DEVELOPMENT OF NOVEL ION MOBILITY-MASS SPECTROMETRY SHIFT REAGENTS FOR PROTEOMIC APPLICATIONS

By

THOMAS J. KERR

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

May, 2011

Nashville, Tennessee

Approved:

Professor John A. McLean

Professor David E. Cliffel

Professor Eva Harth

Professor Chris Janetopoulos

Dedicated to my family, especially my parents,

for their constant support

ACKNOWLEDGEMENTS

I would like to thank Dr. John McLean and the entire McLean research group for their constant advice, help, suggestions, and encouragement. It was a great honor to be able to help get the research group started and work on some of the group's first experiments involving peptide labeling. It was a great experience seeing the lab grow from a handful of students to its present size.

The members of my committee, Dr. David Cliffel, Dr. Eva Harth, and Dr. Chris Janetopoulos, have provided many tips and ideas concerning my research. I am especially grateful for the time and effort spent helping and encouraging me as I completed the requirements for my degree.

I also acknowledge the support of friends and family. They have supported me throughout my education experience, first to school at Wofford in South Carolina, and also through graduate school at Vanderbilt. Without their constant support, graduate school would have been much more difficult.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AQUA	absolute quantitation
CBZ	carboxybenzyl protecting group
CCS	collision cross section
CID	collision induced dissociation
CYC	horse cytochrome c
Da	Dalton
DCM	dichloromethane
DHB	2,5-dihydroxybenzoic acid
DMF	dimethylformamide
DOTA	1,4,7,10-tetra-azacyclododecane-*N*, *N*', *N*'', *N*'''-tetraacetic acid
DTIMS	drift tube ion mobility spectrometry
DTT	dithiothreitol
ECAT	element-coded affinity tags
ESI	electrospray ionization
Fmoc	fluoronylmethyloxycarbonyl protecting group
FMOC-Arg-OH	Fmoc protected arginine residue

HPLC	high performance liquid chromatography
ICAT	isotope coded affinity tags
iTRAQ	isobaric tags for relative and absolute quantitation
IM-MS	ion mobility-mass spectrometry
IMS	ion mobility spectrometry
Ln(III)	lanthanide metal
LYSC	hen egg white Lysozyme
MALDI	matrix assisted laser desorption/ionization
MS	mass spectrometry
MSI	mass spectrometric imaging
MS/MS	tandem mass spectrometry
MYG	horse heart myoglobin
m/z	mass to charge ratio
NHS	*N*-hydroxysuccinimide
NMR	nuclear magnetic resonance spectroscopy
SATP	*N*-succinimidyl *S*-acetylthiopropionate
TCEP	tris-2-carboxyethylphosphine
TOF	time-of-flight

TWIMS	traveling wave ion mobility spectrometry

CHAPTER I

INTRODUCTION

1.1 Mass spectrometry-based proteomics

Contemporary proteomics faces increasing challenges as analysis of samples of smaller sample size and lower abundance proteins is required. Since many biological processes are regulated by changing protein expression profiles, a great deal of current research focuses on the area of relative quantitation. Analysis of protein structure and sequence, including any post-translational modification (PTMs), enhances the understanding of the related biological processes. Mass spectrometry (MS) is well suited to address some of these challenges and the development of novel mass spectrometry-based methods has led to a rapid increase in the knowledge generated from proteomic experiments.

A great deal of current proteomics research focuses on the quantitation of protein expression profiles, identification of PTMs, and the correlation of this information to the specific state of a biological system. Traditionally, separation and quantitation of proteins has been accomplished using 2D gel-based methods which require dyes, fluorophores, or radioactive labels. Typically proteins are separated on a gel, which is then stained using a dye (*e.g.* coomassie blue or silver stain) in order to visualize the proteins. Alternatively, radioactive labels may be covalently attached to proteins prior to gel separation. Subsequent to separation, protein spots are excised and quantified using a radiation counter.[1]

These methods can provide a general indication of protein relative abundance, however gel based methods also face several limitations. Low abundance proteins, for

example, are difficult to quantify using these methods. Further sample preparation can remove high abundance proteins, but low abundance proteins can still be hard to quantify. Furthermore, proteins of similar size and isoelectric point may overlap in the gel spot, leading to misidentification or overestimation of its quantity. Mass spectrometry-based analysis offers several advantages over gel-based methods.

Contemporary MS-based methods for protein quantitation include stable isotope incorporation by enzymatic, metabolic, or covalent modification methods, in addition to label free methods. Stable isotope incorporation allows comparison of a number of control and experimental protein expression profiles while retaining similar structure, and thus ionization efficiency. Enzymatic labeling is accomplished by proteolysis of the protein in a solution of $H_2{}^{18}O$, resulting in the incorporation of an ^{18}O stable isotope in nearly every peptide. Comparison (using MS) of proteins digested normally and using $H_2{}^{18}O$ allows for relative quantitation of the protein expression profiles under experimental conditions.[2-4]

^{18}O labeling provides a relatively straightforward method of quantitation; however this method suffers from several limitations. The primary challenge of ^{18}O labeling is the pH dependence which may result in the incorporation of multiple ^{18}O isotopes in a single peptide. Multiple labeling splits peptide signal into several channels, impeding peak area analysis. Despite these challenges, ^{18}O labeling provides a relatively low cost and easy method for the relative quantitation of peptides. Additionally, labeling may be optimized for peptide length by the use of different proteases.

A similar method, which incorporates stable isotopes metabolically, is termed stable isotope labeling by amino acids in cell culture, or SILAC.[5-6] The SILAC method is similar to ^{18}O labeling in that a stable isotope is incorporated directly into the protein of

interest. In this case, cells are cultured in the presence of amino acids that have incorporated a heavy isotope, commonly ^{13}C or ^{15}N. Control cell samples are grown in normal amino acid media. By mixing purified and proteolyzed samples, relative protein expression profiles may be elucidated. The main limitations of SILAC are the time consuming need for cells to be cultured, and the requirement that samples be biologically viable. Clinical fluid samples that contain proteins, for example, cannot be analyzed by SILAC.

Another method similar to SILAC is that of absolute quantitation (AQUA).[7] The AQUA method of quantitation is similar to SILAC in that peptides incorporating heavy amino acids are used, however in AQUA recombinant peptides are used rather than peptides purified from cell culture. The use of recombinant isotopically heavy peptides removes the need for cell culture prior to experimental analysis and allows for quantitation of non-living samples such as clinical fluid samples.

Typical AQUA experiments proceed by spiking in a known amount of an isotopically heavy synthesized peptide of the same sequence as the peptide of interest (including all PTMs). A comparison of the peak areas of the experimental peptide to the isotopically heavy peptide provides excellent absolute quantitation data since their ionization efficiencies are nearly identical. The main drawback of this method is that the sequence of the peptide of interest must be known *a priori* before quantitation experiments. Additionally, an isotopically heavy version of each peptide to be quantified must be synthesized. For small experiments (a couple of peptides) synthesis is not an issue, however it can quickly become cost prohibitive when considering a proteome wide quantitation experiment.

Incorporation of stable isotopes by covalent modification of peptides is another method of mass spectrometry-based protein quantitation. Isotope coded affinity tagging (ICAT)[8] experiments proceed by the covalent addition of a label incorporating ^{2}H or ^{13}C isotopes or labels of natural isotopic abundance. For example, the isotopically heavy label may be used to modify proteins exposed to experimental conditions while the isotopically light label is used to modify proteins in the control sample. Labeled samples are then combined, proteolyzed, and analyzed by MS. Relative quantitation information is obtained by comparing the peak areas of the heavy versus light labeled peptides.

Current ICAT methods rely on ^{13}C isotopes to overcome the challenge of differing elution profiles of heavy *versus* light labeled peptides in liquid chromatography (LC) methods.[9] The ICAT method (Figure 1.1) also offers the advantage of being able to analyze multiple experimental samples simultaneously, as opposed to the single experimental condition required by metabolic and enzymatic labeling. Traditional ICAT labels (such as Figure 1.1a) are limited to selective modification of the thiol group of cysteine, a relatively rare amino acid (*ca.* 3% of typical proteins in invertebrates) which some protein families lack altogether. However, an analogous label has been developed which provides ICAT like results with the modification of primary amine sites[10] (*i.e.* the N-terminus and lysine residues). Despite these improvements, the ICAT method still suffers from a limited number of simultaneous samples and a small (*i.e.* 8 Da), fixed mass shift, which can hamper the analysis of large peptides.

An analogous method termed isobaric tags for relative and absolute quantitation (iTRAQ) offers similar data in a slightly different manner.[11-12] Using iTRAQ, peptides are labeled with isobaric tags (Figure 1.2) that differ in the incorporation of stable isotopes (^{13}C or ^{15}N) in a diagnostic reporter fragment region, which may be cleaved by tandem

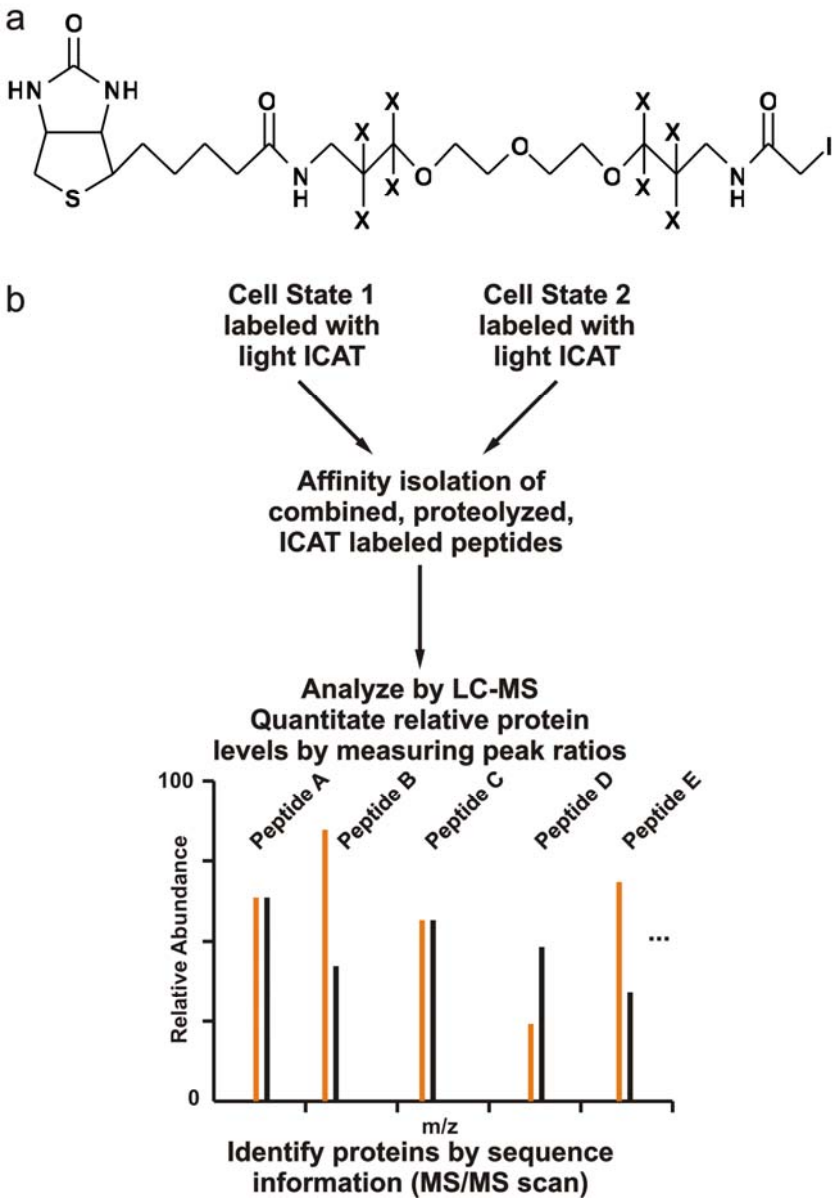


Figure 1.1. A generalized scheme illustrating the ICAT method of relative quantitation. Proteins are covalently tagged at reduced thiol sites using the tag structure shown in (a). Locations marked with an X indicate sites of either hydrogen or deuterium incorporation. Current methods rely on the incorporation of ^{13}C in the structure of the tag instead of deuterium. The progression of relative quantitation experiments is shown in (b). Adapted with permission from *Nature Biotechnology*, **1999**, 17, 994-999. © (2003) Macmillan Magazines Ltd.

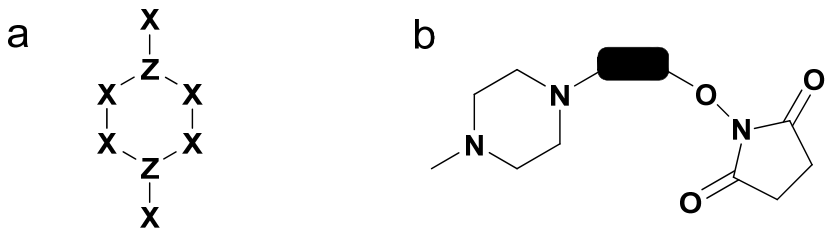


Figure 1.2. An illustration of iTRAQ isobaric tagging reagents used for the relative quantitation of peptides and proteins. (a) A schematic of the structure of the reporter region of the tag (113-121 m/z) which provides relative quantitation data following MS/MS analysis. Atoms indicated by X are either ^{12}C or ^{13}C. Atoms represented by Z bare either ^{14}N or ^{15}N. (b) An illustration of a complete iTRAQ tag which consists of the reporter region (left), a linker region to maintain the tags as isobaric (middle), and a chemically selective region for covalent modification of proteins and peptides (right). Adapted with permission from ref. 12.

MS (MS/MS) analysis. The different isobaric labels incorporate a mass balance region (not analyzed for relative quantitation data) that corrects for the different reporter fragment masses among the different tags. iTRAQ labels typically modify primary amine sites, thus labeling nearly every peptide after proteolysis.

A typical iTRAQ experiment proceeds by proteolysis of the control and experimental samples. Each sample is then labeled at primary amine sites, each with a different iTRAQ reagent (*i.e.* a different reporter ion mass). Samples are subsequently mixed and analyzed by MS. Relative quantitation information on the peptides of interest is obtained by acquiring MS/MS data for each peptide. The peak areas of the reporter ions are compared to give the relative protein expression profile. iTRAQ suffers from several limitations, chiefly that in large scale proteomics experiments MS/MS information is typically only obtained on relatively high abundance peptides. Data dependent scanning often skips over low abundance peptides since fragmentation spectra typically have low signal.

1.2 Fundamentals of Ion Mobility-Mass Spectrometry

Ion mobility spectrometry (IMS) has been around in some form for over a century.[21] It is important to note that IMS has been referred to by several names over the years, including plasma chromatography and ion chromatography. Early in the development of IMS, it was primarily used for applications in fundamental physics in the study of atomic ions.[13] These experiments were typically attempts to elucidate atomic ion properties. Eventually, small molecules were also studied using IMS methods and near mid century analysis of larger molecules with applications in chemistry were becoming a possibility.

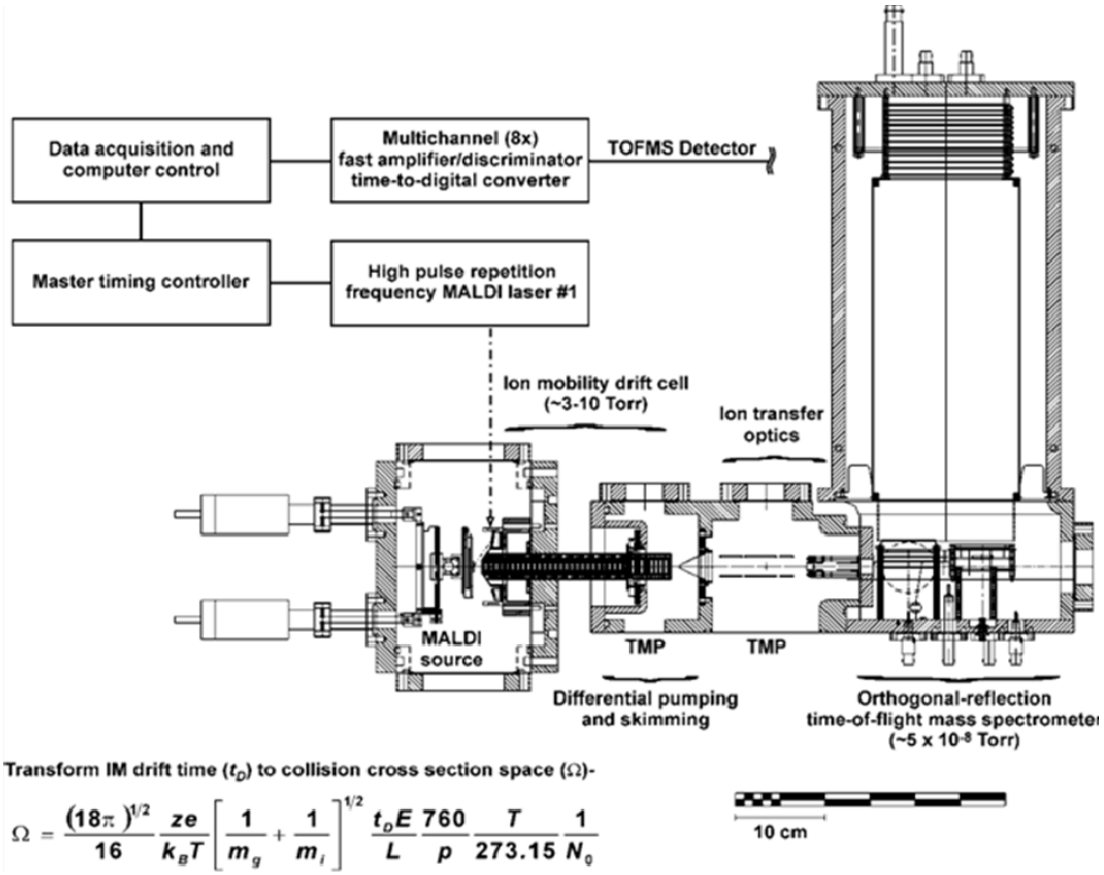


Figure 1.3 A schematic of the direct current (DC) drift tube IM-MS instrument used for collision cross section measurements. Ions are generated in the MALDI source and then travel through the drift tube where they are separated according to size. A differential skimming and pumping region allows for the high pressure differential between the drift cell (~ 4 Torr) and the TOF (~10^{-7} Torr). Ions are then pulsed orthogonally into the TOF for mass separation.[32]

In the 1960s, IMS was integrated with MS as a hyphenated method offering both size selection and mass identification through a TOF or quadrupole. The development of electrospray ionization (ESI) and matrix assisted laser desorption/ionization (MALDI) in the late 1980s and early 1990s made analysis of larger biologically relevant molecules a possibility.[14] Early work on analysis of biomolecules using IM-MS paved the way for the more complex samples seen in the 2000s. Eventually complex samples, such as whole cell lysates, were analyzed using IM-MS.[33]

Drift tube ion mobility instruments[15] (Figure 1.3) consist of an ion source, an IMS drift cell, and a detector. Ions are generated in the source and are then directed into the IMS drift cell. The drift cell consists of a stacked set of conducting rings which have a low voltage gradient from front to back. The drift tube is also pressurized (usually 4-10 Torr) with a background buffer gas, commonly helium or some other non-reactive gas. Ions directed into the drift cell are pulled through the gas under the direction of a weak electrostatic field. As the ions traverse the drift cell they experience collisions with the buffer gas, slowing them down. Larger ions experience more collisions, thus larger ions exit the drift cell later than smaller ions.

Another method of IMS termed traveling wave IMS (TWIMS)[16] is used in the Waters Synapt G2 instrument (Figure 1.4). It operates in a similar fashion except that pulsed voltage waves guide the ions through the drift cell. That is, variable low voltage waves travel down the length of the drift cell. As these waves pass, ions are pushed along for a short distance before rolling over the top of the wave due to collisions with the buffer gas impeding ion progress. The size of the ion, and thus the number of collisions with the buffer gas, determine how readily ions are moved by the traveling

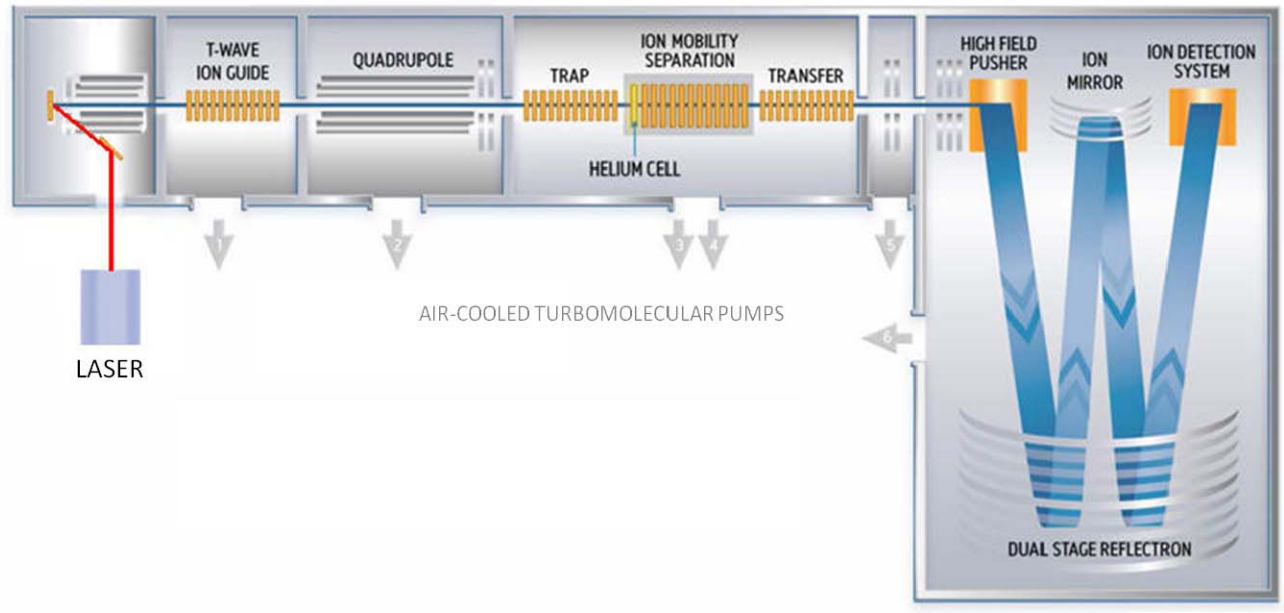


Figure 1.4. A schematic diagram of the MALDI-TWIM-TOFMS instrument (Waters Synapt G2 HDMS, Waters Corp., Manchester, UK).

waves. Larger ions experience more collisions with the buffer gas slowing them down so several wave passes are required to reach the end of the drift cell. Parameters such as the wave height (voltage) and velocity can be optimized to achieve excellent separation in the IM dimension.

Subsequent to IM separation (whether by DTIMS or TWIMS), ions are pulsed orthogonally into a TOFMS for mass analysis. Since the time scale of mass analysis is much smaller than the time scale of IM separation (10s μs versus 10s ms, respectively) several mass spectra may be taken over the IM elution profile of a single drift cell injection. By taking multiple mass spectra for each IM elution profile and offsetting subsequent TOF extractions, two-dimensional IM-MS conformation space data may be obtained using a single detector. Interleaving in this manner allows for rapid separation and mass identification of molecules.

Since DTIMS uses a steady, linear electrostatic gradient for separation of ions, it is possible to empirically measure the collision cross section of those ions. The kinetic theory of gasses[17] and the hard sphere model[18-19] (HSM, Figure 1.5) approximate the consequences of ion-neutral collisions in relation to collision cross section. The HSM treats the ion as a hard sphere that encounters elastic collisions with the buffer gas molecules where only momentum is transferred. Comparisons between HSM calculated cross sections and theoretical cross sections show that the hard sphere model is most relevant at masses greater than *ca*. 500, thus the HSM works well for peptide cross section analysis. Empirically determined cross sections may be calculated using equation 1, if the drift time, temperature, pressure of the buffer gas, mass of the ion, length of the drift cell, and the voltage gradient across the drift cell are known.[17]

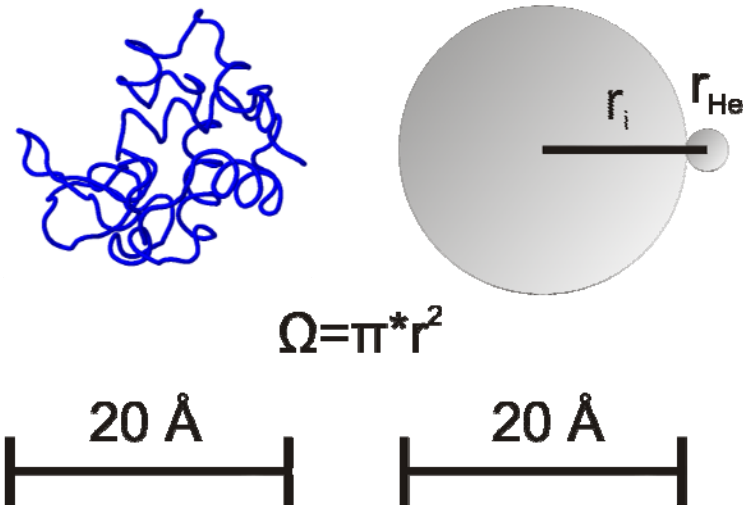


Figure 1.5. A schematic illustrating the calculation of collision cross section. Experimenatlly, the collision cross section may be described as a circular area described by the radius of the ion and buffer gas combined, as shown above.

$$\Omega = \frac{(18\pi)^{1/2}}{16}\frac{ze}{(k_b T)^{1/2}}\left[\frac{1}{m_I}+\frac{1}{m_B}\right]^{1/2}\frac{t_D E}{L}\frac{760}{P}\frac{T}{273.2}\frac{1}{N} \qquad \textbf{Eq. 1.1}$$

The specific buffer gas used can have a dramatic effect on the quality of data obtained and different gasses have advantages and limitations.[18] Helium is the preferred gas for collision cross section measurements since it is light, small, unreactive, and has a low polarizability. Gases with a higher polarizability will interact with ions longer, extending the drift time and reducing the applicability of the hard sphere model since ion-neutral collision are no longer purely elastic. Other gasses such as argon, or methane, as well as changes in the electric field, may be used to optimize IM separation,[20] however helium and nitrogen are used exclusively in this work.

