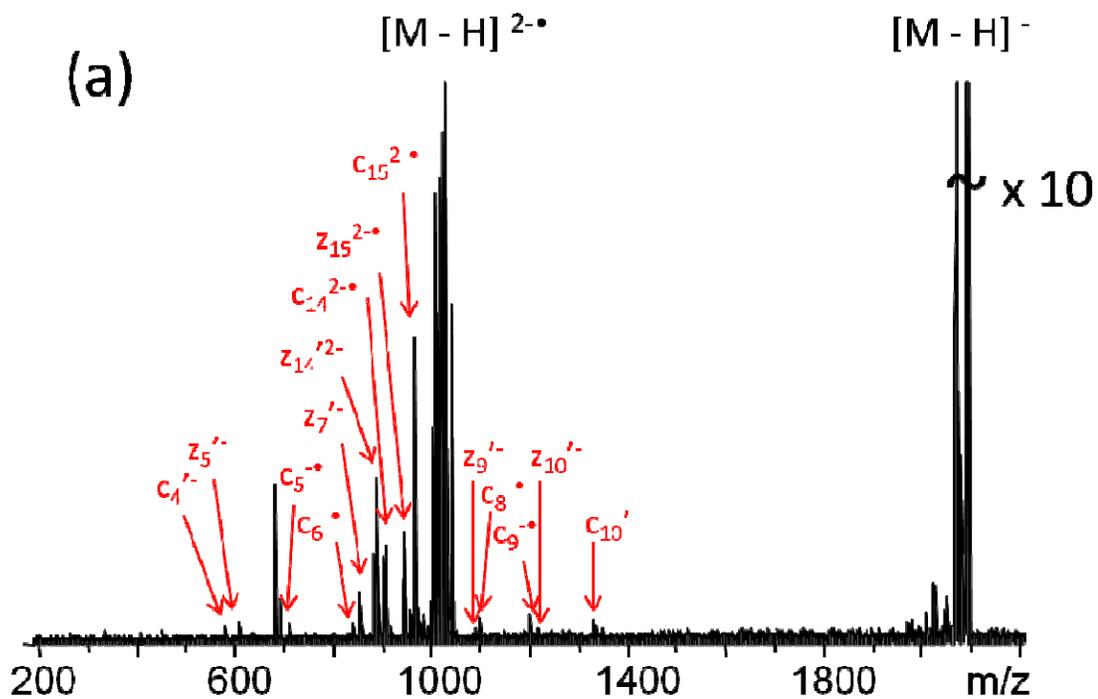
The phosphopeptide BBC (H-FQpSEEQQQTEDELQDK-OH, monoisotopic mass 2060.8211) was electrosprayed in negative ion mode at a concentration of 0.5  $\mu$ M as shown in Figure 6.1.



**Figure 6.1.** Overlay of two negative ion mode ESI mass spectra of the phosphopeptide H-FQ-pS-EEQQQTEDELQDK-OH at 0.5  $\mu$ M. The red spectrum was optimized for the 1- charge state and the blue spectrum was optimized for the 2- charge state. Several ion source and ion transfer parameters were tuned to optimize detection of 2- and 1- ions, respectively, with the time-of-flight between the external hexapole and the ICR cell having the largest influence.

The mass spectrometer can be tuned for optimum detection of either the 1- or 2- charge state of a certain peptide. In order to make a fair comparison of the signal abundance of the 1- and 2- charge states of BBC, two ESI spectra were acquired from the

same sample but with different sets of parameters for mass spectrometric detection, each optimized for detecting one of these charge states. The flight time allowed for ions to be transferred from the external hexapole to the ICR cell had the largest influence. Two ESI spectra resulting from each of these two conditions are overlaid in Figure 6.1. The ion abundance of doubly deprotonated BBC was four times higher than singly deprotonated BBC, following normalization to charge state. Both the 1- and 2- ions were quadrupole isolated, externally accumulated in a hexapole for the same amount of time (3 s) and subjected to niECD in the ICR cell. The corresponding niECD spectra of singly and doubly deprotonated BBC are shown in Figure 6.2.



**Figure 6.2.** niECD of the singly (a) and doubly (b) deprotonated phosphopeptide H-FQpSEEQQQTEDELQDK-OH. The cathode voltage was - 4.1 V and the irradiation time was 4 s.

In the niECD spectrum of singly deprotonated BBC, the  $[M - H]^{2-\bullet}$  radical from electron capture was observed as the most abundant product ion and *c/z*-type ions were also dominant. 14 out of 30 possible *c/z*-type ions were detected corresponding to cleavage of 12 out of 15 backbone amine bonds. Phosphate was retained on all product ions containing the phosphoserine residue. Abundant  $c\bullet/c'$  and  $z'/z\bullet$  ions were observed from hydrogen migration, a phenomenon that also has been observed in conventional ECD.<sup>35</sup> In cation ECD, it has been shown that the extent of hydrogen migration is related to the lifetime of radical ion intermediates,<sup>36</sup> which is affected by the amount of intramolecular non-covalent interactions.<sup>37</sup> For niECD it has been proposed that zwitterionic structures are important for electron capture to occur.<sup>18</sup> Within a zwitterion the proton may be stabilized by a negatively charged group, such as the C-terminus or acidic side chains. Such salt bridges could contribute to the high degree of observed hydrogen transfer. Another observation was the shorter irradiation time (4 s) compared to the previous report<sup>18</sup> (10-20 s). One reason for this time improvement could be that the previous niECD was performed on a different instrument and, thus, electron and precursor ion trajectories can be different. Another reason may be that BBC (2061 Da) is larger than most peptides investigated in the previous work. More structural flexibility of larger peptides could enhance self-solvation of a positive charge. Thus, larger peptides could form more stable zwitterionic structures that are more likely to capture electrons.

Following electron irradiation of doubly deprotonated BBC, an  $[M - 2H]^{3-\bullet}$  radical from electron capture was observed at low abundance but no *c/z* type ions were detected. An electron detachment species,  $[M - 2H]^\bullet$ , and *a/x*-type ions were also observed. Coulomb repulsion between doubly negatively charged precursor ions and electrons is

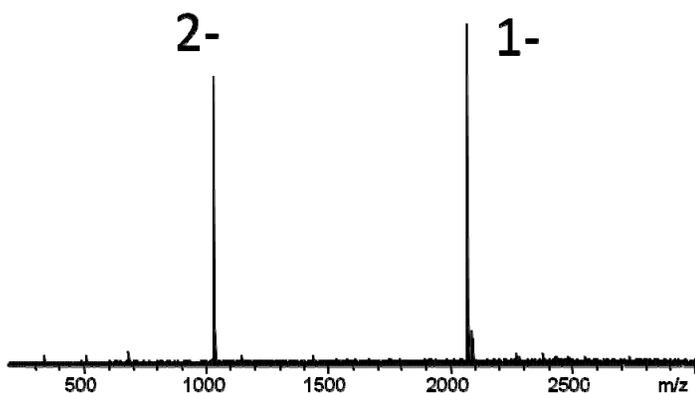
more significant than between singly charged precursor ions and electrons. Thus electron capture is less likely to occur for doubly charged anions. Another reason for this low level of electron capture may be that a hexapole is used to transport ions into the ICR cell, thus focusing the ions into a compressed packet in both the radial and axial directions. The tightness of the ion packet may lower the overlap with the electron beam. In our previous niECD work, a doubly charged peptide was fragmented by niECD into *c/z*-type ions. However, that experiment was performed on a different FT-ICR instrument with a series of ion lenses instead of a hexapole. Ions transported by the series of ion lenses are less axially focused compared with the hexapole and the overlap between ions and the electron beam may be larger than in the hexapole-equipped instrument. One possible explanation for the observation of electron detachment and *a/x*-type ions may also be related to the specific instrument configuration. Electrons from the cathode travel in the opposite direction of the analyte ions. Electrons with an average energy of 4 eV travel through the ICR cell (the trapping plates are at less than -1 V). The “side-kick” electrode pair located on the opposite side of the analyzer cell was set to -3 V and -11 V. Thus, electrons may be turned around by these voltages to be reflected back into the ICR cell at higher than the original energy.<sup>38,39</sup> Such higher energy electrons could cause the observed EDD-type product ions. Electron reflection should also happen in niECD of 1-ions. However, the corresponding charge-reduced product ions would be neutral and thus cannot be detected.

In the ESI spectrum the ion abundance of doubly deprotonated BBC was four times higher than the singly deprotonated form, however, niECD of singly charged BBC provided significantly more sequence information. In other words, a large portion of

peptide anions generated by ESI could not be efficiently utilized for niECD and sequence analysis.

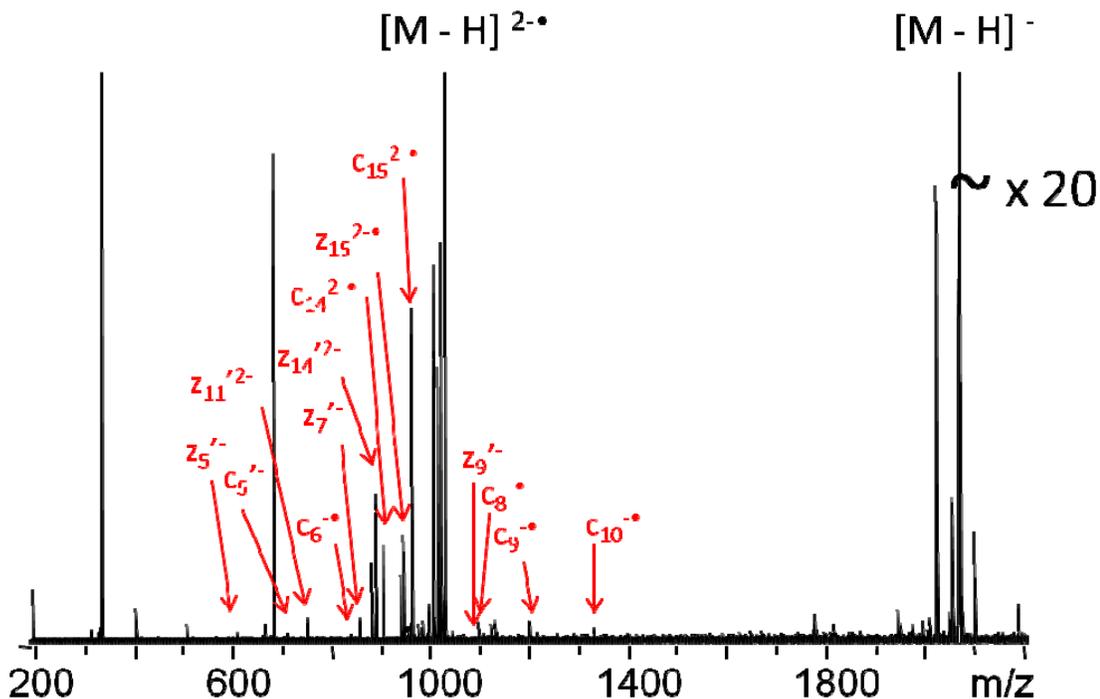
### 6.3.2 PTR-assisted niECD of the Doubly Deprotonated Phosphopeptide BBC

PTR has been utilized to reduce the charge of electrosprayed proteins.<sup>40-42</sup> In our case it is desired to use PTR to convert doubly charged peptide anions to their singly charged form and thus improve niECD efficiency. Doubly charged peptide anions were accumulated in the external hexapole and reacted with singly protonated fluoranthene cations. As shown in Figure 6.3, the abundance of singly deprotonated peptide ions increased after PTR. The proton transfer reaction yield was 30%, based on signal abundance normalized to charge state. The optimum PTR reaction time was found to be 1 s to maximize the abundance of 1- ions. Longer PTR further neutralizes 1- ions and thus reduces their abundance.



**Figure 6.3.** Mass spectrum following PTR between doubly deprotonated peptide anions and protonated fluoranthene.

Reaction products were transferred to the analyzer cell for niECD experiments. Ions were dissociated under the same conditions as in Figure 6.2. The corresponding niECD spectrum is shown in Figure 6.4.

