RADICAL-PROMOTED PEPTIDE CLEAVAGE IN THE GAS-PHASE VIA
PEROXYCARBAMATES

By
Almary Chacon

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
in
Chemistry

December, 2007

Nashville, Tennessee

Approved:
Professor Ned A. Porter
Professor Michael P. Stone
Assistant Professor Brian O. Bachmann
Professor David L. Hachey
To my beautiful daughter Isabella
I would like to give a very special thanks to my advisor Dr. Ned Porter, not only for his guidance in the science of this project, but even more for his incredible patience and generosity. I would also like to thank my committee members for their input in the progression of this project over the years. Another special thanks to the mass spectrometry research core, in particular Dr. David Hachey, Lisa Manier, Dawn Overstreet, and Dr. Wade Calcutt, for all their help with mass spectrometry experiments.

I would also like to thank former and current members of the Porter Lab. A special thanks to the pioneers of this project; Dr. Douglas Masterson for his guidance and friendship and Dr. Huiyong Yin for all his help with many aspects of the research, including the preparation of this manuscript. Dr. Christopher Rector for his friendship and companionship through all the hoops and loops of this journey. Dr. Keri Tallman also for her friendship and scientific advice. Dr. Hye-Young Kim for all her help with the proteomics aspects of the project. And to all other lab members I have worked with throughout the years who have made the lab a very lively place. Another special thanks to Marianne Beebe, who has taken care of us in the lab for many years and has been a great listening ear. I would also like to mention Dr. Clare Kenny-Carney, former member of the Wright group, for her friendship and advice with proteomics experiments.

I am very grateful to my family and friends for all their support. To my parents for believing in me, and encouraging me through the end. They are the reason I am where I am. To my sister Oriana for always being so proud of me.
To all of my husband's family for their love and support. And finally, I could not have done this without my husband Robert, my best friend, thank you for being my rock all these years, and for making sure we remember to have a little fun. And to my adorable daughter Isabella, for bringing so much happiness to our lives.
TABLE OF CONTENTS

Page

DEDICATION ........................................................................................................... ii

ACKNOWLEDGEMENTS ....................................................................................... iii

LIST OF TABLES .................................................................................................. viii

LIST OF SCHEMES ................................................................................................... ix

LIST OF FIGURES .................................................................................................. xi

Chapter

I. INTRODUCTION .................................................................................................... 1

Proteomics ........................................................................................................... 1
  Bottom-up proteomics ..................................................................................... 1
  Mass spectrometry ......................................................................................... 5
  Mass Spectrometry Analysis in Proteomics ....................................................... 9
  Peptide fragmentation and protein identification .............................................. 10
  Database search algorithms and other data manipulation software ................. 13
  Top-Down Proteomics ................................................................................... 15
  Charge state manipulation .............................................................................. 16
  High-resolution product ion mass determination ........................................... 17

Radical Degradation of Proteins ....................................................................... 19
  Radical sources ............................................................................................ 19
  Carbon-centered radicals .............................................................................. 20
  Nitrogen-centered radicals ............................................................................ 21

Research Goals .................................................................................................. 22
References .......................................................................................................... 23

II. FREE RADICAL-PROMOTED PEPTIDE CLEAVAGE VIA LYSINE PEROXYCARBAMATES ........................................................................................................... 28

Introduction ......................................................................................................... 28
Results and Discussion ........................................................................................ 32
  Model Lysine Peroxycarbamate Studies ............................................................... 32
  Mechanism of Fragmentation of Model Lysine Peroxycarbamate ................. 34
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>III-1.</td>
<td>N-terminal peroxycarbamates of amino acid adducts and their CID fragmentation.</td>
<td>72</td>
</tr>
<tr>
<td>III-2.</td>
<td>CID fragmentation of N-terminal peroxycarbamates of peptides.</td>
<td>77</td>
</tr>
<tr>
<td>III-3.</td>
<td>Peptides detected by LC-MS from a modified myoglobin digest.</td>
<td>83</td>
</tr>
<tr>
<td>III-4.</td>
<td>b- and y-ions from MS$^2$ of unmodified peptide VEAKIAGHGQEVLIR compared to b- and y-ions from MS$^3$ of the modified peptide ([C$_4$H$_9$OOC(O)M]).</td>
<td>84</td>
</tr>
<tr>
<td>III-5.</td>
<td>First 6 results from SALSA search for peptides with Histidine at the N-terminus in a myoglobin digest.</td>
<td>86</td>
</tr>
<tr>
<td>III-6.</td>
<td>Analysis of modified peptides (N-terminal peroxycarbamates) by MALDI.</td>
<td>88</td>
</tr>
<tr>
<td>IV-1.</td>
<td>CID fragmentation of model ornithine derivatives.</td>
<td>107</td>
</tr>
<tr>
<td>IV-2.</td>
<td>CID data for peptide ornithine peroxycarbamate derivatives and their unmodified analogues.</td>
<td>111</td>
</tr>
<tr>
<td>IV-3.</td>
<td>CID fragmentation for peptide ornithine derivatives.</td>
<td>111</td>
</tr>
<tr>
<td>Scheme</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>I-1.</td>
<td>Homolytic cleavage reactions: A) decomposition of alkyl peroxides, B) decomposition of azo compounds, C) homolysis of diacyl peroxides, D) decomposition of t-butyl peroxycarbamates.</td>
<td>20</td>
</tr>
<tr>
<td>I-2.</td>
<td>Alkoxyl radical induced peptide backbone cleavage.</td>
<td>21</td>
</tr>
<tr>
<td>I-3.</td>
<td>Mechanism of the Hoffmann-Löffler-Freytag reaction.</td>
<td>22</td>
</tr>
<tr>
<td>II-1.</td>
<td>Mechanism for the Hoffman-Loffler-Freytag reaction.</td>
<td>29</td>
</tr>
<tr>
<td>II-2.</td>
<td>Modification of lysine model as a peroxycarbamate and fragmentation by CID.</td>
<td>30</td>
</tr>
<tr>
<td>II-3.</td>
<td>CID fragmentation of peptides modified at the N-terminus with an azo radical initiator.</td>
<td>31</td>
</tr>
<tr>
<td>II-4.</td>
<td>CID fragmentation of N- and S-nitroso peptides.</td>
<td>31</td>
</tr>
<tr>
<td>II-5.</td>
<td>Peroxycarbamate modification of model lysine.</td>
<td>32</td>
</tr>
<tr>
<td>II-6.</td>
<td>CID fragmentation of model lysine peroxycarbamate.</td>
<td>33</td>
</tr>
<tr>
<td>II-7.</td>
<td>Proposed mechanism of fragmentation of model lysine peroxycarbamate II-3 (m/z = 325).</td>
<td>38</td>
</tr>
<tr>
<td>II-8.</td>
<td>CID fragments of modified Ac-GSAKVSF where lysine was converted to the respective t-butyl peroxycarbamate (m/z 875.3, M+116 + Li⁺).</td>
<td>40</td>
</tr>
<tr>
<td>II-9.</td>
<td>A) Proposed mechanism of radical-promoted peptide fragmentation by lysine aminyl radicals produced from lysine peroxycarbamates. B) Proposed mechanism of radical-promoted peptide fragmentation by ECD.</td>
<td>41</td>
</tr>
<tr>
<td>II-10.</td>
<td>Synthesis of peroxy carbonate reagents.</td>
<td>45</td>
</tr>
<tr>
<td>III-1.</td>
<td>Fragmentation of N-terminal peroxy carbamates by CID.</td>
<td>67</td>
</tr>
<tr>
<td>III-2.</td>
<td>Modification of lysine model as a peroxycarbamate and fragmentation by CID.</td>
<td>68</td>
</tr>
<tr>
<td>III-3.</td>
<td>CID fragmentation of peptides modified at the N-terminus with an azo radical initiator.</td>
<td>68</td>
</tr>
<tr>
<td>III-4.</td>
<td>Peroxycarbamate modification of N-termini and CID fragmentation.</td>
<td>69</td>
</tr>
</tbody>
</table>
III-5. CID fragmentation of N-terminal peroxyacarbamate amino acid ester derivatives.............................................................. 71

III-6. A) Fragmentation of N-terminal peroxycarbamates by CID. B) CID of the [C₄H₉OOC(O)M – 116] fragment obtained by in-source CID of N-terminal peroxycarbamates............................................................... 80

III-7. Preparation of deuterated peroxycarbonate reagent. .............................. 81

III-8. Proposed mechanism for dissociation of N-terminal peroxycarbamate. .. 82

IV-1. Lactamization of ornithine in solution...................................................... 102

IV-2. Possible gas-phase radical pathways for ornithyl peroxycarbamate peptides. ................................................................................................................................. 103

IV-3. Modification of N-α-acetyl-ornithine-methyl ester as a peroxycarbamate. .................................................................................. 106

IV-4. Akabori reaction of a tripeptide. .............................................................. 108

IV-5. Proposed mechanism of asparagine peptide bond cleavage by hydrazine. .................................................................................. 113

V-1. Modification of lysine model as a peroxycarbamate and fragmentation by CID.............................................................. 126

V-2. CID fragmentation of N-terminal peroxycarbamates.............................. 127

V-3. Labeling of b-ion series by MS³ of N-terminal peroxycarbamates. ....... 128

V-4. Lactamization of Ornithyl peptides. ..................................................... 129

V-5. Conversion of arginine to ornithine. ...................................................... 130

V-6. The Akabori reaction for identification of C-terminus. ......................... 131
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-1.</td>
<td>Schematics of shot-gun proteomics.</td>
<td>3</td>
</tr>
<tr>
<td>I-2.</td>
<td>Illustration of the use of on-line ion-exchange and reverse-phase HPLC for</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>shot-gun proteomics.</td>
<td></td>
</tr>
<tr>
<td>I-3.</td>
<td>Schematics of the basic components of an ESI-triple-quadrupole mass</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>spectrometer.</td>
<td></td>
</tr>
<tr>
<td>I-4.</td>
<td>Schematics of an ESI-ion-trap mass spectrometer.</td>
<td>8</td>
</tr>
<tr>
<td>I-5.</td>
<td>Schematics of a reflectron MALDI-TOF mass spectrometer.</td>
<td>9</td>
</tr>
<tr>
<td>I-6.</td>
<td>General illustration of data-dependant scanning with an ESI-ion-trap</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>instrument.</td>
<td></td>
</tr>
<tr>
<td>I-7.</td>
<td>Peptide fragmentation nomenclature.</td>
<td>12</td>
</tr>
<tr>
<td>I-8.</td>
<td>Schematics of a FT-ICR instrument.</td>
<td>17</td>
</tr>
<tr>
<td>II-1.</td>
<td>CID of the lithium adduct of II-3.</td>
<td>34</td>
</tr>
<tr>
<td>II-2.</td>
<td>A) ESI-CID spectrum of the lysine peroxycarbamate model II-3 (m/z = 325).</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>B) ESI-CID spectrum of aminyl radical II-4 (m/z = 208) formed by in-source</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CID of II-3.</td>
<td></td>
</tr>
<tr>
<td>II-3.</td>
<td>A) In-source CID of the lithium adduct of II-3 (m/z = 325), effect of tube</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>lens voltage on the intensity of product ions.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B) CID of aminyl radical II-4 (m/z = 208) produced by in-source CID of</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lithiated II-3, effect of CID off-set voltage on product ion intensities.</td>
<td></td>
</tr>
<tr>
<td>II-4.</td>
<td>A) MALDI spectrum of unmodified peptide Ac-SYMEHFRTGKPVGKKR.</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>B) MALDI spectrum of modified peptide as a mixture of 1, 2, and 3 lysine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>peroxycarbamate adducts.</td>
<td></td>
</tr>
<tr>
<td>II-5.</td>
<td>Amount of reagent versus extent of lysine modification for peptide AcGKASVKF.</td>
<td>47</td>
</tr>
<tr>
<td>II-6.</td>
<td>Amount of reagent versus extent of lysine modification for peptide SYSMEHFRTGKPVGKKR.</td>
<td>48</td>
</tr>
<tr>
<td>II-7.</td>
<td>Amount of 4-nitrophenyl- t-butyl peroxycarbonate (II-2) versus extent of</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>modification for peptide SYSMEHFRTGKPVGKKR.</td>
<td></td>
</tr>
</tbody>
</table>
II-8. ESI-MS spectrum of modified Ribonuclease A showing the 10H$^+$ charge state................................................................. 51

II-9. Cation-exchange chromatogram of modified Ribonuclease A. .............. 52

II-10. LC-MS analysis of singly modified ribonuclease A obtained from cation-exchange separation. .................................................. 53

II-11. Illustration of formation of regioisomers of ribonuclease A due to 1 peroxycarbamate modification. ........................................ 54

III-1. CID spectrum of N-terminal peroxycarbamate valine methyl ester. .......... 71

III-2. Peroxycarbonate reagents. ..................................................................... 74

III-3. LC/MS of modified peptide YRVRFLAKENVTQDAEDNC... ................. 76

III-4. LC-SRM trace of peptide mixture HPKRPWIL and DRVYIHPF modified with reagent 1b at pH 5.5......................................................... 78

III-5. MALDI spectra of modified peptides as N-terminal peroxycarbamates.... 89

III-6. MS/MS MALDI spectrum of the imine ion [M – (R + 1) + 1H]$^+$ of the peptide HGTVVLTALGGILK from the modified myoglobin tryptic digest................................................................. 90

IV-1. CID fragmentation of ornithine peroxycarbamate methyl ester. 
Effect of CID off-set voltage on ion intensities. ........................................ 106

IV-2. CID spectrum of ornithine methyl ester peroxycarbamate derivative IV-3................................................................. 107

IV-3. MS spectrum showing conversion of Arg to Orn for peptide MEHFRWG (m/z = 962, [M + 1H]$^+$) Ornithine derivative m/z = 920 [M – 42 + 1H]$^+$. A) Microwave irradiation for 5 min. B) Microwave irradiation for 10 min..... 109

IV-4. SDS-PAGE of proteins undergoing hydrazinolysis .............................. 114

IV-5. BSA sequence coverage (red letters) for band 5 (top to bottom) from SDS-PAGE. ................................................................. 115
CHAPTER I

INTRODUCTION

Proteomics

Proteomics is the study of the proteome, which contains all of the proteins derived from a genome. Proteomics involves the study of multiprotein systems, the interactions between the components of these systems, and their role as part of a larger network. This includes, among other things, identification of proteins from a known genome, mapping of post translational modifications, and comparison of protein expression in different cell states.\(^1\) There are two general approaches to proteomics analyses, bottom-up and top-down. Bottom-up proteomics is the most widely used and involves enzymatic digestion of the proteins into peptides prior to mass spectrometry analysis. Protein identification is achieved by the use of database search algorithms. “Top-down” protein identification approaches, which have been developed more recently, involve the mass spectrometry analysis of the intact proteins without prior digestion.\(^2\)

Bottom-up proteomics

The first step in a bottom-up method involves isolation of the protein of interest using separation techniques such as 2D polyacrylamide gel electrophoresis (2D-PAGE), ion-exchange chromatography, and reverse phase chromatography.\(^1\) 2D-PAGE consists of a combination of separation by isoelectric point and size. The isoelectric point of a protein is the point at which
there is an equal distribution of positive and negative charges throughout the protein resulting in a zero net charge. In the Isoelectric point separation step, a protein mixture is applied to a strip of polyacrylamide gel containing an immobilized pH gradient. This immobilized pH gradient gel (IPG) contains acrylamide derivatives that have buffering moieties with a specific pKa directly incorporated into the gel. As voltage is applied, the proteins will focus at the points on the gel where they are neutral. In a 2D-SDS-PAGE experiment, the proteins are applied to an IPG strip in the first step, then the strip is treated with the buffer containing sodium dodecyl sulfate (SDS) and jointed to the SDS-PAGE part of the gel, where they are resolved by size. In size separation, the proteins associate with SDS forming micelles of size according to their molecular weights. By this process the proteins are given a net negative charge which increases in magnitude with increasing size. A voltage is applied which causes the negatively charged micelles to migrate through the gel towards the positive electrode, with the low molecular weight proteins moving faster than the high molecular weight proteins. The gel is then stained by conventional techniques such as silver, coomassie, and amido black stains.\textsuperscript{3,5}

Once a mixture of proteins has been resolved, the different proteins are digested with a protease, typically trypsin. This is a serine protease which cleaves at the C-terminus of lysine and arginine, unless followed by proline. The digestion is done either in solution after extraction of the proteins from the gel, or directly in the gel (“in gel digestion”).\textsuperscript{3, 5} Alternatively, the proteins can be digested prior to separation, this is referred to as the “shot-gun” approach (Figure
In this approach the separation is done at the peptide level by tandem liquid chromatographic techniques, involving ion-exchange and reverse-phase chromatography. In ion exchange chromatography, the peptides are separated by charge. A resin with immobilized charged groups is used and elution of the peptides is done by varying the ionic strength of the buffer. For peptides, cation-exchange chromatography is the method of choice. The resins used typically contain an immobilized negatively charged group such as sulfate or carboxylate. Peptides with multiple positive charges will adhere more strongly to the resin and will require a higher ionic strength for elution than singly charged peptides.

Figure I-1. Schematics of shot-gun proteomics.

Reverse-phase liquid chromatography (RP-LC) separates molecules based on hydrophobicity. The packing material for RP-LC consists of large hydrocarbon chains immobilized on silica. Hydrophobic peptides then have a
higher affinity for the packing material and elute later than more polar peptides.³ Tandem LC coupled to mass spectrometry, has been shown to be very effective for the analysis of complex peptide mixtures. Yates et al. describes a tandem LC-MS method that combines ion exchange LC and RP-LC.⁸ First a peptide mixture is applied to a strong cation-exchange microcapillary column. Peptides are eluted by a step gradient where the salt concentration of the buffer is increased. With each step, a group of peptides is eluted which then passes on to a RP microcapillary column for further separation and then to the mass spectrometer (Figure I-2). This approach provides high resolution analysis and allows for a larger number of peptides to be identified in a single run compared to RP-LC alone.⁸,⁹

**Figure I-2.** Illustration of the use of on-line ion-exchange and reverse-phase HPLC for shot-gun proteomics.
Mass spectrometry

In order to understand proteomics analyses it is necessary to have an understanding of the different mass spectrometry techniques and how they are applied to protein identification.

The first step in mass spectrometry analyses requires the ionization of the analyte into the gas-phase. There are numerous methods for ionization, however electrospray ionization (ESI) and matrix assisted desorption ionization (MALDI) are the most widely used in proteomics and will be discussed in such context. In positive ESI, the analyte (usually a peptide mixture), is introduced as an acidic aqueous solution. The sample is passed through a small-diameter needle, where a high, positive voltage is applied causing the sample to spray out into a mist of droplets. These droplets contain peptide ions as well as the sample solvent. The acidic conditions cause the droplets to have a net positive charge which results in their movement towards the negatively charged inlet of the instrument. Evaporation reduces the size of the droplets which causes the splitting of these into smaller charged droplets. This process occurs several times with the aid of a flow of nitrogen and heat until dessolvated gas-phase peptide ions are obtained. These are then directed into the mass analyzer. ESI generally produces multiply charged species, and the number of charges in the case of peptides and proteins, is influenced by the presence of basic amino acid residues such as arginine and lysine. One of the advantages of this ionization method in proteomics is its compatibility with HPLC separations.
ESI can be coupled to several mass analyzers; however, the most commonly used in bottom-up proteomics are the triple-quadrupole and the ion-trap analyzers. The triple-quadrupole analyzer is composed of three quadrupoles, which consist of four metal rods arranged in parallel (Figure I-3). An electric field is established through the quadrupole by applying direct current (DC) and radio frequency (RF) voltages simultaneously. The electric field causes ions of a certain m/z value to have a stable cork-screw like trajectory throughout the field, while other ions of different m/z move outwards and get lost by a variety of mechanisms. Varying the RF/DC potential applied to the rods will determine what ions get through the quadrupole. In a triple-quadrupole analyzer, only two quadrupoles are subjected to DC and RF voltages, one is located at the entrance of the analyzer (Q1) and the other at the exit (Q3). The quadrupole located in the middle (Q2) is used as a collision cell in collision-induced dissociation (CID) where ions can be subjected to collisions with argon gas atoms to induce fragmentation.\textsuperscript{11-13} A full scan can be done to obtain information about all the ions present in a sample at a given time. The instrument achieves this by scanning all the ions passing through Q1. This quadrupole can also be used as a mass filter to allow only one ion to pass through to Q2. For CID experiments, Q1 is used to select the ion of interest, which then passes to the collision cell (Q2) and gets fragmented. The fragment ions are then scanned in Q3.\textsuperscript{11} In an ESI triple-quadrupole instrument, fragmentation of certain ion species can also occur at the source when enough voltage is applied, due to the relatively high gas pressure in the source. This phenomenon is referred to as in-source CID.\textsuperscript{14}
Figure I-3. Schematics of the basic components of an ESI-triple-quadrupole mass spectrometer.

The ion-trap analyzer consists of two conical electrodes and one ring electrode (Figure I-4). The ions are held in the small interior volume between the conical electrodes. By lowering and raising the voltages on the entrance and exit trap electrodes, ions pass through the trap, are stored for a period of time (μs to sec), and are released into the detector. Ions of a certain m/z are selectively released from the trap in the mass selective stability mode. One of the main advantages of the ion-trap is that it allows for multiple CID (MS/MS) events (MS^n). Product ions from a first MS/MS event can be stored, reactivated in another MS/MS event and so on.  