1.3 Applications of IM-MS to proteomics

The development of soft ionization sources (ESI and MALDI) led to increased interest in analysis of biomolecules by mass spectrometric methods.[21-24] The mating of IMS and MS enhanced the data collected from analysis of biological samples. Initially, research focused on peptides, but was expanded to analysis of intact proteins by the late 1990s and early 2000s. During this time it was discovered that biomolecules tend to separate based on class (*e.g.* lipid versus peptide) in IM-MS data, commonly referred to as conformation space.[25-26] Further research outline the boundary of an area specific to peptides, the peptide correlation band. This work illustrated that 99% of peptide signals fall within ±10% of the midline of this correlation band.[27] While ion mobility separations

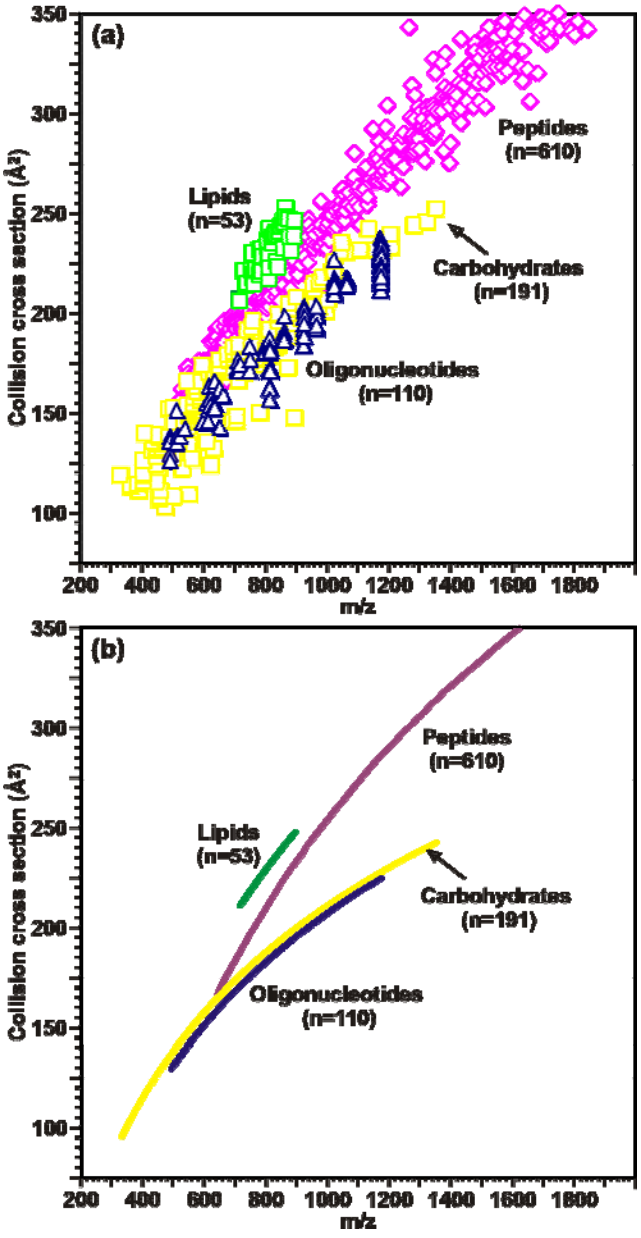


Figure 1.6. A graphical representation of the area different classes occupy in IM-MS conformation space. Note that peptides occupy a relatively narrow band of conformation space and are separated from both lipids and carbohydrates. Adapted with permission from ref. 31.

are possible using either ion source, MALDI was preferred for this work due to its production of singly charged ions nearly exclusively, as well as for its tolerance of salts.

Further analysis showed that each class of biomolecule occupies a specific area of conformation space under a given set of experimental conditions (Figure 1.6). Thus, signals from different classes of biomolecules may be separated in two-dimensional

space. This separation by class is due to the gas phase packing efficiency of these molecules which increases in the order lipids < peptides < carbohydrates < oligonucleotides.[21] Structural information about a specific signal may be inferred from its specific location within the correlation band. For example, phosphopeptides tend to appear in the lower portions of the peptide correlation band;[28] however a signal's location within the correlation band should not be used for confident assignment of signals. Further analysis, by MS/MS for example, should be completed to obtain confident identification information.

The IM-MS separation dimensions are highly correlated, as opposed to highly orthogonal separation dimensions such as isoelectric focusing and gel electrophoresis in 2D gel methods. Dimensional analysis indicates that mass and collision cross section are related since biomolecules are made up of a limited number of different elements (*i.e.* C, H, O, P, S, and N).[13] Thus, biopolymer mass scales as mass (or length) cubed. Surface area (collision cross section, the IM dimension) scales as length squared. This correlation gives biomolecular correlation bands a curved appearance.

Since the biomolecular correlation bands are fairly narrow (±10% from the midline)[27] it is possible to shift specific signals out of their correlation band into an area that is not predicted to contain signals of that particular biomolecular class. Due to the curved nature of the correlation bands, it is readily apparent that there are two options

for the direction of shift. Low density molecules that have high collision cross section compared to mass have been studied previously.[29-30] Molecules such as crown ethers, dendrimers, or polymer based reagents will shift specific signals to an area above the correlation band. Furthermore, the crown ether shift reagents[29] are non-covalent, allowing for dissociation prior to MS detection in order to retain the original peptide mass.

A second option is using high density reagents that have a low collision cross section (compared to mass) to shift specific signals to an area below the correlation band. Development of high density, lanthanide-based shift reagents are discussed in this report.

1.4 Summary and objectives

For my dissertation research, I aimed to develop and explore the application of novel ion mobility shift reagents for proteomic applications. These shift reagents were envisioned to provide a rapid method for screening complex proteomic samples for specific peptide functionalities. In addition, labels would be able to provide both a frame of reference for tandem MS sequence identification, as well as a method of relative quantitation of peptides and proteins. Toward this goal, cysteine and primary amine specific shift reagent applications were explored for their potential in aiding the confident identification of peptide signals. These shift reagents do indeed show great utility in the relative quantitation of peptide samples in both one and two dimensional MS methods. Further development of the shift reagent labeling method included exploration of their utility for *in situ* labeling of tissue sections. Some specific milestones and the chapter in which each is discussed follows:

1. Do lanthanide based shift reagents provide an acceptable method of the relative quantitation of peptides? Does the high degree of flexibility of these reagents offer advantages over less flexible quantitation reagents? Do these labels dissociate under CID conditions? This is addressed in Chapter 2: relative quantitation of peptide functionality using metal chelation labels.

2. Do lanthanide based shift reagents offer a useful degree of shift from the established peptide correlation band? Is this shift significant enough to avoid confusion between labeled and unlabeled peptide signals? This is addressed in Chapter 3: application of metal chelation labels to relative quantitation of peptides using ion mobility-mass spectrometry.

3. Do these shift reagent labels offer a reasonable method for the relative quantitation of proteins? Do these labels provide a reasonable method for the relative quantitation of a proteolyzed mixture of several proteins? How does this compare with current methods of relative quantitation? This is addressed in Chapter 2: Relative Quantitation of Peptide Functionality Using Metal Chelation Labels.

4. Is it possible to do multiplex labeling of multiple functionalities on a single peptide? Does the use of a different lanthanide for each functionality offer advantages or aid in sequence assignment? This is addressed in Chapter 4: Multiplexed Analysis of Peptide Functionality Using Metal Chelation Labels.

The primary goal of this dissertation research was to develop novel IM-MS shift reagents to allow rapid identification of specific peptide functionalities in conformation space. This was achieved using lanthanide coordinating labels. Initial proof of concept experiments using pure peptides indicated that the lanthanide-based labels provided a

significant amount of shift from the peptide correlation band. From there, the project was expanded to include relative quantitation of pure peptides, labeling of complex mixtures (simulated by tryptic protein digests), relative quantitation of protein digests, relative quantitation of proteolyzed mixtures of proteins, and multiplex labeling of different functionality (*i.e.* both the cysteine side chain and primary amines) on the same peptide using different lanthanide metals. In addition, both *in situ* labeling of mouse brain tissue sections and the development of an ionization enhancing label for use in tagging phosphopeptides were explored. Description of these two areas are described in Appendix A and B.

References

1. Aebersold, R.; Mann, M., "Mass spectrometry-based Proteomics." *Nature* **2003**, 422, 198-207.

2. Desiderio, D.M.; Kai, M. "Preparation of Stable Isotope-incorporated Peptide Internal Standards for Field Desorption Mass Spectrometry Quantification of Peptides in Biologic Tissue." *Biomed. Mass Spectrom.* **1983**, 10, 471-479.

3. Yao, X.; Freas, A.; Ramirez, J.; Demirev, P.A.; Fenselau, C. "Proteolytic ^{18}O Labeling for Comparative Proteomics: Model Studies with Two Serotypes of Adenovirus." *Anal. Chem.* **2001**, 73, 2836-2842.

4. Miyagi, M.; Rao, K.C.S. "Proteolytic ^{18}O-Labeling Strategies for Quantitative Proteomics." *Mass Spectrom. Rev.* **2007**, 26, 121-136.

5. Oda, Y.; Huang, K.; Cross, F.R.; Cowburn, D.; Chait, B.T. "Accurate Quantitation of Protein Expression and Site-specific Phosphorylation." *Proc. Nat. Acad. Sci. USA* **1999**, 96, 6591-6596.

6. Paša-Tolić, L.; Jensen, P.K.; Anderson, G.A.; Lipton, M.S.; Peden, K.K.; Martinovis, S.; Tolic, N.; Bruce, J.E.; Smith, R.D. "High Throughput Proteome-wide Precision Measurements of Protein Expression Using Mass Spectrometry." *J. Am. Chem. Soc.* **1999**, 121, 7949-7950.

7. Gerber, S.A.; Rush, J.; Stemman, O.; Kirschner, M.W.; Gygi S.P. "Absolute Quantification of Proteins and Phosphoproteins from Cell Lysates by Tandem MS." *Proc. Nat. Acad. Sci. USA* **2003**, 100, 6940-6945.

8. Gygi, S.P.; Rist, B.; Gerber, S.A.; Turecek, F.; Gelb, M.H.; Aebersold, R. "Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags." *Nat. Biotech.* **1999**, 17, 994-999.

9. Yi, E.C.; Li, X.; Cooke, K.; Lee, H.; Raught, B.; Page, A.; Aneliunas, V.; Hieter, P.; Goodlett, D.R.; Aebersold, R. "Increased Quantitative Proteome Coverage with $^{13}C/^{12}C$-based, Acid-cleavable Isotope-coded Affinity Tag Reagent and Modified Data Acquisition Scheme." *Proteomics* **2005**, 5, 380-387.

10. Hoang, V.M.; Conrads, T.P.; Veenstra, T.D.; Blonder, J.; Terunuma, A.; Vogel, J.C.; Fisher, R.J. "Quantitative Proteomics Employing Primary Amine Affinity Tags." *J. of Biomoll. Tech.* **2003**, 14, 216-223.

11. Ross, P.L.; Huang, Y.N.; Marchese, J.N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D.J. "Multiplexed Protein Quantitation in *Saccharomyces cerevisiae* Using Amine-reactive Isobaric Tagging Reagents." *Moll. Cell. Prot.* **2004**, 3, 1154-1169.

12. Pierce, A.; Unwin, R.D.; Evans, C.A.; Griffiths, S.; Carney, L.; Zhang, L.; Jaworska, E.; Lee, C.; Blinco, D.; Okoniewski, M.J.; Miller, C.J.; Bitton, D.A.; Spooncer, E.; Whetton, A.D. "Eight-channel iTRAQ Enables Comparison of the Activity of Six Leukemogenic Tyrosine Kinases." *Moll. Cell. Prot.* **2008**, 7, 853-863.

13. McLean, J.A.; Schultz, J.A.; Woods, A.S. In *Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicabilities, and Biological Applications*, Richard B. Cole, Ed.; John Wiley & Sons: Indianapolis, IN, 2007; pp 411-439.

14. Gieniec, J.; Mack, L.L.; Nakamae, K.; Gupta, C.; Kumar, V.; Malcolm, D. "Electrospray Mass Spectroscopy of Macromolecules: Application of an Ion-drift Spectrometer." *Biomed. Spectrom.* **1984**, 11, 259-268.

15. Karasek, F.W. "Plasma Chromatography." *Anal. Chem.* **1974**, 46, 710A-720A.

16. Shvartsburg, A.A.; Smith, R.D. "Fundamentals of Traveling Wave Ion Mobility Spectrometry." *Anal. Chem.* **2008**, 80, 9689-9699.

17. Mason, E.A.; McDaniel, E.W. *Transport Properties of Ions in Gases.* John Wiley & Sons: Indianapolis, IN, 1988.

18. Mesleh, M.F.; Hunter, J.M.; Shvartsburg, A.A.; Schatz, G.C.; Jarrold, M.F. "Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential." *J. Phys. Chem.* **1996**, 100, 16082-16086.

19. Shvartsburg, A.A.; Jarrold, M.F. "An Exact Hard-spheres Scattering Model for the Mobilities of Polyatomic Ions." *Chem. Phys. Lett.* **1996**, 261, 86-91.

20. Asbury, G.R.; Hill, H.H. "Evaluation of Ultrahigh Resolution Ion Mobility Spectrometry as an Analytical Separation Device in Chromatographic Terms." *J. Microcolumn Separations.* **2000**, 12, 172-178.

21. McLean, J.A.; Ruotolo, B.T.; Gillig, K.J.; Russell, D.H. "Ion Mobility-mass Spectrometry: A New Paradigm for Proteomics." *Int. J. Mass Spectrom.* **2005**, 240, 301-315.

22. Ruotolo, B.T.; Giles, K.; Campuzano, I.; Sandercock, A.M.; Bateman, R.H.; Robinson, C.V. "Evidence for Macromolecular Protein Rings in the Absence of Bulk Water." *Science.* **2005**, 310, 1658-1661.

23. Wyttenbach, T.; Bowers, M.T. "Gas-phase Conformations: The Ion Mobility/Ion Chromatography Method." *Top. Curr. Chem.* **2003**, 225, 207-232.

24. Jarrold, M.F. "Peptides and Proteins in the Vapor Phase." *Ann. Rev. Phys. Chem.* **2000**, 51, 179.

25. Smith, R.D.; Loo, J.A.; Loo, R.R.O.; Busman, M.; Udseth, H.R. "Principles and Practice of Electrospray Ionization-mass Spectrometry for Large Polypeptides and Proteins." *Mass Spectrom. Rev.* **1991**, 10, 359-451.

26. Wittmer, D.; Chen, Y.H.; Luckenbill, B.K.; Hill, H.H. "Electrospray Ionization Ion Mobility Spectrometry." **1994**, 66, 2348-2355.

27. Tao, L.; McLean, J.R.; McLean, J.A.; Russell, D.H. “A Collision Cross-section Database of Singly-charged Peptide Ions.” *J. Am. Soc. Mass Spectrom.* **2007**, 18, 1232-1238.

28. Ruotolo, B.T.; Verbeck, G.F.; Thomson, L.M.; Woods, A.S.; Gillig, K.J.; Russell, D.H. “Distinguishing Between Phosphorylated and Nonphosphorylated Peptides with Ion Mobility-mass Spectrometry.” *J. Prot. Res.* **2002**, 1, 303-306.

29. Hilderbrand, A.E.; Myung, S.; Clemmer, D.E. “Exploring Crown Ethers as Shift Reagents for Ion Mobility Spectrometry.” *Anal. Chem.* **2006**, 78, 6792-6800.

30. Howdle, M.D.; Eckers, C.; Laures, A.M.F.; Creaser, C.S. “The Use of Shift Reagents in Ion Mobility-Mass Spectrometry: Studies on the Complexation of an Active Pharmaceutical Ingredient with Poly Ethyleneglycol Excipients.” *J. Am. Soc. Mass Spectrom.* **2009**, 20, 1-9.

31. Fenn, L.S.; Kliman, M.; Mahsut, A.; Zhao, S.R.; McLean, J.A. “Characterizing Ion Mobility-mass Spectrometry Conformation Space for the Analysis of Complex Biological Samples.” *Anal. Bioanal. Chem.* **2009**, 394, 235-244.

32. Sundarapandian, S.; May, J.C.; McLean, J.A. “Dual Source Ion Mobility-Mass Spectrometer for Direct Comparison of Electrospray Ionizatoin and MALDI Collision Cross Section Measurements.” *Anal. Chem.* **2010**, 82, 3247-3254.

33. Lee, Y.J.; Hoaglund-Hyzer, C.S.; Srebalus Barnes, C.A.; Hilderbrand, A.E.; Valentine, S.J.; Clemmer, D.E. “Development of High-Throughput LC-injected Ion Mobility-Q-TOF Techniques for Analysis of Complex Peptide Mixtures.” *J. Chromatography B.* **2002**, 782, 343-351.

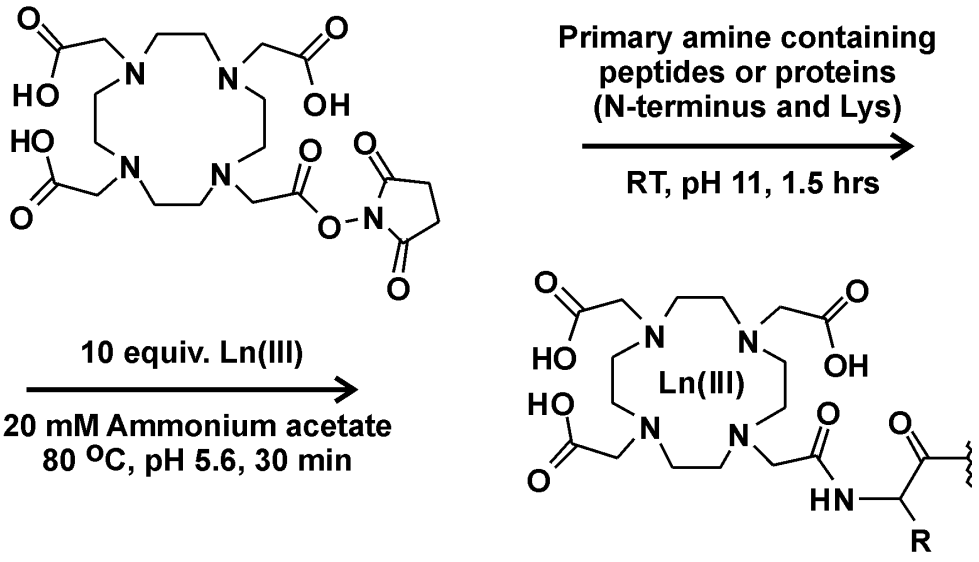
CHAPTER II

RELATIVE QUANTITATION OF PEPTIDE FUNCTIONALITY USING METAL CHELATION LABELS

2.1 INTRODUCTION

Mass spectrometry based relative quantitation of protein expression profiles has experienced a recent period of rapid growth as new instrumental technologies and labeling strategies have been developed.[1] New advances in labeling methods,[1-3] and label free methods,[4-6] have changed the way in which the relative quantitation of protein expression profiles using multiple samples is obtained. Many labeling methods, such as SILAC[1] and ICAT[2] use isotope enrichment of one sample versus another to maintain the chemical composition of the peptide for uniform ionization efficiency in the ratioing of peak areas for relative quantitation. However, the number of experimental conditions that can be analyzed by these methods is limited to typically less than 4 by the nominal isotopic shifts afforded by isotopic enrichment of ^{13}C, ^{15}N, and ^{2}H.

More recently, element coded affinity tags (ECAT) have been described for the relative quantitation of peptides.[7-9] ECAT is analogous to ICAT in that relative quantitation information is obtained based on the comparative ratio between an isotopically heavy and light label, respectively, however the source of the isotopic shift is incorporation of different lanthanide (Ln) metal atoms. ECAT labels are similar to that shown in scheme 2.1, whereby a Ln(III) ion is chelated and is subsequently covalently attached at cysteine sulfhydryl groups via reaction with a maleimido moiety.


Primary amine containing
peptides or proteins
(N-terminus and Lys)
RT, pH 11, 1.5 hrs
10 equiv. Ln(III)
20 mM Ammonium acetate
80 °C, pH 5.6, 30 min
Ln(III)
HO
OH
N
O
HN
R

Scheme 2.1.

The ionic radii of Ln(III) ions are nearly invariant across the series of metals (*ca.* 87-117 pm), which in turn results in minor differences of otherwise very strong stability constants (typical stability constant, log(K), values of 20-24).[10] The stability of such labels is underscored by their use as common MRI contrast enhancing agents.[11] For MS labeling purposes, the ability to select different Ln(III) ions for labeling different experimental conditions allows isotopic shifts of up to 36 Da and can accommodate the simultaneous relative quantitation of experimental samples limited only by the number of lanthanides or isotopically enriched lanthanides used.

In contrast with studies using ECAT for the labeling of cysteine residues (*via* sulfhydryl derivatization),[7-9] the present studies utilizing primary amine derivatization offer three important advantages, Firstly, unless the N-terminus is modified, nearly all peptides will incorporate at least one label. Secondly, labeling at a pH approximately the same as the pK_a of the side chain of lysine (~11.1) results in multiple labeling of the peptide (Figure 2.1). The latter has particular significance when utilizing trypsin proteolysis, because this results in peptides either containing a single label (R-terminated) or two labels (K-terminated) for facile assignments. This is also significant from the standpoint of MS/MS sequencing of peptides because the label(s) only reside at the N-terminus or at both termini, which provides a frame of reference for assigning fragment ion peaks. Similar to previous reports, the ECAT labels described here remain intact through collision induced dissociation (CID) processes using both MALDI (Figure 2.2) and ESI (Figure 2.3) ion sources.[7-9] Thirdly, primary amine labeling can be combined with sulfhydryl labeling using different metals for each (Chapter 4) to enhance the information content for specific peptides. Similar to ICAT strategies, the ratio of the integrated isotope cluster of the peptide corresponding to heavy and light labeling is compared.

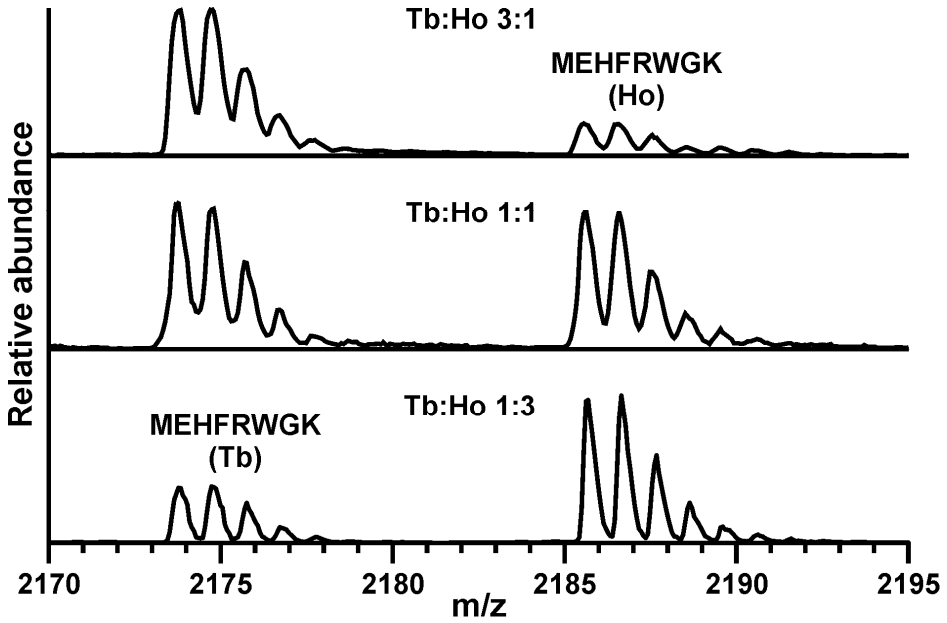


Figure 2.1. Typical spectra in relative quantitation via primary amine labeling. Top, middle, and bottom correspond to expanded mass spectra about the peptide MEHFRWGK labeled with Tb and Ho in ratios of 3:1, 1:1, and 1:3, respectively. The peptide is labeled both at the N-terminus and the C-terminal Lys residue resulting in a 12 Da mass shift.