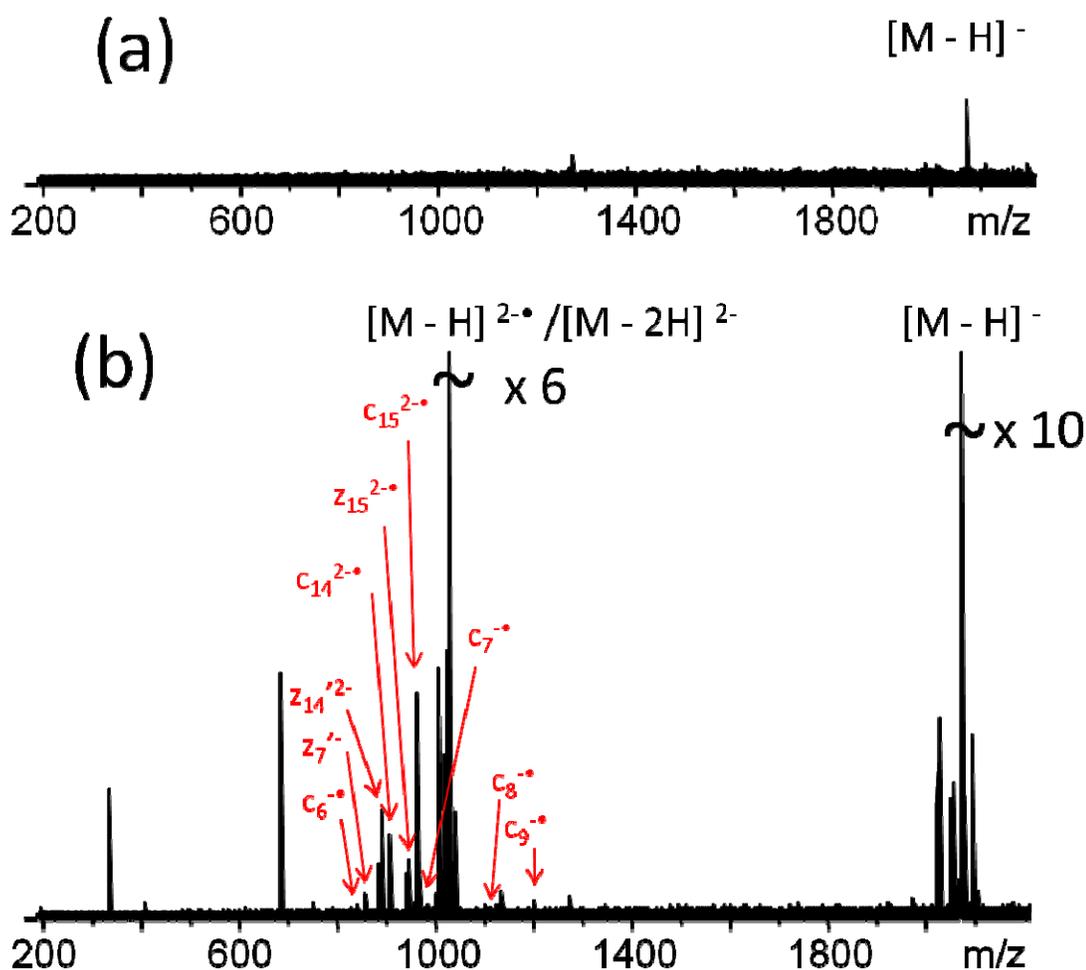


**Figure 6.4.** PTR-niECD spectrum of the doubly deprotonated phosphopeptide H-FQpSEEQQQTEDELQDK-OH. The PTR time was 1 s, the cathode voltage was - 4.1 V and the irradiation time was 4 s.

Similar sequence coverage was obtained in niECD of singly deprotonated H-FQpSEEQQQTEDELQDK-OH and PTR-niECD of the doubly deprotonated form. For PTR-niECD, 13 possible *c/z* ions were detected, corresponding to 11 backbone cleavages, compared to 14 *c/z* ions and 12 backbone cleavages in niECD alone.  $z_{11}^{2-}$  ions were unique to PTR-niECD and a  $c_4$  ion was unique to niECD. Due to the low abundance of the 1- charge state following ESI of the peptide H-FQpSEEQQQTEDELQDK-OH, 5-10 s external hexapole accumulation was required for efficient niECD of this charge state. By contrast, with PTR, the 2- charge state could be accumulated for only 1.5 s. Thus,

PTR shortens the experiment time for this peptide while providing similar sequence coverage.

In ESI, peptides at lower concentration tend to carry more charges, due to the relatively higher concentration of charge carrying molecules in the electrospray solution. Thus, niECD can be more challenging at lower concentration. In order to compare niECD and PTR-niECD performance, the phosphopeptide BBC was electrosprayed at a concentration of 0.1  $\mu\text{M}$  (compared to 0.5  $\mu\text{M}$  in the previous experiments shown in Figures 6.1 to 6.4) to generate 1- and 2- charge states. The niECD spectrum from the 1- charge state is shown in Figure 6.5(a) whereas the PTR-niECD spectrum of the 2- charge state is shown in Figure 6.5(b).



**Figure 6.5.** (a) niECD of the singly deprotonated phosphopeptide H-FQ-pS-EEQQTEDELQDK-OH at 0.1  $\mu$ M. The spectrum is shown at 100% scale. (b) PTR-niECD of the doubly deprotonated form of the same peptide.

For the niECD experiment of BBC in the 1- charge state, precursor ions were accumulated for 5 s. However, the abundance was so low that no fragments were observed. For PTR-niECD of BBC in the 2- charge state, precursor ions were accumulated for the same time, resulting in detection of 10 out of 30 possible *c/z* ions, corresponding to eight out of 15 possible backbone cleavages. It is clear that, in this case, PTR significantly improved the sensitivity of niECD.

### 6.3.3 Effect of PTR on Sensitivity and Product Ion Yield in niECD

As shown above, PTR prior to niECD improved the sensitivity of niECD for the phosphopeptide BBC. We investigated more peptides, mostly phosphopeptides, to evaluate the sensitivity improvement by PTR, as summarized in Table 6.1.

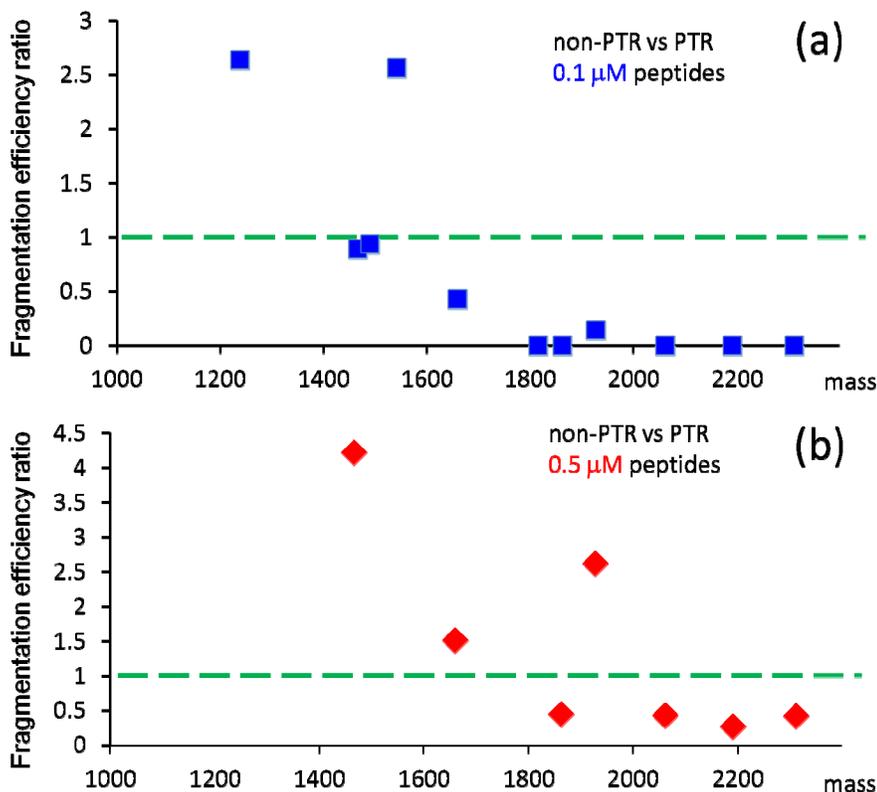
mass	PTR - niECD of M <sup>2-</sup>	niECD of M <sup>-</sup>
2311	S F V L N P T N I G M [pS] K S S Q G <sup>1</sup> H V <sup>1</sup> T <sup>1</sup> K	S F V L N P T N I G M pS K S S Q G H V T K
2191	D L D V P I P G R F D R R V [pS] V A A <sup>1</sup> E	D L D V P I P G R F D R R V pS V A A E
2061	F Q [pS] E E Q Q <sup>1</sup> T <sup>1</sup> E D E L Q D <sup>1</sup> K	F Q pS E E Q Q Q T E D E L Q D K
192/	D I G [pS] E [pS] I E D Q A M E <sup>1</sup> D <sup>1</sup> I <sup>1</sup> K	D I G pS E pS I E D Q A M E D I K
1862	T R D I pY <sup>1</sup> E T D pY pYR <sup>1</sup> K	T R D I pY E T D pY pY R K
1702	T R D I Y E [T] D pY Y R <sup>1</sup> K	T R D I Y E T D pY Y R K
1660	V P [Q] L E L V P N pS A E E R	V P Q L E L V P N pS A E E R
1542	T S T E P Q pY Q P G [E <sup>1</sup> N <sup>1</sup> ] L-NH <sub>2</sub>	T S T E P Q pY Q P G [E <sup>1</sup> N <sup>1</sup> ] L-NH <sub>2</sub>
1491	D F [E] [E] I P [E] [E] S Y L <sup>1</sup> Q	D F [E] [E] I P [E] [E] S Y L <sup>1</sup> Q
1466	T V D M [E] [pS] T E V F T K	T V D M [E] [pS] T E V F T <sup>1</sup> K
1296	D R V Y I H P F H L	D R V Y I H P F H L
1238	K R S [pY] E E H <sup>1</sup> I <sup>1</sup> P	K R S [pY] E <sup>1</sup> E <sup>1</sup> H <sup>1</sup> I <sup>1</sup> P

**Table 6.1.** Comparison of sequence coverage in the form of *c/z*-type fragments from PTR-niECD of doubly deprotonated peptides and niECD of the same peptides in their singly charged form. The total acquisition time for PTR-niECD and niECD was the same for both charge states. All peptides were electrosprayed at 0.1 μM.

Based on Table 6.1, for investigated peptides heavier than 1702 Da, more backbone cleavages were observed when PTR was applied. In several cases (H-SFVLNPTNIGMpSKSSQGHVTK-OH, H-DLDVPIPIGRFDRRVpSVAAE-OH, H-FQpSEEQQQTEDELQDK-OH, H-TRDIpYETDpYpYRK-OH and H-TRDIYETDpYYRK-OH), no backbone cleavages were observed from niECD of singly deprotonated peptides whereas PTR-niECD provided 27-61% sequence coverage. For H-SFVLNPTNIGMpSKSSQGHVTK-OH and H-DLDVPIPIGRFDRRVpSVAAE-OH the phosphorylation sites could be precisely located. For peptides between 1660 Da and 1296 Da, the sequence coverage was the same in PTR-niECD and direct niECD

experiments. In these cases both the 2- and 1- charge states of the electrosprayed peptides were abundant and PTR did not seem to improve the sequence coverage. For the peptide H-KRSpYEEHIP-OH with a mass of 1238 Da, the 1- charge state was the major form in the ESI spectrum. In this case PTR was not advantageous.

One way to evaluate the duty cycle of one MS/MS technique is to compare the product ion yield for a certain acquisition time. If one MS/MS technique yields more backbone fragments within a shorter time period it would improve the performance of LC-MS/MS. We compared the product ion yields for the same acquisition time between PTR-niECD and niECD, as shown in Figure 6.6.



**Figure 6.6.** Comparison of fragmentation efficiencies in non-PTR niECD and PTR-niECD at two peptide concentrations. (a) Peptides were electrosprayed at 0.1 μM. (b) Peptides were sprayed at 0.5 μM.

In Figure 6.6, fragmentation efficiency is defined as the summed abundance of assignable *c/z* type ions divided by the acquisition time of a single scan. The plotted fragmentation efficiency ratio is between non-PTR niECD of singly deprotonated peptides and PTR-niECD of the same peptides but in their doubly deprotonated form. A ratio of 1 means that non-PTR niECD and PTR-niECD generate the same product ion abundance within the same acquisition time, as represented by the dotted line.

In Figure 6.6a, peptides were electrosprayed at 0.1  $\mu\text{M}$ . As the peptide mass increases, the fragmentation efficiency ratio decreases, demonstrating that improvement of niECD from PTR is more significant. For peptides heavier than  $\sim 1800$  Da, product ions generated by niECD alone were limited and, in order to detect significant fragments from these peptides at 0.1  $\mu\text{M}$ , PTR was necessary. For most peptides from  $\sim 1400$  to  $\sim 1800$  Da, niECD provided detectable abundance of product ions. However, PTR-niECD provided more. The only exception is H-TSTEPQpYQPGENL-NH<sub>2</sub> (1542 Da). One possible reason could be that the product ion abundance was very low in both niECD and PTR-niECD and thus the abundance ratio calculations have large error. For the peptide lighter than  $\sim 1400$  Da, niECD alone provided more abundant product ions than PTR-niECD.