Matrix Assisted Desorption Ionization (MALDI) requires the incorporation of the peptide sample into a matrix composed of crystalline UV-absorbing molecules. The most common matrices are α-cyano-4-hydroxycinnamic acid, sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid), and 2,5-dihydroxybenzoic acid. These molecules must absorb at the wavelength of the laser, which is normally 337nm for a nitrogen laser. When the laser hits the matrix crystals, they absorb energy in the form of photons, which is emitted as heat. Rapid heating
causes expansion of the matrix and peptide molecules into the gas phase, where energy is transferred from the photoionized matrix molecules to the peptide molecules, which in turn get ionized. 

This ionization method generates mostly singly charged ions, although some doubly and triply charged protein ions are observed. The mass analyzer typically coupled with MALDI is the time-of-flight (TOF).

![Schematics of an ESI-ion-trap mass spectrometer.]

**Figure I-4.** Schematics of an ESI-ion-trap mass spectrometer.

The time-of-flight mass analyzer separate ions and measures their m/z based on the time they take to pass from the ion source to the detector. Ions are generated in the source and flown down the flight tube under the influence of an accelerating potential. To increase mass resolution, a reflectron is introduced at the end of the tube. Once the ions reach the reflectron, they are gradually repelled back to a second detector (Figure I-5). MS/MS experiments are possible by combining two TOF analyzers (MALDI-TOF-TOF). 

11, 17
Mass Spectrometry Analysis in Proteomics

Mass spectrometry analysis is the most important step in protein identification. The two most common protein identification methods through peptide analysis are peptide mass fingerprinting and peptide sequencing by tandem mass spectrometry. In peptide mass fingerprinting the peptides derived from an enzymatic protein digestion are analyzed typically by a MALDI-TOF for accurate mass measurements. The parent protein then is identified by matching the measured peptide masses to the corresponding peptide masses from a protein database. This database contains a great number of known protein sequences derived from known genomes. By performing a theoretical digestion of all the proteins in the database, a list of peptide sequences with their corresponding m/z's is generated which is compared to the experimental peptide m/z's. Protein identification in this manner is very dependant in the accuracy of the peptide m/z measurements and available databases. In addition, false
positive matches are common due to the complexity of the MS spectra. Nevertheless, this approach allows for fast and relatively simple protein identification. 4, 18, 19

A more useful, yet more complex approach to protein identification involves peptide sequencing by tandem mass spectrometry. This is carried out using ESI-triple quadrupole or ESI-ion-trap instruments. The latter is commonly the best choice for shot-gun proteomics analyses due to better sensitivity and automatization of data acquisition. In this approach, not only peptide masses are measured, but CID fragmentation patterns are obtained which are used to determine the amino acid composition through database search algorithms. ESI-ion trap instruments allow for high throughput analysis with data dependant scanning (FigureI-6). In data dependant mode, first a full spectrum is obtained of all the m/z's present at a certain LC retention time, then the software selects the most abundant ions in the full spectrum to perform MS/MS. A great amount of data is generated in this manner with a single sample injection. This approach has the advantage over peptide mass fingerprinting of better accuracy of protein identification with reduced false positive matches. In addition, its applications are not limited only to protein identification; other information such as sites of posttranslational modifications can be obtained. 6

**Peptide fragmentation and protein identification**

In collision induced dissociation (CID), ions are activated by collisions with neutral gas molecules. Peptide ions dissociate under CID conditions generating a
series of predictable backbone fragments. These fragments are used to deduce the amino acid composition of the peptide.

Figure I-6. General illustration of data-dependant scanning with an ESI-ion-trap instrument.

Peptide fragmentation along the backbone is the most important for sequencing analyses. However, numerous other fragments can be present in an MS/MS spectrum. A diagram showing peptide fragmentation nomenclature is shown in Figure I-7. Cleavage of the bond between the carbonyl carbon and the amide nitrogen to form b and y-ions is the most common in CID. These fragments are the most important for peptide sequencing; therefore algorithms are optimized for this type of fragmentation. In the y-ion series, the positive charge of the C-terminus of the peptide is retained. The b-ion series contains the charged N-terminus. The ion number indicates the number of amino acids present. For example b₁, b₂, and b₃ ions contain 1, 2, and 3 amino acids.
respectively. b- and y-ions are complementary and provide a pattern in which m/z gaps between consecutive fragments correspond to the amino acids in the sequence. Under ESI conditions, multiple charged ions are observed. Multiply charged peptide ions provide more information in CID spectra than singly charged peptide ions, since neutral fragments are lost and not observed in the spectra.\textsuperscript{13, 19}

![Figure I-7. Peptide fragmentation nomenclature](image)

Protein identification using MS/MS spectra can be done in two ways, by de novo sequencing and by the use of search algorithms. In de novo sequencing, peptide sequences are determined by manually inspecting the MS/MS data or with the help of de novo sequencing software such as Lutefisk.\textsuperscript{20} Then the protein sequence is matched to homologous sequences in a genome database using the BLAST search algorithm.\textsuperscript{21} This process is very lengthy and inefficient compared to automated analysis of MS/MS data by search algorithms such as SEQUEST.\textsuperscript{22} However, it can be useful when only a small set of data is being
analyzed, or when the genome of the organism being studied is not known. In SEQUEST algorithm searches, b- and y-ions of peptides are matched with theoretical b- and y-ions of the corresponding peptides in the protein database. This approach allows for analysis of a much greater amount of MS/MS spectra, and it is the most widely used.

**Database search algorithms and other data manipulation software**

SEQUEST is the most widely used database search algorithm. The way this search algorithm works is by comparing MS/MS spectra to theoretical MS/MS spectra available from a protein database. First one must specify the digestion enzyme, typically trypsin. The program then performs a theoretical digestion of all the proteins available for a certain database, for example, human, bovine etc... A list of peptides from all the proteins is generated, then a list of theoretical b- and y-ions for each peptide. The program matches the precursor peptide ion m/z to possible peptides from the list, and the MS/MS spectra are compared to theoretical spectra. A correlation score is calculated that tells how accurate the match is. Peptide sequences are then matched to proteins in the database. Although this software simplifies and speeds up data analysis tremendously, it cannot decide whether the data is significant. However the probability of a legitimate match increases with increasing number of matched peptides with good correlation scores. There are other problems associated with using SEQUEST, and that is the issue of post-translational modifications. Unexpected post-translational modifications may not be accounted for in the matching process and this would disrupt protein identification. However, it is
possible to account for known modifications such as phosphorylation, cysteine oxidation, and glycosylation by modifying search parameters. Nevertheless, SEQUEST has proven to be a great tool for proteomics analyses allowing for manipulation of enormous amount of data in relatively short time.\textsuperscript{13, 24}

Often in proteomics there is an interest not only on identifying proteins in a biological sample, but also specific features such as post-translational modifications and protein adducts. There is spectra manipulation software that has been designed for this purpose such as SALSA\textsuperscript{25} (Scoring Algorithm for Spectral Analysis) and P-Mod\textsuperscript{26}. SALSA is used for finding specific features in the MS/MS spectra. These features can be a product ion having a specific m/z, a neutral loss, which would be a m/z difference between the precursor ion and a fragment present in the MS/MS spectrum, a charged loss, and an ion pair, which are two fragments with a m/z difference equal to the m/z of a certain amino acid. SALSA identifies these characteristics in all of the MS/MS spectra and scores the data according to the number of characteristics matched and peak intensities.\textsuperscript{25} This is a great tool for analyzing mass spectral data and obtaining specific information about the protein that may be missed by SEQUEST.

The p-Mod software was designed specifically for mapping of post-translational modifications. In this case, the protein identity is known. The user inputs the protein sequence and performs a theoretical digestion to generate a list of peptides. The software matches the experimental data to the theoretical peptide list and identifies m/z differences between the MS/MS precursor ions and the m/z of the precursors generated on the theoretical list. It also identifies the
location of the modification in the sequence by detecting the mass shift in the b- and y-ions. A p-value is calculated which reflects the probability of a false positive match.26

Bottom-up proteomics is the most widely used approach for protein identification. On-line multidimensional capillary HPLC methods used for shotgun approaches as described above, allow for high resolution separations and fully automated, high throughput analyses. Moreover, bioinformatics tools are optimized for bottom-up strategies. However there are several limitations to this tactic. First, these methods seldom provide high sequence coverage of the protein. Although only a fraction of the peptides derived form a given protein is necessary for identification, a lot of information is lost in the analysis. For example, post translational modifications may not be represented in the peptide population ultimately identified. Peptides of low abundance are also suppressed by high-abundance peptides in the MS spectra. And although bottom-up analyses are high throughput, on-line multistep chromatography separations still require extended run times (hours).27 Top-down approaches offer solutions to these problems, however, they are not yet well established and new techniques and improvements are still needed for these approaches to become more widely used.28

Top-Down Proteomics

Top-down proteomics involves the direct molecular weight measurement of intact proteins followed by fragmentation of the protein ions in the gas-phase. Fragmentation of intact protein ions was first demonstrated for ribonuclease A
using an ESI-triple quadrupole instrument.\textsuperscript{29} However, the product ion spectra derived from multiply charged protein ions are very complex, and without high mass resolution, assignment of product ion charge states is not possible. To overcome this problem two approaches have been developed. One is the use of more powerful instruments such as fourier transform mass spectrometers (FT-ICR MS)\textsuperscript{30} and the other is charge state manipulation.

\textbf{Charge state manipulation}

Gas-phase ion chemistry makes possible the use of low resolving power mass analyzers such as triple-quadrupoles and ion-traps for obtaining sequence information of whole intact proteins. Ion-Ion proton transfer is the most commonly used charge manipulation technique for reducing multiple charged product ions to their singly charged state. In this reaction a proton from a multiply charged cation is transferred to a singly charged anion to give a reduced charge cation and a neutral species (Equation 1). This ion chemistry is achieved with the modification of the mass spectrometer to include an atmospheric sampling glow discharge ionization source (ASGDI) to introduce negative ions into the mass analyzer. This has been more widely achieved with ESI-ion-traps; however it has also been shown with a triple-quadrupole mass analyzer. This technique has also been useful for charge state manipulation of protein precursor ions. Where charge states can be concentrated and purified for dissociation.
High-resolution product ion mass determination

Very high isotopic mass resolution can be achieved with a Fourier-Transform Ion Cyclotron Resonance mass analyzer (FTICR). This is typically combined with an ESI source and it is commonly used in top-down proteomics. In FTICR, ions are trapped in a magnetic field where they oscillate at a certain frequency, adopting stable cyclotron orbits. The ions then get excited to a greater cyclotron radius by applying an oscillating electric field. The signal is detected as an image current on a pair of plates which the packet of ions passes close to. The ions are excited simultaneously by applying a wide range of excitation energies at once. The signal is a multiple ion wave plot of the amplitude of the frequency over time. Fourier transform is then used to produce a mass spectrum.

Figure I-8. Schematics of a FT-ICR instrument.
High magnetic field strength FT-ICR instruments offer very high resolving power and provide high mass accuracy. With these instruments, it is possible to determine isotope spacing of high mass ions and product ions, therefore charge state determinations are feasible. Also, unit mass resolutions allow for identification of disulfide bridges, deamination, and other post-translational modifications.

Electron capture dissociation (ECD) is used for the dissociation of ions as opposed to CID. In this method, ions pass through a high-current electron beam and are activated by ion-electron collisions. The result is fragmentation of N-Cα bonds along the backbone to form mainly c- and z- type ions. ECD provides greater sequence coverage than CID (more bonds are cleaved) and allows preservation of post-translational modifications. 31

The main advantage of top-down over bottom-up methods is the higher sequence coverage obtained. The entire protein is available for sequencing and not just a small population of abundant peptides. In addition, bypassing all the preparative steps associated with bottom-up methods greatly reduces analysis time. 32 However, this approach is not as mature as bottom-up strategies and is limited by a great number of factors. First of all, FT-ICR instruments are highly sophisticated and very expensive, which limits the access to these to a very few number of research laboratories. Another problem with top-down proteomics is the need to separate small quantities of complex protein mixtures and yet obtain samples compatible with ESI. Therefore direct protein MS analysis is so far limited to isolated proteins or simple mixtures. In addition, mass accuracy and
resolution decrease as molecular weight increases, therefore identification of very large proteins remains challenging. Alternatively, hybrid methods have been shown to provide the best of both worlds. VerBerkmoes et al. reports the identification of 868 proteins from the bacterium *S. oneidensis* using a combination of the shot-gun method and FT-ICR techniques. Nevertheless top-down proteomics methods continue to improve and have become more common in the past few years.

**Radical Degradation of Proteins**

Radicals are formed generally either by homolytic bond cleavage which is induced by the absorption of energy (thermolysis or photolysis), or by electron transfer reactions. Radicals can undergo numerous processes such as substitution, addition, rearrangement, elimination and fragmentation. These reactions are important in chemical as well as in biological systems. In chemistry, radical reactions have been useful in the synthesis of organic molecules and in radical polymerization reactions. In biology, radicals are involved in reactions with biomolecules such as proteins, lipids, and nucleic acids, which can lead to damage and degradation.

**Radical sources**

Chemical bonds are cleaved to form radicals upon absorption of energy such as heat or radiation; however, some bonds do this more readily than others. The most common bonds to undergo homolysis are the peroxy bond and the azo bond. Such linkages serve as initiators for other radical processes.
peroxide bond is a particularly weak bond with bond dissociation energies of 30 to 35 Kcal/mol.\textsuperscript{35} Diacyl peroxides are known to undergo decomposition in solution and in the gas-phase via homolytic or heterolytic pathways to give free radicals.\textsuperscript{36, 37} Homolysis of acylperoxides typically results in carboxyl radicals which further decompose to release CO\textsubscript{2} and a carbon-centered radical. The carbon radical can then couple to another carbon radical or trap another carboxyl radical. Peroxycarbamates have been shown to undergo thermal decomposition in a similar fashion by first cleavage of the O-O bond and loss of CO\textsubscript{2} to form a nitrogen-centered radical and an alkoxy radical.\textsuperscript{38} Scheme I-1 shows general equations for all homolytic cleavages discussed above.

**Scheme I-1.** Homolytic cleavage reactions: A) decomposition of alkyl peroxides, B) decomposition of azo compounds, C) homolysis of diacyl peroxides, D) decomposition of t-butyl peroxycarbamates.

**Carbon-centered radicals**

Radical reactions with peptides and proteins normally give rise to carbon-centered radicals at either an amino acid side chain or a backbone $\alpha$-carbon.
The carbon-centered radicals are most commonly formed by abstraction of a hydrogen by another radical species inter- or intra-molecularly.\textsuperscript{39} The fate of the carbon-centered radicals is controlled by a number of processes, which include fragmentation of the backbone. For example, it has been shown that exposure of proteins to radiolytically generated HO\textsuperscript{-} in the presence of O\textsubscript{2} leads to fragmentation of the proteins.\textsuperscript{39} In another study, it was found that alkoxyl radicals formed at the C-3 position on aliphatic amino acid side chains lead to the formation of carbon-centered radicals that can undergo β-scission, resulting in backbone cleavage and loss of the amino acid side chain (Scheme I-2).\textsuperscript{40}

\[ \text{NH}_2\text{O} \text{R}_1\text{R}_2\text{O} \rightarrow \text{NH} \cdot \text{CH}_2\text{CO} + \text{CHO} \]

\[ \text{O}_2 \]

Backbone cleavage

\textbf{Scheme I-2.} Alkoxyl radical induced peptide backbone cleavage.

\section*{Nitrogen-centered radicals}

Protein-derived nitrogen-centered radicals can be formed at the backbone or at lysine side-chains.\textsuperscript{39} It has been shown that reaction of proteins with hypochlorous acid results in the formation chloroamines which are unstable and decomposed to give nitrogen-centered radicals.\textsuperscript{41} It has also been shown that such reactions can lead to protein backbone cleavage; however the mechanisms of such processes are not well understood. There is evidence that nitrogen-
centered radicals can undergo rearrangements to form carbon-centered radicals. For example, lysine side-chain aminyl radicals undergo intramolecular hydrogen abstraction to form carbon-centered radicals. This process is analogous to the Hoffman-Löffler-Freytag reaction in organic synthesis. The Hofmann-Löffler-Freytag reaction is known to be initiated by the formation of a chloroammonium salt from an N-chloroamine in the presence of acid (Scheme I-3). Then by irradiation or heat, this salt is converted to a nitrogen radical cation. Hydrogen atom abstraction by the nitrogen radical cation at C-5 results in a carbon centered radical which is then captured by a chlorine radical. The 1,5-hydrogen abstraction is favored by the stability of a six-member ring transition state. Ionic cyclization follows to form pyrrolidine derivatives.


Research Goals

Radical degradation of proteins has received a great amount of attention in the context of disease. However, the application of radical processes to proteomics has yet to be explored. The work presented here involves the
investigation of the radical induced backbone cleavage of peptides and proteins and its application to the improvement of top-down proteomics methodologies.

Our research efforts are aimed at the following:

1. The development of a gas-phase peptide fragmentation strategy based on backbone cleavage induced by lysine side-chain nitrogen-centered radicals. This technique has the potential to facilitate fragmentation of intact proteins by mass spectrometry methods.

2. The development of a mass spectrometry technique for identification of the N-terminal amino acid of peptides, based on radical induced amino acid side chain cleavage. Applications of this technique include improvement of peptide mass fingerprinting methods and de novo sequencing strategies.

3. The investigation of chemical-based methods of fragmentation of large proteins to improve top-down methodologies.

References


CHAPTER II

FREE RADICAL-PROMOTED PEPTIDE CLEAVAGE VIA LYSINE PEROXycARBAMATES

Introduction

Because bottom-up proteomics strategies require long multi-step sample purification and analysis, and because of the potential advantages offered by top-down methods, it would be desirable to improve top-down strategies. Despite all the advances in mass spectrometry, in particular the development and improvement of FT-ICR instruments, fragmentation of intact proteins remains challenging.\textsuperscript{1-3} Moreover, access to FT-ICR technology is limited to a small number of research laboratories due to the high cost of purchase and maintenance of the instruments. Therefore mass spectrometry techniques that facilitate gas-phase fragmentation of proteins are very valuable for the improvement of top-down approaches.

Radical reactions of proteins have been shown to result in protein degradation by radical induced backbone cleavage.\textsuperscript{4-6} Nitrogen-centered radicals of lysine side-chains have been shown to induce protein backbone cleavage possibly via intramolecular hydrogen abstraction to form carbon-centered radicals. Precedent for this mechanism can be found in the Hoffman-Löffler-Freytag reaction, where an aminyl radical cation undergoes 1,5 hydrogen abstraction to form a carbon-centered radical (Scheme II-1).
Based on this knowledge, the idea of gas-phase protein backbone cleavage promoted by nitrogen-centered radicals seems feasible. This could be the basis for a mass-spectrometry sequencing technique that offers the possibility of intact protein fragmentation without the need for prior enzymatic digestion. The chemistry would involve the introduction of a labile covalent bond to generate peptide or protein radical species that would facilitate fragmentation. Former laboratory members (Dr. Douglas Masterson and Dr. Huiyong Yin) pioneered this idea by the use of the chemical modification of lysine residues as peroxycarbamates.\textsuperscript{7} These adducts undergo dissociation under ESI-CID and well as MALDI conditions. The result is a lysine aminyl radical that can initiate a fragmentation pathway in the gas-phase to produce peptide backbone cleavage in the form of a-, b-, c-, and z-type ions.\textsuperscript{7, 8} These backbone fragments can be used to obtain peptide sequence information.

The lysine peroxycarbamates are obtained by reaction of a lysine containing peptide or protein with 4-nitrophenyl-\textit{t}-butylperoxycarbonate. The resulting lysine peroxycarbamates contain a labile O-O bond which is cleaved by the CID process. Loss of CO\textsubscript{2} then leads to the formation of the aminyl radical.
Subsequent local and remote hydrogen abstraction followed by β-scission give rise to peptide and side chain cleavage (Scheme II-2).

![Chemical Structure](image-url)

**Scheme II-2.** Modification of lysine model as a peroxycarbamate and fragmentation by CID.

Work by other researchers proceeded with the study of other free-radical fragmentation pathways initiated by dissociation of labile groups. Hodyss et al. reported on work in which the N-terminus of peptides was modified with a free radical initiator reagent containing an azo linkage. Dissociation of the azo bond after CID lead to a- and z-type ions similar to those formed by electron capture dissociation (Scheme II-3). Hao and coworkers studied the effect of NO modification of cysteine and tryptophan residues on peptide fragmentation. They found that dissociation of the S-NO and N-NO bonds resulted in thiol and...
aminyl radicals that promoted peptide cleavage products different from those found by CID of unmodified peptides (Scheme II-4).

Scheme II-3. CID fragmentation of peptides modified at the N-terminus with an azo radical initiator.

Scheme II-4. CID fragmentation of N- and S-nitroso peptides.

The development of low-energy based fragmentation techniques has the potential to be useful in assisting fragmentation of intact proteins by mass
spectrometry without the need of prior enzymatic digestion. Therefore such methods would be very valuable in the improvement of top-down proteomics approaches.