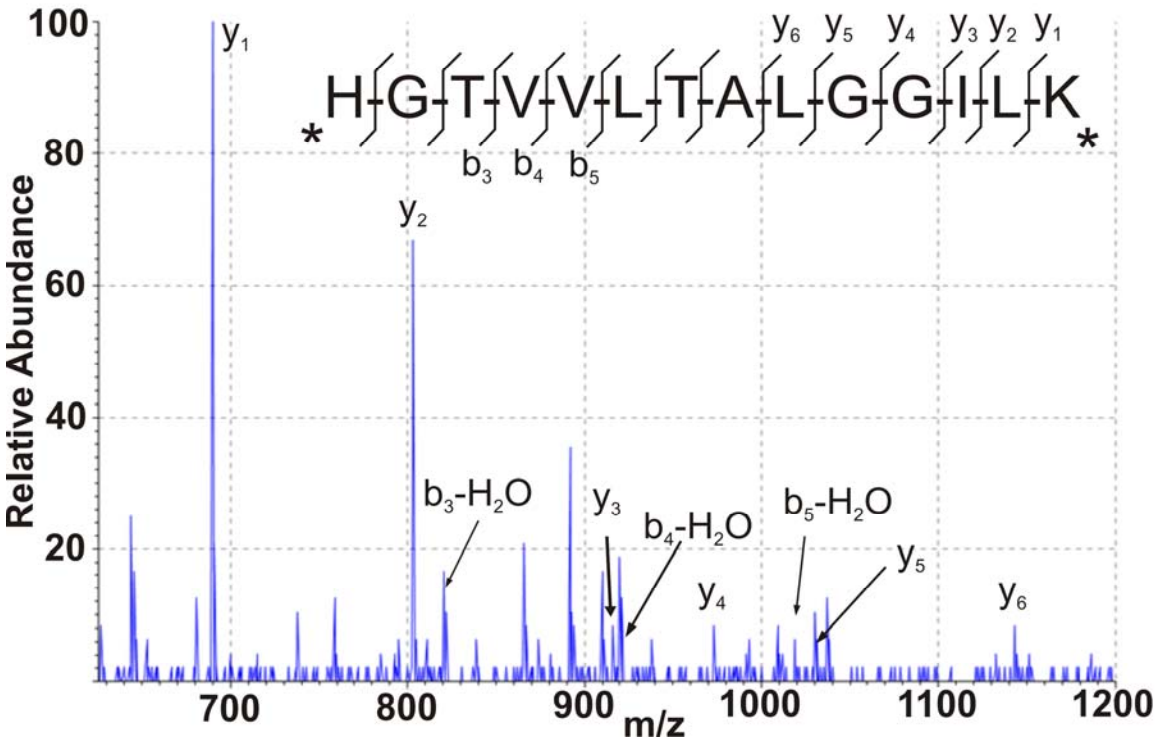


Figure 2.2. Tandem MS spectrum (generated using a MALDI ion source) of a peptide (HGTVVLTALGGILK) from a tryptic myoglobin digest labeled with the primary amine specific shift reagents. The peptide sequence is shown above with an asterisk (*) marking the location of tag addition (*i.e.* at the N-terminus and the Lys residue). Labeled b and y ion peaks correspond to the fragment mass including the additional mass of the label since both termini are labeled.

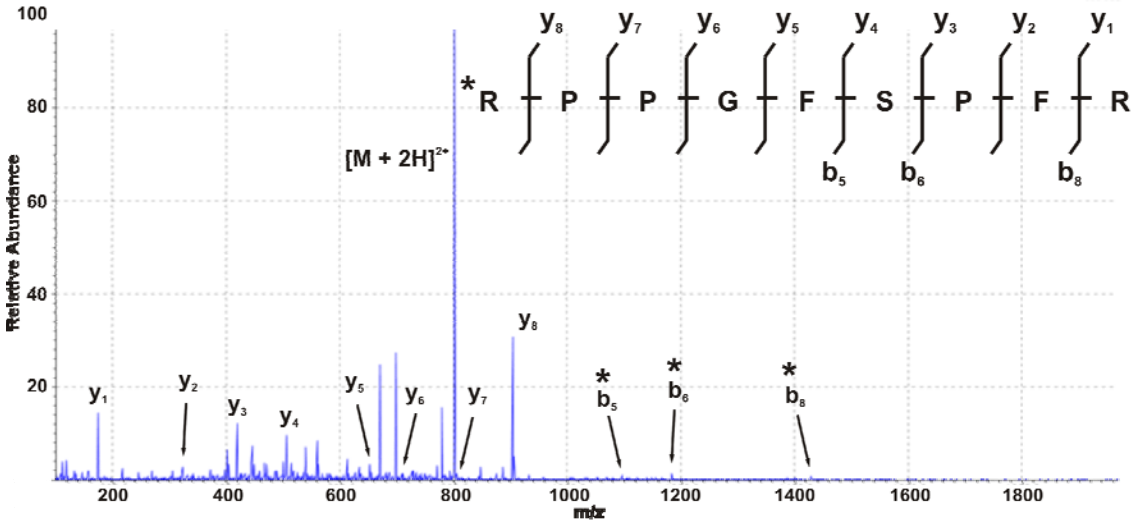


Figure 2.3. Tandem MS spectrum (generated suing an ESI ion source) of bradykinin (RPPGFSPFR) labeled with the primary amine specific shift reagent. The peptide sequence is shown above with an asterisk (*) marking the location of tag addition.

2.2 Experimental

The specific labels used in this work consist of a metal chelation moiety, a linker region that acts as a spacer between the chelation center and the reactive target, and a primary amine reactive *N*-hydroxysuccinimide (NHS) ester group (shown in Scheme 1). The metal chelation moiety is based on 1,4,7,10-tetra-azacyclododecane-*N*, *N*', *N*'', *N*'''-tetraacetic acid (DOTA) obtained from Macrocyclics (Dallas, TX). The lanthanides used in this work include La, Lu, Tm, Ho, and Tb, which were purchased from Strem (Newburyport, MA) or Sigma (St. Louis, MO) as chloride hydrate salts. These lanthanides were selected because of their mass differences and because they are (Tb, Ho, and Tm) or are nearly (La and Lu), monoisotopic (Table 2.1). Covalent labeling of individual peptides and those generated from tryptic proteolysis of proteins was performed in a two-step reaction as depicted in Scheme 1.

Briefly, the first step involves the addition of the NHS-DOTA label to primary amine sites on the peptide (*i.e.* at the N-terminus and any Lys residues) at room temperature and pH 10. This is followed by chelating the chosen Ln(III) metal to the tag by heating the tagged sample in the presence of 10 equivalents of lanthanide metal at pH 5.6, which was optimized based on a previously published protocol.[12] Following labeling, the labeled peptides are stable over a wide pH range (3-10 investigated thus far), however after spotting onto a MALDI target, small amounts of degradation products may be observed after *ca.* 5 hours.

Proteolysis of common proteins was used to generate a mixture of peptides for labeling. Initially, proteins were denatured at 90˚C and any disulfide bonds were reduced using an appropriate reducing agent.[13] For this dissertation *tris*-2-carboxyethylphosphine

	La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
136	-	0.2	-	-	-	-	-	-	-	-	-	-	-	-	-
138	0.1	0.3	-	-	-	-	-	-	-	-	-	-	-	-	-
139	99.9	-	-	-	-	-	-	-	-	-	-	-	-	-	-
140	-	88.5	-	-	-	-	-	-	-	-	-	-	-	-	-
141	-	tr	100	-	-	-	-	-	-	-	-	-	-	-	-
142	-	11.1	-	27.2	-	-	-	-	-	-	-	-	-	-	-
143	-	-	-	12.2	-	-	-	-	-	-	-	-	-	-	-
144	-	tr	-	23.8	-	3.1	-	-	-	-	-	-	-	-	-
145	-	-	-	8.3	-	-	-	-	-	-	-	-	-	-	-
146	-	-	-	17.2	-	-	-	-	-	-	-	-	-	-	-
147	-	-	-	-	-	15.0	-	-	-	-	-	-	-	-	-
148	-	-	-	5.7	-	11.2	-	-	-	-	-	-	-	-	-
149	-	-	-	-	-	13.9	-	-	-	-	-	-	-	-	-
150	-	-	-	5.6	-	7.4	-	-	-	-	-	-	-	-	-
151	-	-	-	-	-	-	47.8	-	-	-	-	-	-	-	-
152	-	-	-	-	-	26.8	-	0.2	-	-	-	-	-	-	-
153	-	-	-	-	-	-	52.2	-	-	-	-	-	-	-	-
154	-	-	-	-	-	22.8	-	2.2	-	-	-	-	-	-	-
155	-	-	-	-	-	-	-	14.8	-	-	-	-	-	-	-
156	-	-	-	-	-	-	-	20.5	-	tr	-	-	-	-	-
157	-	-	-	-	-	-	-	15.7	-	-	-	-	-	-	-
158	-	-	-	-	-	-	-	24.8	-	0.1	-	-	-	-	-
159	-	-	-	-	-	-	-	-	100	-	-	-	-	-	-
160	-	-	-	-	-	-	-	21.9	-	2.3	-	-	-	-	-
161	-	-	-	-	-	-	-	-	-	18.9	-	-	-	-	-
162	-	-	-	-	-	-	-	-	-	25.5	-	0.1	-	-	-
163	-	-	-	-	-	-	-	-	-	24.9	-	-	-	-	-
164	-	-	-	-	-	-	-	-	-	28.3	-	1.6	-	-	-
165	-	-	-	-	-	-	-	-	-	-	100	-	-	-	-
166	-	-	-	-	-	-	-	-	-	-	-	33.5	-	-	-
167	-	-	-	-	-	-	-	-	-	-	-	22.9	-	-	-
168	-	-	-	-	-	-	-	-	-	-	-	27.0	-	0.1	-
169	-	-	-	-	-	-	-	-	-	-	-	-	100	tr	-
170	-	-	-	-	-	-	-	-	-	-	-	14.9	-	3.0	-
171	-	-	-	-	-	-	-	-	-	-	-	-	-	14.3	-
172	-	-	-	-	-	-	-	-	-	-	-	-	-	21.8	-
173	-	-	-	-	-	-	-	-	-	-	-	-	-	16.1	-
174	-	-	-	-	-	-	-	-	-	-	-	-	-	31.8	-
175	-	-	-	-	-	-	-	-	-	-	-	-	-	-	97.4
176	-	-	-	-	-	-	-	-	-	-	-	-	-	12.8	2.6

Table 2.1. Table of lanthanide isotopes and relative abundance values useful for selection of the metal for incorporation in the shift reagent tag.

(TCEP) was used more frequently than dithiothreitol (DTT) to limit quenching and side reactions when labeling using the maleimido label. Upon cooling to room temperature, trypsin was added in a 1:40 w/w ratio at a pH of approximately 8. This reaction is then heated at 37˚C overnight (18-24 hours). The resulting peptides are then desalted prior to labeling and analysis.

A more complex mixture was simulated by proteolyzing multiple proteins in a one-pot reaction. Three proteins, hen egg white Lysozyme, horse cytochrome c, and horse heart Myoglobin, were mixed in an equimolar ratio in 60mM ammonium bicarbonate buffer of pH ~8. TCEP reducing agent was added to cleave the disulfide bonds in Lysozyme, and the protein mixture was denatured at 90˚C for 20 minutes. After cooling to room temperature, an appropriate amount of trypsin (a 1:40 ratio by mass) was added and proteolysis was allowed to proceed for 18-24 hours. After digestion, the resulting peptides were desalted using C18 spin columns. These peptides were then either analyzed using MALDI TOFMS or labeled according to Scheme 2.1.

The change in ionization efficiency of peptide upon labeling was also explored. A series of peptides based on of the sequence of bradykinin (RPPGFSPFR) were synthesized (Genscript) whereby one or both arginines were replaced with an aspartic acid residue. An equimolar aliquot of each of these three peptides was mixed and analyzed using an Applied Biosystems DE STR TOFMS. The peak areas of each peptide were compared. Subsequently, these same three peptides were labeled under the same reaction conditions in an equimolar ratio and analyzed using the same instrument. The peak area distribution of labeled versus unlabeled were compared to determine the effect of the lanthanide label on the ionization efficiency of the peptide.

2.3 Results and discussion

Several proof of concept peptides were chosen (Table 2.2) for simulated relative quantitation experiments. After labeling, the ratio of the monoisotopic peak areas of the tagged species was compared to the theoretical ratio for each experiment. A wide range of theoretical quantitation ratios were analyzed. Percent error for model peptides was generally below 15% with the lowest percent errors occurring near a theoretical ratio of 0.5 (Table 2.2). As the ratio decreases below 0.5, more error between the expected value and the experimental value arises. Therefore, samples of extremely different concentration may require dilution of the high abundance sample so the experimental ratio will be 0.5-1 for the most accurate results. The wide variety of metals and peptide masses used illustrates that the particular metal system chosen does not affect the accuracy of the relative quantitation measurement.

Relative quantitation of a complex mixture was simulated by tagging tryptic peptides from horse heart Myoglobin and horse cytochrome c. Figure 2.4a represents relative quantitation using cytochrome c. It is important to note that singly labeled peptides exhibit a 6 Da shift, while doubly labeled peptides show a 12 Da shift due to the addition of an extra label. The ratio of the monoisotopic peak areas for each peptide pair was compared to the theoretical ratio of 0.5. The average ratio of all Myoglobin peptides analyzed agrees very well with the expected ratio of 0.5. Tryptic Myoglobin peptides were labeled using the primary amine specific ECAT like reagent according to Scheme 2.1. The tagged peptides were desalted, mixed, and spotted on a MALDI target plate for analysis (Figure 2.4b).

Peptide Sequence	Underivatized m/z	Metal System	Theoretical Ratio	Experimental Ratio[b]
CLRRASLG	872.9	Ho / Tb	0.75	0.76 ± 0.05 (3)
VHHQKLVFF	1155.4	Ho / Tb	0.67	0.67 ± 0.02 (2)
EGVYVHPV	900.0	Tm / La	0.75	0.80 ± 0.05 (4)
MEHFRWGK	1091.3	La / Tm	0.50	0.52 ± 0.07 (3)
IHPF	513.6	Tb / Ho	1.00	0.97 ± 0.32 (3)
YLEFISDAIIHVLHSK[a,c]	1886.2	Ho / Tb	0.50	0.46 ± 0.03 (3)
YLEFISDAIIHVLHSK[a,d]	1886.2	Ho / Tb	0.50	0.51 ± 0.05 (3)
HGTVVLTALGGILK[a]	1379.7	Ho / Tb	0.50	0.47 ± 0.03 (3)
ALELFR[a]	748.9	Ho / Tb	0.50	0.47 ± 0.13 (3)
GLSDGEWQQVLNVWGK[a]	1817.0	Ho / Tb	0.50	0.56 ± 0.04 (3)
VEADIAGHGQEVLIR[a]	1607.8	Ho / Tb	0.50	0.54 ± 0.14 (3)

[a] Peptides from a tryptic digestion of horse heart myoglobin.
[b] Error represents ± 1 σ for *n* number of measurements indicated in parentheses.
[c] Incorporation of one tag. [d] Incorporation of two tags

Table 2.2. Relative quantitation data for primary amine tagging.

The ECAT labeling method can also provide primary structural information in addition to relative quantitation information. Structural information can be obtained in the same experiment as relative quantitation data or in a standalone experiment. Figure 2.4b illustrates a relative quantitation experiment on a trypsin digest of horse heart Myoglobin. The metal system chosen, (Ho/Tb) results in an isotopic shift of 6 Daltons for a singly labeled peptide.

Lysozyme provides an interesting illustration where metal system selection and flexibility becomes important to the success of the experiment. Two Lysozyme peptides, GYSLGNWVCAAK and CKGTDVQAWIR (1269.5 and 1277.5 m/z, respectively), are separated by *ca.* 8 Daltons. Using the Ho/Tb metal system used for Myoglobin (a 6 Dalton shift) results in an overlap of the isotope profiles of the two labeled species. Changing the metal system to Tm/Tb (a 10 Dalton shift) greatly reduces spectral overlap. Changing the metal system again to Lu/Tb (a 16 Dalton shift) separates the peaks even further resulting in complete resolution as illustrated in Figure 2.5.

Labeling of complex mixtures is an important step toward a labeling method becoming more useful. Current proteomics problems are usually studied in a relatively complex state. These complex samples may contain many proteins that must be proteolyzed and labeled for either identification in 2D IM-MS conformation space, or for purposes of relative quantitation using one dimensional MALDI TOFMS. Complex samples consisting of three proteins (horse heart Myoglobin, horse cytochrome c, and intensity peptides were identified, however they are not labeled due to space constraints. Panel b in Figure 2.6 illustrates the MALDI TOFMS spectrum of this complex tryptic digest labeled with a primary amine specific shift reagent coordinated to terbium. The same high intensity peaks from panel a are labeled in panel b.

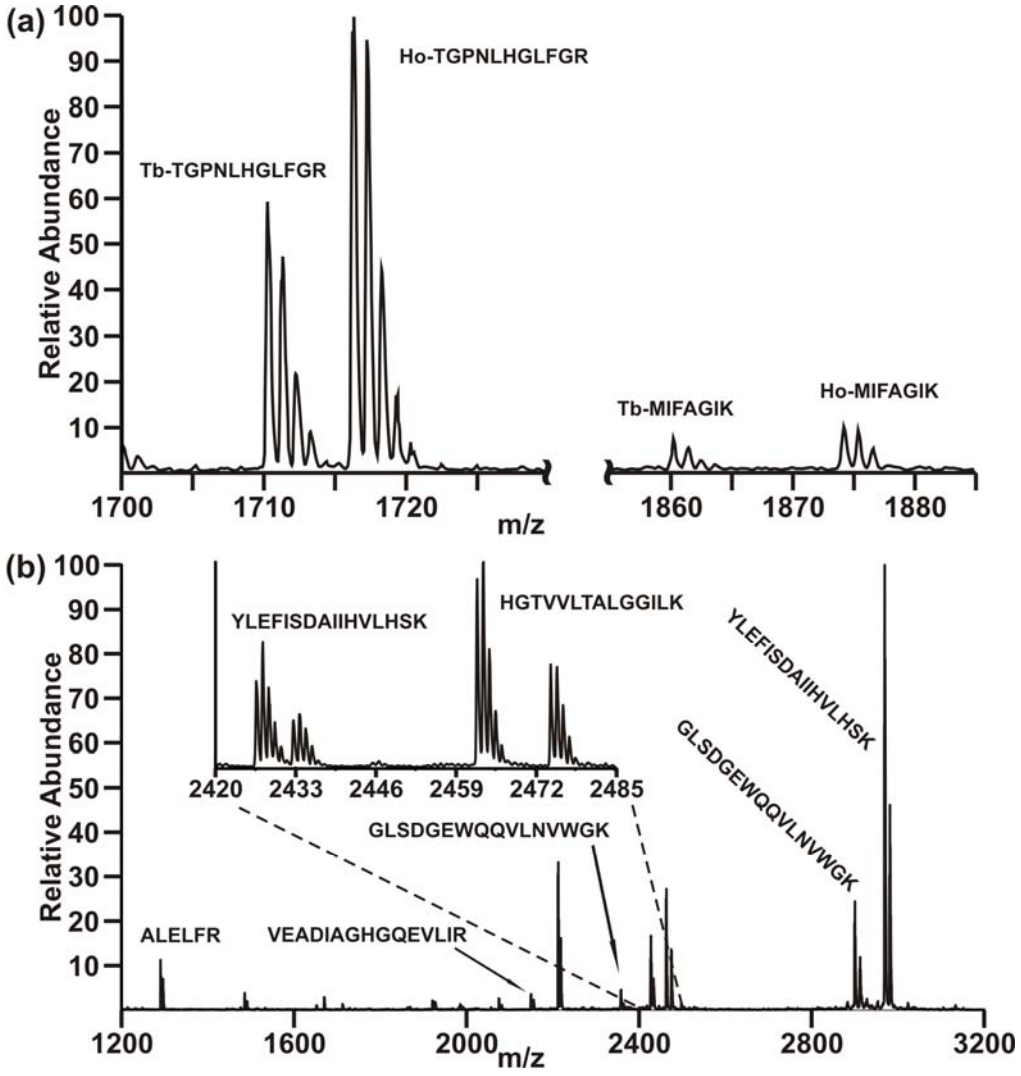


Figure 2.4. (a) An illustration of the mass shift for singly and doubly labeled peptides in a tryptic cytochrome c digestion relative quantitation experiment with a theoretical ratio of 0.5. TGPNLHGLFGR was tagged once (N-terminus) using the Ho/Tb metal system showing a 6 Dalton shift. MIFAGIK was tagged twice (N-terminus and Lys residue) and shows a 12 Dalton shift. (b) A typical spectrum illustrating a relative quantitation (0.5 theoretical ratio) experiment using horse heart Myoglobin and the Ho/Tb metal system. Inset zooms in on two pairs of peaks showing singly versus doubly tagged species.

The ionization efficiency of labeled peptides and unlabeled peptides were compared using a simple ratio.[14-15] Briefly, the peak area associated with each peptide sequence was divided into the total peak area of all three peptides. A plot of this data is show in Figure 2.7. As indicated in panel A (as well as a close up with error bars in panel B), both labeled and unlabeled peptides show similar ionization profiles.

In addition, peptides labeled with different lanthanide metals also show very little deviation. This small deviation supports the conclusion that the lanthanide labels do not significantly affect the relative ionization efficiency of peptides (shown in Figure 2.7, panel b) and thus are useful for relative quantitation purposes. Widely differing ionization profiles when using different metals would complicate relative quantitation experiments and introduce more error into the analysis.

The ECAT method of obtaining relative quantitation data works well for both simple and more complex peptide systems. The ECAT method works best when there is less than a three-fold difference in concentration between the peptides of interest since the percent error increases as the difference in concentration increases. This effect is likely due to ion suppression effects in the source of the mass spectrometer resulting from the low concentration peptide competing with the high concentration peptide for ionization.

2.4 Conclusion

The relative quantitation of peptides and proteins is an integral part of proteomics research. As sample complexity has increased and lower abundance proteins are studied, novel methods of analysis must be developed. Lanthanide based relative quantitation reagents offer several advantages over other stable isotope methods

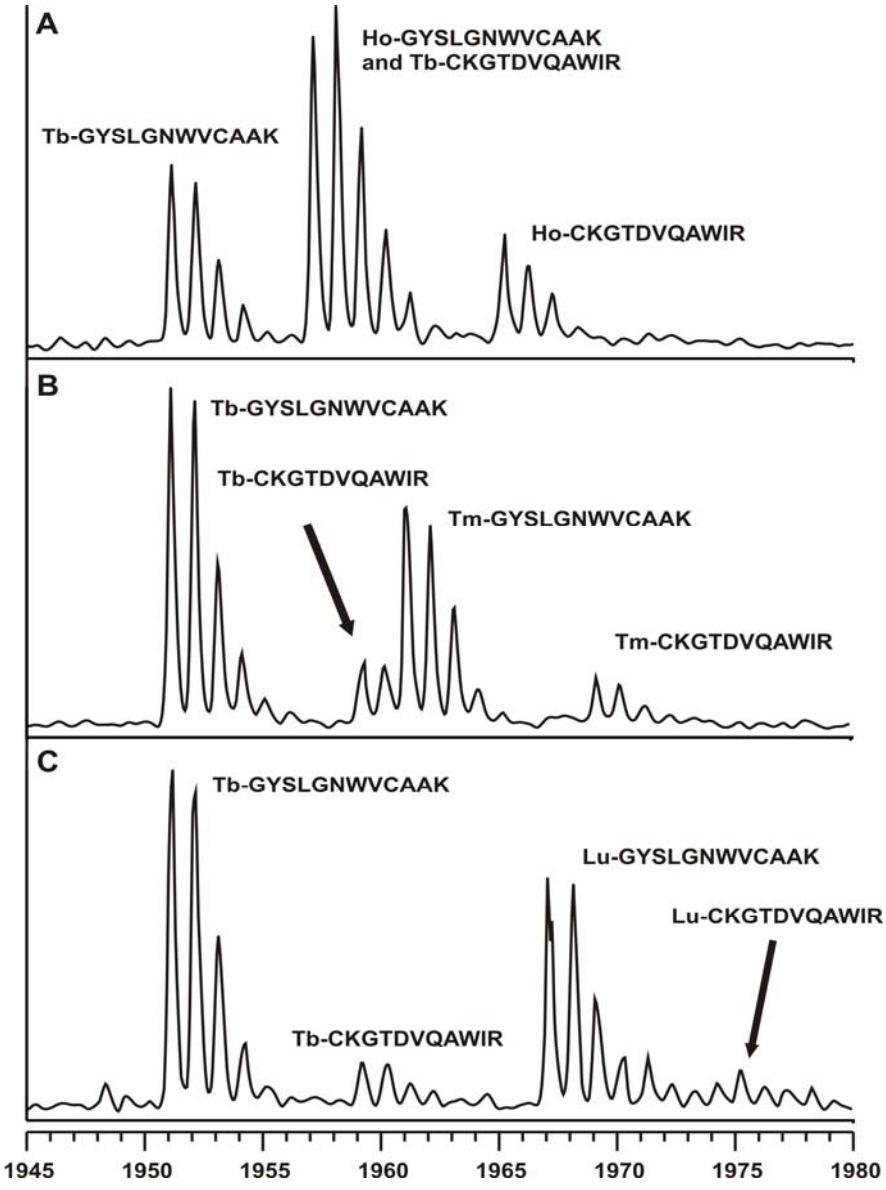


Figure 2.5. An illustration of the optimal metal system being dictated by the specific sample. In this experiment, a hen egg white Lysozyme digestion was tagged using a cysteine specific shift reagent and the various metal systems in a theoretical ratio of 2:1. Panel A shows the Tb/Ho metal system (a 6 Dalton shift) where two peaks overlap, causing error in the relative quantitation. Panel B illustrates use of the Tb/Tm metal system (a 10 Dalton shift) on the same sample. The labeled peaks are still not well resolved. Panel C shows the Tb/Lu metal system (a 16 Dalton shift). The labeled peaks are completely resolved, allowing for more accurate relative quantitation results.