In Figure 6.6b, peptides were electrosprayed at 0.5  $\mu\text{M}$ . Similar to the case of 0.1  $\mu\text{M}$ , as the peptide mass increases, the fragmentation efficiency ratio decreases. For most peptides heavier than  $\sim 1800$  Da, product ions generated by niECD alone were detectable, while PTR-niECD provided approximately twice the abundance. For peptides lighter than  $\sim 1700$  Da, niECD provided more abundant product ions. From the general trend

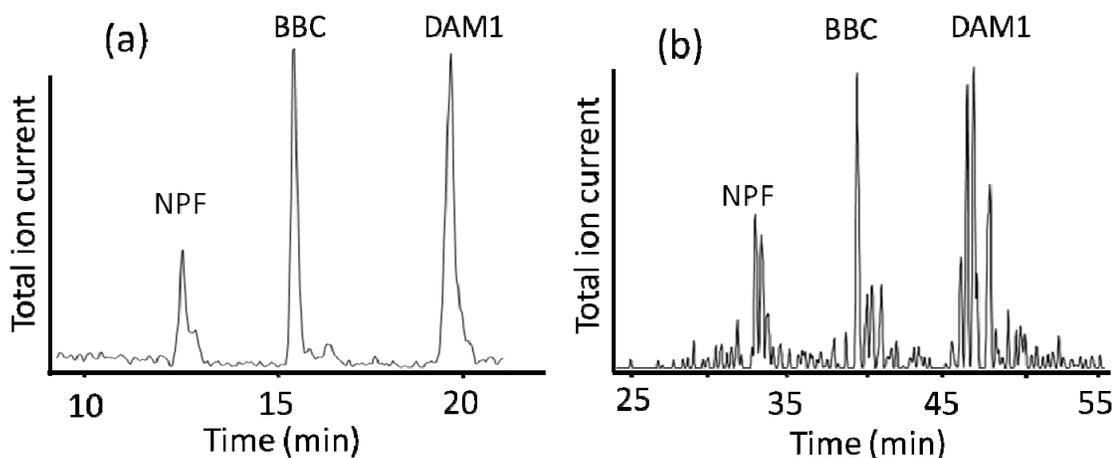
line in Figure 6.6b, PTR-niECD is more efficient for 0.5  $\mu$ M peptides heavier than approximately 1700 Da, while, for 0.1  $\mu$ M peptides, PTR-niECD is more efficient for peptides heavier than approximately 1500 Da. This difference between the threshold masses could arise from the fact that peptides electrosprayed at lower concentration tend to carry more charges.

In proteomic studies trypsin is the most commonly employed protease. The length distribution of tryptic peptides was demonstrated to be centered around 13-15 amino acid residues.<sup>43</sup> From these numbers, the mass distribution can be estimated to be centered around 1445-1667 Da, based on the average mass of natural amino acids.<sup>44</sup> As shown in Figure 6.6 the threshold mass for PTR-niECD to be more effective than niECD alone is around 1500-1700 Da, which overlaps with the center of the tryptic peptide mass distribution. Thus PTR could potentially improve the fragmentation efficiency for roughly half of tryptic peptides in a proteomic study based on niECD, provided that niECD is successful.

#### **6.3.4. Coupling Segmented-flow Fraction Collection and niECD**

Coupling niECD to offline-nanoLC by segmented-flow fraction collection could potentially be advantageous over coupling to online-nanoLC because the infusion flow rate of the collected droplets can be lower than the nanoLC flow rate. A lower flow rate allows more time for the mass spectrometer to perform niECD and signal average over multiple scans. Fraction collection by segmented flow may better preserve temporal resolution than conventional fraction collection techniques due to the small size of each droplet.<sup>33, 34</sup>

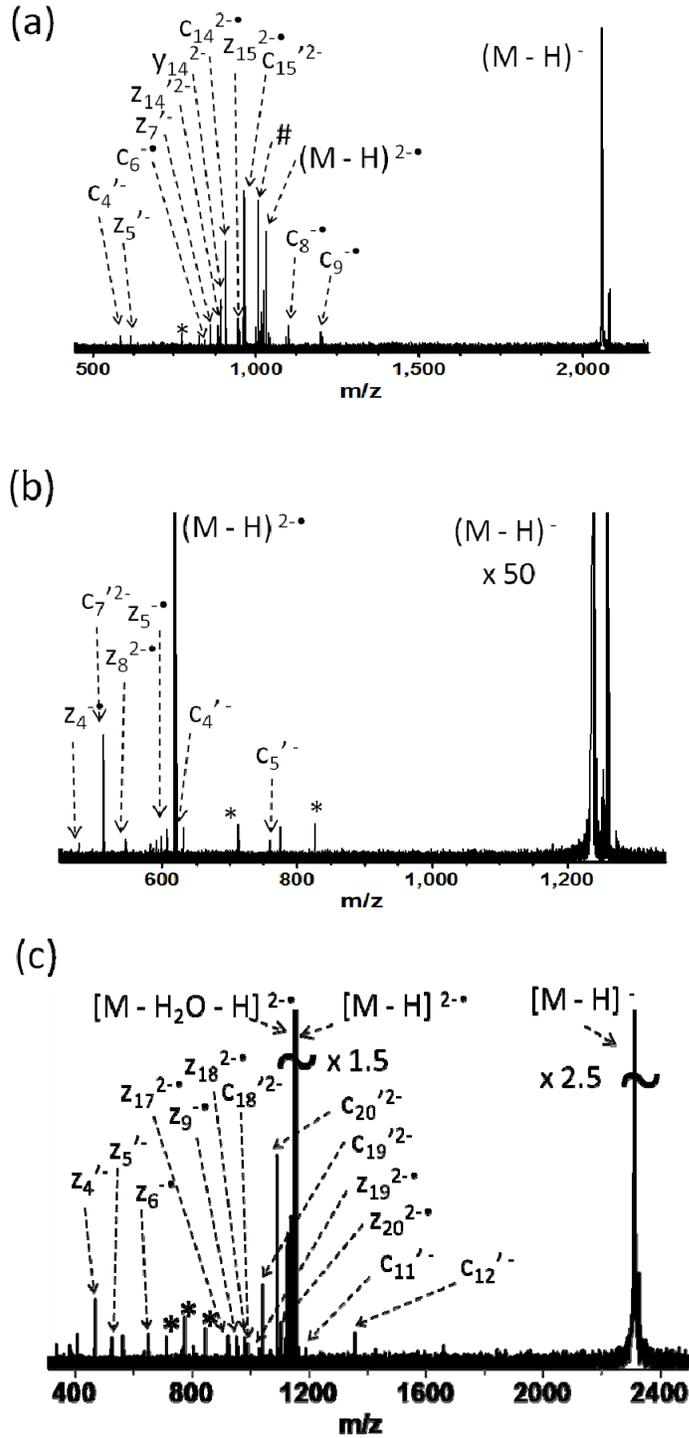
Figure 6.7a shows the total ion chromatogram of a mixture of peptides (NPF, BBC and DAM1) separated by nano-LC and analyzed by online-nanoESI-MS. The nano-LC flow rate was 300 nL/min. Figure 6.7b shows the total ion chromatogram of the same mixture separated by nano-LC, collected by segmented flow and analyzed by offline-nanoESI-MS at a flow rate of 100 nL/min. It can be seen that LC temporal resolution is preserved by the segmented-flow fraction collection.



**Figure 6.7.** Total ion chromatograms of a peptide mixture (NPF, BBC, DAM1) separated by nano-LC. (a) Eluent was analyzed by online-nanoESI-MS at a flow rate of 300 nL/min. (b) Eluent was collected by segmented-flow fraction collection and analyzed by offline-nanoESI-MS at 100 nL/min. The mass spectrometer was a Thermo-Finnigan LCQ ion trap instrument.

The same mixture of peptides were separated by nanoLC, collected by segmented-flow fraction collection and subjected to niECD. The infusion flow rate was 100 nL/min when none of the three peptides was present in the droplets. The flow rate was slowed down to 20 nL/min when one of the three peptides was observed and niECD was performed. Figure 6.8 shows niECD spectra of NPF, BBC and DAM1 from infusion

of the segmented flow into an FT-ICR mass spectrometer.



**Figure 6.8.** niECD spectra of (a) NPF, (b) BBC and (c) DAM1. Peptides were collected by segmented flow fraction collection from nano-LC and infused into a nanoESI-FT-ICR mass spectrometer. Spectra were signal-averaged over 4-8 scans.

In the niECD spectrum of NPF, four out of eight total possible backbone cleavages were observed corresponding to seven out of 16 possible *c/z*-type ions. In the niECD spectrum of BBC, nine out of 15 total possible backbone cleavages were observed corresponding to 10 out of 30 possible *c/z*-type ions. For niECD of DAM1, 12 out of 20 total possible backbone cleavages were observed corresponding to 13 out of 40 possible *c/z*-type ions. In the cases of BBC and DAM1, phosphorylation sites could be directly located.

## 6.4 Conclusion

Peptides tend to carry more charges as their sizes increase or concentrations decrease. However, niECD is more efficient for peptide anions with fewer charges. Precursor ion accumulation can be a significant contributor to the total acquisition time in an niECD experiment. Here we demonstrated that PTR prior to niECD can decrease the ion accumulation time, and improve niECD sensitivity. Among the peptides investigated, niECD performance of most peptides heavier than 1500-1700 Da was improved by PTR. For several peptides electrosprayed at 0.1  $\mu\text{M}$ , niECD of the singly charged form did not provide any detectable fragments, while PTR-niECD of the doubly charged form provided 27-61% sequence coverage. Also, we demonstrated that niECD can be coupled to offline nano-LC by segmented flow fraction collection.

## 6.5 Bibliography

- (1) Hayes, R. N.; Gross, M. L. *Methods Enzymol.* **1990**, *193*, 237-263.
- (2) Carr, S. A.; Huddleston, M. J.; Annan, R. S. *Anal. Biochem.* **1996**, *239*, 180-192.
- (3) McLachlin, D. T.; Chait, B. T. *Curr. Opin. Chem. Biol.* **2001**, *5*, 591-602.

- (4) Mann, M.; Ong, S. E.; Gronborg, M.; Steen, H.; Jensen, O. N.; Pandey, A. *Trends Biotechnol.* **2002**, *20*, 261-268.
- (5) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
- (6) 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.
- (7) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793-1800.
- (8) Shi, S. D. H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, *73*, 19-22.
- (9) Neubauer, G.; Mann, M. *Anal. Chem.* **1998**, *71*, 235-242.
- (10) Kweon, H. K.; Hakansson, K. *Anal. Chem.* **2006**, *78*, 1743-1749.
- (11) Yagami, T.; Kitagawa, K.; Aida, C.; Fujiwara, H.; Futaki, S. *The Journal of Peptide Research* **2000**, *56*, 239-249.
- (12) Nemeth-Cawley, J. F.; Karnik, S.; Rouse, J. C. *J. Mass Spectrom.* **2001**, *36*, 1301-1311.
- (13) Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Mass Spectrom. Rev.* **2002**, *21*, 87-107.
- (14) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.
- (15) Coon, J. J.; Shabanowitz, J.; Hunt, D. F.; Syka, J. E. P. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 880-882.
- (16) Huzarska, M.; Ugalde, I.; Kaplan, D. A.; Hartmer, R.; Easterling, M. L.; Polfer, N. C. *Anal. Chem.* **2010**, *82*, 2873-2878.
- (17) Kweon, H. K.; Hakansson, K. *J. Proteome. Res.* **2008**, *7*, 749-755.
- (18) Yoo, H. J.; Wang, N.; Zhuang, S.; Song, H.; Hakansson, K. *J. Am. Chem. Soc.* **2011**, *133*, 16790-16793.
- (19) Fung, Y. M. E.; Adams, C. M.; Zubarev, R. A. *J. Am. Chem. Soc.* **2009**, *131*, 9977-9985.
- (20) Herron, W. J.; Goeringer, D. E.; McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **1995**, *6*, 529-532.
- (21) Huang, T. Y.; McLuckey, S. A. *Proteomics* **2010**, *10*, 3577-3588.