**Results and Discussion**

**Model Lysine Peroxycarbamate Studies**

Studies on the fragmentation of a model lysine peroxycarbamate by ESI-CID were carried out by Masterson and Yin. Modification the model lysine (N-α-acetyl lysine methyl ester) to the corresponding peroxycarbamate was carried out by reaction with the reagent II-2 (4-nitrophenyl-t-butyl-peroxycarbonate) under basic conditions (1:1 0.1M NH₄CO₃/acetonitrile, pH 8.6) as shown in Scheme II-5. The immediate formation of a yellow solution indicated the release of 4-nitrophenolate from the reaction.

![Scheme II-5. Peroxycarbamate modification of model lysine.](image)

To investigate the fragmentation of this model lysine peroxycarbamate, it was subjected to CID using an ESI-triple-quadrupole instrument. Fragmentation of the lithium adduct of II-3 at low CID energy off-sets (< 20 eV) produced mainly
a fragment which correspond to the loss of the \( t\)-BuOOC(O)- and is attributed to an aminyl radical (Scheme II-6). As the CID voltage was increased, a series of other smaller fragments were observed. The fragments II-5-8 are emphasized as they are the product of chain loss and backbone cleavage. A plot of the effect of CID energy off-set on product ion intensities is shown in Figure II-1. Along with the appearance of these smaller fragments, there was a decrease in the intensity of the aminyl radical ion II-4 with increased CID off-set energy, which suggested this to be an intermediate. None of these fragments were observed when the unmodified N-\( \alpha \)-acetyl-lysine methyl ester lithium adduct was subjected to CID.

![Scheme II-6. CID fragmentation of model lysine peroxycarbamate.](image)

33
Figure II-1. CID of the lithium adduct of II-3. Effect of increasing CID off-set voltage on intensity of product ions.

Mechanism of Fragmentation of Model Lysine Peroxycarbamate

Based on the preliminary studies by Masterson and Yin discussed in the previous section, ESI-CID experiments were carried out in order to confirm the aminyl radical as an intermediate in the side chain and backbone cleavage pathways. An in-source CID experiment was performed on the model lysine peroxycarbamate II-3 lithium adduct (m/z =325). With increased tube lens voltages (> 90 V) we were able to observe the aminyl radical II-4 in the full Q1 scan at m/z = 208, and selected this ion for CID. The resulting product ion spectrum showed ions like those observed in the CID spectrum of the parent
peroxycarbamate adduct II-3 at m/z = 325 (Figure II-2). As the CID off-set voltage was increased, there was an increase in the intensity of side chain and backbone cleavage products and a decrease in the intensity of the parent ion II-4 (Figure II-3). These results suggest that fragments II-5-8 are indeed products of the decomposition of II-4.

The formation of the aminyl radical II-4 is the result of homolytic cleavage of the O-O bond followed by decarboxylation, a reaction known to occur with diacyl peroxides and peroxycabamates.\(^{11-13}\) The aminyl radical intermediate may undergo several radical rearrangement pathways that lead to side chain and backbone fragmentation giving products II-5-8 as well as other other fragments (Scheme II-7). Hydrogen abstraction by the aminyl radical II-4 can occur on the lysine side chain or backbone. A 1,5 hydrogen shift, through a six-member ring transition state, followed by \(\beta\)-scission, leads to fragments II-7, and II-9. This process is well-known in the Hofmann-Löffler-Freytag reaction.\(^{14, 15}\) The 1,5 hydrogen shift may be followed by a 1,2 hydrogen shift and \(\beta\)-scission to give side chain fragment II-6. The aminyl radical II-4 may also undergo 1,4 hydrogen shift through a 5-member ring transition state to give II-8.
Figure II-2. A) ESI-CID spectrum of the lysine peroxycarbamate model II-3 (m/z = 325). B) ESI-CID spectrum of aminyl radical II-4 (m/z = 208) formed by in-source CID of II-3. All species are lithium adducts.
Figure II-3. A) In-source CID of the lithium adduct of II-3 (m/z = 325), effect of tube lens voltage on the intensity of product ions. B) CID of aminyl radical II-4 (m/z = 208) produced by in-source CID of lithiated II-3, effect of CID off-set voltage on product ion intensities.
Other fragmentation pathways can compete with the formation of the aminyl radical. For example, incomplete decomposition of the t-butyl peroxy group of parent II-3 and recombination of the aminyl radical with an alkoxyl radical, leads to product ions II-11-13. The formation of these fragments is supported by studies on a deuterated lysine peroxycarbamate model containing the moiety \((\text{CD}_3)_3\text{OOC(O)}^-\). It was found that fragment II-13 contained the \((\text{CD}_3)_3\text{O-}\) moiety. In addition, a deuterium was found in fragments II-11 and II-12, indicating involvement of the t-butyl group in the formation of these compounds.

\[
\begin{align*}
\text{II-13, m/z 281} & \quad \text{II-8, m/z 137} & \quad \text{II-4a, m/z 208} \\
\text{II-11, m/z 253} & \quad \text{II-3, m/z 325} & \quad \text{II-5, m/z 192} \\
\text{II-12, m/z 225} & \quad \text{II-10, m/z 179} & \quad \text{II-4c, m/z 208} \\
\end{align*}
\]

**Scheme II-7.** Proposed mechanism of fragmentation of model lysine peroxycarbamate II-3 (m/z = 325).
Peptide Peroxycarbamate Studies\textsuperscript{7,8}

Peptides containing one or more lysines were modified with the peroxycarbonate reagent II-2. Singly, doubly, and up to multiply modified peptide-peroxycarbamate adducts were observed by ESI-MS depending on the number of lysines present in the peptide. The peptide Ac-GSAKVSF protected at the N-terminus by an acetyl group, was modified at the free lysine side chain as a t-butyl peroxycarbamate. The ESI-MS spectrum of the lithium adduct showed peaks corresponding to unreacted peptide m/z = 859.3 (M + Li\textsuperscript{+}) and a lysine peroxycarbamate adduct m/z = 875.3 (M + 116 + Li\textsuperscript{+}). CID on this adduct gave the daughter ions shown in Scheme II-8. These ions were not observed in the CID spectrum of the unmodified peptide. These fragments are similar to the CID products derived from the model lysine peroxycarbamate; however, other backbone cleavage fragments of the type a, b, and c were also observed. Replacing Li\textsuperscript{+} with other counter ions such as Na\textsuperscript{+} and Ag\textsuperscript{+} gave similar results.

CID of the protonated adduct did not lead to the expected daughter ions, possibly due to the hydrophobicity of the peptide. However, more polar peptides (containing multiple lysines) showed typical fragmentation patterns with all counter ions including H\textsuperscript{+}. Multiply-modified peptides showed fragments in the CID spectra corresponding to the loss of each t-BuOOC(O)- group, as well as side chain and peptide cleavage fragments derived from each lysine aminyl radical.
Scheme II-8. CID fragments of modified Ac-GSAKVSF where lysine was converted to the respective t-butyl peroxyxycarbamate (m/z 875.3, M+116 + Li$^+$).

The radical process leading to peptide fragmentation in these studies can be correlated with the mechanism of peptide fragmentation by Electron Capture Dissociation (ECD). A comparison of the mechanism of radical-promoted peptide fragmentation by aminyl radicals produced from lysine peroxyxycarbamates and by ECD is shown in Scheme II-9. In ECD, electron capture by a positively charged amino acid leads to a backbone carbon-centered radical, which then results in side chain and backbone cleavage with c-, and z-ions being predominant. ECD has been successfully used in top-down FT-ICR proteomics studies of intact proteins due to the higher sequence coverage obtained compared to CID.
Scheme II-9. A) Proposed mechanism of radical-promoted peptide fragmentation by lysine aminyl radicals produced from lysine peroxyxycarbamates. B) Proposed mechanism of radical-promoted peptide fragmentation by ECD.
MALDI experiments on peptide peroxy carbamates also produced backbone fragmentation initiated by loss of the t-BuOOC(O)- group. The intact peptide peroxy carbamate adducts were not observed since they do not survive the laser desorption process. Instead, decomposition of the t-BuOOC(O)- group occurs at the source and only fragments derived from this process are detected. As an example, the MALDI spectrum of the modified peptide Ac-SYMEHFRWGKPVGKKR (a mixture of 1, 2, and 3 peroxy carbamate modifications) is shown in Figure II-4. Fragments were detected corresponding to backbone fragmentation at each of the lysine residues. In addition, the sequence of the fragments was confirmed by MS/MS analysis.

All of these peptide studies suggest that radical-promoted gas-phase fragmentation of proteins is feasible; however, protein studies face some limitations. First, the presence of multiple lysine modifications on large peptides or proteins results in complex CID spectra, which is a challenge in fragment characterization. In addition, when modified small proteins such as Ribonuclease A are subjected to ESI-CID or MALDI analysis, fragmentation is not clearly detected. In the case of ESI-MS analysis, because the ions are distributed along multiple charge states, only a small population of the protein is represented in a particular charge state. When the protein is modified with II-2, there is also a distribution of multiple lysine peroxy carbamate adducts in each charge state, which means an even smaller amount of protein is represented in a single ion. Therefore when selecting a single adduct ion for CID, the product ions will have very low intensities.
Figure II-4. A) MALDI spectrum of unmodified peptide Ac-SYMEHFRWGKPVGKKR. B) MALDI spectrum of modified peptide as a mixture of 1, 2, and 3 lysine peroxycarbamate adducts. $a_{11}$, Ac-SYMEHFRWGK*, $a_{15}$, Ac-SYMEHFRWGKPVGK*, $a_{16}$, Ac-SYMEHFRWGKPVGKK* (modifications represented by the asterisk). C) MS/MS spectrum of ion $a_{11}$ from spectrum A.

In the case of MALDI analysis of modified ribonuclease A, a broad peak near the $m/z$ of the unmodified protein was observed, possibly due to metastable ion decay. In order to improve ESI-CID spectra, it is necessary to concentrate the protein population to fewer ions. This can be achieved by reducing the extent of lysine modification, therefore obtaining fewer peroxycarbamate adducts. In addition, fewer adducts would improve the resolution of the MALDI spectra. The peroxycarbonate reagent II-2 is very reactive towards the amino group of the lysine side-chain, therefore when modifying peptides or proteins with this reagent most or all lysines will be converted to their peroxycarbamates. In order to control
the extent of lysine modification, less reactive reagents and milder reaction conditions are necessary.

**Synthesis of Other Peroxycarbonate Reagents**

In order to control the extent of lysine modification in peptides and proteins, we synthesized other less reactive peroxycarbonate reagents than II-2. These less reactive reagents differ from II-2 in that the nitro group is replaced by an alkoxy group (ethylene glycol). The reaction of the lysine ε-amino group and II-2 is a nucleophilic acyl substitution\(^{17}\), and the reactivity of II-2 will be greatly affected by the nature of the phenyl substituent group. In II-2, the nitro group has an electron withdrawing effect, making the acyl carbonyl electron deficient and therefore more susceptible to nucleophilic attack. The nitro group also stabilizes the phenolate leaving group through resonance, making this a better leaving group. The alkoxy group has an electron donating effect, making the acyl carbonyl electron rich, and the phenolate a worse leaving group, which slows down the rate of nucleophilic attack. The synthesis of these less reactive reagents is shown in Scheme II-10. The peroxycarbonates II-16 and II-17 were synthesized by the reaction of the phenols II-14 and II-15 with phosgene, followed by reaction with \(t\)-butyl hydroperoxide in the presence of pyridine. Synthesis of the phenol II-14 involved the alkylation of acetophenone with 1-bromo-2-methoxy-ethane in the presence of potassium carbonate, followed by a Baeyer-Villiger oxidation and acid hydolysis.\(^{18, 19}\) The synthesis was also attempted by alkylation of hydroquinone with 1-bromo-2-methoxy-ethane under
basic conditions;\textsuperscript{20, 21} however this method was not successful possibly due to quenching of the reaction by air oxidation of hydroquinone. Synthesis of the phenol II-15 was carried out by alkylation of resorcinol with 1-bromo-2-methoxy-ethane in the same manner as the alkylation of acetophenone; however, an excess of resorcinol was necessary in order to minimize dialkylation.

Scheme II-10. Synthesis of peroxycarbonate reagents.

**Peptide modifications with II-16 and II-17**

The reagents II-16 and II-17 were used to modify peptides with 2 or more lysines and study the extent of modification achieved with these compounds. The peptides used for these experiments were the 7-mer AcGKASVKF (777.4 Da) and the 17-mer SYSMEHFRWGKPGKVR (2093.4 Da). The peptides were modified in 1:1 NH\textsubscript{4}HCO\textsubscript{3}/acetonitrile and using different amounts of reagent.
mass spectrum (ESI-Q1 scan) was obtained for each reaction mixture and the relative peak intensities for the different peroxycarbamate adducts were plotted against the amount of reagent used (Figure II-5 and Figure II-6). With both reagents II-16 and II-17, it is feasible to obtain a significant amount of singly modified peptides while keeping multiply adducts at relatively low levels. However, reaction times with these reagents are longer than with II-2. The reagent II-16 required overnight incubation, and II-17 two-hour incubation, whereas reactions with II-2 only required a few minutes.

Figure II-7 illustrates the high reactivity of II-2 towards lysine modification compared to II-16 and II-17. When the 17-mer was modified with II-2, multiple lysine sites were modified with very low amounts reagent. In addition, all 4 amine sites on the peptide were modified, whereas with II-16, and II-17 only 3 modifications were detected.

In summary, these experiments show that reagents II-16 and II-17 are less reactive than II-2. Therefore we can limit the number of lysines modified by using these reagents. This may be useful in the characterization of CID spectra of modified peptides containing a large number of lysine residues. The extent of lysine modification can then be varied in a systematic way to determine patterns of fragmentation. Reagent II-17 may be more useful in these studies since II-16 requires overnight incubation. Nevertheless, different reagents may suit different systems being studied.
Figure II-5. Amount of reagent versus extent of lysine modification for peptide AcGKASVKF. A) Modification of peptide with II-16. B) Modification of peptide with II-17. Relative abundances include all charge states present in the spectrum for a given adduct. Molar equivalents of reagents were calculated based on moles of peptide used.
Figure II-6. Amount of reagent versus extent of lysine modification for peptide SYSMEFRTWGGKPVGKR. A) Modification with II-16. B) Modification with II-17. Relative abundances include all charge states present in the spectra for a given adduct. Molar equivalents of reagents were calculated based on moles of peptide used.
Figure II-7. Amount of 4-nitrophenyl-\(t\)-butyl peroxycarbonate (II-2) versus extent of modification for peptide SYSMEHFRWGKPGKR. Relative abundances include all charge states for a given peroxycarbamate adduct. Molar equivalents are calculated based on moles of peptide used.

**Lysine Peroxycarbamates of Ribonuclease A**

The protein ribonuclease A (13682 Da) was modified with reagent II-2. Modification of this protein resulted in a mixture of multiple peroxycarbamate adducts. The protein contains 11 lysines, and all lysine side chains and the N-terminus amine are prone to modification. When the modified protein mixture was analyzed by ESI-MS, up to 10 peroxycarbamate adducts were observed. CID experiments on these adducts did not lead to detectable product ions. The ribonuclease A peroxycarbamate adducts were also studied by MALDI. As mentioned above in the peptide peroxycarbamate studies, intact peroxycarbamate adducts are not detected in MALDI due to their decomposition during the laser desorption process. Instead, in the case of peptide peroxycarbamates, only the fragments derived from loss of the \(t\)-BuOOC(O)-
group are seen. In the case of ribonuclease A, the spectrum of the modified protein showed a broad peak around the \( m/z \) of the unmodified protein. However, no fragments derived from peroxycarbamate decomposition were observed. In order to simplify our analyses of the protein peroxycarbamates, and improve the MALDI spectrum resolution, we reduced the extent of lysine modification by adjusting the reaction conditions. We first attempted the modification of ribonuclease A with reagents II-16 and II-17; however, the longer reaction times necessary for significant modification to take place, lead to degradation of the protein. Therefore, we limited our studies to modifications with II-2, and reduced the amount of reagent used from 140 molar equivalents of reagent to 2-10 molar equivalents. Figure II-8 shows a comparison of the extent of modification of ribonuclease A with 140 and 10 molar equivalents of reagent. The portion of the ESI-MS spectrum containing the peroxycarbamate adducts in the \( 10^+ \) charge state is shown. From these spectra it is evident that we are able to obtain mostly singly and doubly modified protein by lowering the amount of reagent used. We note however, that the lower extent of lysine modification is accompanied by a significant amount of unreacted protein.
To further simplify MS studies of these protein adducts, we isolated the protein containing 1 and 2 peroxycarbamate adducts by cation-exchange chromatography. The cation-exchange chromatogram of the protein peroxycarbamate mixture is shown in Figure II-9. Several peaks were observed for singly and doubly modified protein, which is possibly due to modification at different lysine positions. Fractions were
collected from this run and analyzed by ESI-LC-MS. Figure II-10 shows the LC and MS spectrum for a single fraction containing singly modified protein.

![Figure II-9. Cation-exchange chromatogram of modified Ribonuclease A. Fractions were collected every minute for MS analysis.](image)

ESI-CID and MALDI analyses of these isolated protein peroxycarbamate fractions however did not lead to protein fragmentation. In ESI-triple quadrupole and ion trap analyses, although we were successful in isolating ribonuclease A containing only 1 or 2 peroxycarbamate modifications, the fractions obtained from cation-exchange chromatography may still contain a mixture of regioisomers. These regioisomers arise from the adduction of 1 peroxycarbamate at different lysine positions on the protein. CID of this mixture of regioisomers would result in a different set of fragments ions for each regioisomer. These fragments would also be multiply charged, therefore a single fragment ion in the CID spectrum
would represent a very small percentage of the total ion population, thus remaining undetectable in the experiment. A schematic of this process is shown in Figure II-11.

Figure II-10. LC-MS analysis of singly modified ribonuclease A obtained from cation-exchange separation. A) LC chromatogram showing singly modified ribonuclease at 19.24 min. B) MS spectrum under the LC peak at 19.24 min. Numbers in red represent different charge states.

In the case of MALDI we know from peptide peroxycarbamate studies that fragmentation of the lysine peroxycarbamates occurs at the source. When singly modified or doubly modified ribonuclease A is subjected to MALDI, the spectrum shows a single intense peak with m/z equal to that of the unmodified protein. It is possible that this peak corresponds to a mixture of lysine aminyl radicals corresponding to different regioisomers. However, no further fragmentation is observed.
**Figure II-11.** Illustration of the formation of regioisomers of ribonuclease A due to 1 peroxycarbamate modification. Each regioisomer gives rise to different fragments and each fragment may be multiply charged.

**Conclusion**

The modification of lysine as a peroxycarbamate can induce site-specific peptide cleavage in the gas-phase by ESI-CID or MALDI. This fragmentation occurs by a radical mechanism initiated by homolysis of the O-O bond, followed by decarboxylation. The aminyl radical produced in this process undergoes radical rearrangements which lead to backbone and side chain fragments. These backbone fragments are of the type a, b, c, and z, and may be useful in peptide sequencing. Although our preliminary studies on protein
peroxycarbamates have not yet been successful, further studies utilizing more advanced mass spectrometers such as an FT-ICR may offer more insight in the utility of this strategy for top-down approaches.

**Experimental Procedures**

**General**

The model lysine (N-$\alpha$-acetyl lysine methyl ester) was purchased from Aldrich Chemical Co. (Milwaukee, WI, USA) in its hydrochloride salt form. Peptides were purchased from American Peptide Co. (Sunnyvale, CA, USA). Ribonuclease A was obtained from Sigma (Milwaukee, WI, USA). All other reagents were purchased from Aldrich Chemical Co. Solvents were HPLC grade, purchased from Fisher Scientific (Pittsburgh PA, USA). $t$-Butyl-4-nitrophenyl-peroxycarbonate was synthesized according to previous work in our laboratory. The NH$_4$CO$_3$ buffer system used for lysine peroxycarbamate modifications consisted of a 1:1 mixture of 0.1 M NH$_4$CO$_3$ (pH 8.6) and acetonitrile. For cation-exchange chromatography, a carboxymethyl column (Hydrocell CM 1500, 150 x 4.6 mm) was used, purchased from BioChrom Labs, Inc. (Terre Haute, IN, USA). Cation-exchange analysis was carried out on a Waters HPLC system (Milford, MA, USA). LC-MS analysis were done using a C18 column (10 cm x 1 mm x 5um), from Grace Vydac (Hesperia, CA, USA). NMR spectra were acquired on a Bruker Advance 300 MHz instrument. HRMS was carried out at Ohio State University (Columbus, OH, USA).
Peroxycarbamate Modifications

Modification of the model lysine II-1 was carried out by combining 5 µl of the amino acid stock solution (0.1M in 1:1 acetonitrile/water) with 20 µl of NH\textsubscript{4}CO\textsubscript{3} buffer system, and 20 µl of II-2 (0.1 M in acetonitrile). The final concentrations for the model lysine and reagent II-2 were 0.01 M and 0.04 M respectively. Reactions were incubated at room temperature for 15 minutes. Modification of peptides was carried out in a similar manner by combining 10 µl of peptide stock solution in water (1 mg/ml), 20 µl of NH\textsubscript{4}CO\textsubscript{3} buffer system, and 10 µl of reagent II-2 solution (0.1 M in acetonitrile). The molar ratio of reagent to peptide was approximately 100:1. For reagent reactivity studies with II-2, II-16, and II-17, the amount of reagent was adjusted to ratios from 1:1 to 30:1 reagent to peptide. For II-16, the modification reactions were allowed to incubate overnight at room temperature. For II-17, two-hour incubation times were sufficient to obtain enough modified peptide for MS analysis.