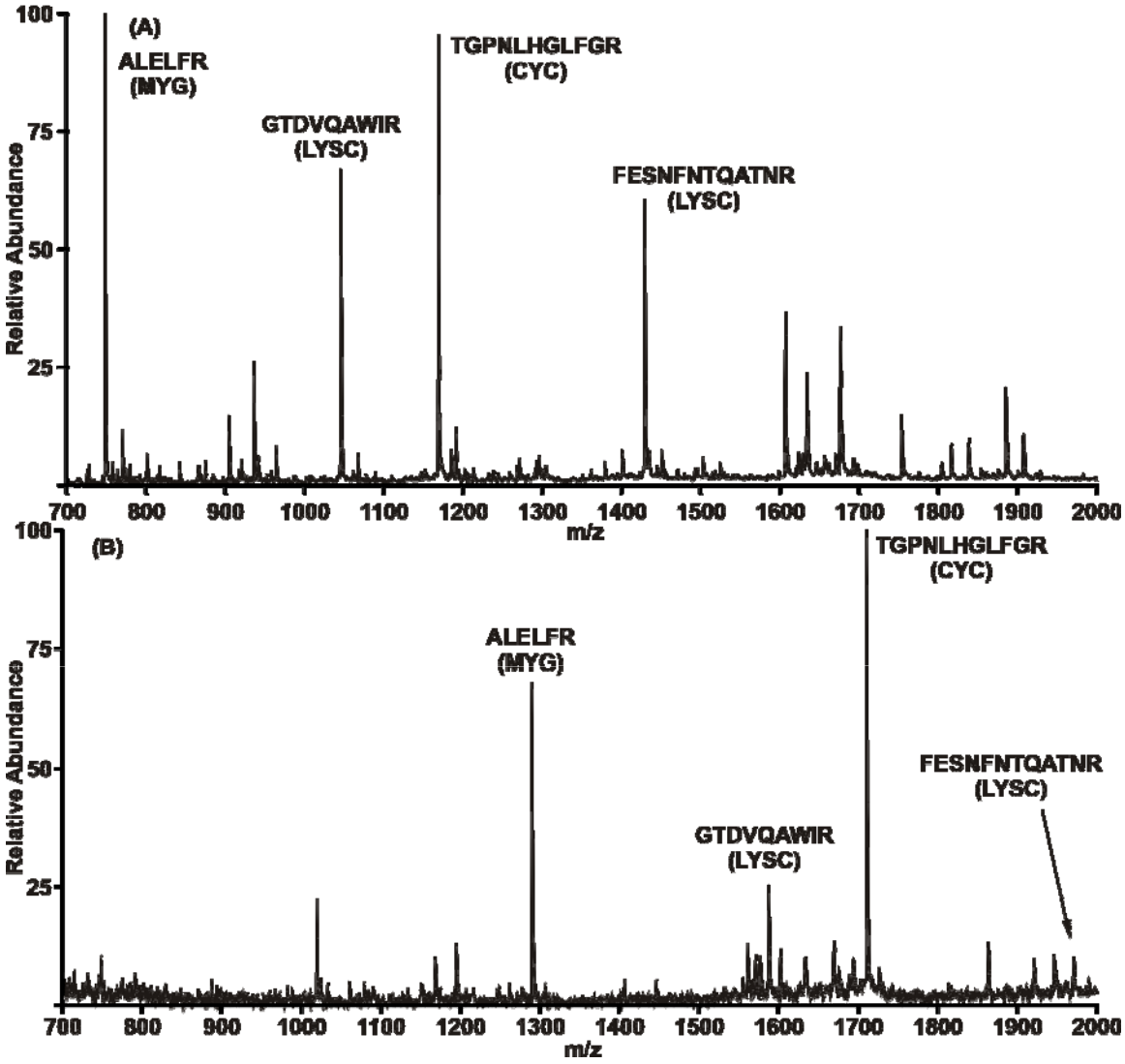


Figure 2.6. An illustration of a complex proteolysis experiment in which horse heart Myoglobin, horse cytochrome c, and hen egg white Lysozyme were digested simultaneously. Panel A illustrates the analysis and identification of unlabeled peptides. The sequence and source (in parentheses, Myoglobin:MYG, cytochrome c:CYC, Lysozyme:LYSC). Panel B shows the analysis of the same sample of proteolyzed peptides labeled with a primary amine specific shift reagent that coordinated terbium.

currently available, namely greater flexibility in the selection of the mass shift between labeled peptide species. This flexibility offers quantitative methods to systems that require a different amount of mass shift, such as Lysozyme. The primary amine specific shift reagents shown in this chapter are preferred for relative quantitation purposes since nearly every peptide in a proteolysis experiment will be labeled. The cysteine specific shift reagent will allow for quantitation, however since the relative frequency of cysteine is lower than the frequency of primary amines, every peptide may not receive a label. Importantly for relative quantitation experiments, the relative ionization efficiencies of different lanthanide metals in incorporated in the label are nearly identical. Thus, when different metals are used, these ECAT like reagents offer the ability to obtain excellent relative quantitation data on complex peptide samples.

We are grateful for financial support provided by the Defense Threat Reduction Agency (HDTRA1-09-1-0013), the National Institute on Drug Abuse (RC2DA028981), the Vanderbilt University College of Arts and Sciences, the Vanderbilt Institutes for Chemical Biology, the Vanderbilt Institute for Integrative Biosystems Research, and Ionwerks, Inc (Houston, TX).

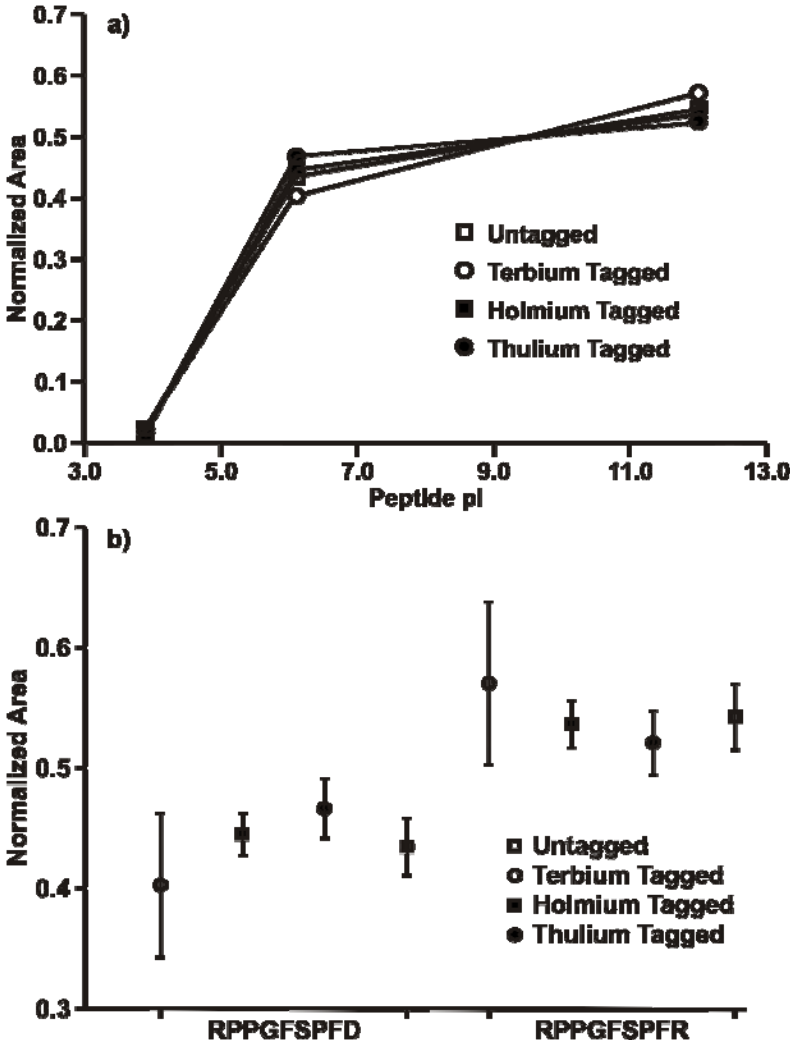


Figure 2.7. A graphical representation of the relative ionization efficiency of a series of peptides related to bradykinin (RPPGFSPFR). Two additional peptides were synthesized whereby one or both arginine residues in bradykinin was replaced with aspartic acid, creating a more acidic peptide. Panel A shows the peak areas of each peptide relative to the combined peak areas of the three peptides. As expected peptides with more basic residues showed better ionization, due to increased gas phase basicity. Panel B illustrates the two more basic peptides (including error bars) showing a significant difference in ionization between the two peptides, however, there is not a significant difference in ionization between the same peptide labeled with different lanthanide metals.

References

1. Aebersold, R.; Mann, M., “Mass spectrometry-based Proteomics.” *Nature* **2003**, 422, 198-207.

2. Gygi, S.P.; Rist, B.; Gerber, S.A.; Turecek, F.; Gelb, M.H.; Aebersold, R. “Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags.” *Nat. Biotech.* **1999**, 17, 994-999.

3. Gerber, S.A.; Rush, J.; Stemman, O.; Kirschner, M.W.; Gygi S.P. “Absolute Quantification of Proteins and Phosphoproteins from Cell Lysates by Tandem MS.” *Proc. Nat. Acad. Sci. USA* **2003**, 100, 6940-6945.

4. Chelius, D.; Bondarenko, P.V. “Quantitative Profiling of Proteins in Complex Mixtures Using Liquid Chromatography and Mass Spectrometry.” *J. Proteome Res.* **2002**, 1, 317-323.

5. Steen, H.; Jebanathirajah, J.A.; Springer, M.; Kirschner, M.W. “Stable Isotope-free Relative and Absolute Quantitation of Protein Phosphorylation Stoichiometry by MS.” *Proc. Nat. Acad. Sci. USA* **2005**, 102, 3948-3953.

6. Old, W.M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K.G.; Mendoza, A.; Sevinsky, J.R.; Resing, K.A.; Ahn, N.G. “Comparison of Label-free Methods for Quantifying Human Proteins by Shotgun Proteomics.” *Moll. Cell. Proteom.* **2005**, 47, 1487-1502.

7. Whetstone, P.A.; Butlin, N.G.; Corneillie, T.M.; Meares, C.F. “Element-coded Affinity Tags for Peptides and Proteins.” *Bioconj. Chem.* **2004**, 15, 3-6.

8. Liu, H.; Zhang, Y.; Wang, J.; Wang, D.; Zhou, C.; Cai, Y.; Qian, X. “Method for Quantitative Proteomics Research by Using Metal Element Chelated Tags Coupled with Mass Spectrometry.” *Anal. Chem.* **2006**, 78, 6614-6621.

9. Ahrends, R.; Pieper, S.; Kühn, A.; Weisshoff, H.; Hamester, M.; Lindemann, T.; Scheler, C.; Lehmann, K.; Taubner, K.; Linscheid, M.W. “A Metal-coded Affinity Tag Approach to Quantitative Proteomics.” *Moll. Cell. Proteom.* **2007**, 6, 1907-1916.

10. Tóth, E.; Brücher, E. “Stability Constants of the Lanthanide (III)-1,4,7,10-tetraazacyclododecane-*N,N’,N’’,N’’’*-tetraacetate Complexes.” *Inorg. Chim. Acta* **1994**, 221, 165-167.

11. Caravan, P. “Strategies for Increasing the Sensitivity of Gadolinium-based MRI Contrast Agents.” *Chem. Soc. Rev.* **2006**, 35, 512-523.

12. Zhu, X.; Lever, S.Z. “Formation Kinetics and Stability Studies on the Lanthanide Complexes of 1,4,7,10-tetra-azacyclododecane-*N,N’,N’’,N’’’*-tetraacetic acid by Capillary Electrophoresis.” *Electrophoresis.* **2002**, 23, 1348-1356.

13. Park, Z.Y.; Russell, D.H. “Thermal Denaturation: A Useful Technique in Peptide Mass Mapping.” *Anal. Chem.* **2000**, 72,2667-2670.

14. Grey, A.; Hoeger, C.A.; Miller, C.L.; Goedken, T.; Rivier, J.E.; Fischer, W.H. “Monitoring Protein Kinase and Phosphatase Reactions with Matrix-assisted Laser Desorption/Ionization Mass Spectrometry and Capillary Zone Electrophoresis: Comparison of the Detection Efficiency of Peptide-Phosphopeptide Mixtures.” *Biological Mass Spectrom.* **1994**, 23, 519-528.

15. Dekker, L.J.; Dalebout, J.C.; Siccama, I.; Jenster, G.; Smitt, P.A.S.; Luider, T.M. “A New Method to Analyze Matrix-assisted Laser Desorption/Ionization Time-of-flight Peptide Profiling Mass Spectra.” *Rapid Comm. Mass Spectrom.* **2005**, 19, 865-870.

CHAPTER III

APPLICATION OF METAL CHELATION LABELS TO RELATIVE QUANTITATION OF PEPTIDES USING ION MOBILITY-MASS SPECTROMETRY

3.1Introduction

IMS has proven to be a useful analytical tool for the detection of suspected drugs, chemical warfare agents, and explosives.[1-2] When IMS is combined with mass spectrometry (IM-MS) it becomes a useful tool for the rapid analysis of complex biological samples.[3] IM-MS provides a rapid alternative separation technique in that separations are performed over the course of 100s of μs to ms, compared to the minutes to hours time scale common to LC-based methods. IM-MS based methods also provide several other advantages: (i) mitigation of ion suppression effects, (ii) less complicated MS spectra facilitate peak identification, and (iii) analyte signals occur separate from chemical noise.

Ion mobility shift reagents allow specific signals (based on chemical functionality) to be isolated from signals of the same class of biomolecule (*e.g.* lipid or peptide) that do not contain that specific functionality. IM-MS shift reagents offer enhanced separation in the IM dimension by significantly altering the gas phase packing efficiency of the ion. Both crown ethers,[4] and PEG-based shift reagents[5] have been previously studied and offer robust methods for the isolation of signals in IM-MS conformation space.

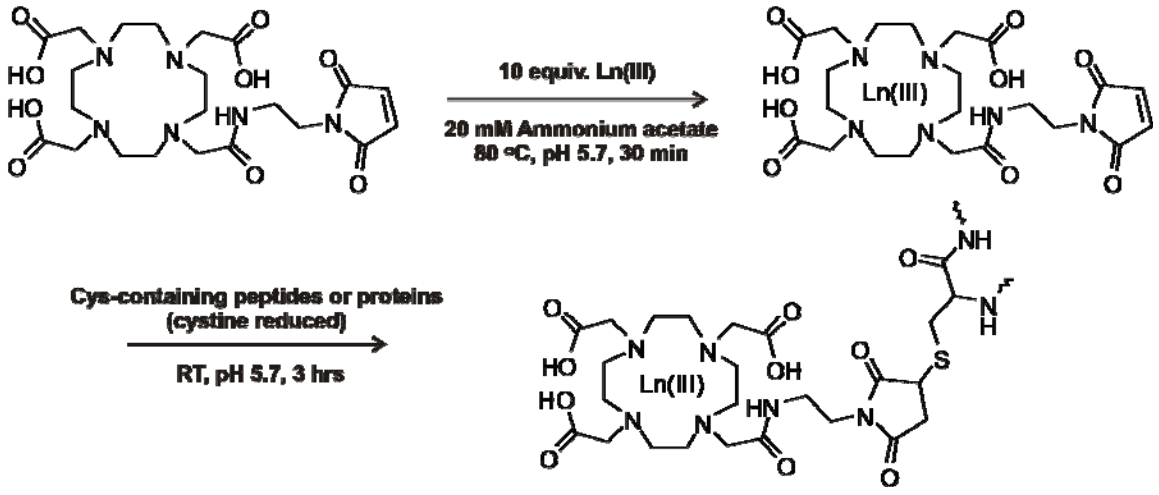
The crown ether and PEG-based shift reagents (studied by the Clemmer and Creaser groups, respectively) function by greatly increasing the collision cross section of

the labeled ion while increasing mass to a much smaller extent. This represents a decrease of the gas phase packing efficiency when compared to unlabeled ions of the same mass. These are termed low density shift reagents since the labeled peptides are shifted to an area above (longer drift times) the peptide correlation band. An interesting property of the crown ether shift reagents is that labeling is a result of non-covalent interaction with basic sites on the peptide.[4] Thus, the shift reagent may be removed by collisional activation after elution from the drift cell so that the parent peptide mass is retained. This chapter focuses on high density shift reagents and their application to isolating specific peptide signals in an area representing a smaller collision cross section than is predicted by the same mass in the peptide correlation band.

3.2 Experimental

All model peptides were purchased from American Peptide Company (Sunnyvale,CA) or Anaspec (San Jose, CA). Proteins were purchased from Sigma (St. Louis, MO). Lanthanide chelating labels were purchased from Macrocyclics (Dallas, TX). Lanthanide metals (La, Tb, Ho, Tm, and Lu) were purchased as $LnCl_3$ hydrates from Strem chemicals for Research (Newburyport, MA). Solvents were also purchased from Sigma and all reagents were used without further purification.

Several cysteine-containing peptides and proteins were chosen for proof-of-concept experiments. These peptides were chosen in order to provide a range of mass (*ca.* 700-1900 Da) and cysteine location within the peptide sequence. Model peptides were labeled at cysteine residues using a label based on 1,4,7,10-tetra-azacyclododecane-*N*, *N*', *N*'', *N*'''-tetraacetic acid (DOTA) as shown in Scheme 3.1.



Scheme 3.1. An illustration of the labeling reaction for specific modification of cysteine residues using an IM-MS shift reagent.

Lanthanide metal atoms were coordinated (prior to labeling) according to a method adapted from Zhu and Lever.[6] Briefly, the cysteine specific label and lanthanide metal (in a 1:50 ratio) were heated at 80˚C for 30 minutes at a pH of 5.7. Subsequent to metal incorporation, the free thiols of cysteine residues were modified by the maleimide functionality of the label (a 1:20 thiol:label ratio) at room temperature for three hours at a pH of 5.7. After modification, labeled peptides were desalted using Ziptips (Millipore) and analyzed using a custom IM-MS instrument (illustrated in Figure 1.3).

Similarly, primary amine sites were labeled using a shift reagent that has a N-hydroxysuccinimide reactive moiety. Briefly, labeling progressed by mixing the primary amine specific label and peptide (in a 20:1 ratio) in borate buffered solution of pH 10. Subsequent to primary amine modification, a lanthanide metal was incorporated by adding lanthanide (in a 50:1 ratio) to the labeled peptide at an acetate buffered pH of 5.7. The solution was heated at 80˚C for 30 minutes. Labeled peptides were desalted using ziptips (Millipore) prior to analysis.

Complex samples were simulated by peptide mixtures derived from proteolysis of model proteins. Hen egg white Lysozyme was chosen since it provided a mixture of several cysteine-containing peptides in addition to many cysteine free peptides. Lysozyme was proteolyzed as described elsewhere.[7] Briefly, disulfide bonds were reduced using *tris* (2-carboxyethyl) phsophine (TCEP) and the protein was heat denatured at 90˚C for 20 minutes prior to digestion. An appropriate amount of trypsin (1:40 trypsin:protein by mass) was added and allowed to digest the protein at 37˚C and pH 8 for 18-24 hours. Subsequent to digestion, the resulting peptides were desalted using C18 spin columns (Pierce) and peptide labeling proceeded as described in Scheme 3.1. Labeled protein digests were analyzed by IM-MS.

DTIM-MS spectra were obtained using a custom IM-time-of-flight MS with a MALDI source (MALDI-IM-TOFMS) operating at 500 Hz with a frequency-tripled Nd:YLF laser with a wavelength of 349 nm. Analysis was aided by custom visualization software (Ionwerks, Inc., Houston, TX). 2,5-dihydroxybenzoic acid (DHB) was used as the MALDI matrix in a matrix:analyte ratio of approximately 100:1. The drift tube was pressurized with helium to *ca.* 4 Torr for all experiments. The collision cross section for each labeled and unlabeled peptide was calculated using equation 1.1, which is based on the kinetic theory of gasses.[8] Collision cross sections for several peptides labeled with each of the five lanthanides were also calculated.

3.3 Results and discussion

Due to the curvature of the peptide correlation function,[9] two possible shift reagent strategies (as illustrated in Figure 3.1) may be imagined, each with their own strengths and challenges: (i) high cross section labels of relatively low mass, as represented by parachutes and (ii) high mass shift reagents of low cross section, illustrated as anchors. The high density nature of the lanthanide-based shift reagents isolates labeled peptide signals in an area below the peptide correlation function. High density labels have enhanced utility on peptides of low mass (*i.e.* peptide mass below *ca.* 2000) since the slope of the peptide correlation function is more vertical than at higher masses.

Initial high density shift reagent experiments were conducted using several model peptides. Cysteine containing peptides, in the presence of several cysteine free peptides, were labeled as illustrated in Scheme 3.1. The unlabeled peptides remain near

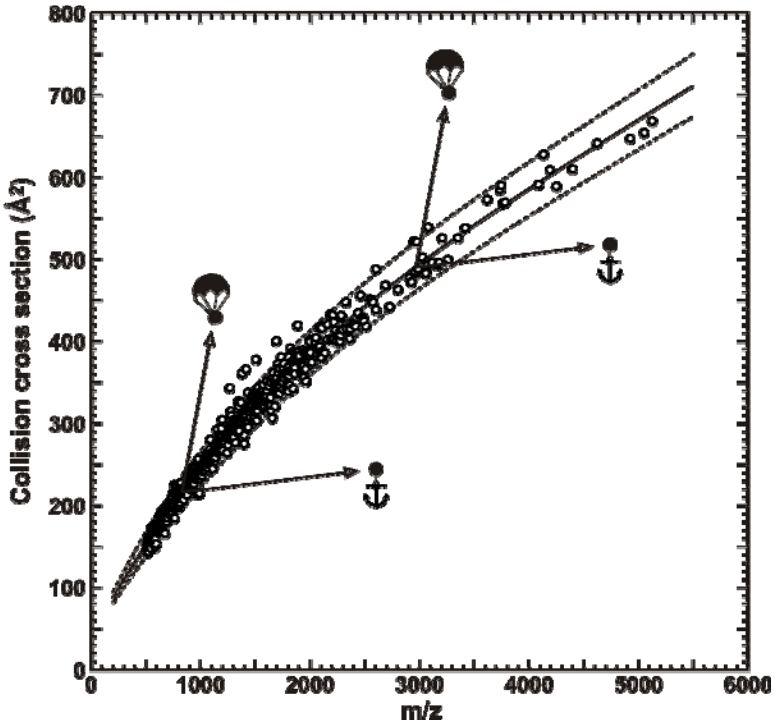


Figure 3.1. A cartoon illustrating the two types of IM-MS shift reagents. Low density (high collision cross section and low mass) shift agents are illustrated as balloons. High density shift agents (high mass and low collision cross section) are shown as anchors.

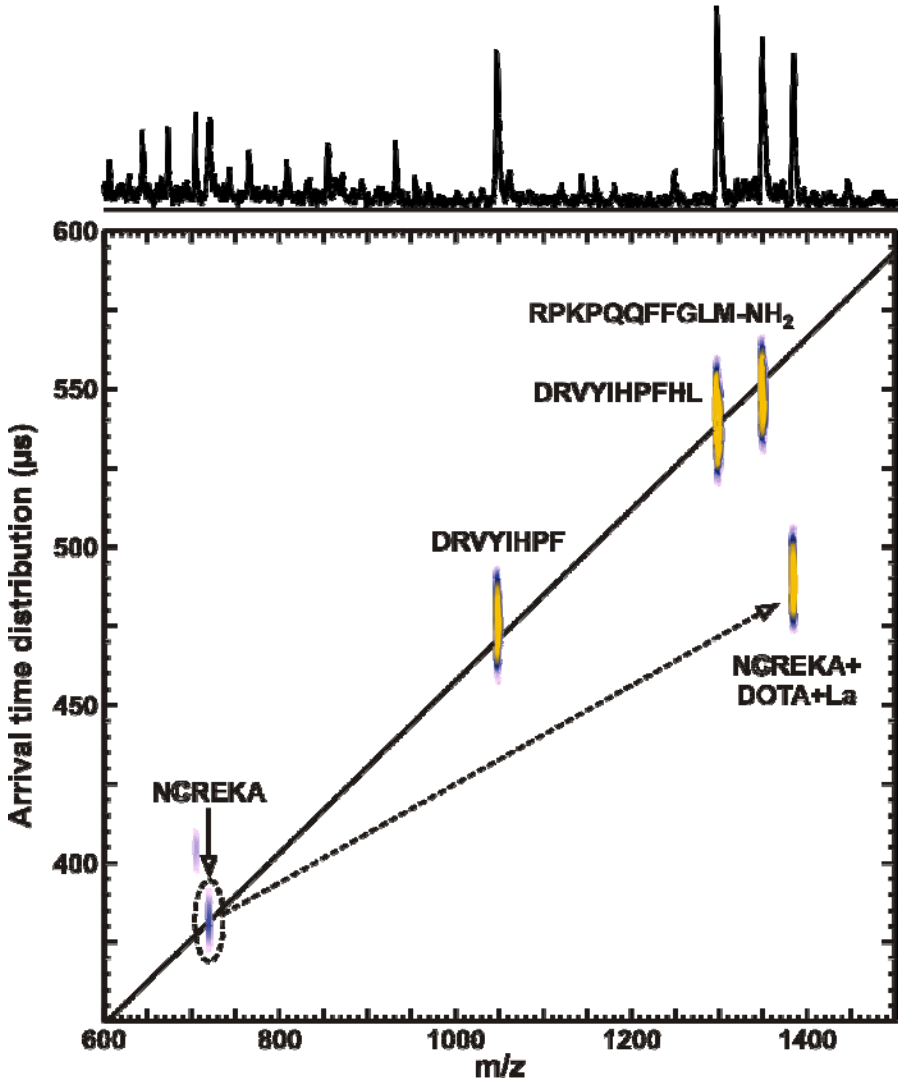


Figure 3.2. An illustration of a peptide (NCREKA) labeled with the cysteine specific IM-MS shift reagent incorporating a lanthanum atom. Several other unlabeled peptides are shown for comparison to the unlabeled peptide correlation band.

the peptide correlation band (shown as a solid black line in Figure 3.2), while the labeled peptide was shifted to an area in conformation space outside of the predicted peptide correlation band. Expanding on initial promising results, a much larger set of peptides of varying structure and sequence were analyzed. Collision cross sections of both labeled and unlabeled peptides were measured. The results of these calculations are shown graphically in Figure 3.3 and in tabular format in Table 3.1.