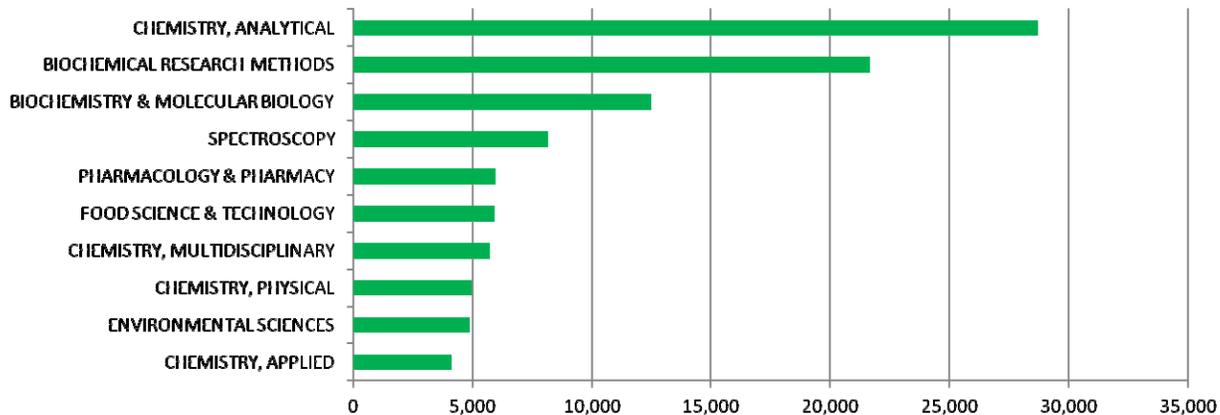
- (22) Stephenson, J. L.; McLuckey, S. A. *Anal. Chem.* **1996**, *68*, 4026-4032.
- (23) Reid, G. E.; Shang, H.; Hogan, J. M.; Lee, G. U.; McLuckey, S. A. *J. Am. Chem. Soc.* **2002**, *124*, 7353-7362.
- (24) He, M.; McLuckey, S. A. *J. Am. Chem. Soc.* **2003**, *125*, 7756-7757.
- (25) Newton, K.; McLuckey, S. *J. Am. Soc. Mass. Spectrom.* **2004**, *15*, 607-615.
- (26) Kennedy, R. T.; Jorgenson, J. W. *Anal. Chem.* **1989**, *61*, 1128-1135.
- (27) Meiring, H. D.; van der Heeft, E.; ten Hove, G. J.; de Jong, A. P. J. M. *J. Sep. Sci.* **2002**, *25*, 557-568.
- (28) Hunt, D. F.; Henderson, R. A.; Shabanowitz, J.; Sakaguchi, K.; Michel, H.; Sevilir, N.; Cox, A. L.; Appella, E.; Engelhard, V. H. *Science* **1992**, *255*, 1261-1263.
- (29) McCormack, A. L.; Schieltz, D. M.; Goode, B.; Yang, S.; Barnes, G.; Drubin, D.; Yates, J. R. *Anal. Chem.* **1997**, *69*, 767-776.
- (30) Van Pelt, C. K.; Zhang, S.; Fung, E.; Chu, I.; Liu, T.; Li, C.; Korfmacher, W. A.; Henion, J. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1573-1578.
- (31) Edgar, J. S.; Milne, G.; Zhao, Y.; Pabbati, C. P.; Lim, D. S. W.; Chiu, D. T. *Angew. Chem. Int. Ed.* **2009**, *48*, 2719-2722.
- (32) Niu, X. Z.; Zhang, B.; Marszalek, R. T.; Ces, O.; Edel, J. B.; Klug, D. R.; deMello, A. J. *Chem. Commun.* **2009**, 6159-6161.
- (33) Pei, J.; Li, Q.; Lee, M. S.; Valaskovic, G. A.; Kennedy, R. T. *Anal. Chem.* **2009**, *81*, 6558-6561.
- (34) Li, Q.; Pei, J.; Song, P.; Kennedy, R. T. *Anal. Chem.* **2010**, *82*, 5260-5267.
- (35) 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.
- (36) O'Connor, P. B.; Lin, C.; Cournoyer, J. J.; Pittman, J. L.; Belyayev, M.; Budnik, B. A. *J. Am. Soc. Mass. Spectrom.* **2006**, *17*, 576-585.
- (37) Tsybin, Y. O.; He, H.; Emmett, M. R.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2007**, *79*, 7596-7602.
- (38) McFarland, M. A.; Chalmers, M. J.; Quinn, J. P.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 1060-1066.

- (39) Tsybin, Y. O.; Quinn, J. P.; Tsybin, O. Y.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 762-771.
- (40) Stephenson, J. L.; McLuckey, S. A. *Anal. Chem.* **1998**, *70*, 3533-3544.
- (41) Hogan, J. M.; McLuckey, S. A. *J. Mass Spectrom.* **2003**, *38*, 245-256.
- (42) He, M.; Reid, G. E.; Shang, H.; Lee, G. U.; McLuckey, S. A. *Anal. Chem.* **2002**, *74*, 4653-4661.
- (43) Yu, L.; Tan, Y.; Tsai, Y.; Goodlett, D. R.; Polfer, N. C. *J. Proteome. Res.* **2011**, *10*, 2409-2416.
- (44) Senko, M. W.; Beu, S. C.; McLafferty, F. W. *J. Am. Soc. Mass. Spectrom.* **1995**, *6*, 229-233.

## Chapter 7

### Conclusions and Future Outlook

Mass spectrometry has been widely studied and developed for various scientific areas. Figure 7.1 illustrates the distribution of the number of journal articles related to mass spectrometry in ten different research areas. Biological sciences are one of the major areas in which mass spectrometry is involved.



**Figure 7.1.** Distribution of journal articles related to mass spectrometry in ten different scientific areas. The numbers were retrieved from Web of Science (Thomson Reuters, New York, NY), a well established database of scientific journal articles on 03/30/2011.

As an analytical chemistry technique, mass spectrometry is one powerful tool to characterize peptides/proteins and nucleic acids. For these molecules, information

regarding identification, quantification, structure, and interaction can possibly be obtained.

This dissertation demonstrates my efforts in mass spectrometry method development for peptide/protein and nucleic acid analysis. Projects are classified into three categories: (1) Applying existing emerging MS techniques to molecules not previously studied by these techniques. (2) Improvement of MS techniques. (3) Applying MS techniques to solve biological problems.

## **7.1 Summary of Results**

First we compared non-acetylated PNA fragmentation in both positive and negative ion mode by CAD, IRMPD, ECD (positive ion mode only) and EDD (negative ion mode only). Also, we employed double resonance to investigate the mechanism of water loss in CAD and IRMPD of PNAs. For investigated MS/MS techniques in positive ion mode, full sequence coverage was obtained from CAD and IRMPD. However, abundant water loss, neutral loss, and unassignable peaks were also observed. Either the N- and C-terminus of PNA could be involved in the water loss pathway, as observed in double resonance experiments. Acetylation of the N-terminus reduced facile water loss and the spectrum was easier to interpret. ECD provided less backbone cleavages than CAD and IRMPD and also abundant neutral loss peaks. For investigated MS/MS techniques in negative ion mode, full sequence coverage was also obtained in CAD and IRMPD but neutral loss was also observed. However, compared with positive ion mode, neutral loss was of lower abundance. Thus CAD spectra in negative ion mode were easier to interpret. EDD provided limited backbone cleavages. Overall, operating CAD

and IRMPD in negative ion mode, as well as acetylation of the PNA N-terminus, reduced the abundance of neutral loss peaks and simplified spectrum interpretation.

Second, we investigated the characterization of the 27-mer RNA R1 and its complexes by top-down mass spectrometry, as well as ion mobility mass spectrometry (IM-MS). Gas-phase structure and folding status of RNA R1 could be probed by IM-MS. More charges on the molecule seemed to correlate with a less compact gas phase structure, as observed from CCS measurements in IM-MS. Competitive experiments revealed that ligand binding specificity was dependent on charge state and further confirmed the correlation between RNA R1 folding status and ligand binding. Sequence coverage from AI-EDD experiments increased as the charge on R1 increased. From AI-EDD experiments a ligand-binding site from non-specific binding was identified from certain charge states. In general, given the correlation between AI-EDD efficiency, binding specificity, precursor charge state, ligand amount in electrospray solution, and the organic solvent content, care should be taken when ESI-MS and AI-EDD are applied to the investigation of RNA-ligand complexes to avoid non-specific binding.

We also extended top-down mass spectrometry of proteins to negative ion mode. IRMPD and AI-EDD were applied to three representative proteins. Results were compared with IRMPD and AI-ECD in positive ion mode. In general, negative ion IRMPD and AI-EDD provided complementary sequence information compared to positive ion IRMPD and AI-ECD. However, less overall sequence coverage was seen in negative ion mode. More complementary fragments were observed in negative ion mode for the acidic proteins  $\beta$ -casein and calmodulin than for the basic protein RNase B. The reason for this behavior could be enhanced “salt-bridge” formation in positively charged

acidic proteins. In particular, negative ion IRMPD of  $\beta$ -casein, with multiple phosphorylated sites close to each other, yielded backbone cleavages between the first four phosphorylated amino acids from the N-terminus. Such cleavages, which are essential to precisely locate phosphorylation sites, were not observed in positive ion IRMPD and only partially in AI-ECD.

A new MS/MS technique, negative ion electron capture dissociation (niECD), was developed in our group. We demonstrated electron capture by singly deprotonated RNA/DNA and multiply charged RNA. Activation before or after the electron capture event, denoted as AI-niECD and niECD-IRMPD, respectively, provided further fragmentation. Fragments from niECD-IRMPD of the RNA/DNA hexamers A6, G6, C6, and U6/T6 were compared with EDD of the same but doubly deprotonated RNA/DNA, and with ECD of the same but doubly protonated RNA/DNA. In general, niECD-IRMPD provided similar fragmentation patterns and fragmentation efficiency as EDD. However, characteristic “ECD-like” radical fragments suggested similarities to ECD as well, supporting the proposed zwitterion mechanism for niECD.

niECD was originally demonstrated for peptides, which tend to carry more charges as their sizes increase or their concentrations decrease. However, niECD is generally more efficient for peptide ions with fewer charges. Precursor ion accumulation time can be a significant contributor to the total acquisition time in an niECD experiment. We demonstrated that PTR prior to niECD can potentially decrease the ion accumulation time, and improve niECD sensitivity. Among the peptides investigated, niECD performance for most peptides heavier than 1500-1700 Da was improved by PTR. For several peptides electrosprayed at 0.1  $\mu$ M, niECD of the singly charged form did not

provide any detectable fragments, while PTR-niECD of the doubly charged form provided 27-61% sequence coverage.