Ribonuclease A Modifications

Modification of ribonuclease A to obtain multiple peroxycarbamate adducts was carried out by combining 100 µl of the protein stock solution (10µg/ml in water) with 100 µl of NH\textsubscript{4}CO\textsubscript{3} buffer system and 100 µl of reagent II-2 stock solution (0.1 M in acetonitrile). The molar ratio of reagent to protein was approximately 140:1. To reduce the number of peroxycarbamate modification, the amount of reagent was reduced to a ratio of 10:1 reagent to protein. Cation-exchange chromatography separation of the modified protein adducts was done.
according to an altered literature procedure. For the separation, an HPLC method was used with a commercially available cation-exchange column (see general section). The elution solvents consisted of A: 95% water, 5% methanol; B: 95% 0.16 M NH₄PO₄, 5% methanol. The elution gradient was 0% to 20% solvent B into solvent A in 60 minutes, at a flow rate of 1ml/min. Fractions were collected using an automated Waters (Milford, MA, USA) fraction collector every 1 min. with a 15 min. start delay. The fractions were reduced to 200 μl under vacuum for LC/MS analysis.

**Synthesis of Peroxycarbonate Reagents**

**Synthesis of 4-nitrophenyl-t-butylperoxycarbonate (II-2)**

A 50 ml round bottom flask was charged with 20 ml of methylene chloride and 2.0g (9.9 mmoles) of 4-nitrophenyl chloroformate. The solution was cooled to 0°C and 1 eq. of t-butyl hydroperoxide was added in decane in one portion. The resulting solution was stirred and 1 eq. of anhydrous pyridine was added drop wise over 10 minutes. The solution was allowed to warm to room temperature and stirred for an additional hour. The resulting solution was then diluted with 50 ml of methylene chloride and washed 2X with 1% H₂SO₄, 1X with brine, and dried over MgSO₄. The solution was then concentrated and purified by silica gel column chromatography with 1:1 hexane/ether to provide 1.0g of pure product as a white solid (3.9 mmoles, 39%). ¹H-NMR (300 MHz, CDCl₃): 1.40 (s, 9H), 7.41 (d, 2H, j=9.3 Hz), 8.29 (d, 2H, j=9.3 Hz). 13C-NMR (75 MHz, CDCl₃): 25.9, 85.3,
Synthesis of peroxycarbonate reagents II-16 and II-17

4-(2-methoxy-ethoxy)-acetophenone \(^{19}\)

The compounds 4-Hydroxy-acetophenone (4.00 g, 29.4 mmol), 1-bromo-2-methoxy-ethane (4.89 g, 35.1 mmol), and potassium carbonate (5.80 g, 35.1 mmol), were added to 50 ml of anhydrous DMF. The reaction was allowed to stir overnight at the solvent reflux temperature in a nitrogen atmosphere. The reaction mixture was cooled to room temperature, water was added (50 ml) and the solution was extracted with ether (3 x 50 ml). The combined ether extract was washed with water (2 x 10 ml) and dried over MgSO\(_4\). Evaporation of the solvent left a caramel color paste (4.14 g, 73 % yield). \(^1\)H NMR (300 MHz, CDCl\(_3\)): \(\delta\) 2.53 (s, 3H), 3.44 (s, 3H), 3.75-3.78 (m, 2H), 4.16-4.19 (m, 2H), 6.90-6.98 (m, 2H), 7.88-7.94 (m, 2H). \(^{13}\)C NMR (300 MHz, CDCl\(_3\)): \(\delta\) 26.27, 59.19, 67.38, 70.70, 114.16, 130.41, 130.47, 162.60, 196.69. HRMS (M+Na\(^+\)) 217.083513 calculated, 217.08371 observed.

4-(2-methoxy-ethoxy)-acetoxybenzoate \(^{18}\)

To a solution of 4-(2-methoxy-ethoxy)-acetophenone (3.68g, 18.6 mmol) in chloroform (60 ml), was added \(m\)-chloroperbenzoic acid (6.4g, 37.2 mmol) and \(p\)-toluenesulfonic acid (60 mg, 1.6 %). The reaction was allowed to stir overnight at room temperature. The mixture was filtered and the mother liquor was washed with 0.1 M pH 7.4 phosphate buffer (4 x 60 ml). The organic layer was washed
with water (2 x 20 ml) and dried over MgSO₄. The solvent was evaporated leaving a yellow-orange paste. The crude product was redissolved in a minimum amount of chloroform (10 ml) and passed through a silica plug (1:1 ether/hexane). Removal of solvent resulted in a cream color solid (3.26g, 83 % yield).

\[ ^1\text{H NMR (300 MHz, CDCl}_3\text{)}: \delta 2.27 (s, 3H), 3.43 (s, 3H), 3.72-3.75 (m, 2H), 4.08-4.28 (m, 2H), 6.89-7.0 (m, 4H). \]

\[ ^{13}\text{C NMR (300 MHz, CDCl}_3\text{)}: \delta 21.00, 59.14, 67.61, 70.92, 115.15, 122.22, 144.29, 156.38, 169.79. \]

**4-(2-methoxy-ethoxy)-phenol (II-14)**

A solution of 4-(2-methoxy-ethoxy)-acetoxy-benzoate (2.75g, 13.0 mmol) in acetone (80 ml) and 4 M HCl (15 ml) was heated to solvent reflux temperature (60 °C) overnight under nitrogen. The acetone was removed in vacuo and the aqueous residue was extracted with ether (3 x 50 ml). The combined ether extract was washed with water (2 x 10 ml) and dried over MgSO₄. The ether was removed in vacuo leaving an off-white solid (2.08 g, 90 % yield).

\[ ^1\text{H NMR (300 MHz, CDCl}_3\text{)}: \delta 3.47 (s, 3H), 3.72-3.75 (m, 2H), 4.04-4.07 (m, 2H), 6.72-6.81 (m, 4H). \]

\[ ^{13}\text{C NMR (300 MHz, CDCl}_3\text{)}: \delta 59.16, 67.89, 71.16, 115.76, 115.99, 149.80, 152.78. \]

HRMS (M+Na⁺) 233.078428 calculated, 233.07925 observed.

**3-(2-methoxy-ethoxy)-phenol (II-15)**

Resorcinol (2.4g, 21.6 mmol) was added to 20 ml of anhydrous DMF. To this mixture, K₂CO₃ (1.8 g, 10.8 mmol) was added, followed by slow addition of 1-bromo-2-methoxy-ethane (1 ml, 10.8 mmol). The reaction mixture was allowed to stir at the solvent reflux temperature, overnight, under nitrogen. The mixture was cooled to room temperature, diluted with 20 ml of water and extracted with ether.
The combined ether extract was washed with water (5 ml) and dried over MgSO₄. The solvent was evaporated in vacuo and the crude liquid was purified by silica-gel column chromatography (10 % methanol in chloroform) yielding a light yellow solid (1g, 54 % yield). \(^1\)H NMR (300 MHz, CDCl₃): \(\delta\) 3.44 (s, 3H), 3.73-77 (m, 2H), 4.07-4.10 (m, 2H), 6.51-6.54 (m, 3H), 7.12-7.19 (m, 1H). \(^13\)C NMR (300 MHz, CDCl₃): \(\delta\) 59.13, 67.13, 70.92, 101.64, 107.04, 129.72, 159.86.

**4-(2-methoxy-ethoxy)-phenyl-t-butylperoxycarbonate (II-16) and 3-(2-methoxy-ethoxy)-phenyl-t-butylperoxycarbonate (II-17)**

Diisoprophylethylamine (2.20 ml, 12.4 mmol) was added slowly to a solution of the phenol (II-14, 2.08g, 12.4 mmol, or II-15, 2.7g, 16.0 mmol) in methylene chloride. The resulting solution was added very slowly to a cold (0 °C) 1.93 M phosgene solution in toluene (13 ml, 25 mmol). The reaction mixture was allowed to reach room temperature and stirred overnight under nitrogen. The excess phosgene was removed by purging with a moderate stream of dry nitrogen and adding solvent as needed for 1 ½ hours. The solvent was evaporated in vacuo and the crude product was used without purification. The crude chloroformate was redissolved in 40 ml of methylene chloride and cooled to 0 °C. A 5 to 6 M solution of t-butyl-hydroperoxide in dodecane (2.4 ml, 12 mmol) was added, followed by drop-wise addition of pyridine (1 ml, 12.4 mmol). The reaction mixture was allowed to stir overnight at room temperature, under nitrogen. The solvent was evaporated and the product purified by silica gel column chromatography (30 % ethyl acetate in hexane) as a light yellow solid II-15, 65 % yield) and a clear liquid (II-16, 64 % yield). \(^1\)H NMR II-16 (300 MHz,
CDCl₃): δ 1.37 (s, 9H), 3.45 (s, 3H), 3.71-3.74 (m, 2H), 4.07-4.10 (m, 2H), 6.89-6.93 (m, 2H), 7.06-7.11 (m, 2H). **II-17** δ1.36 (s, 9H), 3.43 (s, 3H), 3.72-3.75 (m, 2H), 4.08-4.12 (m, 2H), 6.78-6.85 (m, 3H), 7.27-7.30 (m, 1H). **II-17** δ 25.88, 59.12, 67.62, 70.85, 84.59, 115.23, 121.52, 144.44, 153.30, 156.72. **II-17** δ 25.74, 59.10, 67.40, 70.66, 84.60, 107.37, 112.66, 112.94, 129.83, 151.42, 152.72, 159.55. For **II-16** HRMS (M+Na⁺) 307.11520 calculated, 307.11618 observed.

**Mass Spectrometry**

Mass spectrometry analyses of the model lysine peroxycarbamate and peptides were performed on a TheromoFinnigan (San Jose, CA, USA) TSQ 7000 triple-quadrupole instrument, equipped with an ESI source. The capillary temperature was kept at 200 °C, and the voltage was 20V. The electrospray needle voltage was 4.5 kV, and the tube-lens voltage was maintained between 70 and 100 V. The sheath and auxiliary gases (N₂) were adjusted to maximize signal. The off-set voltage in CID experiments was generally between 20 and 30 eV. Samples were introduced into the source for direct infusion experiments either as lithiated (LiCl) or protonated solutions (acetic or formic acid). For in-source CID studies of the model lysine peroxycarbamate, the tube-lens voltage was varied from 20 to 120 V. For CID mechanism studies, the off-set voltage was varied from 5 to 35 eV. LC/MS analysis of ribonuclease A peroxycarbamate adducts was carried out on a ThermoFinnigan (San Jose, CA, USA) LCQ Deca XP ion-trap instrument, equipped with an ESI source, and a HP Agilent (Palo
Alto, CA, USA) HPLC system. Instrument parameters were tuned using angiotensin II peptide as a standard to maximize signal. MALDI analyses were carried out for ribonuclease A on an Applied Biosystems (Foster City, CA, USA) Voyager DE STR instrument. For peptides, MALDI analyses were done with an Applied Biosystems (Foster City, CA, USA) 4700 Proteomics instrument, equipped with a TOF-TOF mass analyzer. Matrices consisted of α-cyano-4-hydroxycinnamic acid for peptides and 3,5 Dimethoxy-4-hydroxycinnamic acid for ribonuclease A (10 mg/ml in 4:6 0.1 % TFA/acetonitrile). Samples were spotted on a gold target plate (0.5 μl sample and 0.5 μl matrix solution).

References


CHAPTER III

N-TERMINAL AMINO ACID SIDE CHAIN CLEAVAGE OF CHEMICALLY MODIFIED PEPTIDES IN THE GAS PHASE

Introduction

The rapid development and improvement of mass spectrometry, along with the availability of large genome databases, have made possible high-throughput characterization of complex protein mixtures. Nevertheless, the presence of unexpected post-translational modifications often leads to unsuccessful or incorrect identifications in database searches. In addition, the study of proteins derived from organisms whose genomes are not complete is limited. In such cases, more labor-intensive data analysis such as de novo sequencing must be used, where the peptide sequence is reconstructed from the MS/MS fragmentation data with the help of sequencing algorithms. Due to the complexity of peptide MS/MS spectra, a number of methods have been developed to facilitate interpretation. N-terminal and C-terminal charged derivatives were developed in the late 1980’s to be used with Fast Atom Bombardment (FAB) analyses in order to simplify fragmentation spectra produced from high energy collision induced dissociation (CID). In addition, these charged derivatives were useful in increasing signal intensity of peptide fragment ions in such experiments. Keough and coworkers derivatized the N-termini of peptides with a sulfonic acid group. Fragmentation of these peptides by CID and Postsource Decay (PSD) Matrix-Assisted Laser Desorption Ionization (MALDI), yielded selectively the y-ion series. Work by Geskell et al. investigated
N-terminus derivatization using phenylthiocarbamoyl derivatives similar to those employed in Edman degradation.\textsuperscript{10} CID experiments in this case produced predominantly $b_1$ and $y_{n-1}$ ions. James and coworkers carried out the modification of N-termini of peptides with nicotinyl-N-hydroxysuccinimide which resulted in suppression of the $y$-ion series in MALDI tandem time of flight (MALDI-TOF-TOF) analyses.\textsuperscript{11} More recently, Reilly and co-workers investigated the effect of acetaminidation of N-termini on fragmentation patterns in CID spectra.\textsuperscript{12} In most cases studied, there was an enhancement of $y_{n-1}$ ions while diminishing other fragmentations, and in some cases contiguous series of $b$-$\text{NH}_3$ ions were observed. Another approach to improving de novo methods involves stable isotopic labeling. Esterification of C-termini with deuterium labeled methanol results in a mass shift of the $y$-ions with respect to the unlabeled peptide ester thus identifying the $y$-ion series.\textsuperscript{13} In a similar manner, $O^{18}$ labeling of C-termini provides information about the $y$-ion series.\textsuperscript{14} Cannon and coworkers presented a computational method to obtain information from the natural isotopic distributions of peptide fragments in MS/MS spectra, in order to facilitate identification of $b$- and $y$-ions without prior derivatization.\textsuperscript{15}

In the work described in this chapter, we have utilized the peroxycarbamate chemistry discussed in chapter II to develop a mass spectrometry based N-terminus identification technique. The conversion of the N-terminal $\alpha$-amino group of amino acids and peptides to a peroxycarbamate adduct, leads to CID fragmentation involving cleavage of the N-terminal amino acid side chain (Scheme III-1). This fragmentation provides information about the
nature of the N-terminal residue and may also serve as a labeling technique for elucidating the b-ion series in de novo analyses.

\[
\text{Scheme III-1. Fragmentation of N-terminal peroxycarbamates by CID.}
\]

Results and Discussion

The chemistry of peroxycarbamates is based on previous work from our laboratory on the derivatization of lysine residues as discussed in chapter II.\textsuperscript{16, 17} The mechanism for the fragmentation of lysine peroxycarbamates is thought to be initiated by the dissociation of the t-butyl peroxy group to form an aminyl radical intermediate, followed by remote hydrogen abstraction and $\beta$-fragmentation of the resulting carbon radicals (Scheme III-2). Introduction of the labile peroxide bond provides a low-energy fragmentation pathway to free radicals that ultimately affect the peptide fragmentation patterns. More recently, Beauchamp and coworkers\textsuperscript{18} made use of reagents that modified peptide N-termini with an azo functional group. Decomposition of the labile azo linkage led to a protocol they dubbed “Free Radical Initiated Peptide Sequencing (FRIPS)” where a- and z-type fragments are obtained by CID instead of the typical b- and y-type ions (Scheme III-3).
Scheme III-2. Modification of lysine model as a peroxycarbamate and fragmentation by CID.

Scheme III-3. CID fragmentation of peptides modified at the N-terminus with an azo radical initiator.
Based on Lysine peroxycarbamate chemistry, we investigated the effect of modification of the N-terminus as a peroxycarbamate on fragmentation patterns upon CID. Amino acids and peptides modified at the N-terminus as peroxycarbamates, and subjected to CID, lead to the cleavage of the N-terminal amino acid side chains (Scheme III-4). This fragmentation allowed the identification of the N-terminal residue and led the way to a mass spectrometry based N-terminus identification technique.¹⁹

Scheme III-4. Peroxycarbamate modification of N-termini and CID fragmentation.

**N-Terminal amino acid side-chain loss**

A series of amino acids and peptides were modified at the N-terminus with reagents III-1a-1c to form N-terminal peroxycarbamates. The amino acids were ester derivatives in the form NH₂CH(R)CO₂R' (M) where R is the amino acid side chain, and R' is either a methyl, ethyl or a t-butyl group and were converted to
the peroxycarbamates $\text{C}_4\text{H}_9\text{OOC(O)NHCH(R)CO}_2\text{R}^\prime$ ($\text{C}_4\text{H}_9\text{OOC(O)M}$). ESI-CID analysis of the lithium adducts of these derivatives lead to fragmentation involving the loss of the $t$-butyl peroxide group ($\text{C}_4\text{H}_8\text{OOC(O)}$, $m/z = 116$), to give a product ion with $m/z$ equal to the underivatized amino acid ester $M$ [$\text{C}_4\text{H}_9\text{OOC(O)M} – 116$], and another product ion which corresponds to the imine NH=CHCOOR’, or [M - (R+1)] with $m/z$ equal to the loss of the $t$-butyl peroxide, 1 hydrogen, and the amino acid side chain, $[\text{C}_4\text{H}_9\text{OOC(O)M} – 117 – R]$ (Scheme III-5). A sample CID spectrum of modified valine methyl ester is shown in Figure III-1. Lithium was used as the counter ion for CID of the modified amino acids due to the unavailability of the N-terminus for protonation. Cleavage of the amino acid side chain was not observed in the CID spectra of unmodified amino acids. In most cases $[\text{C}_4\text{H}_9\text{OOC(O)M} – 116]$ was the predominant fragment for the modified compounds, this suggests that gas phase decomposition of the $t$-butylperoxy group is an important process; however, the imine fragment was present in most spectra. A list of modified amino acids and corresponding CID fragmentations are presented in Table III-1.
Scheme III-5. CID fragmentation of N-terminal peroxyacarbamate amino acid ester derivatives.

Figure III-1. CID spectrum of N-terminal peroxy carbamate valine methyl ester.
Table III-1. N-terminal peroxycarbamates amino acid adducts and their CID fragmentation.

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>[M + Li]^+</th>
<th>[‘BuOOC(O)M + Li]^+</th>
<th>Neutral Loss</th>
<th>R</th>
<th>R m/z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly</td>
<td>96.1</td>
<td>212.0</td>
<td>117.0</td>
<td>H</td>
<td>1</td>
</tr>
<tr>
<td>Ala</td>
<td>110.1</td>
<td>226.1</td>
<td>132.0</td>
<td>CH₃</td>
<td>15</td>
</tr>
<tr>
<td>Val</td>
<td>138.0</td>
<td>254.2</td>
<td>160.1</td>
<td>CH(CH₃)₂</td>
<td>43</td>
</tr>
<tr>
<td>Leu</td>
<td>152.1</td>
<td>268.2</td>
<td>174.1</td>
<td>CH₂CH(CH₃)₂</td>
<td>57</td>
</tr>
<tr>
<td>Ile</td>
<td>152.2</td>
<td>268.2</td>
<td>174.2</td>
<td>CH(CH₃)CH₂CH₃</td>
<td>57</td>
</tr>
<tr>
<td>Glu</td>
<td>210.0</td>
<td>326.0</td>
<td>204.1</td>
<td>(CH₂)₂(O)(CO)OC₂H₅</td>
<td>101</td>
</tr>
<tr>
<td>Asp</td>
<td>168.0</td>
<td>283.9</td>
<td>190.0</td>
<td>CH₂(CO)OCH₃</td>
<td>73</td>
</tr>
<tr>
<td>Asn</td>
<td>194.4</td>
<td>311.1</td>
<td>----</td>
<td>CH₂(CO)NH₂</td>
<td>58</td>
</tr>
<tr>
<td>His</td>
<td>176.1</td>
<td>291.9</td>
<td>198.0</td>
<td>CH₂C₆H₅N₂</td>
<td>81</td>
</tr>
<tr>
<td>Met</td>
<td>170.1</td>
<td>285.9</td>
<td>192.1</td>
<td>(CH₂)₂SCH₃</td>
<td>75</td>
</tr>
<tr>
<td>Phe</td>
<td>186.0</td>
<td>302.0</td>
<td>208.1</td>
<td>CH₂C₆H₅</td>
<td>91</td>
</tr>
<tr>
<td>Ser</td>
<td>126.0</td>
<td>242.0</td>
<td>148.1</td>
<td>CH₂OH</td>
<td>31</td>
</tr>
<tr>
<td>Tyr</td>
<td>202.1</td>
<td>318.0</td>
<td>224.1</td>
<td>CH₂C₆H₅O</td>
<td>107</td>
</tr>
<tr>
<td>Trp</td>
<td>225.0</td>
<td>341.0</td>
<td>247.0</td>
<td>CH₂C₆H₅N</td>
<td>130</td>
</tr>
<tr>
<td>Thr</td>
<td>140.2</td>
<td>256.1</td>
<td>162.1</td>
<td>CH(OH)CH₃</td>
<td>46</td>
</tr>
<tr>
<td>Gln</td>
<td>208.3</td>
<td>325.1</td>
<td>189.0</td>
<td>(CH₂)₂(O)(CO)NH₂</td>
<td>72</td>
</tr>
<tr>
<td>Lys</td>
<td>181.1</td>
<td>297.2</td>
<td>189.0</td>
<td>(CH₂)₂NH₂</td>
<td>72</td>
</tr>
<tr>
<td>Pro</td>
<td>137.0</td>
<td>251.9</td>
<td>----</td>
<td>C₃H₇</td>
<td>43</td>
</tr>
<tr>
<td>Cys</td>
<td>------</td>
<td>305.0</td>
<td>217</td>
<td>(CH₂)₂NHC(NH₂)₂</td>
<td>100</td>
</tr>
</tbody>
</table>

---

a Amino acids are methyl esters derivatives (M = NH₂CH(R)CO₂CH₃) with the exception of Asp and Glu which are the dimethyl ester and diethyl ester derivatives respectively, Lys is the ethyl ester and Gln and Asn are the t-butyl ester derivatives. 

b Neutral loss corresponds to 117 + R where 117 is the m/z of the C₄H₉OOC(O) group + 1 hydrogen and R is the amino acid side chain.

c Cleavage of the side chain was not observed for Pro, only loss of the C₄H₉OOC(O) group was seen. 

d The MS spectrum of modified Cys was complex and did not contain a peak corresponding to [C₄H₉OOC(O)M], suggesting decomposition of the amino acid may occur during the modification reaction. Modified Asn was also an exception where no side-chain loss was observed.