The high mass of the lanthanide metal provides the shift necessary to isolate labeled peptide signals in conformation space (Figure 3.3). While most peptides fall within 5-7% of the peptide correlation function[10] (illustrated as a solid black line), modification by high density shift reagents results in >15% average deviation (Table 3.1). This high degree of deviation from the narrow peptide correlation band ensures that labeled peptide signals (*i.e.* peptides that contain a specific functionality) are easily identifiable in conformation space.

The curving nature of the peptide correlation band causes high density shift reagents to have a larger relative shift for lower mass peptides. Table 3.1 illustrates this effect as the deviation from the peptide correlation band tends to decrease from about 20% for low mass peptides to about 10 % for peptides approaching 2000 Da. The choice of lanthanide metal does not significantly affect the deviation of a labeled peptide. For example, all five of the metals used (La, Lu, Tm, Tb, Ho) offer a similar deviation from the correlation band, as illustrated in Figure 3.4.

IM-MS shift reagents can aid in the rapid identification of peptides that contain specific functionality in a complex mixture of peptides. An unlabled tryptic digestion of Lysozyme was analyzed (*i.e.* a peptide mass mapping experiment[11-12]) on IM-MS instrumentation resulting in many positively identified peptides (Figure 3.5a). The three

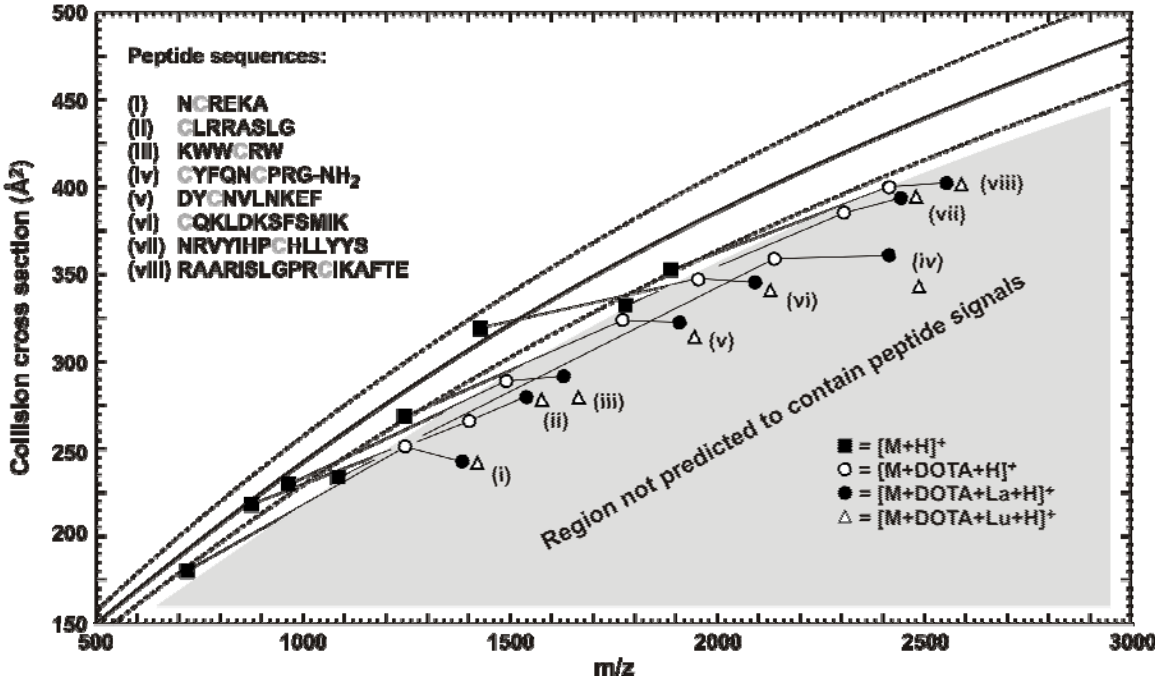


Figure 3.3. An illustration of the collision cross sections of various cysteine-containing peptides labeled with the thiol specific IM-MS shift reagents shown in Scheme 1. Labeled signals (using both lanthanum and lutetium) are shifted to an area in conformation space not predicted to contain peptide signals in the absence of labeling, illustrated as the grey area.

		Collision Cross Section (Å^2)[a]			Relative Deviation (%)	
Peptide Sequence	**Average m/z [M+H]$^+$**	**Untagged [M+H]$^+$**	**Tagged DOTA+La**	**Tagged DOTA+Lu**	**Tagged DOTA+La**	**Tagged DOTA+Lu**
NCREKA	720.80	180.0 ± 0.3	242.4 ± 0.2	241.5 ± 0.4	-20.2	-21.9
CLRRASLG	876.07	218.1 ± 0.1	279.2 ± 0.3	277.7 ± 0.3	-11.1	-16.2
KWWCRW	965.17	229.7 ± 0.1	291.2 ± 0.9	279.0 ± 0.2	-14.0	-18.7
CYFQNCPRG-NH_2[d]	1087.27	233.5 ± 0.1	360.6 ± 0.7	342.6 ± 0.6	-16.4	-21.9
DYCNVLNKEF	1245.40	268.7 ± 0.2	321.9 ± 0.5	313.7 ± 0.8	-14.1	-17.7
CQKLDKSFSMIK	1428.77	319.1 ± 0.8	345.0 ± 0.7	340.3 ± 0.6	-13.0	-15.0
NRVYIHPCHLLYYS	1779.09	332.0 ± 0.8	393.2 ± 0.3	394.0 ± 0.9	-9.5	-10.0
RAARISLGPRCIKAFTE	1890.27	352.5 ± 0.6	401.8 ± 0.2	400.9 ± 0.7	-9.8	-10.7

Table 3.1. Tabular data represented in Figure 3.3. The collision cross sections (average of n=5 measurements) of labeled peptides show a negative deviation of approximately 10-20% from the peptide correlation band.

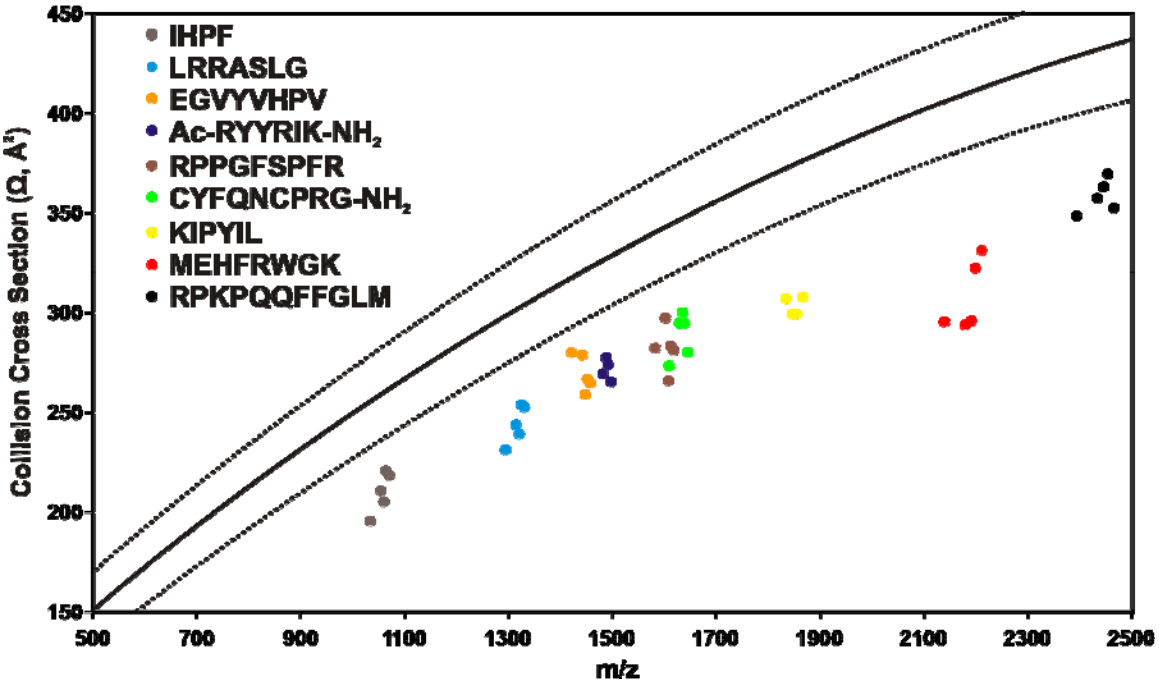


Figure 3.4. A graphic representation showing the collision cross sections of several peptides labeled with lanthanum, terbium, holmium, thulium, and lutetium. The specific lanthanide metal used does not have a significant effect on the average shift of the signal from the peptide correlation band.

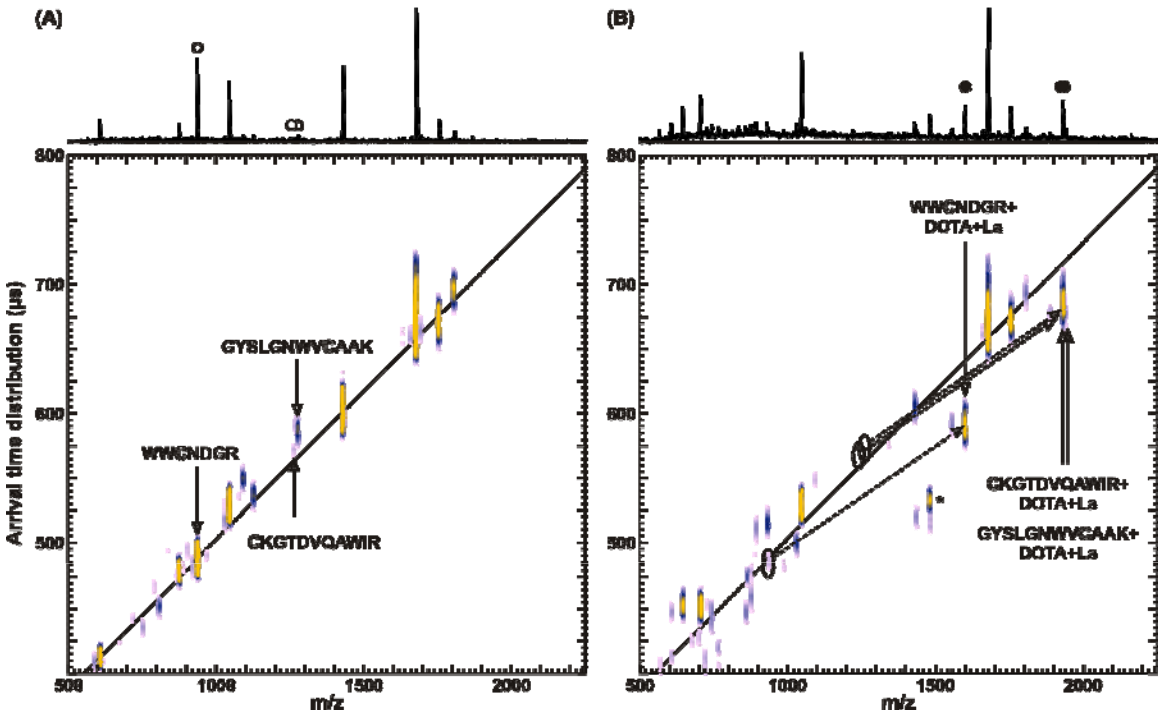


Figure 3.5. (A) A plot of IM-MS conformation space for an underivatized cysteine-reduced tryptic digest of Hen Egg White Lysozyme. In this plot, 12 tryptic peptides were positively identified of which 3 contain cysteine as labeled, which are also indicated in the mass spectrum by open circles. Note that all of the peptides fall within the conformation space predicted for underivatized peptides. (B) A plot of IM-MS conformation space for the same tryptic digest of Lysozyme, whereby cysteine residues are covalently derivatized with the cysteine specific shift reagent coordinated with La. Although the same 12 peptides are identified, the derivatized cysteine containing peptides are mobility shifted to regions where peptide ions are not predicted to occur. For comparison purposes, the ovals indicate where the peptide signals are observed in the absence of labeling (*i.e.* in A) with arrows illustrating where these signals are observed when derivatized, which are also indicated in the mass spectrum by filled circles. The peak labeled "*" corresponds to the cysteine reducing agent DTT which has incorporated two shift reagent+La tags. Peaks observed at -44 Da relative to the derivatized peptides (and the DTT peak) correspond to in-source decay loss of CO_2 from a carboxylic acid moiety on the shift reagent. The solid lines are to assist invisualizing the correlation of collision cross section versus m/z for peptide ions.

		Average m/z ($[M+H]^+$)		Collision Cross Section ($Å^2$)[a]	
Peptide Sequence	**Amino Acid Position**	**Underivatized**	**Derivatized**	**Underivatized**	**Derivatized**
KVFGR	1-5	606.7501	-	168.93 ± 0.27	-
HGLDNYR	15-21	874.9370	-	207.61 ± 0.30	-
WWCNDGR	62-68	937.0305	1600.2645	212.28 ± 0.12	265.28 ± 0.72
GTDVQAWIR	117-125	1046.1780	-	237.39 ± 0.80	-
GYSLGNWVCAAK	22-33	1269.4737	1932.7077	277.08 ± 2.43	342.87 ± 3.65
CKGTDVQAWIR	115-125	1277.4968	1940.7308	271.03 ± 1.45	-[a]
FESNFNTQATNR	34-45	1429.4977	-	278.12 ± 0.35	-
IVSDGNGMNAWVAWR	98-112	1676.9059	-	318.41 ± 0.48	-
NTDGSTDYGILQINSR	46-61	1754.8621	-	317.98 ± 0.79	-

Table 3.2. A comparison of the m/z and collision cross sections determined for cysteine containing peptides that are underivatized and derivatized with a cysteine specific shift reagent incorporating lanthanum . All peptides were generated from a tryptic digestion of hen egg white Lysozyme. The boldfaced peptides each contain a cysteine residue. [a]CKGTDVQAWIR was positively identified by mass and by use of the shift reagents, however the labeled peptide abundance was too weak (since it has a missed cleavage) to obtain an accurate collision cross section measurement.

identified cysteine containing peptides are marked for clarification. Note that all peptides fall within the peptide correlation area, whose midline is illustrated as a solid black line. A separate Lysozyme digest was labeled and analyzed, yielding many peptides within the predicted correlation band in addition to three labeled cysteine containing peptides that were shifted to a location where peptide signals are not predicted to occur (Figure 3.5b). This significant shift in conformation space permits a rapid and simple identification of cysteine containing peptides from a complex mixture of peptides. Table 3.2 summarizes the identification of cysteine containing peptides from a Lysozyme digestion. Collision cross sections are shown for comparison purposes.

In addition to identification of specific functionality in peptides and proteins, lanthanide-based labels may also be used for relative quantitation purposes, as previously described[13-15] in Chapter 2. Analysis of relative quantitation experiments may be completed on IM-MS instrumentation in addition to one dimensional MALDI TOFMS.

Figure 3.6 illustrates the similarity of data between the two instrument platforms. Panel A shows the type of data acquired using MALDI TOFMS (Chapter 2). Panel B represents both the extracted mass spectrum (upper portion) and a view of IM-MS conformation space (lower portion) acquired using a custom IM-MS instrument described previously in this chapter.

3.4 Conclusions

This chapter has focused on the development and use of IM-MS shift reagents for the rapid identification of specific chemical functionality in a complex mixture of peptides. The shift reagents described here consist of a cysteine or primary amine

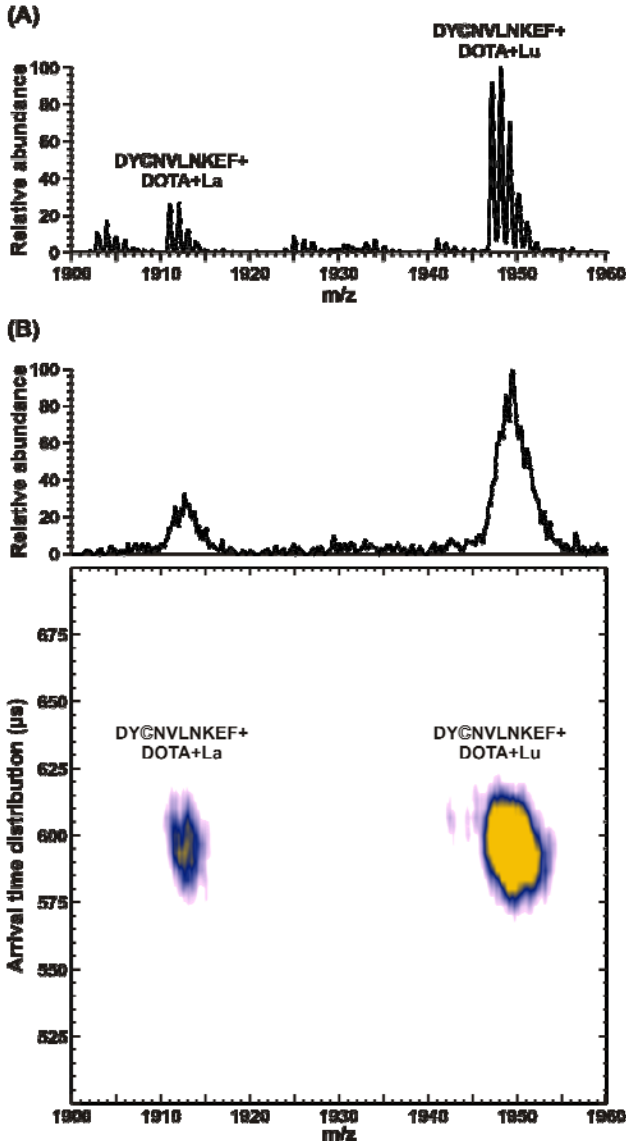


Figure 3.6. A comparison of the types of data collected using MALDI TOFMS (panel A) and drift tube IM-MS instrumentation. The spectra illustrate a relative quantitation experiment whereby a peptide (DYCNVLNKEF) was labeled using the cysteine specific label and La/Lu metal system in a 1:5 ratio. Panel B illustrates the extracted mass spectrum as well as the 2D conformation space.

specific label that is also capable of incorporating a lanthanide metal atom. The lanthanide metal provides the necessary properties of high mass and relatively low collision cross section required to shift peptides signals to an area in conformation space outside of the predicted area for unlabeled peptides. The relatively large mobility shift afforded by these lanthanide-based labels facilitates rapid identification of cysteine containing peptides.

Since the label can incorporate any lanthanide metal, relative quantitation experiments are possible by tagging various experimental states, each with a different lanthanide. Relative quantitation information may also be obtained using IM-MS instrumentation. An advantage of this type of analysis is instant identification of labeled peptides and rapid quantitation analysis. The use of multiple labels and lanthanides offers the possibility of multiplexed experiments whereby multiple functionalities are identified simultaneously.

We are grateful for financial support provided by the Defense Threat Reduction Agency (HDTRA1-09-1-0013), the National Institute on Drug Abuse (RC2DA028981), the Vanderbilt University College of Arts and Sciences, the Vanderbilt Institutes for Chemical Biology, the Vanderbilt Institute for Integrative Biosystems Research, and Ionwerks, Inc (Houston, TX).

References

1. Li, F.; Xie, Z.; Schmidt, H.; Sielemann, S.; Baumbach, J.I. “Ion Mobility Spectrometer for Online Monitoring of Trace Compounds.” *Spectrochim. Acta B.* **2002**, 57, 1563-1574.

2. Miller, R.A.; Nazarov, E.G.; Eiceman, G.A.; King, A.T. “A MEMS Radio-frequency Ion Mobility Spectrometer for Chemical Vapor Detection.” *Sensors and Actuators A.* **2001**, 91, 301-312.

3. McLean, J.A.; Ruotolo, B.T.; Gillig, K.J.; Russell, D.H. “Ion Mobility-Mass Spectrometry: A New Paradigm for Proteomics.” *Int. J. Mass Spectrom.* **2005**, 240, 301-315.

4. Hilderbrand, A.E.; Myung, S.; Clemmer, D.E. “Exploring Crown Ethers as Shift Reagents for Ion Mobility Spectrometry.” *Anal. Chem.* **2006**, 78, 6792-6800.

5. Howdle, M.D.; Eckers, C.; Laures, A.M.F.; Creaser, C.S. “The Use of Shift Reagents in Ion Mobility-Mass Spectrometry: Studies on the Complexation of an Active Pharmaceutical Ingredient with Poly Ethyleneglycol Excipients.” *J. Am. Soc. Mass Spectrom.* **2009**, 20, 1-9.

6. Zhu, X.; Lever, S.Z. “Formation Kinetics and Stability Studies on the Lanthanide Complexes of 1,4,7,10-tetra-azacyclododecane-*N,N’,N’’,N’’’*-tetraacetic acid by Capillary Electrophoresis.” *Electrophoresis.* **2002**, 23, 1348-1356.

7. Park, Z.Y.; Russell, D.H. “Thermal Denaturation: A Useful Technique in Peptide Mass Mapping.” *Anal. Chem.* **2000**, 72,2667-2670.

8. Mason, E.A.; McDaniel, E.W. *Transport Properties of Ions in Gases.* John Wiley & Sons: Indianapolis, IN, 1988.

9. McLean, J.A.; Schultz, J.A.; Woods, A.S. In *Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicabilities, and Biological Applications*, Richard B. Cole, Ed.; John Wiley & Sons: Indianapolis, IN, 2007; pp 411-439.

10. Tao, L.; McLean, J.R.; McLean, J.A.; Russell, D.H. “A Collision Cross-section Database of Singly-charged Peptide Ions.” *J. Am. Soc. Mass Spectrom.* **2007**, 18, 1232-1238.

11. Yates, J.R.; Speicher, S.; Griffin, P.R.; Hunkapiller, T. “Peptide Mass Maps: A Highly Informative Approach to Protein Identification.” *Anal. Biochem.* **1993**, 214, 397-408.

12. Jensen, O.N.; Podtelejnikov, A.V.; Mann, M. “Identification of the Components of Simple Protein Mixtures by High-accuracy Peptide Mass Mapping and Database Searching.” *Anal. Chem.* **1997**, 69, 4741-4750.

13. Whetstone, P.A.; Butlin, N.G.; Corneillie, T.M.; Meares, C.F. “Element-coded Affinity Tags for Peptides and Proteins.” *Bioconj. Chem.* **2004**, 15, 3-6.

14. Liu, H.; Zhang, Y.; Wang, J.; Wang, D.; Zhou, C.; Cai, Y.; Qian, X. “Method for Quantitative Proteomics Research by Using Metal Element Chelated Tags Coupled with Mass Spectrometry.” *Anal. Chem.* **2006**, 78, 6614-6621.

15. Rappel, C.; Schaumlöffel, D. “Absolute Peptide Quantification by Lutetium Labeling and NanoHPLC-ICPMS with Isotope Dilution Analysis.” *Anal. Chem.* **2009**, 81, 385-393.

CHAPTER IV

MULTIPLEXED ANALYSIS OF PEPTIDE FUNCTIONALITY USING METAL CHELATION LABELS

4.1 Introduction

IM-MS shift reagents take advantage of the narrow peptide correlation bands in conformation space by shifting specific signals to an area in which peptides signals are not ordinarily predicted to occur in the absence of labeling. The curvature of the peptide correlation band ensures that the mass (and small cross section increase) added offer a significant shift for rapid identification of labeled peptides.

Analysis of multiple protein samples simultaneously has been completed previously. Most of these experiments have centered around the use of gels in a differential gel electrophoresis (DIGE) experiment.[1] Classically 2D gels are used for this purpose. The first dimension of separates proteins based on their isoelectric point. Separation in this manner separates proteins that vary in native solution-phase charge. The second dimension of separation is a typical gel electrophoresis experiment whereby smaller molecules travel faster (and thus further through the gel) than larger molecules. Although not perfect, higher mass proteins tend to adopt larger structures, restricting their movement through the gel compared to small molecules.

Multiplex labeling is accomplished through the use of cyanine dyes. Briefly, several dyes (each with a different excitation wavelength) are used to modify protein samples, one dye for each protein sample. For example, a control will be labeled with

one dye while an experimental sample is labeled with another dye. The samples are mixed and analyzed by 2D gel electrophoresis. The gel is analyzed using different wavelengths of excitation light to visualize the control and experimentally labeled samples independently. Using cyanine dyes allows for a general indication of changes in the protein expression level and corrects and is not limited by gel drift since the two samples are analyzed on the same gel.

The multiplex labeling described in this chapter is envisioned as similar to the classical multiplex dye labeling utilized for 2D gel analysis; however peptides are labeled at multiple orthogonal functionalities using two different lanthanide metals. The use of two labels with two different metals provides for several interesting experiments. For example, MS/MS analysis would allow for site identification of the two functionalities labeled. Additionally, ICP-based analysis could provide a relative ratio of the functionalities labeled.