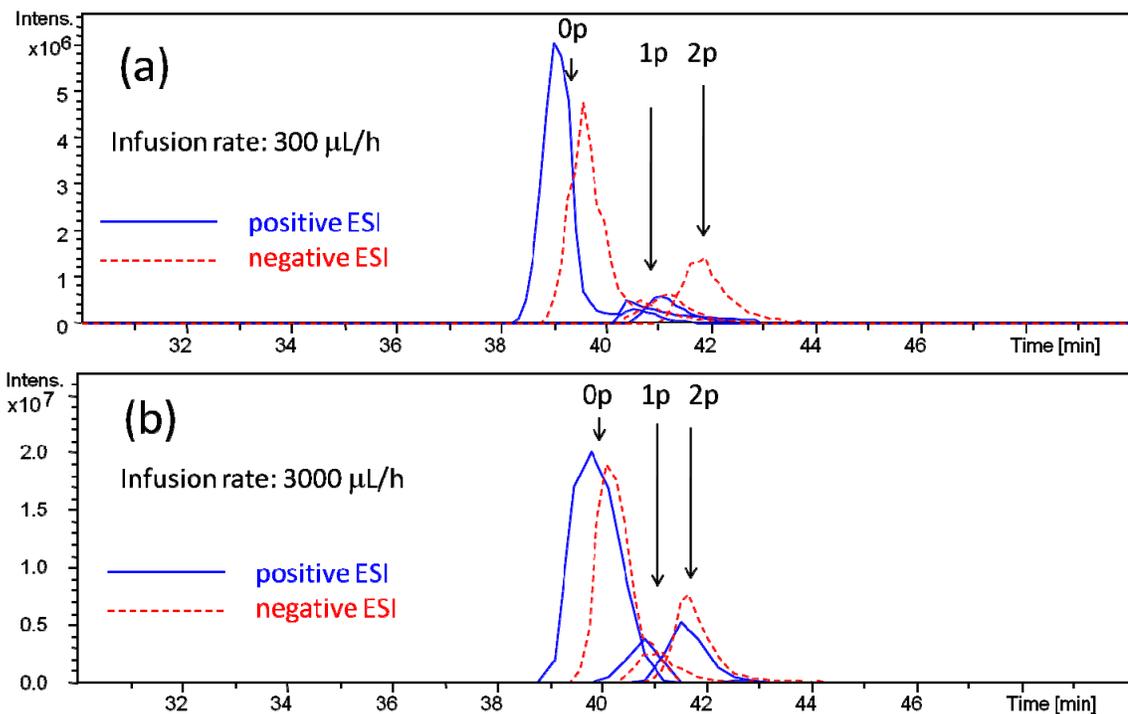
## **7.2 Prospects for Future Work**

### **7.2.1 Comparison of Ionization Efficiency for Detecting Acidic Peptides in Positive and Negative Ion Mode**

In proteomic studies LC-MS is a widely applied technique to analyze proteins or proteomes after digestion by proteases such as trypsin. LC separation prior to mass spectrometric analysis provides numerous advantages including, but not limited to, (1) separation of nonvolatile contaminants, which interfere with ESI, from the peptides to be analyzed; (2) enrichment of each peptide to improve detection sensitivity; (3) reduction of sample complexity and ion suppression. Reverse-phase chromatography is a commonly employed technique to separate peptides based on hydrophobicity<sup>1</sup> in which acids are usually added to the mobile phase to improve the separation,<sup>2</sup> ESI-MS coupled with LC is commonly operated in positive ion mode as the LC eluent is acidified.<sup>3</sup> Tryptic peptides contain at least two basic groups, one at the N-terminus and the other at the C-terminal lysine or arginine side chain. Thus, theoretically, most tryptic peptides in acidified solution should ionize well in positive ion electrospray ionization. Peptides can also be ionized in negative ion mode, which should be particularly favorable for peptides containing carboxylic acid groups at their C-termini and acidic side chains such as aspartic and glutamic acid. Also, PTMs such as phosphorylation and sulfonation can dramatically decrease the pI of a peptide.<sup>4</sup> Acidic peptides show improved ionization efficiency in negative ion mode. For example, it has been demonstrated that phosphorylated peptides from a tryptic digest of a phosphorylated protein were detected with higher sensitivity in negative ion mode.<sup>5</sup> However, systematic comparison of

ionization efficiency for acidic and basic peptides in LC-MS is less well addressed. Such a comparison may provide guidance to the analysis of acidic peptides, particularly in phospho-proteomics where phosphorylated peptides are the analytes of major interest.

In phospho-proteomics, phosphorylated peptides are generally of low abundance compared to non-phosphorylated peptides.<sup>6</sup> During ESI in positive ion mode, regular peptides compete with phosphopeptides for available protons in the solution. The ionization of acidic phosphorylated peptides may thus be suppressed by other more basic peptides,<sup>6</sup> which have higher proton affinity.<sup>7</sup> Figure 7.2a shows extracted ion chromatograms for the peptide H-TPQKPSQNLVPVTPSTTK-OH in unphosphorylated (0p), singly (1p), and doubly (2p) phosphorylated forms. These peptides were from a protein digest, separated on a 1 mm (i.d.) C18 RPLC column with an acidified mobile phase, electrosprayed at 300  $\mu\text{L}/\text{h}$  and detected in positive and negative ion mode, respectively. MS signal for the non-phosphorylated form was higher in positive ion mode whereas MS signal for the singly phosphorylated form was similar in positive and negative ion mode. MS signal for the doubly phosphorylated form was higher in negative ion mode.

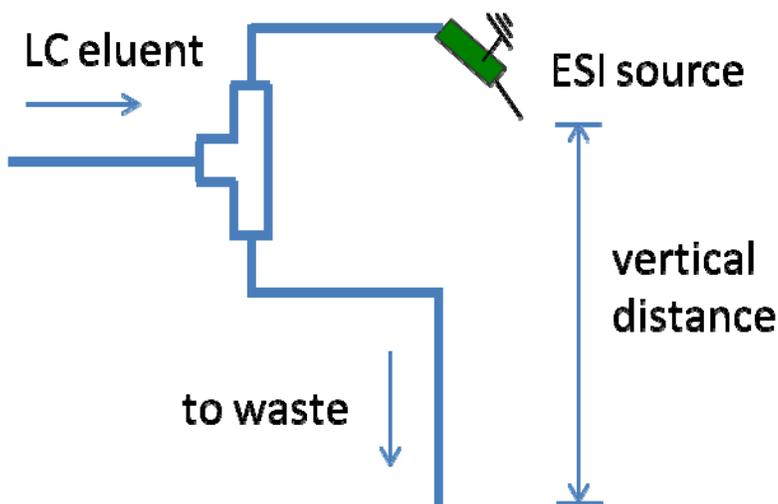


**Figure 7.2.** Comparison of extracted ion chromatograms for the peptide H-TPQKPSQNLVPVTPSTTK-OH in unphosphorylated (0p), singly (1p), and doubly (2p) phosphorylated forms in positive and negative ion mode. ESI flow rates were (a) 300  $\mu\text{L/h}$  and (b) 3000  $\mu\text{L/h}$ .

Figure 7.2 shows extracted ion chromatograms of the same peptides under the same chromatographic conditions, except for the flow rate, which differs by a factor of 10. In this case MS signal of doubly phosphorylated peptides were still smaller in positive ion mode. However, the difference between positive and negative ion mode was smaller, compared to a flow rate of 300  $\mu\text{L/h}$ . In our experiment, the Apollo II ESI source was optimized for a flow rate around 0.1 mL/h. At such flow rate phosphorylated peptides from a tryptic digest of a phosphorylated protein were detected with higher sensitivity in negative ion mode.<sup>5</sup> However the flow rate during online LC-MS experiments was 3 mL/h. At such relatively high flow rate, the total ion current may saturate the ESI source in both positive and negative ion mode. Thus the total numbers of ionized peptides in

both modes may be capped by the capability of the ESI source and a difference can therefore be more difficult to observe.

The flow rate manipulation was achieved by splitting the eluent before the ESI spray tip. Under typical RPLC conditions, the optimum flow rate for a 1 mm (i.d.) RPLC column is around 3000  $\mu\text{L}/\text{h}$ . The Apollo II ESI source on our mass spectrometer was optimized for direct infusion of samples, which typically operates at the 100  $\mu\text{L}/\text{h}$  range. LC flow splitting is commonly achieved by T-shaped splitters. The split ratio is calibrated by the i.d. and length of two pieces of outlet tubing. However, in practice, we found that, in the flow rate range we operate our LC, gravity also plays an important role in controlling the flow rate.



**Figure 7.3.** Systematic diagram of the principle for fine-tuning LC-MS flow rate. Devices are not accurate to scale.

Figure 7.3 shows a systematic diagram of a method for fine-tuning LC-MS flow rate. LC eluent is split by a T-junction from which one portion is transferred to the ESI source. The other portion is transferred to waste or fraction collection. All pieces of tubing are of the same size. By adjusting the vertical distance between the electrospray

tip and the outlet of the waste line the flow rate into the ESI source can be controlled. 300  $\mu\text{L}/\text{h}$  as shown in Figure 7.2a was achieved when the vertical distance was 50 cm.

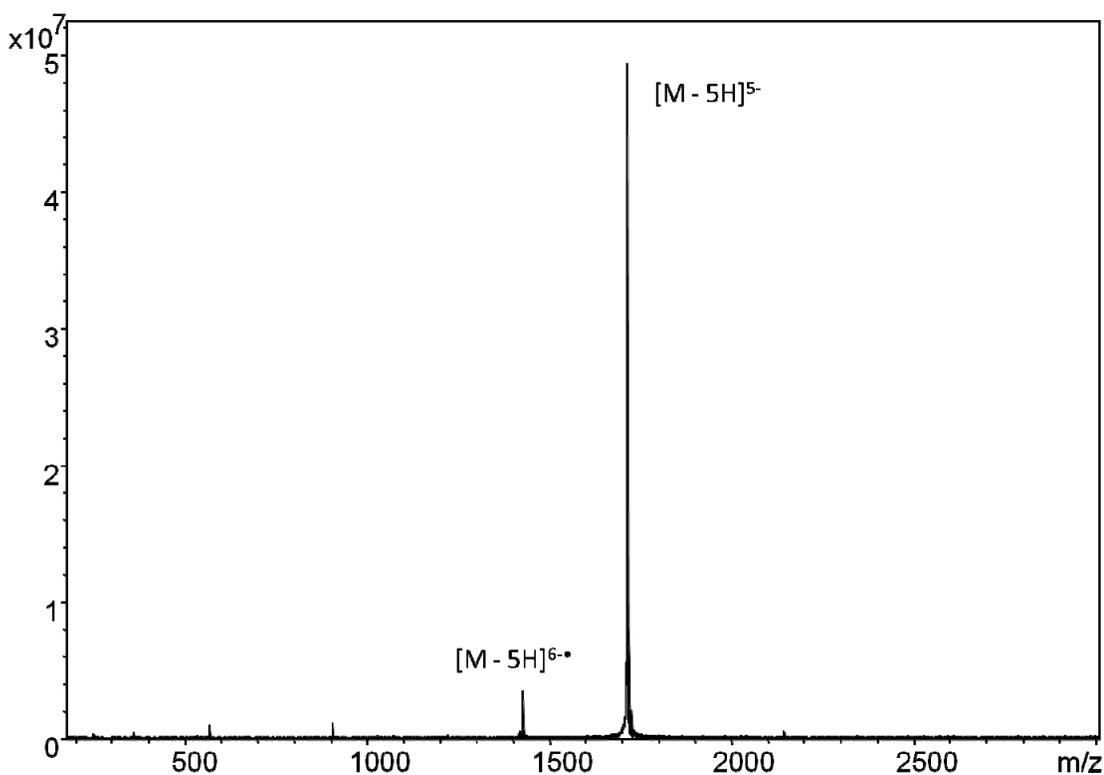
With this splitting approach we could investigate the correlation between ion suppression and flow rate. Then we could systematically compare the ionization efficiency of acidic peptides in positive and negative ion mode. Also we could obtain improved understanding of which flow rate range ion suppression for acidic peptides is more significant.

### **7.2.2 niECD of Proteins and Larger RNAs**

niECD was first demonstrated for singly and doubly charged peptides.<sup>8</sup> One significance of niECD is that it provides opportunities to dissociate peptide anions with acidic labile PTMs, such as phosphorylation and sulfonation, into backbone fragments with PTMs retained. The modification site can be directly located and spectra may be simpler to interpret than commonly applied vibrational activation techniques such as CAD and IRMPD, in which labile PTMs are frequently lost. ECD<sup>9</sup> and ETD<sup>10</sup> have been shown to preserve labile PTMs such as phosphorylation,<sup>11,12</sup> sulfonation,<sup>13</sup> and glycosylation.<sup>14-16</sup> However, ECD and ETD require multiply positively charged ions and only operate in positive ion mode. Due to ion suppression, detection of acidic peptides may be less sensitive in positive ion mode than negative ion mode. Thus niECD combines certain advantages of ECD, ETD and negative ion detection and may be beneficial for PTM detection and localization.

ECD<sup>9</sup> and ETD<sup>17</sup> have been shown to have the ability to effectively dissociate intact protein cations in the gas phase. EDD<sup>18</sup> has been shown to fragment protein anions.

However the sequence coverage in EDD was not as extensive as in ECD or ETD. Until now niECD has been primarily applied to peptides up to 2.3 kDa with backbone cleavages observed. It would be beneficial to investigate niECD for protein anions. If niECD could be developed to dissociate intact proteins it would be possible to directly localize labile PTMs, similar to ECD and ETD, while detecting acidic proteins in negative ion mode at higher sensitivity.



**Figure 7.4.** AI-niECD of ubiquitin with 5 negative charges. Precursor ions were irradiated with an IR laser at 12.5 W for 0.3 s, then with electrons for 5 s. The cathode voltage was set to - 10 V.

Figure 7.4 shows preliminary results for AI-niECD of ubiquitin (8.5 kDa) 5- anions. Major product ions are the radicals formed from electron capture. No backbone cleavages are observed. However, only one charge state of ubiquitin was investigated. It has been shown that fragmentation efficiency in ECD,<sup>19</sup> ETD<sup>20</sup> and EDD<sup>18</sup> is affected by

precursor ion charge state. Also, a denaturing step prior to ESI may unfold the protein and facilitate fragmentation.<sup>18</sup> Investigation of more charge states and sample preparation procedures may be valuable future work.