An imine ion was not observed for modified Cys and Pro. In the case of Cys the peroxycarbamate adduct was not observed in the Q1 spectrum. This adduct may not form or may not be stable due to the reactive thiol group nearby.

The observation that Met does not present this problem supports the hypothesis of the interfering effect of the free thiol group. In most cases in “real” proteomics analyses, Cys residues will be alkylated (e.g., with iodoacetamide) to produce S-carboxamidomethyl derivatives. These, similarly to methionine, will likely be better behaved than a free Cys in the peroxycarbamate reaction. In the case of
modified Pro, the N-terminus is part of a five-member ring structure, which may prevent side chain cleavage. Modified Gly did not appear to yield N-terminal loss of the side chain hydrogen by CID, but instead, a fragment consistent with the \( m/z \) of the N-terminal aminyl radical was observed \([C_4H_9OOC(O)M - 117]\). Asn also does not undergo side chain cleavage for reasons not well understood; however, this was not the case for peptides containing Asn at the N-terminus, as will be discussed in the next section.

**Peptide modifications**

In order to selectively modify peptides at the N-terminus and not at lysine, the pH and reagent were varied (a list of modified peptides and their corresponding fragmentation is shown in Table III-1) which was modified under different pH conditions, ranging from 5.2 to 8.6. In this way the pH optimum for N-terminal modification over \( \varepsilon \)-amino modification of Lys could be established. Previous studies have shown that selectivity for reactions at the N-terminal \( \alpha \)-amino group of peptides can be achieved by pH control.\(^{20-23}\) Lowering the pH below the pKa of the terminal \( \alpha \)-amino group (approximate pKa = 8) increases the selectivity for reaction at the N-terminus with respect to the \( \varepsilon \)-amino group of lysine (approximate pKa = 10.7). In addition to pH control, a less reactive reagent than **III-1a (III-1b)** offers higher selectivity for reaction at the N-terminus (Scheme III-1).
Reagent **III-1b** is less reactive towards nucleophilic attack by the amine nitrogen than **III-1a**, therefore modification occurs at a slower rate with this reagent. The reaction can be accelerated by increasing the temperature to 37 °C, conditions under which reagent **III-1a** undergoes decomposition. LC/MS analysis of the modified peptide revealed that under basic conditions (pH 7.5 to 8.5) and with reagent **III-1a**, modification occurs at both amino group sites, but with greater efficiency on the lysine side chain. Slightly acidic conditions (pH 5.5) and reagent **III-1b** gave better selectivity for the N-terminus. Figure III-3 shows the LC/MS trace of the modified peptide for reactions carried out at pH 8.6 and 5.5. Peak 1 in the chromatogram corresponds to \(m/z = 758.2\) ([M+3H] \(^{3+}\)) where M = YRVRFLAKENVQDAEDNC. Peaks 2 and 3 indicate a single modification with \(m/z = 796.8\) ([C\(_4\)H\(_9\)OOC(O)M + 3H] \(^{3+}\)). Peak 2 corresponds to modification at the N-terminal group, whereas peak 3 corresponds to \(\varepsilon\)-amino modification of lysine. This was confirmed by the CID fragmentation of peak 2, also shown in Figure III-3, where the ion with \(m/z = 722.2\) ([C\(_4\)H\(_9\)OOC(O)M − 117 − 107 + 3H] \(^{3+}\)) indicates the neutral loss of C\(_4\)H\(_9\)OOC(O) and the N-terminal tyrosine side chain (-CH\(_2\)C\(_6\)H\(_4\)OH, m/z = 107). This fragment was absent in the CID spectrum of
peak 3. At lower pH and with the use of reagent III-1b, peak 3 is diminished, which implies a preference for N-terminal modification under these conditions.

Several peptides containing different N-terminal amino acids were modified using reagent III-1b at pH 5.5 and the N-termini were identified in CID experiments (Table III-2). Thus, CID spectra of peptide peroxyxycarbamates bearing alanines at the N-terminus give product ions that indicate the loss of the side chain -CH$_3$ ([$C_4H_9OOC(O)M - 117$-$15$]), whereas terminal valines show a similar loss of the side chain -CH$_3$CHCH$_3$ and so on. In cases where Cys and Pro were the terminal residues (entries 18 and 20 in Table III-2), the CID spectra did not lead to an imine fragment. As seen for the model amino acids, N-terminal Cys did not give a peroxyxycarbamate adduct, due to interference by the free thiol group, and the CID spectrum of N-terminal proline-peroxycarbamate only showed an $[C_4H_9OOC(O)M - 116]$ ion and no side chain cleavage. Nevertheless, the presence of these amino acids anywhere else in the sequence did not affect the fragmentation chemistry at the N-terminus of modified peptides. CID of a modified peptide bearing glycine at the N-terminus (entry 9 in Table III-2) yielded similar fragmentation to the CID of the modified model amino acid, where the N-terminal aminyl radical $[C_4H_9OOC(O)M - 117]$ was observed instead of the expected imine fragment $[C_4H_9OOC(O)M - 117 - 1]$. Results from the CID of a modified peptide bearing Asn at the N-terminus (entry 5 in Table III-2) indicated loss of the Asn side-chain, which is in contrast to results obtained with the modified Asn amino acid derivative.
Figure III-3. LC/MS of modified peptide YRVRFLAKENVTQDAEDNC. (A) Selected Ion Monitoring (SIM) $m/z = 758.2$ ([M+3H]$^{+3}$) where M is the unmodified peptide YRVRFLAKENVTQDAEDNC. (B) SIM $m/z = 796.8$ ([C$_4$H$_9$OOC(O)M + 3H]$^{+3}$) where M was modified with reagent III-1b at pH 5.5. Peaks 2 and 3 correspond to N-terminus and lysine modifications, respectively. (C) SIM $m/z = 796.8$ ([C$_4$H$_9$OOC(O)M + 3H]$^{+3}$) where M was modified with reagent 1a at pH 8.6. (D) CID spectrum of peak 2, parent $m/z = 796.8$ ([C$_4$H$_9$OOC(O)M +3H]$^{+3}$). The product ion $m/z = 722.2$ corresponds to the loss of C$_4$H$_9$OOC(O) and the N-terminal Tyrosine side chain, CH$_2$C$_6$H$_4$OH ([C$_4$H$_9$OOC(O)M – 117 – 107 + 3H]$^{+3}$).
<table>
<thead>
<tr>
<th>Entry</th>
<th>Sequence</th>
<th>( m/z )</th>
<th>([C_4H_9OOC(O)M]^+ )</th>
<th>Imine Fragment</th>
<th>( R )</th>
<th>( R )</th>
<th>( m/z )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>VYIHPF</td>
<td>775.4(1H⁻)</td>
<td>891.5(1H⁺)</td>
<td>731.1(1H⁺)</td>
<td>CH(CH₃)₂</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>HPKRPWIL</td>
<td>523.4(2H⁺)</td>
<td>581.8(2H⁺)</td>
<td>483.1(2H⁺)</td>
<td>CH₂C₆H₅N₂</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>MEHFRW</td>
<td>962.0(1H⁺)</td>
<td>1078.0(1H⁺)</td>
<td>887.2(1H⁺)</td>
<td>(CH₂)₃SCH₃</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>SRVSRKRR</td>
<td>335.2(3H⁺)</td>
<td>373.8(3H⁺)</td>
<td>324.5(3H⁺)</td>
<td>CH₂OH</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>NRYYVHPF</td>
<td>516.4(2H⁺)</td>
<td>574.4(2H⁺)</td>
<td>486.9(2H⁺)</td>
<td>CH₂(CO)NH₂</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>FHPKRPWIL</td>
<td>597.4(2H⁺)</td>
<td>655.4(2H⁺)</td>
<td>551.9(2H⁺)</td>
<td>CH₂C₆H₅</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>KNPYIL</td>
<td>768.4(1Na⁺)</td>
<td>885.2(1Na⁺)</td>
<td>679.7(1Na⁺)</td>
<td>(CH₂)₃NH₂</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>GFLRRI</td>
<td>761.4(1H⁺)</td>
<td>877.5(1H⁺)</td>
<td>760.3(1H⁺)</td>
<td>H</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>WHWLQL</td>
<td>882.5(1H⁺)</td>
<td>998.5(1H⁺)</td>
<td>752.1(1H⁺)</td>
<td>CH₂C₆H₄N</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>QRPRLSHKGPMFPF</td>
<td>518.0(3H⁺)</td>
<td>556.4(3H⁺)</td>
<td>592.6(3H⁺)</td>
<td>(CH₂)₃(CO)NH₂</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>LSRLFDNAE</td>
<td>468.5(2H⁺)</td>
<td>526.5(2H⁺)</td>
<td>877.7(1H⁺)</td>
<td>CH₃CH(CH₃)₂</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>ISRPPGFSFPR</td>
<td>630.9(2H⁺)</td>
<td>689.0(2H⁺)</td>
<td>602.0(2H⁺)</td>
<td>CH₃CH₂CH₃</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>EDKLD</td>
<td>795.2(1Na⁺)</td>
<td>911.5(1Na⁺)</td>
<td>722.0(1Na⁺)</td>
<td>(CH₂)₂(CO)OH</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>YGSLPQKSRQSDEN</td>
<td>868.6 (2H⁺)</td>
<td>926.8(2H⁺)</td>
<td>815.1(2H⁺)</td>
<td>CH₂C₆H₄OH</td>
<td>107</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>TPRK</td>
<td>251.1(2H⁺)</td>
<td>308.9(2H⁺)</td>
<td>205.8(2H⁺)</td>
<td>CH(OH)CH₃</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>APRLRFYSL</td>
<td>561.9(2H⁺)</td>
<td>619.8(2H⁺)</td>
<td>553.1(2H⁺)</td>
<td>CH₃</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>PPQGSFPR</td>
<td>904.3(1H⁺)</td>
<td>1020.4(1H⁺)</td>
<td>---</td>
<td>C₂H₇</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>RPKFQOF</td>
<td>960.3(1H⁺)</td>
<td>1016.5(1H⁺)</td>
<td>801.2(1H⁺)</td>
<td>(CH₂)₂(NH)₂H₂</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>CNLAVAASHIHYONQFQ</td>
<td>660.1(3H⁺)</td>
<td>---</td>
<td>---</td>
<td>CH₃SH</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>YRVRFLAKENVTQDDEC</td>
<td>758.2(3H⁺)</td>
<td>796.8(3H⁺)</td>
<td>722.2(3H⁺)</td>
<td>CH₂C₆H₄OH</td>
<td>107</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>DRVYIHPF</td>
<td>523.5(2H⁺)</td>
<td>582.1(2H⁺)</td>
<td>493.8(2H⁺)</td>
<td>CH₂(CO)OH</td>
<td>59</td>
<td></td>
</tr>
</tbody>
</table>

a \( m/z \) of unmodified peptide. b \( m/z \) of peptide modified with reagent III-1b at pH 5.5. c \([C₄H₉OOC(O)M – 117 – R] \) fragment or imine fragment in CID spectrum resulting from the neutral loss of the C₄H₉OOC(O) group +1 hydrogen (117) and the N-terminal amino acid side chain (R).

When a mixture of the two isobaric peptides HPKRPWIL and DRVYIHPF (entries 2 and 22 in Table III-2) was modified at pH 5.5 using reagent III-1b, the two peptides were distinguishable in LC-MS selected reaction monitoring (SRM) and neutral loss experiments by N-terminal side-chain cleavage (Figure III-4). Thus, the peptide bearing an N-terminal histidine can be distinguished from the peptide bearing aspartic acid without further sequencing. Figure III-4(B) shows the SRM of \( m/z = 582 ([C₄H₉OOC(O)M + 2H]^{2+}) \rightarrow m/z = 483 ([C₄H₉OOC(O)M – 117- 81 + 2H]^{2+}) \) where M is HPKRPWIL and the side chain of histidine (CH₂-C₃H₅N₂) has \( m/z = 81 \). Figure 3(C) shows the SRM of \( m/z = 582 ([C₄H₉OOC(O)M + 2H]^{2+}) \rightarrow m/z = 494 ([C₄H₉OOC(O)-M – 117- 59 + 2H]^{2+}) \) where M is DRVYIHPF and the side chain of aspartic acid (-CH₂COOH) has \( m/z = 77 \).
These results illustrate the use of N-terminal side chain loss as a quick diagnostic peptide identification method without the need for a complete sequencing experiment. This technique may be valuable in peptide mass fingerprinting studies, where N-terminus identification can be used to eliminate false positive matches in database searching.24

**Figure III-4.** LC-SRM trace of peptide mixture HPKRPWIL and DRVYIHPF modified with reagent III-1b at pH 5.5. (A) SIM $m/z = 582 ([C_4H_9OOC(O)M + 2H]^{2+})$ where M is either peptide. (B) SRM of $m/z = 582 ([C_4H_9OOC(O)M + 2H]^{2+}) \rightarrow m/z = 483 ([C_4H_9OOC(O)M – 117 - 81 + 2H]^{2+})$ where M is HPKRPWIL, and the histidine side chain ($\text{CH}_2\text{-CH}_3\text{N}_2$) has $m/z = 81$. (C) SRM of $m/z = 581 ([C_4H_9OOC(O)M – 117 - 81 + 2H]^{2+}) \rightarrow m/z = 494 ([C_4H_9OOC(O)M – 117 - 59 + 2H]^{2+})$ where M is DRVYIHPF and the aspartic acid side chain (-CH$_2$COOH) has $m/z = 59$. 
Mechanism of N-terminal Side Chain Cleavage

CID spectra of N-terminal peroxycarbamate adducts of amino acids and peptides showed fragments with \( m/z \) equal to the \( m/z \) of the unmodified species \([C_4H_9OOC(O)M – 116]\) derived from decomposition of the \( C_4H_9OOC(O) \) and addition of a hydrogen. In most cases this fragment \([C_4H_9OOC(O)M – 116]\) is predominant and it is accompanied by another fragment corresponding to an imine derived from the neutral loss of the N-terminal amino acid side chain \([C_4H_9OOC(O)M – 117 – R]\). In order to determine whether the two fragments are part of the same radical decomposition pathway, we carried out an in-source CID experiment using an ESI-triple quadrupole instrument to obtain the ion \([C_4H_9OOC(O)M – 117 – R]\) in Q1 and isolate it for further CID. The CID spectrum of this ion did not generate an imine ion \([C_4H_9OOC(O)M – 117 – R]\), instead the spectrum was comparable to that obtained for the unmodified species. This suggests that the \([C_4H_9OOC(O)M – 116]\) fragment can not be an intermediate in the side chain loss, \([C_4H_9OOC(O)M – 117 – R]\). Two mechanisms must be in play, with one leading to formation of M from \( C_4H_9OOC(O)M \) and the other giving side chain cleavage (Scheme III-6).
It is unclear how the peroxycarbamate adducts decompose by CID to provide the unmodified peptide or amino acid and to address this question we prepared compound III-2 shown in Scheme III-7. Based on CID studies with a model lysine peroxycarbamate containing the moiety \( \text{C}_4\text{H}_9\text{OOC(O)} \) (See chapter II), we suspected that the t-butyl group of the peroxycarbamate would be a possible source of an H atom during the CID process leading to the peptide or amino acid fragment, \( [\text{C}_4\text{H}_9\text{OOC(O)}\text{M}] \rightarrow \text{M} \). Compound III-2 was prepared from the commercially available 2-bromo-2-methylpropane-\( \text{d}_9 \) by reaction with anhydrous hydrogen peroxide in the presence of silver triflate.\textsuperscript{25} The resulting hydroperoxide was treated with 3-nitrophenyl chloroformate providing III-2 (See experimental section).
Compound **III-2** was then used to modify the N-terminus of the peptide WHWLQL (entry 10 in Table III-2) using conditions for selective N-terminus modification. The deuterium labeled peroxycarbamate C₄H₉-d₉OOC(O), **III-3** was subjected to CID and the resulting product ions included the N-terminal side chain loss fragment, and a fragment corresponding to M+1, suggesting that deuterium transfer to the peptide attends peroxycarbamate dissociation. This indicates that the decomposition of the peroxycarbamate **III-3** follows a pathway in which a hydrogen (deuterium) atom is transferred from the t-butyl group to the α-amino of the N-terminus resulting in a CID fragment identical to the unmodified peptide (Scheme III-8). The mechanistic details of this hydrogen (deuterium) atom transfer remain unclear, although an intermediate carbamic acid could be involved, as could either radical or ion pair species.

A reasonable mechanism for loss of R, the N-terminal amino acid side chain, proceeds through the aminyl radical, which is formed by homolysis of the weak O-O bond followed by decarboxylation. Loss of R by β-fragmentation of the aminyl radical appears to be mechanistically reasonable based on studies on lysine peroxycarbamates.
Scheme III-8. Proposed mechanism for dissociation of N-terminal peroxycarbamate.

**Peroxycarbamate Modification of a Tryptic Digest**

A myoglobin tryptic digest standard was modified using reagent ***III-1c*** at pH 5.5. Reagent ***III-1c*** appeared to provide better conversion to peroxycarbamates than either ***III-1a*** or ***III-1b***. Reagent ***III-1c*** is more reactive than ***III-1a*** and ***III-1b***, which would suggest that this reagent would not be as selective for the N-terminus as reagent ***III-1b***; however, both ***III-1a***, and ***III-1b*** required longer reaction times at pH 5.5 (over 4 hours). The modification reactions described here are pseudo 1<sup>st</sup> order with rate = $k_{obs}[\text{peptide}]$ due to the large excess of reagent used. Because peptides are present in the digest mixture at much lower concentrations (μM) compared to the peptide concentrations in the model modification reactions shown in Table III-2 (mM), the reaction rate is slower with reagents ***III-1a*** and ***III-1b*** in this case and a more reactive reagent is needed. LC-MS/MS analysis was carried out on the unmodified as well as the modified digest. Visual inspection of the MS data identified 12 most abundant peptides, 7 of which were modified (Table III-3).
Identification of the peptides from MS/MS spectra was confirmed by database search algorithm SEQUEST\textsuperscript{27} and the P-Mod algorithm.\textsuperscript{28} N-terminal side chain cleavage was observed in data-dependent MS\textsuperscript{2} experiments for entries 2, 4, 5, 7, and 8. MS\textsuperscript{3} of the imine ion of entry 2 $[C_4H_9OOC(O)M - 117 - R]$, exhibited a labeled b-ion series indicating the loss of the N-terminal valine side chain + 1 hydrogen with respect to the unmodified peptide M (mass shift = -44), while the y-ion series remained unchanged (Table III-4). Thus the fragmentation of the N-terminal side-chain serves to label the b-ion series, which can aid de novo sequencing analyses.