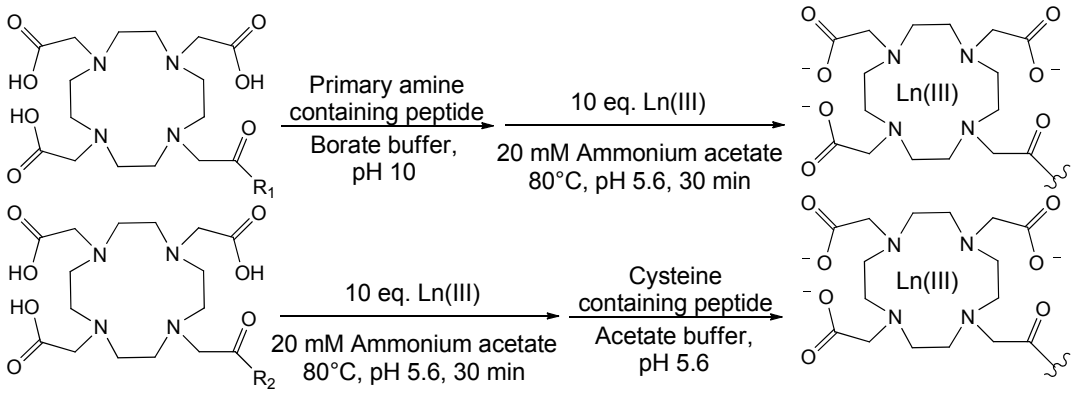
4.2 Experimental

All peptides were purchased from American Peptide Company (Sunnyvale, CA) and proteins, as well as matrices, were purchased from Sigma Aldrich (St. Louis, MO). Lanthanide metals (as $LnCl_3$ hydrates) were purchased from Strem Chemical for Research (Newburyport, MA) and Sigma Aldrich. Functionally specific lanthanide-chelating labels were purchased from Macrocyclics (Dallas, TX). All common solvents were also purchased from Sigma Aldrich. Deionized water was obtained from a Millipore (Bedford, MA) water purification system. All chemicals purchased were used without further purification.

Peptides with differing numbers of primary amine sites (*e.g.* N-terminal primary amine and Lys residue side chains) were selectively labeled by an IM-MS shift reagent consisting of a lanthanide-chelation moiety, a short linker region, and an N-hydroxysuccinimide moiety for the specific modification of primary amine sites. Labeling reactions proceeded as illustrated in Scheme 4.1 whereby primary amine sites were modified by adding a ten-fold molar excess of amine specific label at a borate buffered pH of 10. Subsequent to amine modification, a lanthanide (Ln(III)) metal atom is chelated within the macrocycle by adding a ten-fold molar excess of Ln(III) and heating to 80˚C at an acetate buffered pH of 5.7.[2]

All labeled peptides were desalted using ziptips (Millipore) prior to analysis. MS analysis was completed on an Applied Biosystems (Carlsbad, CA) Voyager DE MALDI TOFMS using 2,5-dihydroxybenzoic acid (DHB) as the matrix. Ion mobility analysis was completed on a Waters Synapt G2 ion mobility-mass spectrometry instrument. Labeled peptide samples were spotted onto a MALDI target plate using the dried droplet method with 2,5-dihydroxybenzoic acid (DHB) as the matrix in a matrix:analyte ratio of approximately 100:1.

Model cysteine-containing peptides were labeled using a shift reagent consisting of a lanthanide-chelation moiety, a short linker region, and a maleimide group for the selective modification of the thiol side chain of cysteine residues. Labeling reactions proceeded as shown in Scheme 4.1. Briefly, a Ln(III) metal was chelated with the macrocycle by heating a ten-fold excess of metal and label to 80˚C at an acetate buffered pH of 5.7. Subsequent to metal chelation, cysteine-containing peptides were modified by adding a ten-fold excess of lanthanide-shift reagent complex to peptides at an acetate buffered pH of 5.7.[2] All labeled peptides were desalted using ziptips (Millipore) prior to MS analysis as detailed above.

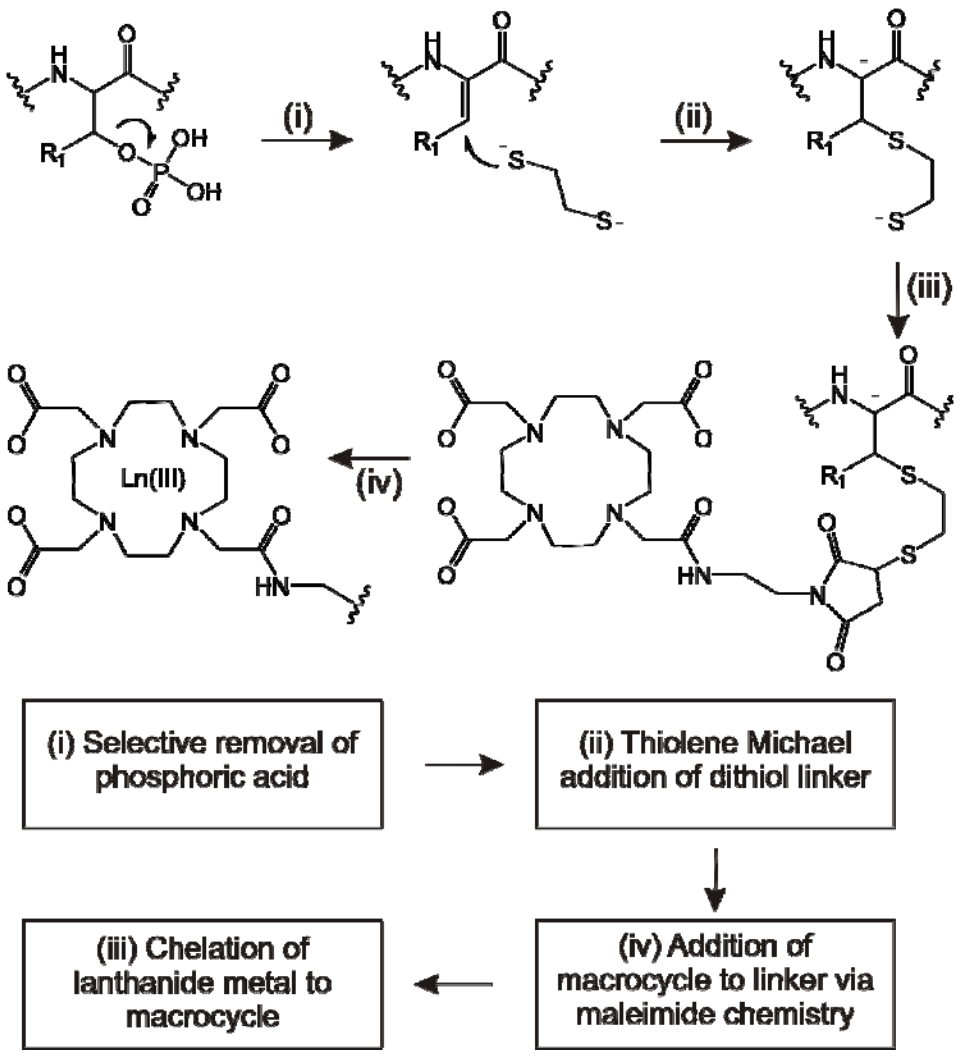


Scheme 4.1. An illustration of the labeling reactions for primary amine specific (R_1= *N*-hydroxysuccinimide) and cysteine specific (R_2=maleimide) shift reagents. The incorporation of the lanthanide metal (Ln III) provides the necessary properties of high mass and the relatively small increase in collision cross section required for high density shift reagents.

Proteins used for simulation of complex mixtures were proteolyzed according to previously published methods.[3] Briefly, proteins were reduced (if required) and heat denatured at 90˚C. Ammonium bicarbonate buffer (pH ~8) was added followed by trypsin in a ratio of 1:40 w:w. Digestion was allowed to proceed overnight (approximately 18-24 hours) at 37˚C. Upon completion, the resulting peptides were desalted using C18 spin columns (Pierce, Rockford, IL). The resulting peptides were then labeled with cysteine specific or primary amine specific shift reagents as shown in Scheme 4.1. Labeled peptides were desalted once again using ziptips prior to analysis.

Post-ionization IM-MS separations selective for phosphorylation residues were demonstrated using bovine beta-casein in a proof-of-concept experiment. The protein was dissolved in ammonium bicarbonate buffer (pH ~8) and thermally denatured at 90˚C for 20 minutes and quenched at -20˚C.[3] Reduction and alkylation was not performed as there are no dicysteinyl bonds in this protein. The protein was subsequently digested with trypsin in a 1:40 w:w ratio for 16-20 hours at room temperature and desalted using C18 spin columns prior to derivatization.

Phosphopeptide specific labeling reaction proceeded as illustrated in Scheme 4.2 whereby tryptic peptides were subjected to a one-pot beta elimination (Scheme 4.2i)/ Michael addition[4] (Scheme 4.2ii) reaction whereby the site of phosphorylation loses phosphoric acid, followed by the addition of an ethanedithiol linker. The beta elimination/ Michael addition reaction mixture consisted of 2.5 mM EDTA, 0.2M ethanedithiol, 0.5M NaOH, 1.5M acetonitrile, 1.5M ethanol, 5M DMSO, and water. The reaction was allowed to progress for 1-2 hours under a nitrogen atmosphere at 55˚C in a manner similar to reaction conditions described previously.[5-7] The samples were then neutralized and purified by gel filtration (Sephadex G-10, Sigma) and reaction completion was confirmed by MALDI TOFMS. Subsequently, the thiolated peptides were labeled with a ten-fold



Scheme 4.2. A schematic diagram of the labeling reactions utilized for selective labeling at sites of phosphorylation.

excess of maleimido-DOTA (Scheme 4.2iii) in a mixture containing acetate buffer (pH 5.7) and DMSO in a 1:1 ratio (v/v), resulting in a covalent modification of the free sulfhydryl group. Finally, a lanthanide metal was chelated to the macrocycle portion of the tag by adding a 100-fold molar excess of metal to peptide and heating to 80˚C for 45 minutes (Scheme 4.2iv).[2]

Differentially labeled samples were then combined and desalted, using C18 spin columns, and analyzed by MALDI IM-MS. Peaks exhibiting a negative deviation from the peptide correlation function were selected for fragmentation in the transfer portion of the Synapt G2 instrument.

Multiplexed labeling was accomplished by labeling primary amines first (including lanthanide incorporation). In a separate reaction, a different lanthanide metal was chelated with the cysteine specific label. The lanthanide-cysteine label complex was then used to modify cysteine residues on the labeled peptide. Different metals can be used to optimize the total mass shift of the multiply labeled peptide or enhance ICP analysis. Samples subjected to multiplex labeling were desalted using ziptips prior to MS and IM-MS analysis as described above.

The fragmentation efficiency of both unlabeled and labeled peptides was also investigated. Peptides (either unlabeled or Tb labeled) were fragmented under collision induced dissociation conditions (CID) using a Waters Synapt G2 instrument. The collision energy was titrated from very low, which did not fragment the peptides to a significant amount, to high energy until complete fragmentation of the parent ion was observed. The fragmentation efficiency (a ratio of all signal excluding the parent ion to all signal, including the parent ion) was plotted as a function of collision energy for analysis.

4.3 Results and discussion

High density shift reagents function by shifting specific peptide signals to an area in IM-MS conformation space where the parent signal is not predicted to occur in the absence of labeling. The high mass of the lanthanide and the minimal branching of the label ensures that labeled peptide signals are shifted to an area below the peptide correlation band. Qualitatively, shift reagents allow for rapid, confident identification of peptides since it is known that all shifted signals contain a specific functionality.

Labeling primary amines ensures that peptides from a complex mixture (*e.g.* a tryptic digest) each receive at least one label. Peptides that contain a lysine residue are labeled twice (*i.e.* the N-terminus and Lys side chain). While the correlation band is shifted since each peptide is labeled, the addition of subsequent labels shift the signal even further from the unlabeled peptide correlation band. In addition, lanthanide-based shift reagents are useful in the site localization of specific functionality since the labels remain intact during typical tandem MS fragmentation conditions.

Protein phosphorylation constitutes a large portion of biologically important post-translational modifications, and elucidating the site of phosphorylation presently remains challenging. Challenges in phosphoproteomics continue to lie in the modification's lability and the substoichiometric amounts available. By selective labeling of this PTM with a high density shift reagent followed by IM-MS analysis, phosphorylated peptides and proteins are shifted to regions of IM-MS conformation space where signal is not predicted to occur. In this manner, phosphorylated peptides may be identified for fragmentation and further analysis (*i.e.* site identification) in a manner more rapid and efficient than traditional data-dependent tandem MS methodologies.

Bovine beta-casein, a model phosphorylated protein, was selectively labeled with a holmium chelated mobility shift label to visually identify phosphorylated peptides in 2D IM-MS conformation space prior to fragmentation. The resulting underivatized tryptic peaks that establish the peptide correlation function and the derivatized phosphorylated peptides are shown in figure 4.1a. Upon visualization of a negative deviation from the peptide correlation function, mobility shifted phosphorylated peptides were selected for fragmentation to elucidate the site of phosphorylation. One mobility-shifted peak was identified as a holmium-labeled tryptic peptide FQpSEEQQQTEDELQDK (where p indicates phosphorylation of the following amino acid, in this case serine), was identified.

The tandem spectrum is shown in figure 4.1b. The expected b/y series ions were observed, including b ion coverage from the b_3 to the b_{15} ion (b_{14} was observed but was not considered significant as it did not have a sufficient S/N value), labeled y_{14} and y_{15} ions, and several labeled water loss ions with little evidence for fragmentation of the label. The stability of this label enables predictable mass shifts of anticipated b and y ions[8-10] which gives an indication of the site of modification previously phosphorylated. Furthermore, because this tag is not labile and does not show phosphorylation site rearrangement, it has an added advantage of more confident phosphorylation site assignment.[11]

Multiplexed labeling (*i.e.* labeling multiple functionalities on a single peptide) is possible using lanthanide-based shift reagents. Using a different lanthanide metal for each label affords the ability to quantify the relative ratio of multiple functionalities on a single peptide using an ICP based instrument. The relative ratio of functionalities would allow for higher confidence in assigning peptide identification. Since the labels remain intact through collision induced dissociation, tandem MS experiments (Figure 4.2) for peptide identification and functionality site localization are possible.

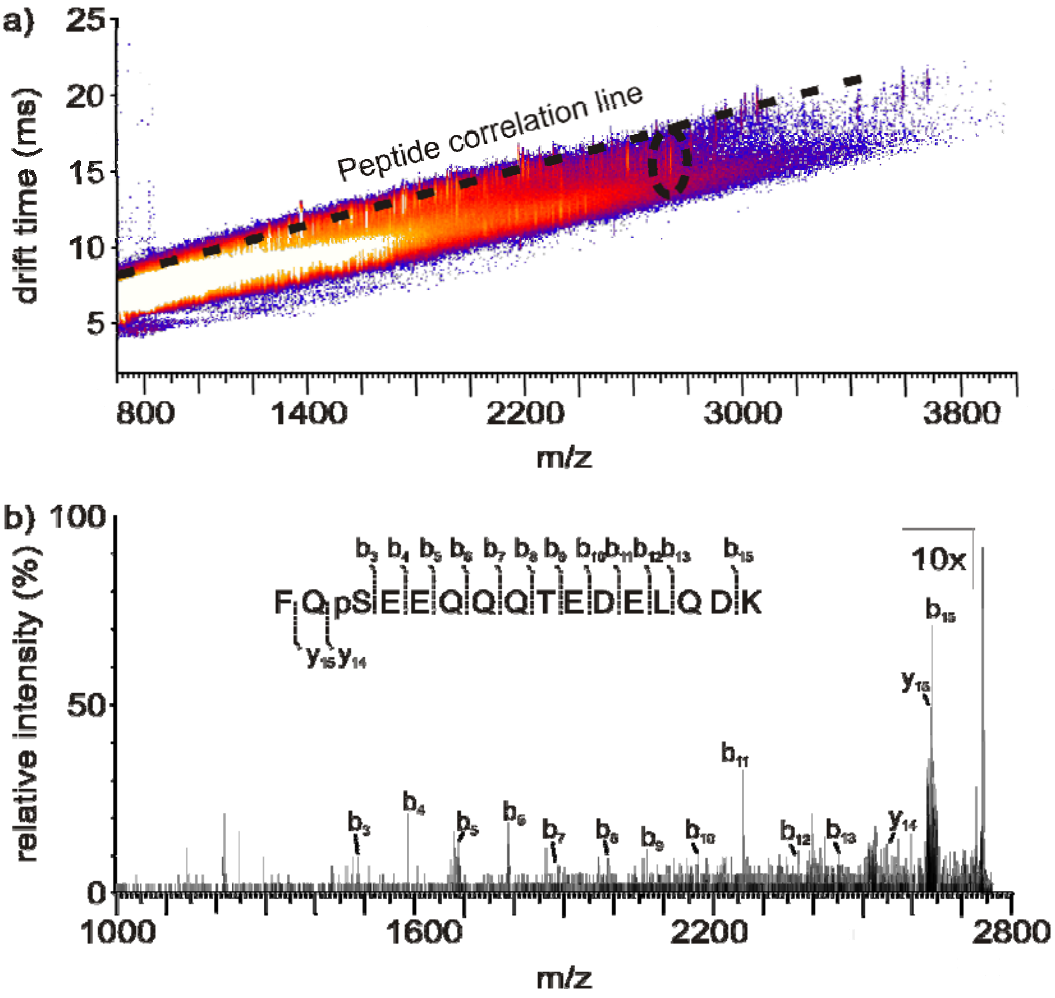


Figure 4.1. (a) 2D IM-MS plot of derivatized tryptic beta-casein. The signal corresponding to derivatized FQpSEEQQQTEDELQDK (indicated in a white dashed circile) exhibits a negative deviation from the peptide correlation line (indicated by a black dashed line), facilitating visual identification prior to fragmentation. (b) Fragmentation spectrum of a Ho-labeled tryptic beta-casein phosphorylated peptide having the sequence FQpSEEQQQTEDELQDK. Spectral peaks from 1000 m/z to 2730 m/z were enhanced 10x to increase visibility of b and y spectral assignments, as the $[M+H]^+$ ion dominated the spectrum. Observed fragmentation peaks are indicated on the peptide sequence above the spectrum as well as in the spectrum. Fragmentation coverage of 80.0% and 13.3% of the labeled peptide was observed for b and y ions, respectively. Importantly, all of the anticipated ions corresponding to labeled positions are observed, demonstrating the utility of this label for phosphopeptide site identification as well as relative quantitation.

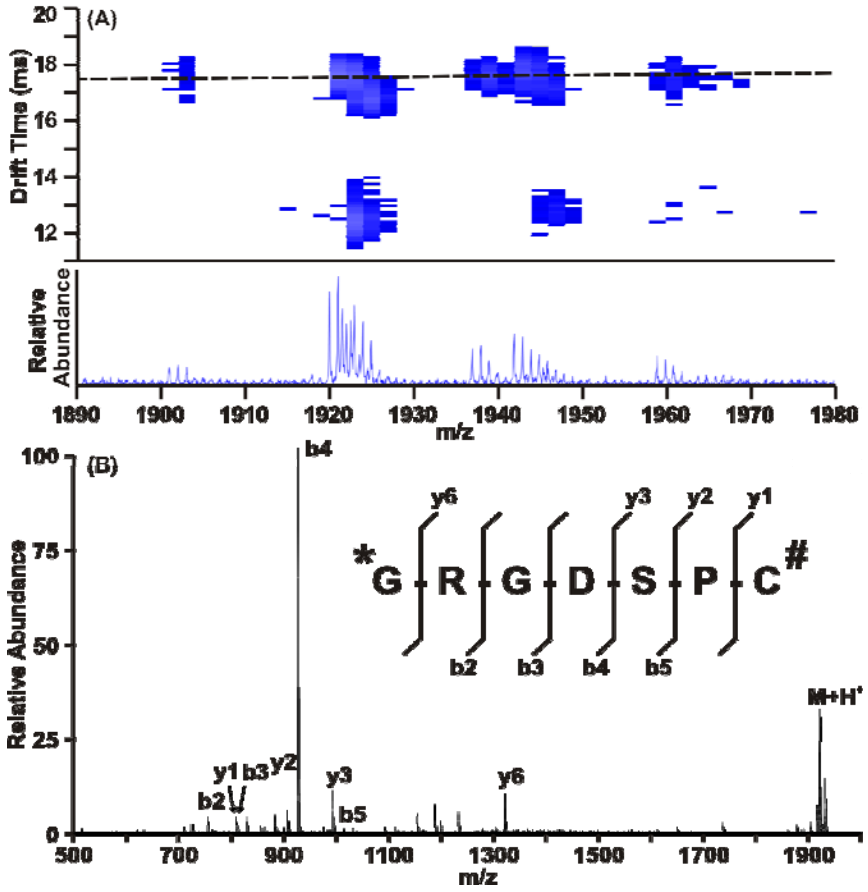


Figure 4.2. (A) A representation of IM-MS analysis of multiplexed labeling. Briefly, a peptide (sequence shown in part (B)) was labeled at both the N-terminal primary amine (Tb labeled) and a cysteine residue (Ho labeled) according to Scheme 4.1. Panel (A) illustrates the peptide correlation band (illustrated as a dashed line) showing a tryptic digest of bovine serum albumin. The lower peaks are due to the multiply labeled peptide shown in (B). The labeled peptide was subjected to MS/MS to ensure confident identification and label stability.

Labeled and unlabeled peptides were fragmented using MALDI TOFMS (Waters Synapt G2). Each peptide was fragmented at various energies under CID conditions. The fragmentation energy (Σ Fragment signal / Σ Total signal) of the unlabeled peptide was plotted versus collision energy in arbitrary units, as described previously.[12] Fragmentation data for labeled peptides were also collected; however the collision energy was corrected using a linear degrees of freedom (DOF) correction (Equation 4.1) in order to offset the increase in collision energy required due to higher mass.[13]

$$E_{DOF} = \left[\frac{DOF_{unlabeled}}{DOF_{labeled}}\right](E_l)_{labeled} \qquad \textbf{Eq. 4.1}$$

Fragmentation data was plotted for both labeled and unlabeled peptides. The curves (Figure 4.3, 4.4, and 4.5) illustrate that a differing amount of energy (relative to the corresponding unlabeled peptide) is required for fragmentation of 50% of the parent signal. For example, some labeled peptides require more energy than the unlabeled peptide (Figure 4.3 for example) while some labeled peptides require less energy (Figure 4.4 for example). This shift in required collision energy is attributed to sequence effects since more acidic peptides tend to fragment more easily. The removal of a basic site by the addition of the label is thought to increase the acidic nature of the peptide sequence. If there are several other basic sites that remain unlabeled, the effect seems to be small, however if one of the last basic sites is removed by the addition of the label, the acidic nature of the peptide increases more.

The mobile proton model[13-15] specifies that a peptide may be protonated at specific locations (*e.g.* basic residues) more frequently than other locations within the peptide. In order for fragmentation to occur, however, this ionizing proton must shift to other portions of the peptide for rearrangement reactions leading to peptide fragments.

Peptides that contain arginine tend to require more fragmentation energy.[16] The fragmentation data presented here confirms this idea.

Upon labeling, peptides that contain arginine (which is not labeled) retain a highly basic site while losing a less basic site (*i.e.* Lys residue side chain or the N-terminus) to the label. These labeled amines become amides. Alternatively, peptides that do not contain arginine are labeled at their most basic site, making the entire peptide more acidic. Acidic (or neutral) peptides do not localize the ionizing proton as well since it binds to the amide nitrogen or oxygen atoms. Since many of the possible locations for the proton fall within a narrow range of proton affinity (compared to Lys or Arg), acidic peptide fragment more easily.

4.4 Conclusion

IM-MS shift reagents can be used to gain several types of information from samples. These reagents can be used to identify several relevant chemical functionalities, including primary amine sites, cysteine sites, and phosphorylation sites. In addition, the use of different metals for different samples affords the utility of relative quantitation experiments. Primary amine sites work well as targets for relative quantitation labels since nearly all tryptic peptides have at least one (or more) primary amine sites. Cysteine labeling can be used to locate cysteine residues within peptides, thus increasing the confidence of peptide identification. Furthermore, phosphorylation sites may be elucidated in a manner more rapid and with great confidence than traditional data-dependent tandem MS methods.

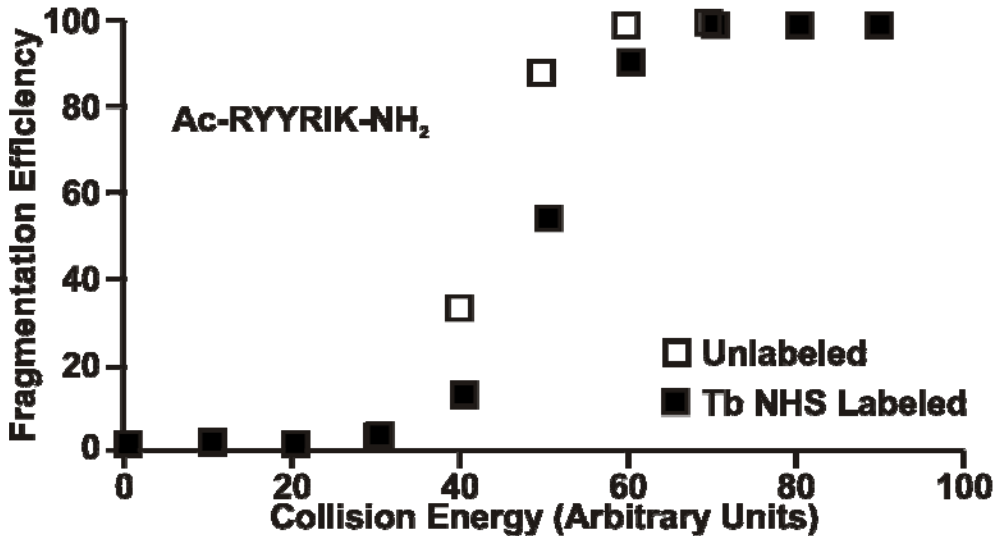


Figure 4.3. A fragmentation efficiency curve for a peptide with the sequence Ac-RYYRIK-NH_2. In this particular case, the labeled peptide required more energy for fragmentation of 50% of the parent signal.