In Chapter 5 it was demonstrated that AI-niECD of DNA/RNA hexamers provided “ECD-like” and “EDD-like” fragments. An 8.4 kDa 27-mer RNA (5'-GGCGUCACACCUUCGGGUGAAGUCGCC-3') was demonstrated to capture electrons as well. It may be valuable to investigate AI-niECD of larger RNAs such as this 27-mer. Similar to AI-niECD of proteins, investigation of more charge states and sample preparation procedures may be helpful in the future. In Chapter 3 it was demonstrated that higher ligand binding specificity was favored at lower precursor ion charge states, which did not undergo efficient EDD. niECD, on the other hand, favors lower charge states, thus the latter technique may be suitable for specifically bound RNA-ligand complexes.

### 7.3 Bibliography

- (1) Chen, H.; Horvath, C. *J. Chromatogr. A* **1995**, *705*, 3-20.
- (2) Mant, C. T.; Hodges, R. S. In *High-performance liquid chromatography of peptides and proteins: Separation, analysis, and conformation* Mant, C. T., Hodges, R. S., Eds.; CRC Press: Boca Raton, Florida, 1991, pp 327-341.
- (3) Shi, Y.; Xiang, R.; Horvath, C.; Wilkins, J. A. *J. Chromatogr. A* **2004**, *1053*, 27-36.
- (4) Gauci, S.; van Breukelen, B.; Lemeer, S. M.; Krijgsveld, J.; Heck, A. J. R. *Proteomics* **2008**, *8*, 4898-4906.
- (5) Kweon, H. K.; Hakansson, K. *Anal. Chem.* **2006**, *78*, 1743-1749.
- (6) Mann, M.; Ong, S. E.; Gronborg, M.; Steen, H.; Jensen, O. N.; Pandey, A. *Trends Biotechnol.* **2002**, *20*, 261-268.

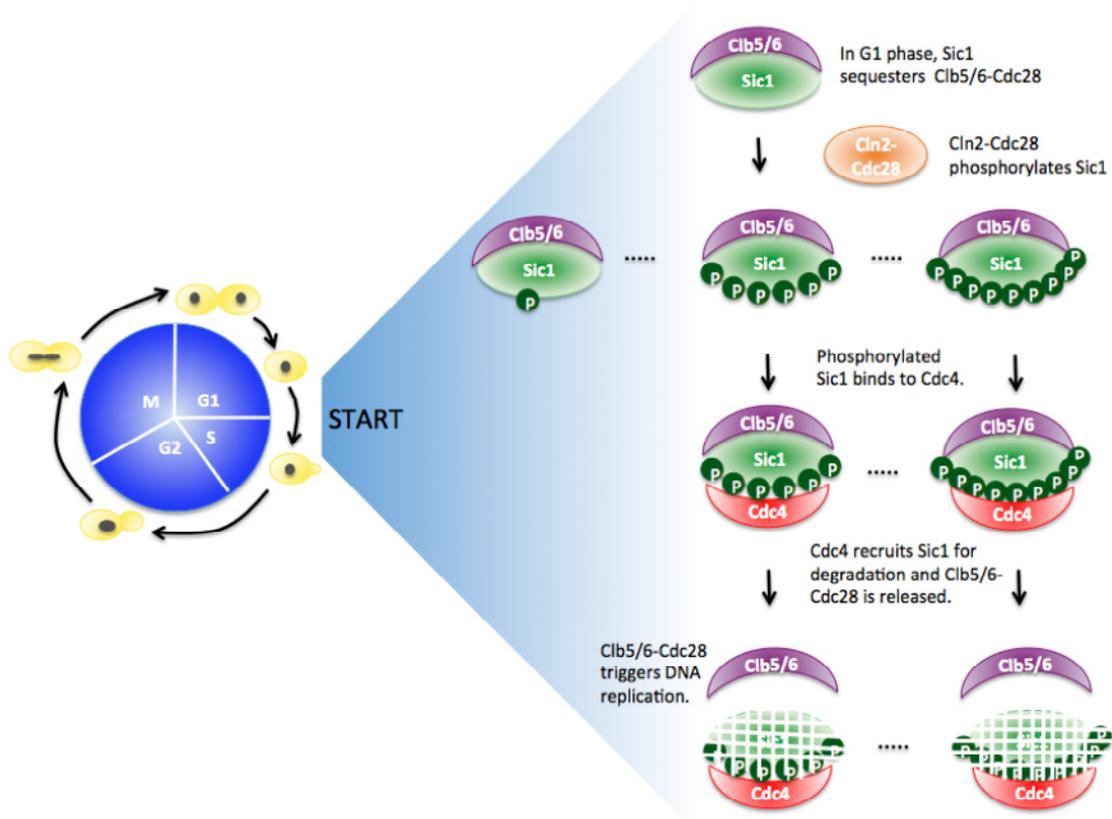
- (7) Garcia, M. C. *J. Chromatogr. B* **2005**, 825, 111-123.
- (8) Yoo, H. J.; Wang, N.; Zhuang, S.; Song, H.; Hakansson, K. *J. Am. Chem. Soc.* **2011**, 133, 16790-16793.
- (9) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, 120, 3265-3266.
- (10) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, 9528-9533.
- (11) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, 14, 1793-1800.
- (12) Shi, S. D. H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, 73, 19-22.
- (13) Kelleher, R. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, 71, 4250-4253.
- (14) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, 71, 4431-4436.
- (15) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, 73, 4530-4536.
- (16) Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnsen, A. H.; Zubarev, R. A. *Anal. Chem.* **2001**, 73, 2998-3005.
- (17) Coon, J. J.; Ueberheide, B.; Syka, J. E. P.; Dryhurst, D. D.; Ausio, J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 9463-9468.
- (18) Ganisl, B.; Valovka, T.; Hartl, M.; Taucher, M.; Bister, K.; Breuker, K. *Chem. Eur. J.* **2011**, 17, 4460-4469.
- (19) Kalli, A.; Hakansson, K. *J. Proteome. Res.* **2008**, 7, 2834-2844.
- (20) Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. *Mol. Cell. Proteomics* **2007**, 6, 1942-1951.

## Appendix

### Quantitative Investigation of the Degree of Phosphorylation on the Yeast Protein Sic1

#### A.1 Introduction

Sic1 is one of the important regulatory proteins controlling the yeast cell cycle.<sup>1</sup> The level of phosphorylation on this protein serves as a “switch” for the yeast cell to enter from G1 phase to S phase. As yeast cells enter G1 phase, a B-type cyclin, Cln-Cdc28 kinases, is activated and phosphorylate Sic1. Multiply phosphorylated Sic1 is targeted by the SCF<sup>Cdc4</sup> ubiquitin ligase complex for ubiquitination and subsequent degradation.<sup>1</sup> Upon degradation of Sic1, Clb5/6-Cdc28 is no longer inhibited,<sup>2-5</sup> thus stimulating DNA replication. This regulatory mechanism is illustrated in Figure A.1.



**Figure A.1.** Scheme of Sic1 regulatory activity in the cell cycle of *S. cerevisiae*. Adapted from Varedi *et al.*<sup>6</sup>

Establishing a model to simulate the phosphorylation/dephosphorylation process of Sic1 may facilitate understanding of the regulatory activity of Sic1. Varedi, *et al.* developed mathematical models to simulate the steady-state binding curves of Sic1-Cdc4 as well as Sic1 phosphorylation/dephosphorylation/binding.<sup>6</sup> Experimental determination of several parameters in the phosphorylation kinetics is crucial to evaluate these models. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) coupled with high performance liquid chromatography (HPLC) can provide high mass accuracy and high throughput for quantification of phosphopeptides in kinetic studies. With mass spectrometry it is also possible to determine the overall phosphorylation state at the intact protein level.

Absolute quantification of peptides by mass spectrometry has been widely explored. Quantification based on extracted ion chromatograms (EICs) is one of the most common approaches. In HPLC-ESI-MS experiments, the MS signal abundance from targeted peptides is summed over the elution time. However, due to variance in ionization efficiency and ion suppression, ion abundance alone can be inaccurate and irreproducible for quantifying peptides across different samples or for comparing different peptides. Abundant non-changing peptides in the LC-MS runs could serve as “landmark” or “background”.<sup>7</sup> After normalization to the abundances of the landmark peptides, run-to-run variation can be reduced. One example is the investigation of phosphorylation response. Under the assumption that the total amount of the phosphorylated peptide of interest and its non-phosphorylated form is constant during the phosphorylation time course, ionization efficiency can be compared between the two forms and absolute quantification of stoichiometry can be achieved.<sup>8</sup> However, to achieve absolute quantification, spiking of reference materials is required. The chemical identity of the reference material should be similar to the peptides of interest, to reduce ion suppression and ionization efficiency variance. Isotopic labeling is one common technique to introduce appropriate standards.<sup>9</sup> Isotopically labeled peptides with the same amino acid sequence as peptides of interest are chemically virtually identical but have different mass. However, isotopic labeling requires isotope-incorporated reagents or peptides. Other reference materials such as valine-leucine substituted standard peptides are more affordable and feasible.<sup>10</sup> Valine-leucine substituted standard peptides provide a 14 Da mass shift, which may reduce the probability of overlapping isotopic distributions.

Here we demonstrate absolute quantification of mutant Sic1 phosphorylation with valine-leucine substituted standard peptides. Also, we monitored steady-state phosphorylation response of wild type Sic1 and mutated Sic1 with various total available sites for phosphorylation with top-down mass spectrometry.

## **A.2. Experimental Section**

Sic1 protein preparation and phosphorylation reactions were described elsewhere.<sup>6</sup> In short, wild-type and mutated His-tagged *S. cerevisiae* Sic1 proteins were expressed in *E. coli* and purified with nickel-NTA columns (Qiagen). Cln2-Cdc28 kinase was expressed in Hi5 cells at the Protein Expression Center of Cal Tech.<sup>11</sup> The Cln2-Cdc28 kinase was purified using glutathione-sepharose beads (GE Healthcare). All phosphorylation reactions were performed in kinase buffer (Cell Signaling) with ATP (Cell Signaling) at 30 °C for various times for dynamic studies or three hours to reach steady state.

In the study of phosphorylation dynamics, phosphorylated Sic1 proteins were digested by trypsin (Promega). Valine-leucine substituted standard peptides (GenScript) were phosphorylated by Cln2-Cdc28. Phosphorylated and non-phosphorylated forms were separated by an Agilent Zorbax XBD C8 reverse phase column and quantified by UV absorbance at 214 nm. Mixtures of phosphorylated and non-phosphorylated forms were spiked into Sic1 tryptic digests as internal standards. Sic1 tryptic digest was subjected to LC-MS analysis. HPLC was performed by an Agilent 1100 series system (Agilent Technologies) on a Phenomenex Synergi Hydro reverse phase C18 column (1.0 x 15 mm, Phenomenex) at a flow rate of 0.05 mL/min. The gradient was solvent B