### Table III-3. Peptides detected by LC-MS from a modified myoglobin digest.

<table>
<thead>
<tr>
<th>Entry $^a$</th>
<th>Peptide Sequence</th>
<th>Unmodified $m/z$</th>
<th>Modified $^b$ $m/z$</th>
<th>Mass Shift $^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HLKTEAEMK</td>
<td>371.8 (4H$^+$)</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>2</td>
<td>ALELFR</td>
<td>374.8 (2H$^+$)</td>
<td>864.3 (1H$^+$)</td>
<td>116</td>
</tr>
<tr>
<td>3</td>
<td>YKELGFQG</td>
<td>472.3 (2H$^+$)</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>4</td>
<td>LFTGHPETLEK</td>
<td>424.8 (3H$^+$)</td>
<td>694.5 (2H$^+$)</td>
<td>116</td>
</tr>
<tr>
<td>5</td>
<td>HGTVVLTAIGGILK</td>
<td>460.4 (3H$^+$)</td>
<td>748.0 (2H$^+$)</td>
<td>116</td>
</tr>
<tr>
<td>6</td>
<td>HPGDFGADAQGAMTK</td>
<td>752.0 (2H$^+$)</td>
<td>810.0 (2H$^+$)</td>
<td>116</td>
</tr>
<tr>
<td>7</td>
<td>VEAKIAGHQVEQVLR</td>
<td>536.6 (3H$^+$)</td>
<td>862.1 (2H$^+$)</td>
<td>116</td>
</tr>
<tr>
<td>8</td>
<td>GLSDGEWQVVLNVWGK</td>
<td>906.6 (2H$^+$)</td>
<td>966.5 (2H$^+$)</td>
<td>116</td>
</tr>
<tr>
<td>9</td>
<td>GHHEAEILKPLAQHATK</td>
<td>464.6 (4H$^+$)</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>10</td>
<td>YLEFISDI/HVLHSK</td>
<td>943.3 (2H$^+$)</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>11</td>
<td>LFTGHPETLEKFDK</td>
<td>416.2 (4H$^+$)</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>12</td>
<td>ELGFQG</td>
<td>650.2 (1H$^+$)</td>
<td>766.1 (1H$^+$)</td>
<td>116</td>
</tr>
</tbody>
</table>

$^a$Entries in blue are peptides for which modification at the N-terminus is observed. $^b$m/z of modified peptides ([C4H9OOC(O)M]). $^c$Mass shift is adjusted for singly charged species.
Table III-4. $b$- and $y$-ions from MS$^2$ of unmodified peptide VEAKIAGHGQEVLIR compared to $b$- and $y$-ions from MS$^3$ of the modified peptide ([C$_4$H$_9$OOC(O)M]).

<table>
<thead>
<tr>
<th>Position</th>
<th>$b$-ions</th>
<th>$y$-ions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MS$^2$ unmodified peptide</td>
<td>MS$^2$ modified peptide</td>
</tr>
<tr>
<td>2</td>
<td>228.79</td>
<td>----</td>
</tr>
<tr>
<td>3</td>
<td>299.92</td>
<td>371.26</td>
</tr>
<tr>
<td>4</td>
<td>414.86</td>
<td>483.72</td>
</tr>
<tr>
<td>5</td>
<td>528.21</td>
<td>401.14</td>
</tr>
<tr>
<td>6</td>
<td>599.25</td>
<td>757.22</td>
</tr>
<tr>
<td>7</td>
<td>656.19</td>
<td>814.34</td>
</tr>
<tr>
<td>8</td>
<td>793.06</td>
<td>951.52</td>
</tr>
<tr>
<td>9</td>
<td>850.39</td>
<td>1008.49</td>
</tr>
<tr>
<td>10</td>
<td>978.37</td>
<td>1079.59</td>
</tr>
<tr>
<td>11</td>
<td>1108.16</td>
<td>1192.40</td>
</tr>
<tr>
<td>12</td>
<td>1206.36</td>
<td>1307.56</td>
</tr>
<tr>
<td>13</td>
<td>1319.99</td>
<td>1378.14</td>
</tr>
<tr>
<td>14</td>
<td>1432.56</td>
<td>1388.61</td>
</tr>
</tbody>
</table>

$^a$ Parent ion m/z = 804.1 (M + 2H)$^{2+}$ where M is the peptide VEAKIAGHGQEVLIR. $^b$Parent ion m/z = 782.1 ([C$_4$H$_9$OOC(O)M – 117- 43 + 2H]$^{2+}$), Where valine (-CH(CH$_3$)$_2$) has m/z = 43. $^c$The difference between the $b$-ions corresponds to the valine side chain +1 hydrogen.

The SALSA algorithm$^{29}$ was used to search through tandem MS data for spectra indicating the N-terminal side-chain losses. The SALSA algorithm was designed to find characteristics in the MS/MS spectra such as charged and neutral losses, a particular product ion, or a sequence motif. It can also search for a combination of two or more of these characteristics and it does not require a peptide sequence input. SALSA scores do not indicate the probability of false positives, but instead rank spectra based on intensity and number of ions found that match a search criteria. We searched our modified myoglobin digest data for the neutral loss of 116, which leads to the fragment [C$_4$H$_9$OOC(O)M - 116] or simply [M]. This generated a list of precursor ions that match the m/z of peptide-
peroxycarbamates in the digest mixture. To find peptides bearing a certain amino acid at the N-terminus, we added a secondary search characteristic corresponding to the neutral loss of $117 + R$, which leads to the imine fragment $[C_4H_9OOC(O)M - 117 - R]$ or $[M - (R+1)]$. Thus, ions corresponding to neutral loss of $117 + R$ were scored only in CID spectra also containing the primary loss of $116$. This scoring scheme reduced the number of false positive matches. For example, a SALSA search for MS/MS spectra of peptides bearing histidine at the N-terminus scored a primary neutral loss of $116$ linked to a secondary neutral loss of $198$ ($116 + 1 + 81$, where $81 = m/z$ of His side chain). To improve matching accuracy, we also specified doubly charged ions, since these dominate the spectra. The first 6 of 24 results in the SALSA output for this search are shown in Table III-5. Given the scoring scheme described above, the highest scoring spectra met both the primary and secondary criteria. The top score matched the $m/z$ of modified peptide $C_4H_9OOC(O)HGTVVLTALGGILK$ (entry 2 in Table III-5). This result was validated by visual inspection of the MS/MS data. Other peptides bearing histidine at the N-terminus (entries 4 and 6) exhibited low scores due to lower relative intensity of product ions corresponding to the scored losses. The peptide in entry 3 does not contain histidine at the N-terminus; however, this precursor ion leads to an intense $[C_4H_9OOC(O)M - 116]$ product ion (neutral loss of 116), therefore it scores higher than entries 4 and 6. It addition, this precursor ion also leads to a neutral loss of 198, which is unrelated to the N-terminal side chain loss; however, this ion is present in lower intensity. Entry 5 has a higher score than entry 6 even though the peptide does not bear
histidine at the N-terminus, because this precursor ion also gives an intense [C₄H₉OOC(O)M - 116] product ion; however, it does not match the neutral loss of 198. Nevertheless, this analysis can be useful in filtering large amount of data in order to find peptides of interest.

Table III-5. First 6 results from SALSA search for peptides with histidine at the N-terminus in a myoglobin digest.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Score</th>
<th>Precursor m/z (2H⁺)a</th>
<th>Retention Time (min)</th>
<th>Modified Peptide (tBuOO(CO)M)</th>
<th>Loss = 116b</th>
<th>Loss = 198c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19.25</td>
<td>747.75</td>
<td>27.01</td>
<td>HGTVLTLAGEILK</td>
<td>689.93</td>
<td>--------</td>
</tr>
<tr>
<td>2</td>
<td>14.46</td>
<td>748.09</td>
<td>27.13</td>
<td>HGTVLTLAGEILK</td>
<td>689.92</td>
<td>648.71</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>966.63</td>
<td>27.70</td>
<td>GLSDGEWQQLNWGGK</td>
<td>908.54</td>
<td>867.27</td>
</tr>
<tr>
<td>4</td>
<td>6.11</td>
<td>601.27</td>
<td>22.80</td>
<td>HLKTEAEKM</td>
<td>542.79</td>
<td>502.26</td>
</tr>
<tr>
<td>5</td>
<td>4.56</td>
<td>739.04</td>
<td>27.95</td>
<td>AELFRNDIAAK</td>
<td>681.11</td>
<td>--------</td>
</tr>
<tr>
<td>6</td>
<td>4.4</td>
<td>817.95</td>
<td>21.14</td>
<td>HPGDFGADQGAMTK</td>
<td>759.87</td>
<td>718.79</td>
</tr>
</tbody>
</table>

a Doubly charged precursor ions which lead to product ions in the MS/MS spectra matching a neutral loss of 116 (loss of C₄H₉OOC(O)) alone or accompanied by the neutral loss of 198 (loss of C₄H₉OOC(O) +1 hydrogen –R of histidine).
b Doubly charged product ions corresponding to the neutral loss of 116 with respect to the precursor ion.
c Doubly charged product ions corresponding to the neutral loss of 198 with respect to the precursor ion.

N-terminal side-chain cleavage by MALDI

A series of peptides containing each 20 amino acids at the N-terminus were modified with conditions selective for N-terminal modification (pH 5.5, reagent III-1b). The adducts were analyzed by a MALDI-TOF (Table III-6). As discussed in Chapter II, intact peroxycarbamate adducts are not detected by MALDI. The laser desorption process results in dissociation of the t-butyl peroxy group at the source. Therefore, only the fragments derived from the adduct dissociation are detected. In the analysis of lysine side-chain peroxycarbamates, fragments corresponding to backbone cleavage at each lysine position are observed. With N-terminal peroxycarbamates, a fragment with m/z equal to the
unmodified peptide is detected, along with a fragment corresponding to the imine \([\text{C}_4\text{H}_9\text{OOC(O)M – 117 – R}]\). These results are consistent with the ESI studies discussed above. In the case of the peptide with the N-terminal glycine, only the loss of the \(\text{C}_4\text{H}_9\text{OOC(O)-}\) group was observed, and no imine fragment was detected (entry 8 in Table III-6). Peptides with proline and cysteine at the N-terminus also did not produce an imine fragment (entries 18 and 20 in Table III-6), as expected based on ESI results. Examples of MALDI spectra are shown in Figure III-5 for entries 4 and 12. The spectra contain ions corresponding to the \(m/z\) of the intact peptides (M), derived from the loss of the \(\text{C}_4\text{H}_9\text{OOC(O)-}\) group, and the corresponding imine ions derived from the cleavage of the N-terminal side chain and 1 hydrogen (\(M – (R + 1)\)). Thus the \(m/z\) difference between the two ions identifies the N-terminal side chain.
Table III-6. Analysis of modified peptides (N-terminal peroxycarbamates) by MALDI.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Sequence</th>
<th>$[^{11}] M + 1H[^{11}] $</th>
<th>$[^{11}] M – (R + 1) + 1H[^{11}] $</th>
<th>R</th>
<th>m/z</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>VSSNISEDPVPV</td>
<td>1265.3(Na$^+$)</td>
<td>1221.6 (Na$^+$)</td>
<td>CH(CH$_3$)$_2$</td>
<td>43</td>
</tr>
<tr>
<td>2</td>
<td>HPKRWPIL</td>
<td>1045.5</td>
<td>965.3</td>
<td>CH$_2$CH$_2$N$_2$</td>
<td>81</td>
</tr>
<tr>
<td>3</td>
<td>MEHFRW</td>
<td>963.3</td>
<td>887.3</td>
<td>(CH$_2$)$_3$SCH$_3$</td>
<td>75</td>
</tr>
<tr>
<td>4</td>
<td>SRVSRRSR</td>
<td>1004.5</td>
<td>972.5</td>
<td>CH$_2$OH</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>NVIEISNDENLRE</td>
<td>1032.1</td>
<td>972.9</td>
<td>CH$_2$(CO)NH$_2$</td>
<td>58</td>
</tr>
<tr>
<td>6</td>
<td>FHPKRWPIL</td>
<td>1194.7</td>
<td>1102.7</td>
<td>CH$_2$C$_6$H$_5$</td>
<td>91</td>
</tr>
<tr>
<td>7</td>
<td>KNPYIL</td>
<td>748.1</td>
<td>675.2</td>
<td>(CH$_2$)$_3$NH$_2$</td>
<td>72</td>
</tr>
<tr>
<td>8</td>
<td>GFLRRRI</td>
<td>1762.3</td>
<td>------</td>
<td>H</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>WHMLQL</td>
<td>883.3</td>
<td>752.3</td>
<td>CH$_2$C$_6$H$_5$N</td>
<td>130</td>
</tr>
<tr>
<td>10</td>
<td>QRPRLSHKGMPF</td>
<td>1551.7</td>
<td>1479.1</td>
<td>(CH$_2$)$_3$(CO)NH$_2$</td>
<td>72</td>
</tr>
<tr>
<td>11</td>
<td>LSRLFDNA</td>
<td>936.4</td>
<td>878.3</td>
<td>CH$_2$CH(CH$_3$)$_2$</td>
<td>57</td>
</tr>
<tr>
<td>12</td>
<td>ISRPPGFSPFR</td>
<td>1261.5</td>
<td>1203.5</td>
<td>CH$_2$CH$_2$CH$_3$</td>
<td>57</td>
</tr>
<tr>
<td>13</td>
<td>EDKOLD</td>
<td>774.3</td>
<td>700.0</td>
<td>(CH$_2$)$_3$(CO)OH</td>
<td>73</td>
</tr>
<tr>
<td>14</td>
<td>DRVYIHPF</td>
<td>1047.3</td>
<td>987.3</td>
<td>CH$_2$(CO)OH</td>
<td>59</td>
</tr>
<tr>
<td>15</td>
<td>YGSLPQKSQRSQDEN</td>
<td>1730.3</td>
<td>1628.1</td>
<td>CH$_2$C$_6$H$_4$OH</td>
<td>107</td>
</tr>
<tr>
<td>16</td>
<td>TPRK</td>
<td>502.2</td>
<td>456.2</td>
<td>CH(OH)CH$_3$</td>
<td>46</td>
</tr>
<tr>
<td>17</td>
<td>APRLRFSYL</td>
<td>1123.6</td>
<td>1107.6</td>
<td>CH$_3$</td>
<td>15</td>
</tr>
<tr>
<td>18</td>
<td>PPGFSFPR</td>
<td>905.3</td>
<td>------</td>
<td>C$_2$H$_7$</td>
<td>43</td>
</tr>
<tr>
<td>19</td>
<td>RPKQQOF</td>
<td>901.4</td>
<td>801.2</td>
<td>(CH$_3$)$_3$NH(NH$_2$)$_2$</td>
<td>100</td>
</tr>
<tr>
<td>20</td>
<td>CNLAVAAASHIYQNQFVQ</td>
<td>------</td>
<td>------</td>
<td>CH$_2$SH</td>
<td>47</td>
</tr>
</tbody>
</table>

$^{a}$ m/z of intact peptide ion derived from the dissociation of the C$_4$H$_9$OOC(O)$^-$ group, followed by hydrogen abstraction. $^{b}$ m/z of imine fragment derived from dissociation of the C$_4$H$_9$OOC(O)$^-$ group and cleavage of the N-terminal side chain.

Analysis of modified myoglobin digest

The modified myoglobin tryptic digest used in the ESI-LC-MS analysis was analyzed by MALDI-TOF as a crude mixture. The modified peptides were identified by the presence of the intact peptide ion (M) and its corresponding imine ion (M – (R + 1)). The MALDI spectrum was complex and interpretation of the data in this manner proved to be time-consuming. LC separation prior to MALDI analysis led to cleaner spectra; however the lack of data manipulation software such as SALSA, compatible with MALDI data, impeded faster interpretation. The peptide mixture was also analyzed by MALDI-TOF-TOF to obtain sequence information on the imine ions. MS/MS spectra on the imine ions showed an m/z shift on the b-ion series compared to the intact peptide ions. This
shift or label was equal to the m/z of the N-terminal side chain + 1H, as seen for ESI experiments. The MS/MS spectrum for the imine ion of the peptide HGTVVLTALGGILK is shown in Figure III-6. The b-ions are shifted by Δm/z = 82, which corresponds to the histidine side chain + 1H. The y-ions remain the same as those for the unmodified peptide, which serves as a fingerprint.

Figure III-5. MALDI spectra of modified peptides as N-terminal peroxycarbamates. A) modified peptide 4 in Table III-6. B) Modified peptide 12 in Table III-6. [M + 1H]⁺ corresponds to the loss of the C₄H₉OOC(O)- group, followed by addition of a hydrogen. [M – (R+1) + 1H]⁺ corresponds to the imine ion derived from loss of the C₄H₉OOC(O)- group, and N-terminal side chain cleavage.

Despite the need for manual interpretation of MALDI data, this type of analysis has the advantage over ESI in that only singly charged species are observed, which facilitates identification of the N-terminal amino acid. The fact that both the imine ion and an ion corresponding to the intact peptide are
detected allows for the analysis to be carried out in one run, without the need to compare spectra of unmodified and modified peptides. In addition, because dissociation of the peroxycarbamates occurs at the source, b-ion labeling can be obtained in a single MS/MS event. Nevertheless, the intact peroxycarbamate adducts are not observed, thus distinguishing intact peptide ions derived from modified peptides and those from unreacted peptides is a difficult task. A combination of both ESI and MALDI techniques would provide complementary information and hence a more complete analysis.

**Figure III-6.** MS/MS MALDI spectrum of the imine ion \([M – (R + 1) + 1H]^+\) of the peptide HGTVVLTALGGILK from the modified myoglobin tryptic digest. The b-ion series has an \(m/z\) shift of -82, which corresponds to \(R + 1\) where \(R\) is the histidine side chain.
Conclusion

Results shown in Figure III-4 illustrate the use of N-terminal side chain loss as a quick diagnostic peptide identification method without the need for a complete sequencing experiment. These results lead the way to the use of this technique in peptide mass fingerprinting studies, where N-terminus identification can be used to eliminate false positive matches in database searching. Because MS/MS spectra of unmodified peptides lack information about the N-terminus ($b_1$ and $y_{n-1}$ ions), identification of N-terminal residues would aid in obtaining complete sequences in de novo analyses. In addition, results in Table III-4 demonstrate that gas phase cleavage of the N-terminal side chain by CID, labels the N-terminus so that a subsequent CID process (MS$^3$) leads to a mass shift in the b-ions when compared to the b-ions from the intact peptide. This mass shift corresponds to the $m/z$ of the N-terminal side chain, and serves to elucidate the b-ion series in de novo sequencing. MALDI results correlate well with ESI analyses. In MALDI however, fragmentation of the N-terminal peroxycarbamates occurs at the source, and the fragmentation products are observed in the spectra without an MS/MS experiment. Therefore, a labeled b-ion series of a particular peptide imine ion can be easily obtained by a single MS/MS event using a TOF-TOF mass analyzer. Peptide mass fingerprinting studies are typically done with MALDI, thus these results further support the utility of the N-terminus identification technique in such studies.

The analysis of a myoglobin tryptic digest by ESI-MS and MALDI demonstrates that this N-terminus identification method is feasible in a
proteomics setting. In ESI analyses, data manipulation software such as SALSA can be used to sort through large amounts of MS/MS spectra and find peptides of interest by N-terminal side chain loss. MALDI analysis can facilitate interpretation of the data by providing singly charged species.

**Experimental Procedures**

**General**

Amino acids were purchased from Aldrich Chemical Co, Inc. (Milwaukee, MA) as either methyl, ethyl or t-butyl ester derivatives in their hydrochloride salt forms. Peptides were purchased from American Peptide Company (Sunnyvale, CA). Solvents were HPLC grade, obtained from Fisher Scientific (Pittsburgh, PA). All other reagents were purchased from Aldrich Chemical Co, Inc. (Milwaukee, MA). Acetate buffers (10 mM) of various pH values were prepared from ammonium acetate titrated to pH 5.2, 5.5, and 5.8 with acetic acid. Tris buffers were 0.1M solutions titrated to pH 7.5 and pH 8 respectively. Ammonium bicarbonate buffer was a 0.1 M solution at pH 8.6. Buffer systems used to modify peptides and amino acids consisted of a 1:1 mixture of the appropriate buffer and acetonitrile. The myoglobin tryptic digest standard mixture (500 pmol) was purchased from Alltech-Life Sciences (Deerfield, IL).
Modification Reagents

The peroxycarbonate reagents **III-1a-1c** were prepared according to the procedure in reference 16. The details of the preparation of **III-1a** were presented in chapter II. Reagents **III-1b** and **III-1c** were synthesized from their chloroformates as described below.

Preparation of chloroformates

3-Nitrophenyl chloroformate and 2-chloro-4-nitrophenyl chloroformate were prepared from 3-nitrophenol and 2-chloro-4-nitrophenol respectively. The phenols and 1 equivalent of diisopropylethylamine as a solution in methylene chloride, were added slowly (over 10 min.) to 5 equivalents of phosgene (20% solution in toluene) at 0°C under argon gas. The reactions were stirred for 3 hours at room temperature. The excess phosgene was removed by sparging with argon gas for 2 hours. The solvent was removed using under vacuum and the crude material was used in the preparation of the peroxycarbonates (**III-1b** and **III-1c**).