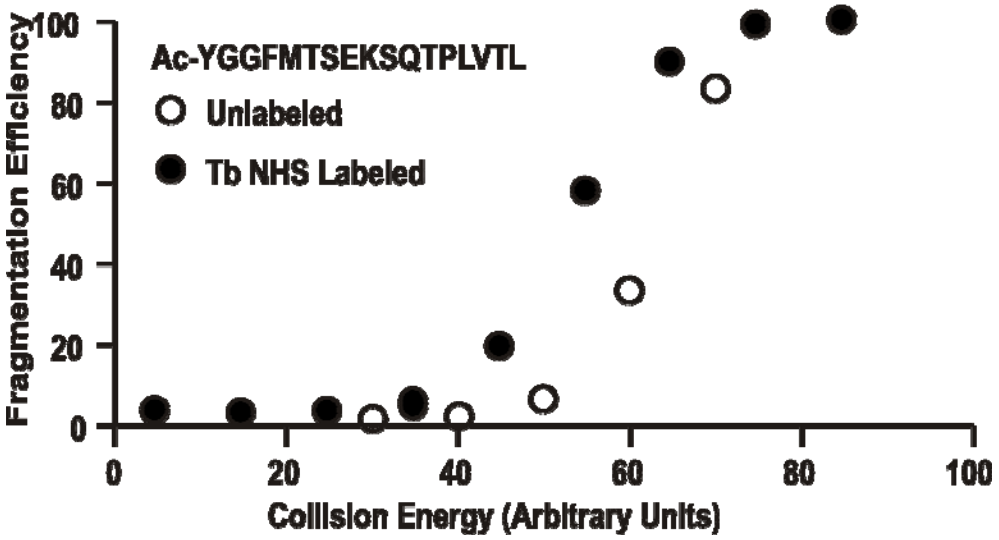


Figure 4.4. A fragmentation efficiency curve for a peptide with the sequence Ac-YGGFMTSEKSQTPLVTL. Unlike figure 4.3, the labeled peptide in this case required less energy for fragmentation of 50% of the parent signal.

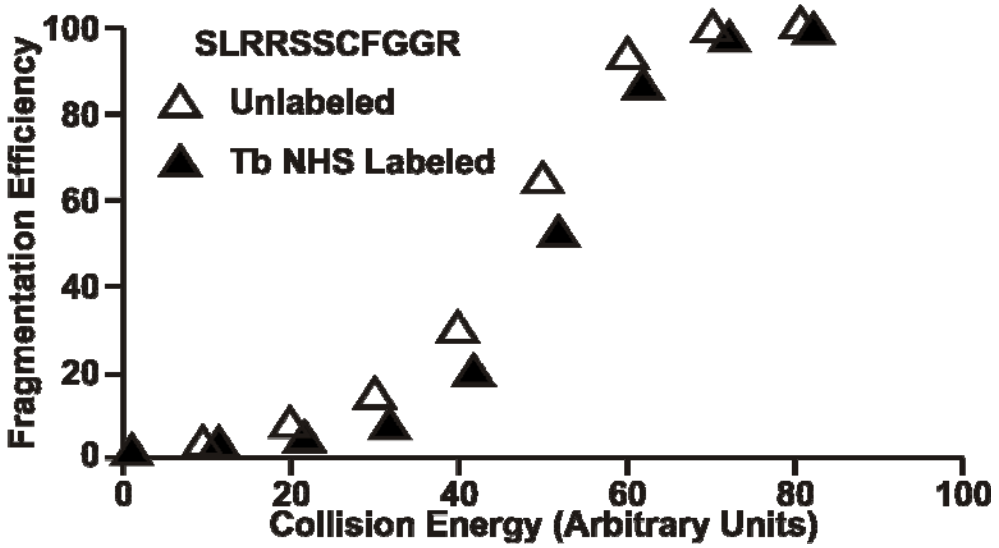


Figure 4.5. A fragmentation efficiency curve for a peptide with the sequence SLRRSSCFGGR. This labeled peptide required nearly the same energy for fragmentation of 50% of the parent signal compared with the unlabeled curve for the same sequence.

Multiplexed labeling can also be used to rapidly identify specific peptides that contain both primary amine sites (*e.g.* Lys residues) and cysteine residues. Further, these multiply labeled functionalities can be quantified using ICP based instruments. For example, one lanthanide can be chelated to the primary amine label and a different lanthanide could be chelated to the cysteine specific label. ICP experiments can elucidate the relative ratio of these two functionalities.

We are grateful for financial support provided by the Defense Threat Reduction Agency (HDTRA1-09-1-0013), the National Institute on Drug Abuse (RC2DA028981), the Vanderbilt University College of Arts and Sciences, the Vanderbilt Institutes for Chemical Biology, the Vanderbilt Institute for Integrative Biosystems Research, and Ionwerks, Inc (Houston, TX).

References

1. Alban, A.; David, S.O.; Bjorkesten, L.; Andersson, C.; Sloge, E.; Lewis, S.; Currie, I. "A Novel Experimental Design for Comparative Two-dimensional Gel Analysis: Two-dimensional Difference Gel Electrophoresis Incorporating a Pooled Internal Standard." *Proteomics* **2003**, 3, 36-44.

2. Zhu, X.; Lever, S.Z. "Formation Kinetics and Stability Studies on the Lanthanide Complexes of 1,4,7,10-tetra-azacyclododecane-*N,N',N'',N'''*-tetraacetic acid by Capillary Electrophoresis." *Electrophoresis.* **2002**, 23, 1348-1356.

3. Park, Z.Y.; Russell, D.H. "Thermal Denaturation: A Useful Technique in Peptide Mass Mapping." *Anal. Chem.* **2000**, 72,2667-2670.

4. Hoyle, C.E.; Bowman, C.N. "Thiol-ene Click Chemistry." *Angewandte Chemie Int. Ed.* **2010**, 49, 1540-1573.

5. Meyer, H.E.; Hoffmann-Posorske, E.; Korte, H.; Heilmeyer Jr., L.M.G. "Sequence Analysis of Phosphoserine-containing Peptides: Modification for Picomolar Sensitivity." *FEBS Letters* **1986**, 204, 61-66.

6. Mega, T.; Hamazume, Y.; Nong, Y.M.; Ikenaka, T. "Studies on the Methods for the Determination of Phosphorylation Sites in Highly Phosphorylated Peptides or Proteins: Phosphorylation Sites of Hen Egg White Riboflavin in Binding Protein." *J. Biochem.* **1986**, 100, 1109-1116.

7. Fadden, P.; Haystead, T.A.J. "Quantitative and Selective Fluorophore Labeling of Phosphoserine on Peptides and Proteins: Characterization at the Attomole Level by Capillary Electrophoresis and Laser-induced Fluorescence." *Anal. Biochem.* **1995**, 225, 81-88.

8. Whetstone, P.A.; Butlin, N.G.; Corneille, T.M.; Meares, C.F. "Element-coded Affinity Tags for Peptides and Proteins." *Bioconj. Chem.* **2004**, 15, 3-6.

9. Ahrends, R.; Pieper, S.; Kuhn, A.; Weisshoff, H.; Hamester, M.; Lindemann, T.; Scheler, C.; Lehmann, K.; Taubner, K.; Linscheid, M.W. "A Metal-coded Affinity Tag Approach to Quantitative Proteomics." *Moll. Cell. Prot.* **2007**, 6, 1907-1916.

10. Kerr, T.J.; McLean, J.A. "Peptide Quantitation Using Primary Amine Selective Metal Chelation Labels for Mass Spectrometry." *Chem. Comm.* **2010**, 46, 5479-5481.

11. Palumbo, A.M.; Reid, G.E. "Evaluation of Gas-phase Rearrangement and Competing Fragmentation Reactions on Protein Phosphorylation Site Assignment Using Collision Induced Dissociation-ms/ms and ms^3." *Anal. Chem.* **2008**, 80, 9735-9747.

12. Jones, J.L.; Dongré, A.R.; Somogyi, Á; Wysocki, V.H. "Sequence Dependence of Peptide Fragmentation Efficiency Curves Determined by Electrospray Ionization/Surface-induced Dissociation Mass Spectrometry." *J. Am. Chem. Soc.* **1994**, 116, 8368-8369.

13. Dongré, A.R.; Jones, J.L.; Somogyi, Á; Wysocki, V.H. ”Influence of Peptide Composition, Gas-phase Basicity, and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model.” *J. Am. Chem. Soc.* **1996**, 118, 8365-8374.

14. Johnson, R.S.; Krylov, D.; Walsh, K.A. “Proton Mobility Within Electrosprayed Peptide Ions.” *J. Mass Spectrom.* **1995**, 30, 386-387.

15. Vaisar, T.; Urban, J. “Gas-phase Fragmentation of Protonated Mono-*N*-Methylated Peptides. Analogy with Solution-phase Acid-catalyzed Hydrolysis.” *J. Mass Spectrom.* **1998**, 33, 505-524.

16. Wysocki, V.H.; Tsaprailis, G.; Smith, L.L.; Breci, L.A. “Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation.” *J. Mass Spectrom.* **2000**, 35, 1399-1406.

CHAPTER V

CONCLUSIONS AND FUTURE DIRECTIONS

5.1 Summary and conclusions

In this dissertation, high density IM-MS shift reagents were described and their utility explored through a variety of samples and experimental conditions. Firstly, the use of IM-MS provides a rapid means of distinguishing biomolecular classes in a complex mixture of biomolecules. The shift reagents described here offer further information in that signals containing specific peptide functionalities (*e.g.* primary amine sites, cys residues, phosphorylation) may be shifted out of the peptide correlation band where they are predicted to occur to an area of conformation space not ordinarily predicted to contain peptide signals in the absence of labeling.

Initial development of lanthanide base IM-MS shift reagents proceeded using model peptides. Peptides were chosen in order to provide a range of masses as well as location of the target moiety within the peptide sequence, the number of targeted moieties in the peptide, and the overall sequence of the peptide. Early studies focused on the relative quantitation (Chapter II) and outlining the shift of labeled peptides from the unlabeled peptide correlation band in IM-MS conformation space (Chapter III).

Furthermore, the ability to label multiple types of functionalities on a single peptide with different lanthanide metals was also explored (Chapter IV). For example, peptides containing both primary amine and cysteine sites were chosen. Primary amine sites were labeled with one lanthanide metal while cysteine sites were labeled with

another metal. The multiply labeled peptides were analyzed by MS and tandem MS. The different labels were determined to be suitable for site identification of multiple functionalities on a single peptide. Additionally, labeling multiple functionalities (each with a different lanthanide metal) would allow for experiments to determine the relative ratio of labeled functionalities upon analysis using inductively coupled plasma instrumentation.

More complex samples were required in order to shift from proof of concept experiments toward biologically relevant samples. Toward this goal, the next development was labeling complex mixtures of peptides. More complex samples were simulated using tryptic proteolysis of model proteins. Digestion in this manner gave samples with peptides of varying length, sequence, number of targeted moieties, and ionization efficiency. Another step toward more complex samples was through proteolysis of multiple proteins in a single reaction. Using multiple proteins gave many more peptides, each different in sequence, length, and number of targeted moieties.

The ability to use different lanthanide metals to provide the density required for an IM-MS shift allows for relative quantitation using these same shift reagents. The different lanthanides used function similar to different isotopes used by other stable isotope methods (*i.e.* ICAT or SILAC). One advantage of using lanthanides, however, is the larger mass shift between differentially labeled samples and the flexibility to tune the mass shift to the specific sample. The larger mass shift and flexibility of the metal system avoids overlapping of the isotope cluster of high mass peptides, a potential limiting factor of ICAT and SILAC. Additionally, the lanthanide based label remains intact through collision induced dissociation (CID) processes, enabling confident sequence identification since the labels provides a frame of reference (*i.e.* doubly primary amine labeled tryptic peptides contain a Lys reside at the c-terminus).

	Ln Labels	AQUA	ICAT	SILAC	iTRAQ
Relative Cost	$	$$$	$$	$$	$$
Labeled Functionality	Cysteine and Primary Amines	N/A	Cysteine	N/A	Primary Amines
# of Simultaneous Samples	ca. 12	2	4	2	8
MS/MS Experiments	Yes	Yes*	Yes	Yes*	Yes**
Labeled Site Identification	Yes	No	Yes	No	No

* Depending on label location
** Required for quantitation

Table 5.1. A comparison of the different properties of each metabolic or chemical method of stable isotope incorportation.

The lanthanide labels described here offer several advantages over other stable isotope quantitation methods in cost, targeted functionality, and ease of use (Table 5.1). Lanthanide based labels provide relative quantitation capability by modifying either primary amines or cysteine residues. The ability to label primary amines allows for the relative quantitation of proteins that do not contain cysteine (such as the histones) where ICAT would not work. Additionally, it provides good accuracy at a much lower cost than the AQUA method. Lower cost allows for the relative quantitation of proteins and peptides on a proteome wide scale without great expense. Another final advantage is that while the label remains intact through MS/MS conditions (allowing for site identification), tandem MS is not required as it is using iTRAQ.

Overall, this dissertation describes the development of IM-MS shift reagents, which may be used to extract more information from biochemical analyses than allowed by one dimensional MS experiments. The separation of biological classes by IM-MS allows for easy identification of signals arising from peptides, however specific sequence information still requires time consuming analysis by tandem MS. The use functionally

specific IM-MS shift reagents allows for the rapid visualization of specific chemical functionalities present in a complex sample.

5.2 Future directions

Through the experiments described here, progress was made on the development, characterization, and analysis properties of high density IM-MS shift reagents. There exists, however, many opportunities for further development and application of this class of reagents. Several future directions are described here.

Further development of the use of the primary amine and cysteine shift reagents described in this dissertation could proceed through analysis of more complex and biologically relevant samples. For example, development of analysis procedures for whole cell lysate would open a new area of research. The ability to use different metals for relative quantitation purposes would be useful for the analysis of different cell states under control and experimental conditions. The multitude of lanthanides available (and the ability to choose a metal system that best suits the analysis) would allow for relative quantitation of multiple samples in a single mass spectrum.

Furthermore, analysis of clinically relevant samples is possible, since sample viability is not a requirement as it is with SILAC. Blood or other fluid analyses could be analyzed for relative quantitation of the protein component in response to drugs or other external stimuli. It may be possible to use these shift reagents for analysis of clinical samples before and after disease treatment, for example.

In addition to fluid samples, analysis of tissue samples (progress described in Appendix B) could also be developed further. Currently mouse brain tissue has been studied for development of *in situ* labeling reactions. Further development of this type of reaction could proceed through mouse or other animal tissue toward the goal of analysis of human tissue suspected to be affected by diseases such as cancer. Use of the shift reagents to aid in the analysis and identification of the proteins present in these types of samples would be beneficial. Furthermore, the use of the label allows for site identification of the tagged functionality, allowing more confidence in the sequence identification.

Another route of development is through synthesis of new classes of shift reagents. Progress toward the synthesis of a new shift reagent that incorporates a MS ionization enhancing moiety in the linker region of the label is described in Appendix A. The incorporation of ionization enhancing moieties (*e.g.* an arginine residue, as in Appendix A) would allow for analysis of less abundant peptides and proteins. The addition of a charge enhancing moiety would also allow for easier analysis of more acidic peptides. Furthermore, ionization efficiency studies would be required to determine the extent of the effect – *i.e.* are all peptides enhanced equally or does the addition of an arginine residue help to even out the ionization of peptides across a wide variety of pI values.

Synthesis of a different ionization enhancing label consisting of a linker region (a CR dipeptide), a thiol terminus, and a macrocycle is in progress through a collaboration with the Harth research lab. This label is being synthesized concurrently with the label described in Appendix A in order to arrive at a functioning shift reagent label. Currently, the dipeptide has been synthesized and tentatively confirmed by NMR

analysis. Further NMR analysis would be required for a positive identification. The addition of the macrocycle is ongoing.

The specific structure described here (containing a thiol terminus) is intended for use in labeling phosphopeptides. Many phosphopeptides show low ionization and are phosphorylated in substoichiometric amounts, creating challenges in their analysis. The additional requirement for small sample size and analysis of proteins of low abundance presents further challenges in their analysis. Great progress was made in the synthesis of this label; however the last deprotection and addition of the macrocycle still presented challenges.

Further development in the use of shift reagents could also come from development of shift reagents of a modular design, as shown in Figure 5.1. This modular design would allow for easy insertion or deletion of functionality from the linker region of the label during synthesis. For example, a betaine charge enhancing moiety may be substituted for the arginine residue. Alternatively, a biotin moiety may be substituted, allowing for rapid enrichment of labeled species from a complex mixture.

The terminal reactive moiety is also envisioned to be easily changed within the modular design. For example, the thiol moiety could be exchanged for a primary amine specific moiety in order to provide enhanced signal for labeled primary amines. Alternatively, other post translational modifications or amino acid residues could be targeted using the appropriate terminal functionality.

Development of shift reagents for ubiquitination, for example could be an area of beneficial research. Further modifications of the shift reagent structure could provide other benefits such as reduction of steric hindrance leading to shorter reaction times. In short, the future development of high density IM-MS shift reagents could proceed along

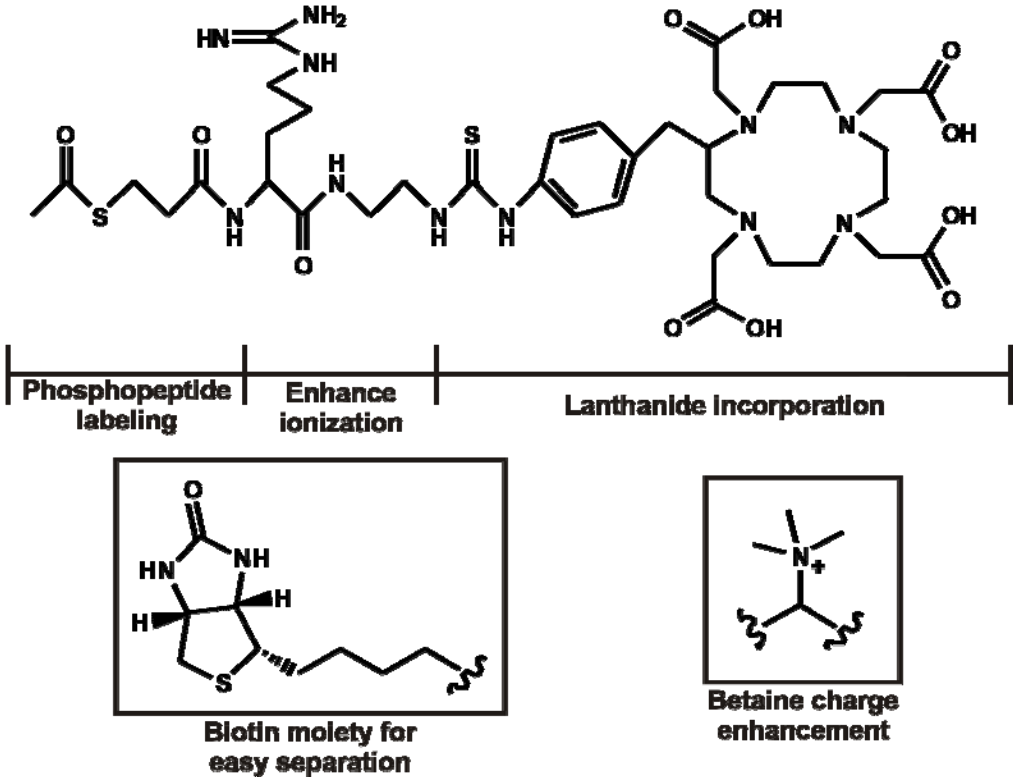


Figure 5.1 A representation of the intended modular design of future IM-MS shift reagents. A class of shift reagents is envisioned whereby synthesis proceeds in a modular fashion. For example, different charge enhancing moieties (*e.g.* arginine residue or betaine) may be used. Alternatively, a biotin moiety may be substituted, allowing for enrichment of the labeled species. Such a modular design allows for great flexibility in the mixing or exchange of different functionality within the linker region or the terminal functionality.

several routes toward becoming a more useful and developed method for analysis of biological samples.

Development of a class of easily synthesized shift reagents as described here would allow great flexibility in the analysis of many different chemical functionalities. The addition of one (or multiple) charge enhancing moieties within the linker region would ensure that even acidic species (which have low ionization efficiency) can be analyzed. The ability to exchange terminal reactive moieties allows for rapid flexibility and application toward different labeling targets as the field of mass spectrometry based proteomics progresses.

APPENDIX A

THIOL TERMINATED SHIFT REAGENT SYNTHESIS PROGRESS

The shift in proteomics toward analysis of smaller samples and less abundant proteins has pushed the limits of current instrument sensitivity. While MS offers much better sensitivity than other methods like NMR or 2D gels, further gains in sensitivity are sought. These sensitivity gains could come from several sources, namely optimization of instrument hardware and settings, or better ionization efficiency at the source.

Methods of enhancing ESI ionization were developed nearly concurrently with the development of the source itself. Using an ESI source, higher charge states allow for mass measurement using quadrupole instruments, more efficient fragmentation of protein analytes, and more signal in tandem MS spectra due to the higher likelihood that each fragment will retain some charge. Fragments of low charge state ions (especially singly protonated ions common to MALDI) are often neutral since only one fragment half retains the parent charge.

Supercharging agents, first developed about a decade ago, serve to increase the ionization state of large proteins during analysis by ESI. These reagents, commonly small molecules mixed in with the analysis solution, can significantly increase the charge state of a protein, thereby lowering its m/z value in the spectrum. The increased charge has been thought to be due to the additive reducing the surface tension of the water droplets in the spray. According to the charge residue model (CRM), lower surface tension within the droplets produces ions of higher charge state.

The charge residue model is one of two prevailing theories concerning the generation of ions using electrospray. Briefly, it states that charged droplets (containing ions) emitted from the source experience several cycles of evaporation of solvent followed by fission of the droplet, effectively reducing the concentration of ions in each droplet below the Rayleigh limit, whereby evaporation of solvent continues. The process repeats with the new daughter droplets until all of the solvent has been evaporated, leaving only ions in the gas phase. Lower surface tension, provided by the supercharging reagents, allows the charged droplet to evaporate to a smaller size before reaching the Rayleigh limit.

The second model of electrospray ionization, the ion evaporation model, is slightly different than the charge residue model. The ion evaporation model states that after emission of charged droplets from the source, ions are pulled individually out of the droplet, similar to evaporation. The remaining solvent is eventually evaporated due to low pressure and heating in the desolvation region of the source. While these two models answer some of the questions remaining about ionization, neither is perfect. It is thought that both processes play a role in the production of ions.

These supercharging reagents are traditionally used for ESI samples; however other methods of enhancing ionization exist for use with MALDI. Ionization in MALDI relies on the gas phase basicity of the analyte for coordination of a proton. Peptides containing arginine residues (high gas phase basicity, or alternatively, proton affinity) ionize much more readily than peptides that contain many acidic residues. Furthermore, permanently charged labels consisting of a quaternary amine (betaine moiety) or a sulfonium moiety greatly enhance ionization efficiency.

Application of these charge enhancing methods to proteomics analyses using MS or IM-MS was envisioned through the addition of ionization enhancing moieties (such as an arginine residue) within the structure of the shift reagent label. Toward this goal, a thiol terminated shift reagent (Figure A1) was envisioned consisting of a macrocycle for lanthanide metal chelation, a short spacer region containing an arginine residue (for enhancing ionization), and a thiol terminus for the modification of beta eliminated phosphopeptides through Michael addition. Phosphopeptides do not ionize well in positive ion mode due to common sequence motifs and the phosphate group.[1-2] Beta elimination of the phosphate group enhances ionization somewhat, however further enhancement was desired in order to improve the ease of analysis and make it possible to study lower abundance phosphoproteins.

Toward this end, a label was envisioned whereby it would function as a shift reagent, but also increase ionization. An arginine residue was chosen as the charge carrier in the linker region since it has high gas phase basicity and is a common amino acid, having two reactive termini – the N-terminus and C-terminus. Starting with arginine, a thiol group would be added, followed by a macrocycle to give the complete label shown in Figure A1.

The synthesis was envisioned as proceeding through a series of amide bond forming reactions (Scheme A1). The first reaction uses a diamino compound to modify the c-terminus of arginine to a protected primary amine, thus creating an arginine residue with two orthogonally protected primary amine groups. The reaction chosen for this step was an EDC coupling reaction.[3-5] The first primary amine protecting group was cleaved and a N-hydroxysuccinimide group was used to modify it. The next group would be deprtected and an isothiocyanate group would be used to add the macrocycle.