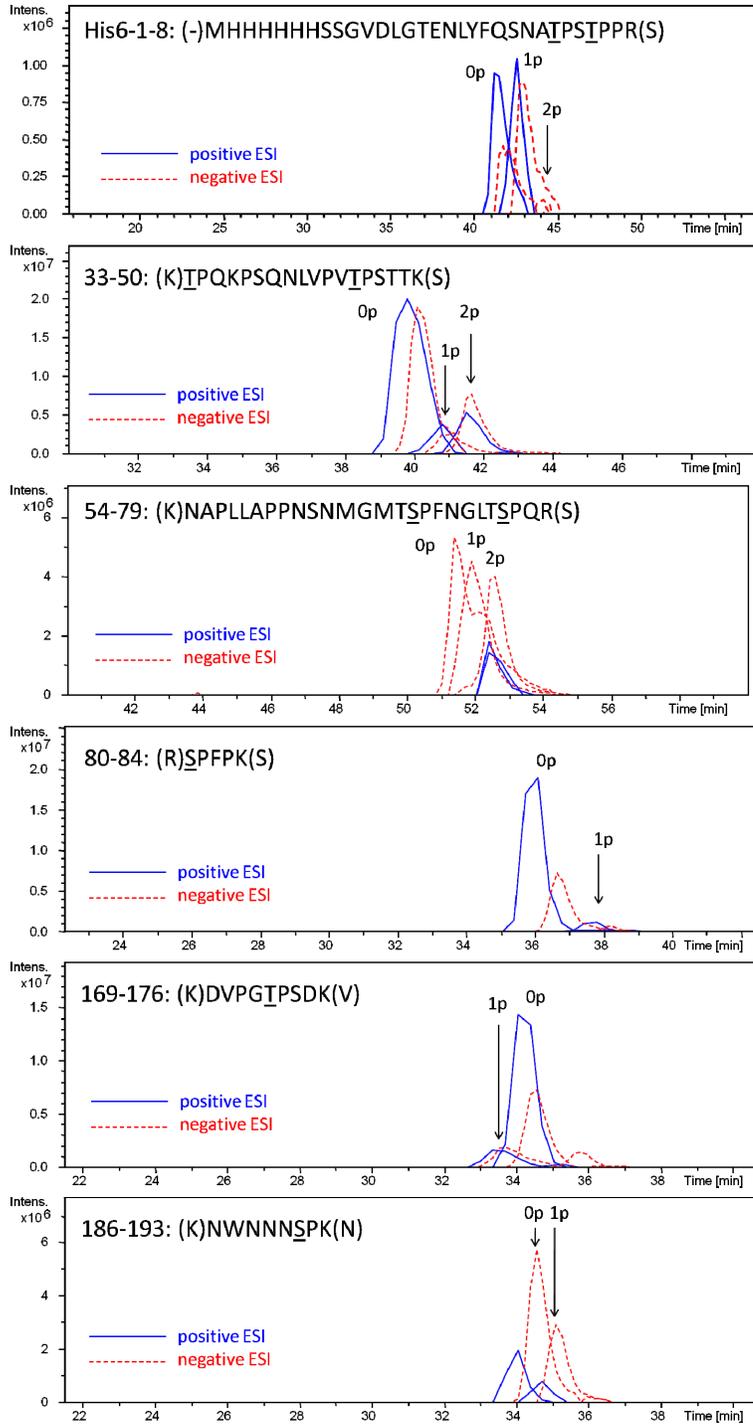
(acetonitrile with 0.1% formic acid) ramping from 2% to 60% over 60 min. Solvent A was water with 0.1% formic acid. The HPLC eluent was directly introduced online to a Bruker 7-T Apex-Q ESI-FT-ICR mass spectrometer (Bruker Daltonics) operating in positive or negative ion mode. LC-MS data were acquired by Hystar and ApexControl software (Bruker Daltonics). Extracted ion chromatograms (EICs) were calculated based on the accurate masses of peptides of interest and a 20 ppm error window. This relatively high error tolerance was used due to the variability in peptide abundance and thus number of ions in the ICR cell. In separate LC runs, fractions were also collected every 1 min and infused into a Bruker 7-T SolariX ESI-Q-FT-ICR mass spectrometer operating in positive ion mode for electron capture dissociation (ECD). ECD experiments were performed in the ICR cell by irradiating precursor ions with electrons for 0.9 s. The cathode bias voltage was - 0.9 V and the lens voltage was set to 0.5-1 V.

In the study of steady-state phosphorylation response, phosphorylated proteins were purified with an Agilent Zorbax XBD C8 reverse phase column. The fractions containing proteins were concentrated and reconstituted in 1:1 (v/v) water/acetonitrile with 0.1% formic acid for positive ion mode ESI. Samples were electrosprayed via an Apollo II (Bruker Daltonics) electrospray source. Mass spectra were acquired on a Bruker 7-T SolariX ESI-Q-FT-ICR mass spectrometer in broad band mode. Spectra were processed by DataAnalysis software (version 4.0, Bruker Daltonics).

### **A.3. Results and Discussion**

#### **A.3.1. Comparison of Ionization Efficiency of Sic1 Peptides in Positive and Negative Ion Mode**

In proteomics RPLC-MS is almost exclusively operated in positive ion mode. First, acidic ion pairing reagents are commonly added to the mobile phase to improve chromatographic resolution.<sup>12, 13</sup> Widely used acids include trifluoroacetic acid, acetic acid, and formic acid. Thus peptides are likely more protonated than deprotonated and show higher detection sensitivity in positive ion mode. Second, trypsin is one of the most common proteases due to its high efficiency and specificity. Tryptic peptides contain at least two basic groups: one at the N-terminus and the other one at the C-terminal lysine or arginine side chains. However, the pKa of phosphorylated amino acid side chains can be lower than 1.5<sup>14</sup> and thus phosphate groups can still be deprotonated in acidified mobile phase. For example, for water with 0.1% (v/v) formic acid the pH is 2.7. The acidity of phosphate groups can dramatically decrease the pI of peptides.<sup>15</sup> Thus, the ionization of acidic phosphorylated peptides may be suppressed by other more basic peptides,<sup>16, 17</sup> which have higher overall proton affinity.<sup>18</sup> It has been demonstrated that phosphorylated peptides from a tryptic digest of a phosphorylated protein were detected with higher sensitivity in negative ion mode.<sup>19</sup> We compared the ionization efficiency in positive and negative ion mode to ensure the phosphorylated peptides from Sic1 protein provided more signal intensity in negative ion mode. It has been demonstrated that nine phosphorylation sites on Sic1 are responsible for the regulatory activity.<sup>1</sup> These nine sites are T2, T5, T33, T45, S69, S76, S80, T173, and S191. In a Sic1 tryptic digest, three peptides contain two proposed phosphorylation sites each and three peptides contain one site each.



**Figure A.2.** Comparison of extracted ion chromatograms for selected tryptic peptides from the protein Sic1 in positive and negative ion mode. Sic1 was phosphorylated to maximum extent. Selected peptides contained phosphorylation sites as proposed.<sup>1</sup>

Figure A.2. illustrates the comparison of selected tryptic peptides from Sic1 in positive and negative ion mode. For H-NAPLLAPPNSNMGMTSPFNGLTSPQR-OH (residues 54-79 in Sic1) and HNWNNNSPK-OH (Sic1 186-193), ion abundance for all phosphorylated forms was higher in negative ion mode compared with positive ion mode. For H-MHHHHHSSGVDLGTENLYFQSNATPSTPPR-OH (Sic1 His6-1-8) ion abundance of the doubly phosphorylated form was higher in negative ion mode than in positive ion mode, while ion abundances of non- and singly phosphorylated forms were similar. For H-SPFPK-OH (Sic1 80-84), H-DVPGTPSDK-OH (Sic1 167-176) and H-TPQKPSQNLVPVTPSTTK-OH (Sic1 33-50), ion abundance in negative ion mode was not significantly higher than in positive ion mode. The observation that not all phosphorylated peptides showed higher ion abundance in negative ion mode may be due to the following reasons: First, in our experiment, the Apollo II ESI source was optimized for a flow rate around 0.1 mL/h. At such flow rate phosphorylated peptides from a tryptic digest of a phosphorylated protein were detected with higher sensitivity in negative ion mode.<sup>19</sup> However the flow rate during online LC-MS experiments was 3 mL/h. At such relatively high flow rate the total ion current may saturate the ESI source in both positive and negative ion mode. Thus the total numbers of ionized peptides in both modes may be capped by the capability of the ESI source and a difference can therefore not be observed. Second, ion suppression experienced by one peptide is affected by the abundance of other peptides. H-SPFPK-OH, H-DVPGTPSDK-OH and H-TPQKPSQNLVPVTPSTTK-OH were among the most abundant peptides in their respective fractions. Thus they may experience less ion suppression.

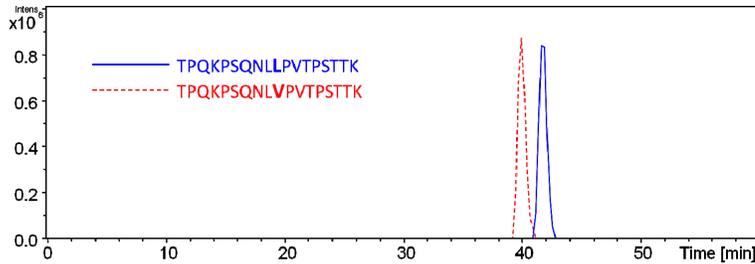
In general, ion abundances of phosphorylated peptides are equal or higher in negative ion mode compared with positive ion mode. The doubly phosphorylated form of H-MHHHHHHSSGVDLGTENLYFQSNATPSTPPR-OH was only detected in negative ion mode. Non-phosphorylated forms were observed in both modes. Thus negative ion mode was selected for subsequent analysis.

### **A.3.2. Validation of Valine-Leucine Substituted Standard Peptides**

Valine and leucine residues differ by one methyl group. Because the ionization efficiency of peptides often depend on their structure, an ideal standard peptide should have the same structure as the analyte. Incorporation of isotopes minimizes the impact on peptide structure. On the other hand, we propose that valine-leucine substitution introduces or removes only one methyl group and may therefore have a small impact as well. However, validation is required to confirm the correlation between valine-leucine substituted standard peptides and peptides to quantify.

To test our hypothesis, we investigated a peptide from wild type Sic1 containing two possible phosphorylation sites, H-TPQKPSQNLVPVTPSTTK-OH and its reference peptide H-TPQKPSQNLLPVTPSTTK-OH (Sic1 33-50-V42L) with one valine substituted by leucine. To mimic the background interference from other Sic1 peptides, a Sic1-2p mutant tryptic digest was used. In this mutant seven out of nine possible phosphorylation sites in Sic1 wild type were mutated to alanine. In Sic1-2p mutant TPQKPSQNLVPVTPSTTK was modified to APQKPSQNLVPVTPSTTK (corresponding to the peptide Sic1 33-50-T33A). 10 pmoles of H-

TPQKPSQNLVPVTPSTTK-OH and H-TPQKPSQNLLPVPSTTK-OH were mixed and spiked into 15 pmoles of Sic1-2p digest and subjected to LC-MS analysis.



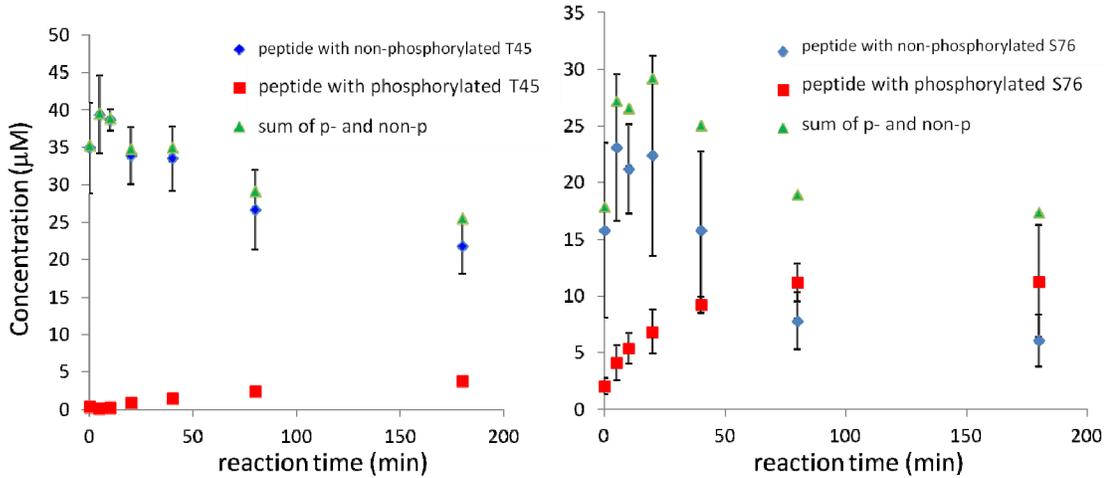
**Figure A.3.** Comparison of EICs for H-TPQKPSQNLVPVTPSTTK-OH and its valine-leucine substituted reference peptide H-TPQKPSQNLLPVPSTTK-OH. 10 pmoles of H-TPQKPSQNLVPVTPSTTK-OH and TPQKPSQNLLPVPSTTK were mixed and spiked into 15 pmoles of Sic1-2p digest. Sic1-2p was phosphorylated to maximum extent.

Figure A.3. shows the comparison of EICs of H-TPQKPSQNLVPVTPSTTK-OH and its valine-leucine substituted reference peptide H-TPQKPSQNLLPVPSTTK-OH. Peak areas of both peptides are similar ( $104\% \pm 10\%$ ). Similar experiments were performed without the Sic1-2p tryptic digest. Peak areas for both peptides were similar ( $101\% \pm 11\%$ ) in those spectra as well. It thus appears that valine-leucine substitution is an acceptable method to generate reference peptides.

### A.3.3. Investigation of Phosphorylation Kinetics of Sic1

In Sic1-2p mutant seven serine or threonine residues were mutated to alanine except for T45 and S76. We used the Sic1-2p mutant to develop an LC-MS work flow. Phosphorylation kinetics could be measured by monitoring the concentrations of H-APQKPSQNLVPVTPSTTK-OH (Sic1 33-50-T33A) and H-NAPLLAPPNSNMGMTAPFENGLTSPQR-OH (Sic1 54-79-S69A) following phosphorylation of Sic1-2p for different lengths of time. The corresponding reference

peptides spiked into the Sic1-2p tryptic digest were H-APQKPSQNLLPVTPTTK-OH (Sic1 33-50-T33A-V42L) and H-NAPVLAPPNSNMGMTAPFNGLTSPQR-OH (Sic1 54-79-S69A-L57V) as well as their singly phosphorylated forms.