Preparation of reagents **III-1b** and **III-1c**

The crude chloroformates (approximately 10 mMoles) were dissolved in 20 ml of methylene chloride. The solutions were cooled to 0°C under nitrogen gas, and 1 equivalent of t-butyl hydroperoxide was added (5 M solution in dodecane). To the reaction mixtures, 1 equivalent of pyridine was added
dropwise over 10 min. The reactions were stirred at 0°C for 1 hour and allowed to reach room temperature. The solution was concentrated and purified by silica gel column chromatography (1:10 methanol/chloroform). The products \textbf{III-1b} and \textbf{III-1c} were obtained as light yellow solids in 40-50% yield. \textbf{III-1b} \textsuperscript{1}H-NMR (300 MHz, CDCl\textsubscript{3}): 1.43 (s, 9H), 7.63 (m, 2H), 8.17 (m, 2H). \textsuperscript{13}C-NMR (75 MHz, CDCl\textsubscript{3}): 26.33, 85.73, 117.06, 121.73, 127.57, 130.77, 145.2, 151.27, 152.70. \textbf{III-1c} \textsuperscript{1}H-NMR (300 MHz, CDCl\textsubscript{3}): 1.41 (s, 9H), 7.47 (d, 2H, J = 8.7 Hz), 8.21 (ddd, 2H, J\textsubscript{1} = 9.0 Hz, J\textsubscript{2} = 2.7 Hz, J\textsubscript{3} = 0.6 Hz), 8.37 (d, 2H, J = 3.0 Hz). \textsuperscript{13}C-NMR (75 MHz, CDCl\textsubscript{3}): 26.30, 86.14, 123.80, 124.26, 126.63, 128.64, 146.5, 151.64, 162.74. Low resolution ESI mass spectrometry analysis for \textbf{III-1b}: calculated mass: 278.06 (M+Na\textsuperscript{+}), observed mass: 278.00. Attempts to obtain a molecular ion for reagent \textbf{III-1c} in mass spectrometry analyses were unsuccessful. Nevertheless, reactions involving this reagent and the N-terminal amino groups of peptides led to adducts identical to those derived from reagents \textbf{III-1a} and \textbf{III-1b}.

**Preparation of reagent III-2**

2-bromo-2-methylpropane-d\textsubscript{9} was dissolved in anhydrous methylene chloride and an excess of anhydrous hydrogen peroxide in ether was added (~3 fold excess). The resulting solution was cooled to 0°C at which point 1 equivalent of silver triflate was slowly added and the reaction mixture was allowed to stir at 0°C for 4 hours. The mixture was filtered to remove the solid silver bromide and the filtrate was washed with fresh methylene chloride. The organic layer was
washed with water (2X), brine (1X) and dried over anhydrous MgSO₄. The resulting solution was concentrated to approximately ½ of the original volume. Caution is advised since removing all solvent may result in explosion. The material was verified by TLC (SiO₂, CH₂Cl₂) against t-butyl hydroperoxide (peroxide stain development). The resulting solution was added slowly to a methylene chloride solution of 3-nitrophenyl chloroformate at 0°C followed by 1 equivalent of triethylamine. The reaction was allowed to reach room temperature and stirred for an additional hour. The solution was washed with 1% H₂SO₄ (2X), brine (1X) and dried over anhydrous MgSO₄. The product was purified by silica gel column chromatography (1:1 hexane/ether) to give a white solid. ¹H-NMR (300 MHz, CDCl₃): 7.60 (m, 2H), 8.15 (m, 2H). ¹³C-NMR (125 MHz, CDCl₃): 24.95 (septet, J = 19.4 Hz), 85.7, 116.71, 121.83, 127.40, 130.28, 151.19, 152.65. Calculated exact mass: 264.1308, HRMS (FAB) observed mass: 264.1049.

Modification Reactions

Reactions with all 20 amino acids as methyl, ethyl, or t-butyl ester derivatives were carried out by combining 5 μl of amino acid solution (0.1M in 50/50 acetonitrile and water), 20 μl of ammonium bicarbonate buffer system (50/50 acetonitrile/buffer), and 20 μl of III-1a (0.1 M in acetonitrile). Final concentrations for amino acids and reagent were 0.01M and 0.04M respectively. Reactions were incubated at room temperature overnight. The crude reaction mixtures were diluted to 1 ml with a 0.4 mM solution of LiCl in methanol for MS
The peptide YRVRFLAKENVTQDAEDNC (entry 21 in Table III-2) was modified at pH range 5.2 to 8.6 by addition of 10 μl of buffer system (50/50 acetonitrile/buffer) to 10 μl of peptide solution (1 mg/ml in water), followed by 5 μl of III-1a or III-1b (0.1 M in acetonitrile). Final concentrations of peptide and reagent were 0.7 μM and 33 M respectively. The reaction was incubated at 37 °C for 1 hr. The crude mixtures were analyzed by reverse phase LC-MS to determine selectivity for N-terminus modification. A series of peptides bearing each of the 20 amino acids at the N-terminus (Table III-2) were modified using the conditions which gave best selectivity for N-terminus modification (pH 5.5 buffer system, reagent III-1b). All peptide reaction mixtures were diluted to 1 ml with a 0.1 % acetic acid solution and analyzed by ESI-MS/MS as direct liquid infusions and by MALDI-TOF.

**Mass spectrometry**

Mass spectrometry analyses of modified single amino acid derivatives and peptides were performed on a ThermoFinnigan (San Jose, CA) TSQ 7000 triple-quadrupole instrument, equipped with an Electrospray Ionization (ESI) source and a Waters (Milford, MA) Alliance HPLC system. The capillary temperature was kept at 200 °C with a voltage of 20 V. The electrospray needle voltage was 4.5 kV, and the tube-lens voltage was maintained between 70 and 100 V. The sheath and auxiliary gases (N2) were adjusted to maximize signal. The offset voltage in CID experiments varied between 20 and 35 eV. Samples were introduced into the ESI source as either lithium solutions or acidic solutions at a
rate of 10 to 20 μl/min in direct liquid infusion experiments. For LC-MS analyses of peptides, a Supelco (Bellefonte, PA) Discovery C18 column was used (25cm x 4.6mm, 5 μm). The solvent conditions consisted of a 20 min. gradient from 5% to 50% solvent B (95% acetonitrile, 5% water, 0.05% TFA) into solvent A (95% water, 5% acetonitrile, 0.05% TFA). LC-MS/MS analyses of peptide derivatives were performed on a ThermoFinnigan (San Jose, CA) LCQ Deca XP ion trap instrument, equipped with an ESI source, and a HP Agilent (Palo Alto, CA) HPLC system. Instrument parameters were tuned using angiotensin II peptide (5 pmol) to maximize signal. LC-MS/MS analyses of the myoglobin digests were performed using a Grace Vydac (Hesperia, CA) C18 column (10cm x 1mm x 5 μm). Solvent conditions consisted of a 30 min. gradient from 100% solvent A (5% acetonitrile, 95% water, 0.5% acetic acid) to 100% solvent B (5% water, 95% acetonitrile, 0.5% acetic acid). MALDI studies for the peptides in Table III-6 were carried out on an Applied Biosystems (Foster City, CA, USA) Voyager DE STR instrument. The myoglobin tryptic digest was analyzed by MALDI using an Applied Biosystems (Foster City, CA, USA) 4700 Proteomics instrument, equipped with a TOF-TOF mass analyzer. The matrix for all analyses was α-cyano-4-hydroxycinnamic acid (a solution in 4:6 0.1%TFA and acetonitrile). Samples were spotted on a gold target plate (1 μl 1:1 matrix/sample mixture).

**Modified myoglobin tryptic digest**

Modification of a myoglobin tryptic digest standard was achieved by combining 50 μl of digest solution (500 pmol in 200 μl of water), 50 μl of acetate
buffer system (pH 5.5), and 25 μl of reagent \textbf{III-1c} (0.1M in acetonitrile). The reaction was incubated at room temperature for 2 hours and analyzed by reverse phase LC-MS/MS. MS/MS data were analyzed with the SALSA algorithm\textsuperscript{29} to search for modified peptides and neutral losses of N-terminal amino acid side chains. The modified peptide mixture was also analyzed by MALDI-TOF-TOF to obtain MS/MS spectra for imine peptide ions.

\textbf{References}


CHAPTER IV

GAS-PHASE FRAGMENTATION OF ORNITHINE PEROXYCARBAMATES AND CHEMICAL CLEAVAGE OF PROTEINS VIA HYDRAZINOLYSIS

Introduction

With the goal of developing a gas-phase protein digestion technique to facilitate top-down proteomics approaches, we investigated the effect of introducing ornithine peroxycarbamates into peptide sequences on CID fragmentation patterns. Ornithine is not a natural occurring amino acid; however, it is an essential intermediate in the biosynthesis of L-arginine and plays an important role in the disposal of excess nitrogen through the urea cycle.\(^1\) In the urea cycle, arginine is converted to ornithine by the enzyme arginase, resulting in the release of urea. Ornithine can undergo lactamization under basic conditions, the stable six-member ring system formed makes this process favorable (Scheme IV-1).\(^2\)

![Scheme IV-1. Lactamization of ornithine in solution.](image-url)
In peptides, lactamization of ornithine may lead to backbone cleavage in solution. Lysine however, having an extra methylene group in the side chain, would lactamize through a seven-member ring, which is an unfavorable process. Based on this knowledge, we hypothesize that ornithine may offer site-specific peptide cleavage in the gas-phase. In addition, ornithine peroxycarbamates may promote more selective peptide backbone cleavage than their lysine analogues. A schematics of possible gas-phase radical-promoted fragmentation pathways for orninyl peroxycarbamate peptides is shown in Scheme IV-2.

Scheme IV-2. Possible gas-phase radical pathways for ornithyl peroxycarbamate peptides.
Ornithine can be introduced into peptide sequences via hydrazinolysis.\textsuperscript{4, 5} Hydrazinolysis was first reported as a method for identifying the C-termini of proteins by Akabori and his collaborators.\textsuperscript{6, 7} It was also found that in this reaction, the guanidino group of arginines was hydrolyzed to the amine, resulting in the formation of ornithine. In addition, hydrazinolysis of proteins has been found to preferably cleave at the C-terminal side of asparagine and at glycine-glycine linkages.\textsuperscript{8, 9} Based on these studies, not only is hydrazinolysis valuable as a mean to introduce unnatural ornithine in peptide sequences, but also it's application as a protein digestion technique deserves attention.

In this chapter, we report on the effect of introducing ornithine and ornithine peroxycarbamates into peptide sequences on backbone fragmentation by CID. Moreover, we explore the use of hydrazinolysis as a protein degradation technique for top-down proteomics approaches.

**Results and Discussion**

**Model Ornithine Peroxycarbamate Studies**

N-\(\alpha\)-acetyl ornithine, and its methyl and ethyl ester derivatives were modified as peroxycarbamates under conditions used for lysine modifications (1:1 0.1 NH\(_4\)HCO\(_3\)/ acetonitrile) as shown in Scheme IV-3. The methyl and ethyl ester derivatives were prepared by treatment of N-\(\alpha\)-acetyl ornithine with tosic acid and either methanol or ethanol. All peroxycarbamate derivatives were subjected to CID as lithiated solutions using an ESI-triple quadrupole mass
The fragmentation of the model ornithine peroxycarbamate methyl ester is shown in Figure IV-1 and Figure IV-2. This fragmentation pattern was consistent with the other ornithine derivatives and was independent of counter ion (Na\(^+\), H\(^+\) and Ag\(^+\) gave similar CID spectra). The formation of the aminyl radical IV-4 is reminiscent of lysine peroxycarbamate studies. Fragments IV-6 and IV-7 are also expected based on lysine peroxycarbamate fragmentation; however, fragment IV-5 is unique to ornithine peroxycarbamates. This fragment has precedent in the autohydrolysis of peptides containing ornithine and \(\alpha\)-\(\gamma\)-diaminobutyric acid (DABA) in solution.\(^{10}\) Under basic conditions the side chain amino group can attack the backbone carbonyl group to form a stable 5-member ring in the case of DABA, and a six-member ring in the case of ornithine, resulting backbone cleavage. Such intramolecular substitutions to form five and six-member rings are favorable. In the case of lysine, this process is not likely due to the formation of a seven-member ring. Thus the introduction of DABA or ornithine in peptide sequences has the potential for site specific peptide cleavage in solution, and this has been recognized by several researches in the past.\(^{2, 11}\)

Figure IV-1 shows that at low CID off-set voltages the predominant species is the aminyl radical IV-4, as expected based on lysine peroxycarbamate studies. As the off-set increases, this aminyl radical diminishes, and fragments IV-5-7 increase in intensity, suggesting the aminyl radical is likely to be an intermediate to IV-5-7. We note that none of the fragments in Figure IV-1 however are present in the CID spectra of ornithine derivatives not bearing a peroxycarbamate. CID
fragmentation data for all ornithine peroxycarbamate derivatives is shown in Table IV-1.

![Scheme IV-3. Modification of N-α-acetyl-ornithine-methyl ester as a peroxycarbamate.](image)

**Figure IV-1.** CID fragmentation of ornithine peroxycarbamate methyl ester. Effect of CID off-set voltage on ion intensities.

![Graph](image)
Figure IV-2. CID spectrum of ornithine methyl ester peroxycarbamate derivative IV-3.

Table IV-1. CID fragmentation of model ornithine derivatives.

<table>
<thead>
<tr>
<th>Orn Derivative</th>
<th>M + Li⁺</th>
<th>IV-4</th>
<th>IV-5</th>
<th>IV-6</th>
<th>IV-7</th>
</tr>
</thead>
</table>
| ![Image](image)

R(PC) = -(CH₂)₃NH(CO)OOfBu.

Introduction of Ornithine in Peptide Sequences

In order to study the CID fragmentation of ornithine peptide peroxycarbamates we introduced ornithine into peptides by hydrolysis of arginines via a modified Akabori reaction. The Akabori reaction was introduced in the 50’s as a method for identifying C-terminal amino acids of linear peptides.\(^6\) In
this reaction, peptides are treated with anhydrous hydrazine at 125 °C for several hours. The result is the degradation of the peptides, where the C-terminal amino acid is liberated as the free acid and the rest are converted to hydrazides (Scheme IV-4). The hydrazides can then be removed from the mixture as water-insoluble compounds by treating with benzaldehyde. Other researchers employed microwave technology to reduce the reaction time.\(^5\) In these later studies it was shown that by treating peptides with 50-70% aqueous hydrazine and microwave irradiation, Akabori cleavage products were obtained in 30 min. In addition, in the first few minutes of microwave irradiation, arginines were converted to ornithines. We used this method to introduce ornithine into commercially available peptides containing 1 or 2 arginines. We found that arginines were converted to ornithine in the first 10 minutes of microwave irradiation. An example is shown in Figure IV-3.

\[\text{Scheme IV-4. Akabori reaction of a tripeptide.}\]
Ornithine Peptide Derivatives

A series of peptides modified as ornithine derivatives were converted to peroxycarbamates by reaction with IV-2. However, in order to assure that only ornithine and not lysines or N-termini were converted to peroxycarbamates, the peptides were first alkylated at these positions. Thus the peptides were first alkylated via treatment with NaCNBH₃ and formaldehyde, which results in all amino groups converted to dimethyl amines. The alkylated peptides were treated with 70% hydrazine and irradiated for 10 min. in a domestic microwave (600 watts), and finally, after removal of hydrazine, the peptides were treated with IV-
2. The ornithine peroxycarbamate peptides were then analyzed by ESI-LC-MS/MS using a triple quadrupole and an ion-trap mass spectrometer. The patterns of fragmentation obtained in MS/MS experiments did not resemble the fragmentation pattern observed for the model ornithine peroxycarbamates. The fragmentation for these peptides was inconsistent and greatly dependant on peptide sequence. MS/MS fragmentation data for these peptides is given in Table IV-2. Although the CID spectra did differ from unmodified analogues, site-specific sequence information was only obtained in very few cases. For example the peptide *MEHF(OrnPC)WG, were the asterisk represents an alkylated N-terminus, and OrnPC refers to the ornithine peroxycarbamate derivative, lead to selective fragmentation at the C-terminal side of ornithine (fragments \( b_5 \) and \( y_2 \)) and at the N-terminal side of ornithine (fragment \( a_4 \)). However this was not the case for most peptides, and this fragmentation was not unique to the peroxycarbamate derivative, the unmodified peptide *MEHF(Orn)WG gave a similar fragmentation pattern. In addition for several peptides, the fragments were difficult to assign. Based on these studies, we concluded that ornithine peroxycarbamates did not offer any advantages over lysine peroxycarbamates in terms of promoting site specific peptide cleavage.

We also investigated fragmentation patterns of ornithine peptide derivatives not bearing peroxycarbamates compared to their arginine analogues. MS/MS data for these peptides is given in Table IV-3. The CID fragmentation of these ornithine peptide derivatives in very few cases exhibited preferred cleavage at ornithine compared to their arginine analogues; however, similarly to
ornithine peroxycarbamates, the fragmentation patterns were dependant on peptide sequence.

Table IV-2. CID data for peptide ornithine peroxycarbamate derivatives and their unmodified analogues.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>CID Fragments</th>
<th>CID Fragments of Unmodified Peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>*MEHF(OmPC)WG</td>
<td>$y_2, b_5, a_4$</td>
<td>$y_2, b_5^{+2}, b_3, b_6^{-2}, a_4, b_4, b_5$</td>
</tr>
<tr>
<td><em>Q(OmPC)P(OmPC)LSHK</em>GMPF</td>
<td>--------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>*YGGFL(OmPC)KYPK</td>
<td>--------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td><em>FHPK</em>(OmPC)PWIL</td>
<td>--------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>*D(OmPC)VYVHPF</td>
<td>$NH_2^{+2}, NH_2^{+rad}, y_6, y_7 - H_2O$</td>
<td>$y_6, y_6{^+2}$</td>
</tr>
<tr>
<td><em>HPPK</em>(OmPC)PWIL</td>
<td>$y_7 - NH_3^{+2}$</td>
<td>$b_7^{-}, y_7{^+2}$</td>
</tr>
<tr>
<td>*N(OmPC)VVVHPF</td>
<td>$y_7 - NH_3^{+2}, y_7 - NH_3, NH_2^{+rad}^{+2}$</td>
<td>$y_3, y_9$</td>
</tr>
</tbody>
</table>

(ornPC) represents ornithine peroxycarbamate. The asterisk represents site of alkylation (dimethyl amine) at the N-termini and lysine side chains. The unmodified peptides are those not bearing a peroxycarbamate ($tBuOO(CO)$-) moiety.

To test whether ornithine could provide a site for specific peptide cleavage in solution, we treated the ornithine containing peptide MEHF(Orn)WG with triethylamine at 65 °C and analyzed the crude reaction mixture by LC-MS. Previous studies on ornithine containing tri-peptides such as Orn-Gly-Gly treated under these conditions have shown lactam formation leading to peptide cleavage at ornithine after 36 hours. However, we found no evidence of cleavage products for our ornithine containing peptide. We note that in the mentioned literature study, it was also found that the tendency for peptide cleavage via lactam formation decreased from a model ornithine methyl ester to the ornithine amide to the tri-peptide Orn-Gly-Gly. Which may explain why we did not see this peptide cleavage with a 7-mer.
Table IV-3. CID fragmentation for peptide ornithine derivatives.

<table>
<thead>
<tr>
<th>CID Fragments</th>
<th>CID Fragments of Ornithine Derivatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEHFRWG</td>
<td>y2, b5</td>
</tr>
<tr>
<td>HPKRPWIL</td>
<td>y2, y3, y4, y5, y6</td>
</tr>
<tr>
<td>DRVYIHPF</td>
<td>y3, y4, y5, y6, y7</td>
</tr>
<tr>
<td>NRVYVHPF</td>
<td>y4, y5, y6, y7, [M – NH₃]⁺</td>
</tr>
</tbody>
</table>

Ornithine derivatives are those where argenines have been converted to ornithine by treatment with hydrazine.

Hydrazinolysis of Proteins

The hydrazinolysis of proteins and its utility in protein sequencing has literature precedent in earlier decades (early 50’s to late 80’s). However, with the development of powerful mass spectrometry based protein sequencing techniques in the 90’s, the research in hydrazinolysis ceased. With the current need for improvements in top-down proteomics approaches, methods which allow partial degradation of proteins prior to top-down mass spectrometry analyses are of great value. Studies on the specificity of hydrazine cleavage have shown that asparagine and glycine-glycine linkages are the most susceptible to hydrazinolysis. The susceptibility of Asn linkages to hydrazinolysis is proposed to be due to the conversion of the side chain amide to the hydrazide, which leads to cyclization to form an acylaminosuccinyl hydrazide (Scheme IV-5). Based on these findings, we studied the applications of this specificity as well as the use of milder hydrazinolysis conditions to obtain partial degradation of large proteins such as Bovine Serum Albumin (BSA) and Human Serum Albumin (HSA) to facilitate top-down sequencing. In order to study the
Hydrazinolysis of BSA and HSA we used the microwave reaction method described by Bose et al.\textsuperscript{5} used for conversion of arginines to ornithines as described in the previous section. Such conditions allow for slower degradation of peptides and therefore the opportunity to control the extent of degradation.

Scheme IV-5. Proposed mechanism of asparagine peptide bond cleavage by hydrazine.

Hydrazinolysis Reactions

BSA and HSA were treated with different amounts of hydrazine (10% to 80%) and irradiated in a domestic microwave (600 watts) for different time intervals (2 min. to 60 min.). After hydrazine removal, the protein was analyzed by SDS-PAGE to determine the extent of protein cleavage. We found that conditions of 35% aqueous hydrazine, and irradiation for 10 min. allowed for partial degradation of the protein into large fragments (>10 KDa). However, the presence of trace amounts of hydrazine in the protein samples caused the samples to form a streak pattern on the gels, leading to poor resolution. Attempts were made to resolve this issue by treating the protein samples with 1% mercaptoethanol and SDS prior to hydrazinolysis as suggested in the literature\textsuperscript{5}; however, there was no increase in gel resolution. Cleaning the protein samples
with C-18 spin micro columns or zip-tips prior to SDS-PAGE analysis resulted in significant loss of material. Subjecting the hydrazine treated protein samples to long periods of in vacuo hydrazine removal afforded the best results. The gel band pattern for the best analysis for hydrazine treated BSA and HSA is shown in Figure IV-4. The BSA and HSA fragment bands were excised, digested with trypsin and analyzed by LC-MS/MS to determine their sequence.

![Figure IV-4](image)

**Figure IV-4.** SDS-PAGE of proteins undergoing hydrazinolysis with 35% hydrazine irradiated for 10 min. M) Protein standards, a) Control BSA, b) Control HSA, c) BSA after hydrazinolysis, d) HSA after hydrazinolysis.

**LC-MS/MS Analysis of Partially Degraded BSA and HSA**

In order to determine the sites of cleavage afforded by hydrazinolysis, bands from the best SDS-PAGE gels were excised, and an in-gel tryptic digestion was carried out, followed by LC-MS/MS analysis. The MS/MS data was
analyzed by SEQUEST\textsuperscript{16} and P-Mod\textsuperscript{17} to identify the different protein fragments. In both cases (HSA and BSA) this analysis led to the identification of peptides belonging to many different sections of the protein sequence, instead of discrete sections that could be assigned to the large fragments observed by SDS-PAGE. It is possible that the resolution of the gels was still not sufficient to allow for excision of single bands, thus leading to samples containing multiple protein fragments. Figure IV-5 shows an example of sequence coverage from analysis of the 5\textsuperscript{th} band (from top to bottom) on the SDS gel for hydrazine treated BSA. From this result it is evident that more than one fragment was included in the sample, since the majority of the protein sequence is represented in this band.

\begin{verbatim}
MKWVTFLLLLFSAYSRGVFRRDTKSEIAHRFKLGEEHFKGLVLIASFQYL
QQCQPFDEHVKLVELETFAKTCVADESHAGCEKLISLHTLFGDELCVKASLRETY
GDMACCEKEQEPERNECFLSHKDSDLPKLKDLPNTLCDELFKIDSKKWGGK
YLYEIAHRHPYFYAPELLYYANKYNQVFQECQAOEDKGPSLKLIIADEF
SSARQRLRCASIQKFGERALKWANWSARLSQKFKAEEFVEVTKLVTDLTKVH
CHGDLLCADDRAKLAKYICDNQTDISSKLKECCDKPKLEKSHIAAEVKDAI
PENLPLTAFDSEDKDVCKNYQAKDAFLGSLYEYSSRHPEYAVSVLLRAK
EEKLECCAKDDPHACYSTVFDKLLHLVDEPQNLKQNCQDFEKLGEYFGQNA
LVRRTRKVPQVSTPTLVEVSRLKVGTRCTKPECERMPCTEDYLSSLILNLRC
VLHEKTPVSEKVTRCTESLNVRRPFSALTPDEYVPKADEKLFTHADICTL
POTEQKIKQTALVEVLLHKPKKATEEQLKTVMENFVAVDKCAADDKEACFA
VEGPKLVSTQTALA
\end{verbatim}

Figure IV-5. BSA sequence coverage (red letters) for band 5 (top to bottom) from SDS-PAGE.

**Conclusion**

Although studies on model ornithine peroxycarbamates indicate lactam formation and backbone cleavage in the gas-phase, peptide studies did not follow this trend. Peptide ornithine peroxycarbamates did not appear to lead to selective backbone fragmentation at ornithine, and did not offer any advantages in this respect compared to lysine peroxycarbamates. Thus future efforts should
continue to focus on lysine peroxycarbamate chemistry. The hydrazinolysis of proteins via microwave irradiation appears to result in partial protein degradation, leading to large protein fragments. The SDS-PAGE analysis resulted in poor reproducibility and band resolution. 2D-SDS-PAGE may offer better band separation for proteomics analyses; however, a better method for hydrazine removal should be used to avoid streaky gel patterns and improve resolution. Nevertheless, hydrazinolysis has the potential to be used as a way to achieve partial cleavage of proteins prior to top-down sequencing analyses in order to facilitate protein fragmentation in the gas-phase, and yet retain the entire protein sequence.

Experimental Procedures

General

N-α-acetyl ornithine was purchased from Aldrich Chemical Co. (Milwaukee, WI, USA). Peptides were purchased from American Peptide Co. (Sunnyvale, CA, USA). BSA and HSA were obtained from Sigma (Milwaukee, WI, USA). All other reagents were purchased from Aldrich Chemical Co. Solvents were HPLC grade, purchased from Fisher Scientific (Pittsburgh PA, USA). t-Butyl-4-nitrophenyl-peroxycarbonate was synthesized according to previous work in our laboratory (See chapter II). The NH₄CO₃ buffer system used for Ornithine peroxycarbamate modifications consisted of a 1:1 mixture of 0.1 M NH₄CO₃ (pH 8.6) and acetonitrile. SDS-PAGE was performed using a NuPAGE® Novex Bis-Tris gel kit from Invitrogen (Carlsbad, CA, USA). LC-MS
analyses were carried out using a C-18 column (10 cm x 1 mm x 5um), from Grace Vydac (Hesperia, CA, USA). In some cases, peptide and protein samples were purified using PepClean™ C-18 Spin columns purchased from PIERCE (Rockford, IL, USA) prior to MS analysis. Microwave irradiation was performed using a domestic Emerson microwave operated on high setting (600 watts).

Model Ornithine Derivatives

The N-α-acetyl ornithine methyl and ethyl esters were prepared by combining N-α-acetyl ornithine (60 mg) with 1.2 equivalents of tosic acid in methanol (for the methyl ester) or ethanol (for the ethyl ester) and refluxing for 2 hours. After removal of excess alcohol, the esters were analyzed by ESI-MS. Preparation of the model ornithine peroxycarbamates was carried out by addition of 0.1 g of ornithine derivative to 2 equivalents of the peroxycarbonate IV-2 in 1ml of 1:1 0.1 NH₄HCO₃/acetonitrile.

Peptide Ornithine Derivatives

Conversion of arginine to ornithine

Peptides (0.2 mg) were dissolved in 1 ml of 70% aqueous hydrazine in glass vials. The vials were placed in a domestic microwave and irradiated at full power (600 watts) for 30 min. Aliquots of 100 µl were obtained at intervals of 5 min. The hydrazine was removed in vacuo and the samples were redisolved in
0.1% formic acid for LC-MS analysis. Complete conversion of arginine to ornithine was achieved after 10 minutes in most cases.

**Alkylation of N-termini and Lysines**

Alkylation of free amino groups prior to hydrazinolysis was carried out following literature procedures. Peptides (0.2 mg) were combined with 200 μl of 0.1M HEPES buffer (N-(2-hydroxyethyl)-piperazine-N'-2-ethanesulfonic acid), and 100 μl of 0.1 NaCNBH$_3$ in HEPES. Formaldehyde (100 μl, 0.1M) was added and the reaction was allowed to stir at room temperature for 2 hours. Solvent was reduced to 100 μl under vacuum and the samples were desalted using C-18 spin columns. Samples were dried under vacuum and subjected to treatment with hydrazine as described above. The excess hydrazine was removed by drying under vacuum followed by treatment with C-18 silica gel (C-18 spin columns). The eluted samples were then dried under vacuum for further derivatization.

**Peptide Ornithine Peroxycarbamates**

The alkylated peptide ornithine derivatives were dissolved in 20 μl of water and combined with 20 μl of 1:1 0.1 NH$_4$HCO$_3$/acetonitrile, followed by addition of the peroxycarbonate IV-2 (20 μl 0.1M in acetonitrile). The reactions were allowed to incubate for 20 minutes and the crude mixtures were used in LC-MS analyses.
Hydrazinolysis of Proteins

Several reactions were prepared with HSA and BSA (1 mg) dissolved in varied amounts of aqueous hydrazine (10 to 70% in 5% increments). The reactions were carried out in glass vials and these were subjected to microwave irradiation (600 watts) for 15 min. Because 35% aqueous hydrazine gave the best results (larger degradation fragments), another set of reactions were prepared with 35% aqueous hydrazine and irradiated at different time intervals (5 to 60 min, 5 min. intervals). We found that 35% aqueous hydrazine and irradiation for 10 min. gave the best partial degradation of both BSA and HSA. All reactions were subjected to vacuum evaporation over night to remove excess hydrazine. The samples were then redisolved in SDS buffer for PAGE analysis.

SDS-PAGE

Gel analysis was carried out using a Bis-Tris gel kit from Invitrogen (Carlsbad, CA, USA). Protein stock samples were approximately 1 mg/ml in SDS running buffer and 10 μl (10 μg) were loaded on to the gel. The gels were developed at a constant voltage of 200V for 30 min. Visualization was done with coomassie blue stain.

In-gel digestion

The different bands present on the gel were excised and cut into small pieces. These gel pieces were transferred to 1.5 ml eppendorf tubes for tryptic
digestion. First the samples were incubated with 100 μl of 100 mM NH₄HCO₃ buffer (pH = 7.0) for 15 min. They were then treated with 150 μl of 5 mM dithiolthreitol (DDT) in NH₄HCO₃ buffer for 15 min. at 55 °C to reduce disulfide linkages. Iodoacetamide was then added to alkylate free sulfhydryl groups (10 μl, 100 mM in NH₄HCO₃ buffer) and the samples were incubated in the absence of light for 15 min. at room temperature. After removal of reagent solution, the gel samples were washed 2X with a 1:1 acetonitrile/ NH₄HCO₃ buffer solution and dehydrated with 100μl of acetonitrile. After removal of solvent, the gels were covered with 50 μl of a trypsin solution (0.01 mg/μl) in 25 mM NH₄HCO₃ buffer. The samples were incubated at 37 °C over night. The liquid phase was collected and the gel was washed with 6:4 acetonitrile/0.1% TFA. The combined liquid phase was dried under vacuum, and the digest peptides resuspended in 50 μl of 0.1% formic acid for LC-MS analysis.

**Mass Spectrometry**

Mass spectrometry analyses of model ornithine derivatives and peptides were performed on a ThermoFinnigan (San Jose, CA) TSQ 7000 triple-quadrupole instrument, equipped with an Electrospray Ionization (ESI) source and a Waters (Milford, MA) Alliance HPLC system. The capillary temperature was kept at 200 °C with a voltage of 20V. The electrospray needle voltage was 4.5 kV, and the tube-lens voltage was maintained between 70 and 100 V. The sheath and auxiliary gases (N₂) were adjusted to maximize signal. The offset voltage in CID experiments varied between 20 and 35 eV. Samples were
introduced into the ESI source as either lithium solutions or acidic solutions at a rate of 10 to 20 μl/min in direct liquid infusion experiments. For LC-MS analyses of peptides, a Supelco (Bellefonte, PA) Discovery C-18 column was used (25cm x 4.6mm, 5 μm). The solvent conditions consisted of a 20 min. gradient from 5% to 50% solvent B (95% acetonitrile, 5% water, 0.05% TFA) into solvent A (95% water, 5% acetonitrile, 0.05% TFA). LC-MS/MS analyses of peptide derivatives were performed on a ThermoFinnigan (San Jose, CA) LCQ Deca XP ion trap instrument, equipped with an ESI source, and a HP Agilent (Palo Alto, CA) HPLC system. Instrument parameters were tuned using angiotensin II peptide (5 pmol) to maximize signal. LC-MS/MS analyses of the hydrazine treated BSA and HSA tryptic digests were performed using a Grace Vydac (Hesperia, CA) C-18 column (10cm x 1mm x 5 μm). Solvent conditions consisted of a 30 min. gradient from 100% solvent A (5% acetonitrile, 95% water, 0.5% acetic acid) to 100% solvent B (5% water, 95% acetonitrile, 0.5% acetic acid).

References


10. Blodgett, J. K.; Loudon, G. M. "Direct Cleavage Versus Transpeptidation in the Autodecomposition of Peptides Containing 2,4-Diaminobutanoic Acid (Daba) and 2,3-Diaminopropanoic Acid (Dapa) Residues - Specific Cleavage of Dapa-Containing Peptides." *Journal of the American Chemical Society* **1989**, 111, (17), 6813-6821.


The development of new and improved mass spectrometry techniques over the last decade has allowed the rapid improvement of proteomics analyses. Bottom-up shot-gun proteomics has become the most automated, widely used technique in protein identification. The use of multi-dimensional chromatographic separations online with mass spectrometry, allows for hightroughput analyses. In addition, the availability of large protein databases and sophisticated database search algorithms make possible identification of multiple proteins in a biological system. However, protein identification is only a part of proteomics research, a great part of proteomics is focused on protein function and interactions, which are significantly influenced by post-translational modifications. Protein identification in bottom-up proteomics requires only a few matching peptide sequences. However to obtain information on post-translational modifications it is advantageous to have access to the entire protein sequence. This is seldom possible using bottom-up approaches since a great part of the protein may be lost in the analysis. In order to solve this problem, top-down strategies have been developed more recently, where the entire protein sequence is interrogated by mass spectrometry. This requires fragmentation of intact proteins using powerful mass spectrometers such as an FT-ICR. Although such techniques have proven to be extremely useful in retaining the
intact protein and therefore allowing 100% sequence coverage, fragmentation of large proteins remains challenging. Thus techniques which facilitate this process and yet minimize protein degradation prior to mass spectrometry analysis are very valuable. The work presented in this dissertation has the goal of improving top-down proteomics by developing techniques to enhance protein fragmentation in the mass spectrometer.

**Lysine Peroxycarbamates**

Research on lysine peroxycarbamates involves the chemical modification of lysine residues with a labile peroxy group to induce fragmentation in the gas-phase at or near the site of modification. CID studies carried out on a model lysine peroxycarbamate showed that this dissociated to form an aminyl radical, which leads to backbone and side-chain fragmentation (Scheme V-1). The study was extended to peptides containing lysine peroxycarbamates. Where CID resulted in the formation of a-, b-, c-, and z-type ions. These ions however are accompanied by fragments resulting from side chain degradation which do not result in backbone cleavage.

MALDI studies on peroxycarbamate modified peptides differed from ESI studies in that peptide fragmentation occurred at the source without the need for CID. However, the intact peptide peroxycarbamate adducts were not observed. MALDI studies offer the advantage of working with singly charged species, which facilitates data interpretation. In addition, peptide fragments are obtained without an MS/MS experiment. However a combination of both ESI and MALDI would
provide a more complete analysis by allowing interrogation of the intact peroxycarbamate adducts and yet obtain clean fragmentation spectra.

![Chemical structure and fragmentation diagram]

**Scheme V-1.** Modification of lysine model as a peroxycarbamate and fragmentation by CID.

Studies on lysine peroxycarbamates of ribonuclease A (~ 14 kDa) were more complex than peptide analyses due to a greater number of lysine modifications. In order to simply interpretation of fragmentation data, we adjusted modification conditions to allow for only 1 to 2 lysine peroxycarbamate adducts, and isolated these by cation-exchange chromatography. ESI-CID of these adducts did not result in detectable protein fragmentation. This is possibly due to the presence of multiple lysine peroxycarbamate regioisomers. MALDI studies also did no lead to fragmentation, however dissociation of the t-butyl
peroxy group was observed. More powerful mass spectrometers such as an FT-ICR may be necessary to achieve protein fragmentation in this manner.

N-terminal peroxycarbamates

Despite all the advances in shot-gun proteomics, de novo protein identification continues to be a useful tool for the study of proteins derived from unknown genomes, as well as for highly post-translationally modified proteins. De novo interpretation of MS/MS spectra can be cumbersome, and for this reason several techniques have been developed to manipulate peptide fragmentation patterns. Our work involves the extension of the lysine peroxycarbamate chemistry to the modification of N-termini. Fragmentation of these N-terminal peroxycarbamates would lead the way to a mass spectrometry based N-terminus identification technique that would facilitate de novo analyses.

We found that CID experiments on peptides modified as peroxycarbamates at the N-terminus lead to cleavage of the N-terminal side chain. This neutral loss can then be used to determine the identity of the amino acid (Scheme V-2). Isobaric peptides were distinguishable in this manner without the need to elucidate the entire sequence.

![Scheme V-2. CID fragmentation of N-terminal peroxycarbamates.](image-url)
Furthermore, cleavage of the N-terminal side chain of peptides led to a labeled b-ion series in MS\(^3\) experiments (Scheme V-3). N-terminal amino acid identification was also achieved in the analysis of modified peptides from myoglobin tryptic digest, which suggested this method was suitable in a proteomics setting.

![Diagram of peptide cleavage and labeling]

Scheme V-3. Labeling of b-ion series by MS\(^3\) of N-terminal peroxycarbamates.

N-terminal side chain loss was also observed in MALDI studies. However, as for the lysine peroxycarbamates, MALDI did not produce intact peptide peroxycarbamate adducts, and N-terminal side chain loss occurred at the source. The presence of singly charged ions facilitated identification of N-terminal amino acids. All of these findings suggest this N-terminus identification technique has the potential to be used as a tool in de novo sequencing. In addition, the distinction of isobaric peptides by the N-terminus is a useful technique in peptide mass fingerprinting, where identification of the N-termini can be used to narrow database searches.\(^{14}\)
Ornithine Peroxycarbamates

Although lysine peroxycarbamates lead to peptide backbone fragmentation upon CID, there is a significant amount of fragmentation that does not lead to backbone cleavage. With the goal of obtaining more selective peptide cleavage, we investigated the CID fragmentation patterns of ornithine peroxycarbamates. Ornithine is an unnatural amino acid which plays an important role in the urea cycle.\textsuperscript{18} It has been found in the past that introduction of ornithine and 2,4-diaminobutanoic acid (DABA) into peptides can lead to peptide cleavage in solution by lactamization of the side chain (Scheme V-4).\textsuperscript{19-21}

With this in mind, we examined the effects of introducing ornithine and ornithine peroxycarbamates in peptides on CID fragmentation.

\textbf{Scheme V-4.} Lactamization of Ornithyl peptides.

We introduced ornithine in peptides by hydrazinolysis, where arginine side chains are hydrolyzed to give ornithine (Scheme V-5).\textsuperscript{22} In CID experiments of ornithine peroxycarbamate model derivatives we found significant fragmentation at the C-terminus by lactamization of the side chain. However these results were unique to the model amino acid and did not carry to the study of peptides containing ornithine peroxycarbamates. In the case of peptides, CID
fragmentation patterns were greatly dependant on the peptide sequence and we did not observe tendency for cleavage at the ornithine or ornithine peroxycarbamate sites. Thus the introduction of ornithine peroxycarbamates in peptides did not provide any advantages in peptide backbone cleavage by CID compared to lysine peroxycarbmates.

![Scheme V-5. Conversion of arginine to ornithine.](image)

**Hydrazinolysis of Proteins**

Although fragmentation of intact proteins by FTICR has come a long way in the past few years, direct fragmentation of large proteins remains challenging.\(^5\) For this reason, techniques that facilitate this process are necessary. Hydrazinolysis of proteins has been investigated in the past by Akabori and coworkers as a C-terminus identification technique, leading to naming of this process the Akabori reaction (Scheme V-6).\(^{23, 24}\) Other researchers have found that asparagine linkages are particularly susceptible to hydrazinolysis.\(^{25, 26}\) In this work we explore the use of hydrazinolysis as a technique for partially degrading large proteins with the goal of improving top-down proteomics approaches.
We found that using a modified akabori reaction procedure involving microwave irradiation, allowed for better control of the extent of protein cleavage.\textsuperscript{22} By adjusting reaction time and hydrazine concentration, we were able to obtain partial degradation of BSA and HSA into large protein fragments that can then be used for top-down analyses. Complete characterization of these protein fragments proved to be challenging; Nevertheless, hydrazinolysis has the potential to be used as a chemical protein degradation technique to facilitate FTICR analyses, and yet retain access to the entire protein sequence.

\textbf{Scheme V-6}. The Akabori reaction for identification of C-terminus.
Concluding Remarks

The research presented in this dissertation introduces valuable concepts for the improvement of proteomics. We have shown that lysine peroxycarbamates induce site-specific peptide fragmentation in the gas-phase. We have also shown that N-termini peroxycarbamates can be used to identify N-terminal amino acids and label b-ion series in MS/MS spectra of peptides. Although the lysine peroxycarbamate chemistry at the protein level remains to be further explored, preliminary results show that it is feasible to control the extent of lysine modification in order to aid MS data interpretation. Future studies that may take this technique further should involve the analysis of modified proteins by FT-ICR. This mass spectrometry technique has been successful in obtaining MS/MS spectra of proteins up to 60 KDa.\textsuperscript{5, 27}

Our studies on hydrazinolysis of BSA and HSA show that this reaction can be used to break down the proteins into smaller, yet still protein-size sections. FT-ICR offers best MS/MS fragmentation efficiency for small proteins (5 – 20 KDa).\textsuperscript{5, 27} Therefore by pre-treating large proteins with hydrazine we can greatly increase the amount of sequence information obtained in a top-down analysis. Future work in this area will require an extensive proteomics study to determine consistent patterns of protein cleavage by hydrazine.
References


21. Blodgett, J. K.; Loudon, G. M. "Direct Cleavage Versus Transpeptidation in the Autodecomposition of Peptides Containing 2,4-Diaminobutanoic
Acid (Daba) and 2,3-Diaminopropanoic Acid (Dapa) Residues - Specific Cleavage of Dapa-Containing Peptides. "Journal of the American Chemical Society 1989, 111, (17), 6813-6821.