**Figure A.4.** Phosphorylation kinetics of T45 (left) and S76 (right) in a Sic1-2p mutant. Concentrations were calculated by normalizing EIC peak areas of 33-50-T33A and 54-79-S69A to 33-50-T33A-V42L and 54-79-S69A-L57V.

Figure A.4 shows phosphorylation kinetics of T45 and S76 in the Sic1-2p mutant. Concentrations were calculated by normalizing EIC peak areas of 33-50-T33A and 54-79-S69A to 33-50-T33A-V42L and 54-79-S69A-L57V, respectively. Both non-phosphorylated and singly phosphorylated forms of 33-50-T33A and 54-79-S69A were quantified. Phosphorylation on S76 was observed to be faster than on T45. It has been demonstrated that the cyclin-B kinase Cdc2 strongly prefers to phosphorylate serine in a SPXR motif.<sup>20</sup> S76 in 54-79-S69A is contained in an SPQR motif and could be preferred for phosphorylation by Cdc28, which also belongs to the cyclin-B kinase family. T45 is not contained in such a motif. On the other hand, it has been shown that T45 contributes more than S76 to Sic1 *in vivo* instability, inferring that T45 contributes more for Cdc4

recognition.<sup>1</sup> If T45 is both preferred for phosphorylation and recognition for ubiquitination, its regulatory activity would primarily be controlled by T45 rather than by multiple phosphorylated sites, which were demonstrated to participate in the regulatory activity.

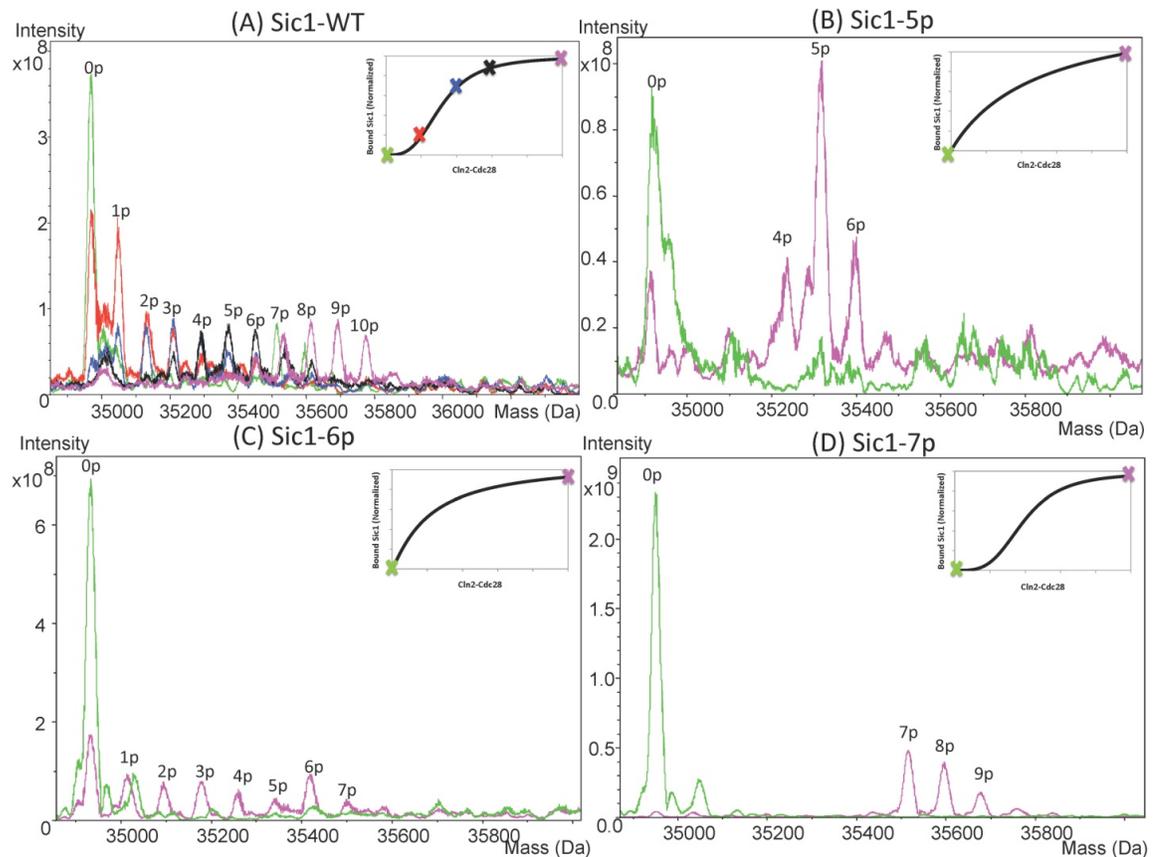
Another observation was that the total peptide concentration (0p + 1p form) decreased over the reaction time. There are four other serine and threonine residues in the APQKPSQNLVPVTPSTTK peptide and three in the NAPLLAPPNSNMGMTAPFNGLTSPQR peptide. Even though T45 and S76 are preferred to be phosphorylated by Cln-Cdc28, it is possible that other sites could be phosphorylated as well. Phosphorylation at more than one site per peptide would cause the total peptide concentration to decrease over the phosphorylation time as we only monitored the 0p and 1p forms. Further experiments are required for confirmation of this hypothesis.

In a wild type Sic1 tryptic digest, three peptides contain two proposed phosphorylation sites. Following Sic1 phosphorylation to maximum extent, the protein was digested and separated by RPLC. We collected singly phosphorylated H-SNATPSTPPR-OH (Sic1 1-8 after removal of the His6 tag by TEV protease) and H-TPQKPSQNLVPVTPSTTK-OH (Sic1 30-50) from the LC separation and performed electron capture dissociation (ECD) to localize the phosphorylation sites. For H-SNATPSTPPR-OH, the phosphorylation site was localized to T5 T2 was not phosphorylated. For H-TPQKPSQNLVPVTPSTTK-OH, phosphorylation on both T33 and T45 was observed:  $c_3'$ ,  $c_5'$ ,  $c_6'$ ,  $c_7'$ ,  $c_8'$ ,  $c_9'$ ,  $c_{11}'$ ,  $c_{12}'$ ,  $z_6^\bullet$ ,  $z_{10}^\bullet$ ,  $z_{11}^\bullet$ , and  $z_{12}^\bullet$  ions were observed in forms with and without phosphate. From this product ion list, the sum of  $c'$

ions without phosphate and  $z^+$  ions with phosphate should roughly represent the abundance of T45 phosphorylation whereas the sum of  $c^+$  ions with phosphate and  $z^+$  ions without phosphate should roughly represent the abundance of T33 phosphorylation. The ratio between the two summed abundances is 0.96, which indicates that the phosphorylation level on T45 and T33 is similar.

#### **A.3.4. Investigation of Distribution of Phosphorylation States on Sic1 wild type and Mutants**

By top-down mass spectrometry we monitored steady-state phosphorylation response of wild type Sic1 and mutated Sic1 with various total available sites for phosphorylation. Wild type and mutated Sic1 proteins were incubated with various amounts of Cln-Cdc28 for three hours.



**Figure A.5.** Distribution of Sic1 phosphoforms obtained from top down mass spectrometry. Mass spectra show the relative distribution of phosphoforms at various levels of Cln-Cdc28 in wild-type Sic1 (A), Sic1-5p (B), Sic1-6p(C), and Sic1-7p (D). Insets show the binding affinity of Sic1 protein to Cdc4 as a function of Cln-Cdc28 amount used to phosphorylate Sic1.<sup>6</sup>

Figure A.5 shows the distribution of Sic1 phosphoforms after phosphorylation by various amounts of Cln-Cdc28. As the amount of Cln-Cdc28 increased, Sic1 proteins with more phosphates were observed. The trend was consistent with the trend of Sic1 binding affinity to Cdc4, which was quantified by immunoblotting with a C-terminal Sic1 antibody. This trend indicates that the degree of Sic 1 phosphorylation is correlated with binding affinity to Cdc4 and subsequent degradation. When Sic1 was phosphorylated on less than four sites (Figure A.5(A)), binding affinity of Sic1 to Cdc4 was low, which

indicates less subsequent degradation and initiation of DNA replication. When Sic1 was phosphorylated on four to six sites, binding affinity of Sic1 to Cdc4 increased dramatically and when Sic1 was phosphorylated to the maximum extent the binding affinity was at a maximum. This correlation may indicate that the appearance of Sic1 phosphorylated on six sites is crucial in the mechanism of Sic1 regulatory function. It should be noted that Sic1 with ten phosphates was observed, which could be due to phosphorylation of sites other than the proposed SP or TP motifs. The latter observation may correlate with the observation that the sum of non-phosphorylated and singly phosphorylated peptide decreased along the reaction course, as mentioned in the previous section.

#### **A.4. Conclusions**

Here we demonstrated method development for quantitative investigation of the degree of phosphorylation on Sic1. LC-MS with valine-leucine substituted peptides was developed to quantitatively study the phosphorylation kinetics of a Sic1 2p mutant. ECD was applied to confirm three phosphorylation sites and semi-quantitatively compare the phosphorylation preference between two pairs of possible phosphorylation sites. Intact Sic1 proteins were also investigated without protease digestion. Phosphoform distributions combined with the binding affinity of Sic1 to Cdc4 revealed that the number of phosphates correlates with the mechanism of Sic1 regulatory activity. The developed methods will be applied in further investigations of Sic1 functionality.

## A.5. Bibliography

- (1) Nash, P.; Tang, X. J.; Orlicky, S.; Chen, Q. H.; Gertler, F. B.; Mendenhall, M. D.; Sicheri, F.; Pawson, T.; Tyers, M. *Nature* **2001**, *414*, 514-521.
- (2) Schwob, E.; Bohm, T.; Mendenhall, M. D.; Nasmyth, K. *Cell* **1994**, *79*, 233-244.
- (3) Verma, R.; Annan, R. S.; Huddleston, M. J.; Carr, S. A.; Reynard, G.; Deshaies, R. J. *Science* **1997**, *278*, 455-460.
- (4) Weinreich, M.; Liang, C.; Chen, H.-H.; Stillman, B. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 11211-11217.
- (5) Cross, F. R.; Schroeder, L.; Bean, J. M. *Genetics* **2007**, *176*, 1541-1555.
- (6) Varedi, M.; Song, H.; Hale, W. A.; Hakansson, K.; Lin, X. N. *Mol. Cell* **submitted**.
- (7) Ong, S. E.; Mann, M. *Nat. Chem. Biol.* **2005**, *1*, 252-262.
- (8) Steen, H.; Jebanathirajah, J. A.; Springer, M.; Kirschner, M. W. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 3948-3953.
- (9) de Leenheer, A. P.; Thienpont, L. M. *Mass Spectrom. Rev.* **1992**, *11*, 249-307.
- (10) Callahan, J. H.; Shefcheck, K. J.; Williams, T. L.; Musser, S. M. *Anal. Chem.* **2006**, *78*, 1789-1800.
- (11) Petroski, M. D.; Deshaies, R. J.; Raymond, J. D. In *Methods Enzymol.*; Academic Press, 2005; Vol. Volume 398, pp 143-158.
- (12) Mant, C. T.; Hodges, R. S. In *High-performance liquid chromatography of peptides and proteins: Separation, analysis, and conformation* Mant, C. T., Hodges, R. S., Eds.; CRC Press: Boca Raton, Florida, 1991, pp 327-341.
- (13) Shi, Y.; Xiang, R.; Horvath, C.; Wilkins, J. A. *J. Chromatogr. A* **2004**, *1053*, 27-36.
- (14) Hung, C.-W.; Kübler, D.; Lehmann, W. D. *Electrophoresis* **2007**, *28*, 2044-2052.
- (15) Gauci, S.; van Breukelen, B.; Lemeer, S. M.; Krijgsveld, J.; Heck, A. J. R. *Proteomics* **2008**, *8*, 4898-4906.
- (16) Mann, M.; Ong, S. E.; Gronborg, M.; Steen, H.; Jensen, O. N.; Pandey, A. *Trends Biotechnol.* **2002**, *20*, 261-268.
- (17) Neubauer, G.; Mann, M. *Anal. Chem.* **1998**, *71*, 235-242.

- (18) Garcia, M. C. *J. Chromatogr. B* **2005**, 825, 111-123.
- (19) Kweon, H. K.; Hakansson, K. *Anal. Chem.* **2006**, 78, 1743-1749.
- (20) Songyang, Z.; Blechner, S.; Hoagland, N.; Hoekstra, M. F.; Piwnica-Worms, H.; Cantley, L. C. *Curr. Biol.* **1994**, 4, 973-982.