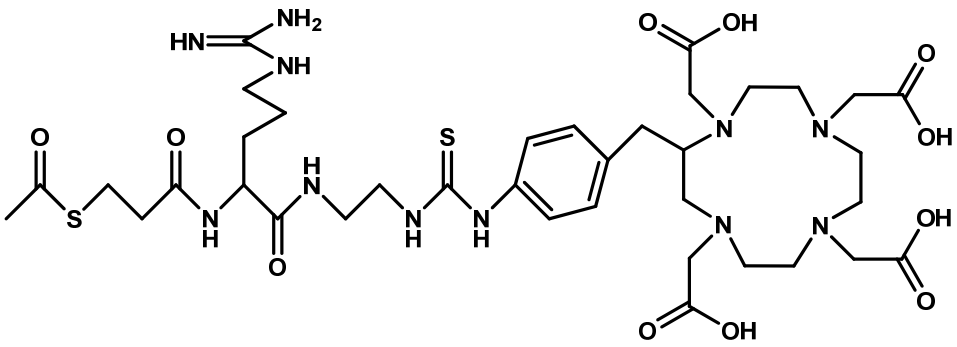
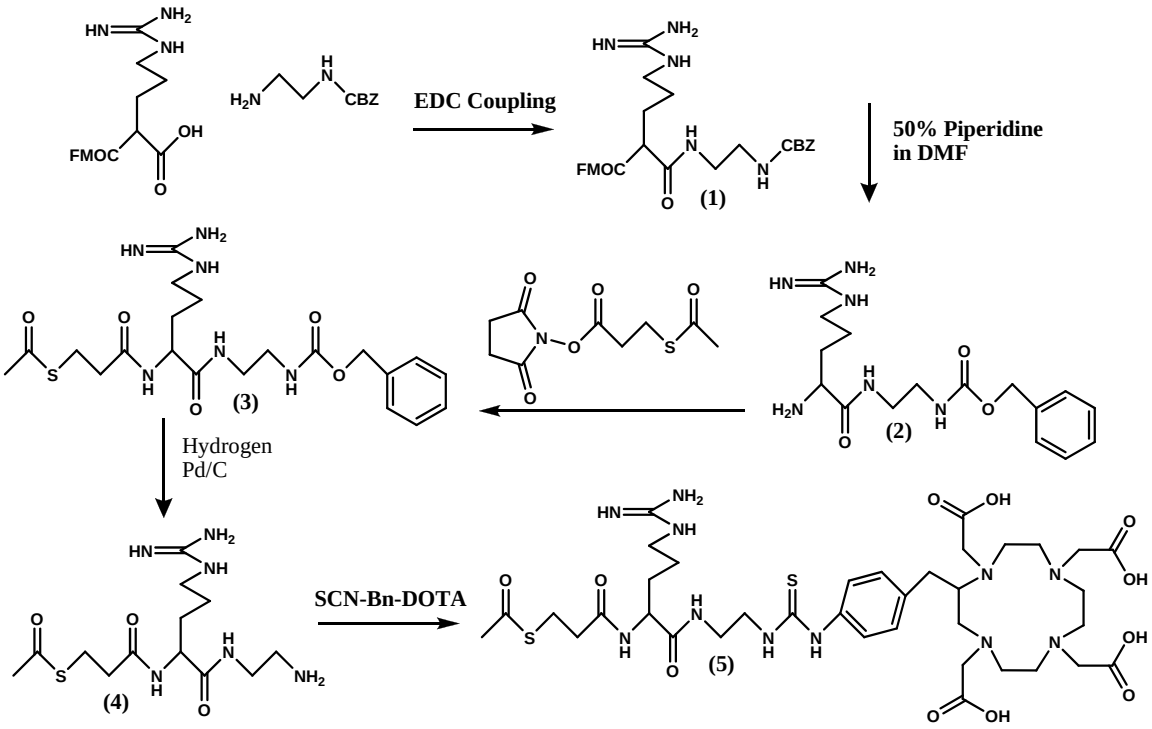


Figure A1. A representation of the proposed IM-MS shift reagent intended to enhance ionization. An arginine residue is incorporated into the liner region for this purpose. The macrocycle incorporates the lanthanide metal, while the thiol group is reactive towards beta eliminated phosphopeptides. The acetyl protecting group on the thiol may be cleaved using hydroxylamine.

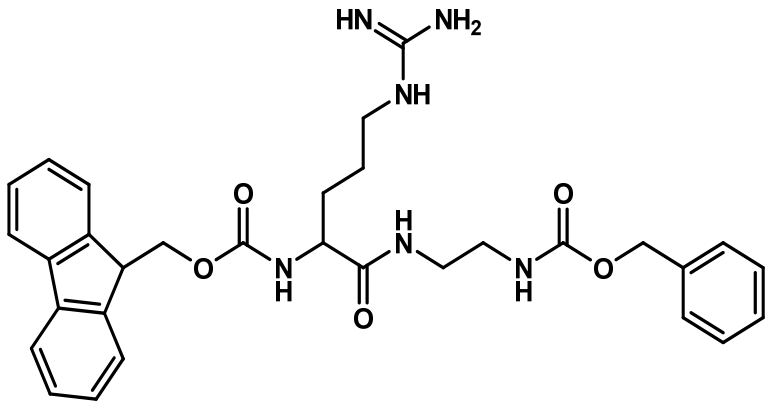


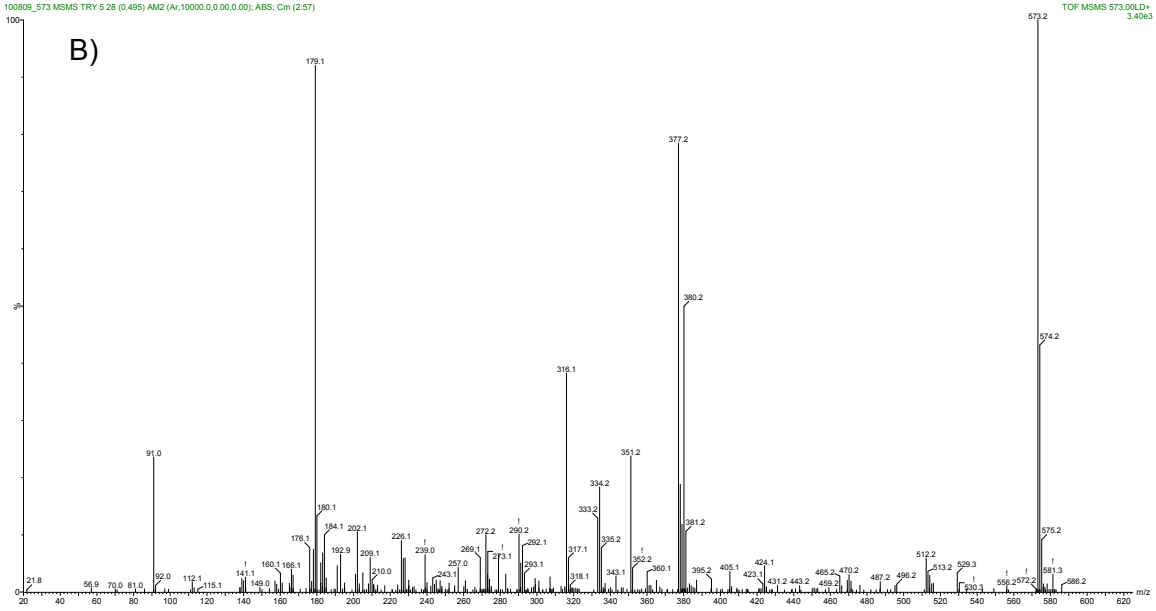
Scheme A1. A representation of the proposed synthesis of the ionization enhancing phosphopeptide shift reagents (5). The first reaction involves modification of the c-terminus of arginine to give an orthogonally protected di-primary amine version of arginine (1). The Fmoc protecting group is cleaved (2) and a protected thiol is added via SATP, resulting in compound (3). The CBZ protecting group is cleaved (4), and the macrocycle is added via isothiocyanate functionality, resulting in compound (5).

The synthesis progressed starting from fluoronylmethyloxycarbonyl protected arginine (FMOC-Arg-OH). A typical synthesis used 75mg of this starting material, 1.31g N-Z-ethylenediamine, 168µL undiluted 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC), 180 mL DI water, and 120 mL absolute ethanol. This mixture was then adjusted to pH 4.2 using concentrated HCl. The reaction was allowed to proceed at room temperature for 3.5-4 hours. Ziptips (Millipore) were used to desalt an aliquot to check reaction progress using MALDI TOFMS. The yield (represented as 573/ 396+573) at this point was usually 30-40%. The mixture is dried using a rotary evaporator (rotovap) and frozen overnight. The yield increases dramatically after drydown and freezing. The lyophillized yield should be approximately 1.8g.

Flash chromatography purification[6] is used to enrich the desired product (1). A contaminant form of the carbodiimide also coelutes to some extent, however this contaiminant does not interfere with the next reactions. Briefly, a column of 1.5 inch diameter is packed with flash grade silica gel (40-63µm particle size, 60Å pore size) to a column height of approximately six inches.[6] The solvent system used is 15% methanol in dichloromethane (DCM). The intermediate product is dissolved in a small portion of the mobile phase and run through the column. Using this solvent system, the desired product should elute at approximately three column volumes of collected fractionated mobile phase. Upon fractionation, MALDI TOFMS is used to locate the fractions of interest for further reactions. The structure of the desired intermediate was tentatively confirmed by MS/MS analysis[7] as shown in Figure A2, however further analysis by NMR would be required for confident structure assignment.

A)





556.3 529.3 514.2 495.2 472.2 465.2 422.2 379.2 377.2 351.2

334.2 238.1 221.1 195.1 193.1 179.1 178.1 165.1 150.1 91.1

86.1

Figure A2. (A) A representation of the structure of the first intermediate in the shift reagent synthesis. The theoretical mass of $[M+H]^+$ is 573.3. (B) A tandem MS spectrum of the first intermediate shown in (A). The masses listed below the spectrum are identified fragmentation peaks in the spectrum. Many of the high intensity peaks are accounted for, however confident structure confirmation requires NMR.

After isolation of the intermediate, the fluoronylmethyloxycarbonyl (FMOC) protecting group is removed using 50% piperidine in dimethylformamide (DMF),[8-9] yielding (2). Approximately 3mL of this deprotection solution is used for 10mg of intermediate up to about 15mL for 100mg of intermediate. The deprotection proceeds at room temperature for approximately 10 minutes before being dried down on a rotovap. Completion of the deprotection is confirmed using MALDI TOFMS.

After deprotection of the N-terminus of the arginine residue, it is thiolated using N-succinimidyl S-Acetylthiopropionate (SATP), yielding product (3). A protected thiol is used since any free thiols may interfere with the next reactions, especially deprotection of the carboxybenzyl (CBZ) protecting group. The acetyl group may be quickly deprotected using hydroxylamine at pH 7.2-7.5.[10] The thiolation reaction proceeds by dissolving 100mg of the deprotected intermediate in methanol and then adding an equal portion of water to bring the solution to 50% methanol. Borate buffer at pH 10 is added to increase the total volume by approximately 50-75%. 55mg of SATP is dissolved in DMF and the aqueous reaction solution is slowly mixed with the SATP solution. After complete mixing, the reaction is allowed to proceed at room temperature and constant stirring for 1.5 hours. The progress of the reaction is then checked by desalting using ziptips followed by analysis using MALDI TOFMS.

After thiolation, the reaction solution is once again dried on a rotovap. The remaining solid (*ca.* 300mg) is separated using a flash column of approximately 0.5 inch diameter and a stationary phase length of 5-6 inches.[7] The solvent system used is 20% methanol in DCM. The second intermediate elutes at roughly two column volumes. MALDI TOFMS is used to locate the fractions of interest. Tandem MS analysis was completed to tentatively confirm the structure, however further purification and NMR analysis would be required for positive identification.

NMR was attempted on both intermediates; however the purity from the flash column was not high enough. Both 1D and 2D NMR experiments resulted in numerous peaks that could not be reconciled with the predicted structure. A possible source of contaminants (which explain some of the NMR peaks) was found to originate from the plastic eppendorf tubes used for sample handling.[11] Confirmation experiments were conducted whereby several eppendorf tubes were placed in pure high grade methanol. After resting overnight, a film was visible on the surface of the methanol, despite the container being covered. Further syntheses using only glass culture tubes showed less contamination in the NMR spectra, however it was still not possible to unambiguously prove the structure of the label using NMR.

The mass spectrum of the same sample seemed to indicate that the isolated intermediate was relatively pure, aside from a few peaks attributed to the MALDI matrix. It is thought that a contaminant is coeluting but since the label intermediates readily ionize (due to the arginine group) label signal is swamping contaminant signal, which ionizes to a lower degree under any circumstances. The intermediates were analyzed using HPLC under isocratic elution to try to increase the purity level, but these efforts also failed.

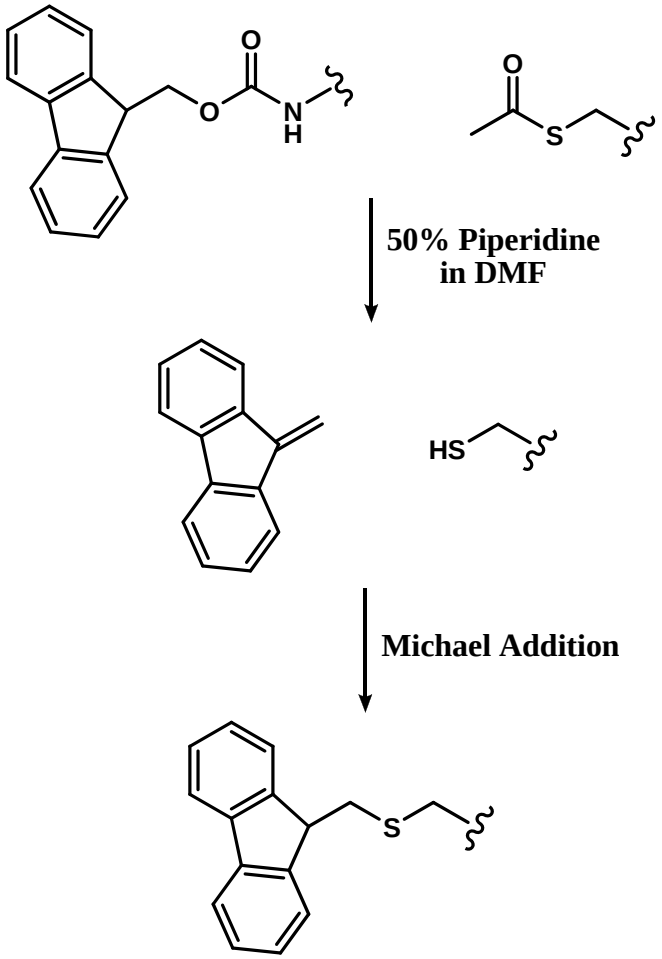
The next step in the synthesis (not completed yet) would be to deprotect the carboxybenzyl (CBZ) group from the other primary amine site.[12-13] Briefly, this is accomplished by completely drying down the product and then dissolving in methanol. It is important that no water remain since it will not allow the reaction to proceed. After dissolving the protected intermediate in methanol, an amount of Pd/C (10%) equal to 10% of the mass of the intermediate is added to the solution. Care should be exercised as this is done since the Pd/C may ignite when added to the methanol solution. The headspace is then purged with hydrogen and a septum is used to seal the flask. The

hydrogen filled balloon is stretched over the end of a syringe and the needle is pushed through the septum to provide a steady hydrogen supply at a partial pressure. The reaction proceeds for 1-1.5 hours under constant stirring at room temperature.

Upon completion, the Pd/C catalyst is filtered off over a bed of celite (diatomaceous earth from Sigma). Reaction progress may be checked using MALDI TOFMS. Further care should be taken when filtering since the Pd/C that has been exposed to hydrogen may ignite if it dries out. Once done filtering, the entire filtering material (including Pd/C) should be sealed in a vial filled with water for disposal.

This CBZ deprotection reaction works well – quickly and quantitatively – when performed on the first intermediate (1), however efforts to deprotect the second intermediate (after thiolation by SATP) have been largely unsuccessful. This is thought to be due to the presence of the protected thiol, a small amount of which may be deprotected, poisoning the Pd/C catalyst. Reversing the order of deprotection (*i.e.* deprotecting the CBZ group first, followed by FMOC) is not possible since the protected thiol (added to the CBZ protected primary amine) deprotects under the deprotection conditions required for FMOC (*i.e.* basic conditions). The acetyl group of the thiol is cleaved allowing the free thiol to modify the dibenzofulvene derivative of the FMOC deprotection through a Michael addition (Figure A4). Efforts to remove this group from the thiol also failed.

The addition of the macrocycle would follow CBZ deprotection. The macrocycle would be added using an isothiocyanate moiety for the modification of the deprotected primary amine.[14-15] The reaction conditions required were investigated using model peptides. The optimal conditions required mixing the isothiocyanate macrcycle and the peptide (a 10:1 molar ratio) at an ammonium bicarbonate buffered pH of 8.6.



Scheme A2. A representation of the difficulty in FMOC deprotection after thiolation using SATP. The FMOC deprotection conditions also cleave the acetyl group protecting the thiol leaving it available for Michael addition to the dibenzofulvene byproduct of the FMOC deprotection.

The reaction was allowed to proceed at 55˚C for 1.5 hours. Reaction progress was checked by desalting using a ziptip, and analysis using MALDI TOFMS.

Had the synthesis of the shift reagent been completed, it would be used to label phosphopeptides for analysis by MALDI-IM-TOFMS. This would be accomplished by beta elimination of the phosphate group (as described in Chapter IV)[16] under basic conditions, followed by addition of the label through Michael addition of the thiol terminus to the peptide. Reaction conditions would have to be optimized since the label represents a significant increase in mass and size over the ethanedithiol linker currently used by groups doing phosphopeptide analysis.

Furthermore, this label would eliminate the steps of gel purification to remove unreacted ethanedithiol, as well as a labeling reaction to modify the new thiol with a shift reagent. The ability to skip these steps through the use of a thiol terminated shift reagent would significantly decrease the time commitment required for phosphopeptide specific labeling. The addition of a highly basic site (the arginine residue) would also enhance ionization of the resulting labeled phosophopeptide, allowing for analysis of less abundant phosphoproteins.

Enhanced ionization would be beneficial for analysis of smaller sample sizes and lower abundance proteins. It would especially aid in the analysis of phosphopeptides in biological matrices since phosphorylation can be a time-dependent PTM and is not completely stoichiometric, meaning that not all copies of a particular protein are phosphorylated at a given time. Additionally, a thiol terminated label such as this could also be used labeling of O-linked glycans, since glycans can also undergo beta elimination. Some *a priori* knowledge of the sample is required in order to avoid confusion between glycopeptides and phosphopeptides.

Other possibilities exist whereby different functionality could be covalently bound to the macrocycle through reaction with the thiol terminus of this tag, or through the other moieties mentioned in this report. Oligonucleotide sequences could be used to create a DNA sequence probe. A biotin moiety could be inserted in the linker region of the label to provide a method of affinity purification prior to IM-MS analysis. Means of analysis of other amino acid residues or PTMs could be added to a macrocycle. In short, the future of IM-MS shift reagents appears bright, with many different applications surely to be developed in the coming years.

We are grateful for financial support provided by the Defense Threat Reduction Agency (HDTRA1-09-1-0013), the National Institute on Drug Abuse (RC2DA028981), the Vanderbilt University College of Arts and Sciences, the Vanderbilt Institutes for Chemical Biology, the Vanderbilt Institute for Integrative Biosystems Research, and Ionwerks, Inc (Houston, TX).

References

1. Bennett, K.L.; Stensballe, A.; Podtelejnikov, A.V.; Moniatte, M.; Jensen, O.N. “Phosphopeptide Detection and Sequencing by Matrix-assisted Laser Desorption/ionization Quadrupole Time-of-flight Tandem Mass Spectrometry.” *J. Mass Spectrom.* **2002**, 37, 179-190.

2. Craig, A.G.; Hoeger, C.A.; Miller, C.L.; Goedken, T.; Rivier, J.E.; Fischer, W.H. “Monitoring Protein Kinase and Phosphatase Reactions with Matrix-assisted Laser Desorption/Ionization Mass Spectrometry and Capillary Zone Electrophoresis: Comparison of the Detection Efficiency of Peptide-Phosphopeptide Mixtures.” *Bio. Mass Spectrom.* **1994**, 23, 519-528.

3. Khorana, H.G. “The Chemistry of Carbodiimides.” *Chem. Rev.* **1953**, 53, 145-166.

4. Yamada, H.; Imoto, T.; Fujita, K.; Okazaki, K.; Motomura, M. “Selective Modification of Aspartic Acid-101 in Lysozyme by Carbodiimide Reaction.” *Biochem.* **1981**, 20, 4836-4842.

5. Ho, G.; Emerson, K.M.; Mathre, D.J.; Shuman, R.F.; Grabowski, E.J.J. “Carbodiimide-mediated Amide Formation in a Two-phase System. A High-yield and Low-racemization Procedure for Peptide Synthesis.” *J. Org. Chem.* **1995**, 60, 3569-3570.

6. Still, W.C.; Kahn, M.; Mitra, A. “Rapid Chromatographic Technique for Preparative Separations with Moderate Resolution.” *J. Org. Chem.* **1978**, 43, 2923-2925.

7. Cohen, L.H.; Gusev, A.I. “Small Molecule Analysis by MALDI Mass Spectrometry.” *Anal. Bioanal. Chem.* **2002**, 373, 571-586.

8. Larsen, B.D.; Holm, A. “Incomplete Fmoc Deprotection in Solid-phase Synthesis of Peptides.” *Int. J. Peptide Protein Res.* **1994**, 43, 1-9.

9. Wenschuh, H.; Beyermann, M.; Krause, E.; Brudel, M.; Winter, R.; Schümann, M.; Carpino, L.A.; Bienert, M. “Fmoc Amino Acid Fluorides: Convenient Reagents for the Solid-phase Assembly of Peptides Incorporating Sterically Hindered Residues.” *J. Org. Chem.* **1994**, 59, 3275-3280.

10. Duncan, R.J.S.; Weston, P.D.; Wrigglesworth, R. “A New Reagent Which May Be Used to Introduce Sulfhydryl Groups Into Proteins, and Its Use in the Preparation of Conjugates for Immunoassay.” *Anal. Biochem.* **1983**, 132, 68-73.

11. Xia, Y.; Patel, S.; Bakhtiar, R.; Franklin, R.B.; Doss, G.A. “Identification of a New Source of Interference Leached from Polypropylene Tubes in Mass-selective Analysis.” *J. Am. Soc. Mass Spectrom.* **2005**, 16, 417-421.

12. Miller, J.J.; Rajaram, S.; Pfaffenroth, C.; Sigman, M.S. “Synthesis of Amine Functionalized Oxazolines with Applications in Asymmetric Catalysis.” *Tetrahedron* **2009**, 65, 3110-3119.

13. Schmuck, C.; Dudaczek, J. “Ion Pairing Between the Chain Ends Induces Folding of a Flexible Zwitterion in methanol.” *Eur. J. Org. Chem.* **2007**, 3326-3330.

14. Mujumdar, R.B.; Ernst, L.A.; Mujumdar, S.R.; Waggoner, A.S. “Cyanine Dye Labeling Reagents Containing Isothiocyanate Groups.” *Cytometry* **1989**, 10, 11-19.

15. Brinkley, M. “A Brief Survey of Methods for Preparing Protein Conjugates with Dyes, Haptens, and Cross-linking Reagents.” *Bioconjugate Chem.* **1992**, 3, 2-13.

16. Hoyle, C.E.; Bowman, C.N. “Thiol-ene Click Chemistry.” *Angewandte Chemie Int. Ed.* **2010**, 49, 1540-1573.

APPENDIX B

MOUSE BRAIN TISSUE LABELING PROGRESS

The development of mass spectrometric imaging (MSI) methods by the Caprioli lab[1-2] in the late 1990's has enabled detailed analysis of intact tissue sections, allowing elucidation of the location of specific molecules within tissue sections. Biomolecules like lipids and peptides can be highlighted to determine not only their location, but also their identity through MS/MS analysis.[2,3-4]Furthermore, non-natural molecules, such as drugs (or their metabolites) may be localized to study the distribution kinetics and ultimate fate of drugs as they are metabolized by the body.[5]

Development of on tissue (*in situ*) reactions has occurred concurrently with widespread use of MSI. Trypsin proteolysis has been used to cleave the proteins present in tissue sections,[6] making analysis and identification by MS methods much easier.Current staining and labeling reactions can provide qualitative information about the tissue section under study.[7-9] DNA labeling, for example, has been used to determine if a significant amount of apoptosis has taken place.[8] Application of *in situ* shift reagent labeling is proposed in this chapter as a way to extract more information from IM-MS spectra of tissue sections. Toward this goal, the cysteine specific label (containing a maleimido reactive moiety) and a primary amine label (containing an isothiocyanate reactive moiety) were tested for their applicability to *in situ* tissue labeling.

Imaging of tissue sections usually proceeds by flash freezing the extracted tissue in liquid nitrogen immediately after dissection. This frozen tissue is sliced on a cryostat so to a thickness of 10-15 μm. The frozen tissue sections are thaw mounted

onto a chilled MALDI target plate (or glass microscope slide) by allowing it to fall onto the plate after slicing.[10] The plate (with tissue slice) is subsequently allowed to equilibrate to room temperature, thereby mounting the tissue onto the plate. For the mouse brain sections described here, mounting was completed by a collaborating lab.

MSI analysis methods eventually recognized great potential in the development of *in situ* reactions (*e.g.* proteolysis using trypsin) that enhance the data collected from tissue sections. Using these methods, proteins in tissue sections can be proteolyzed on site to avoid losing location information. *In situ* proteolysis involves spotting a tissue section with a trypsin (in a buffer of pH ~8) using a hand-held pipet in a low density droplet array, or by using a nanoliter robotic spotter in high density droplet array. Proteins in the tissue section are proteolyzed at room temperature (or alternatively the sample may be gently heated) for several hours prior to matrix deposition. Matrix may be applied using the same method (hand-held pipet or robotic spotter), or by spray coating the entire tissue section. Subsequent to matrix deposition, analysis is completed on a MALDI based instrument platform.

The use of IM-MS for tissue section imaging analysis provides enhanced data over one dimensional MS methods.[11] MSI allows analysis of the tissue in such a way that proteins and other biomolecules are in nearly the same location as they are *in* vivo. IM-MS can separate the biomolecular classes, in addition to chemical noise, in conformation space allowing for rapid identification of the class of molecule responsible for each signal. In addition, mass spectra of each class may be extracted independently for further detailed analysis. This may be useful when only changes specific to the lipid profile are sought, rather than the change in expression profile of different classes of biomolecules. An example of IM-MS analysis of a mouse brain prior to *in situ* trypsin proteolysis is shown in Figure B1.

Typical experiments involving *in situ* proteolysis of mouse brain samples by trypsin proceeded as follows. The brain sections (mounted on a MALDI target plate) were stored in the -80˚C freezer until used. The MALDI plate was allowed to equilibrate to room temperature in a dessicator prior to manipulation. After the plate warmed to room temperature, a 0.5μL aliquot of trypsin solution (2μg/mL in 100mM ammonium bicarbonate, pH ~8) was spotted by hand onto a brain tissue section. The spot was allowed to react at room temperature until dry. Trypsin was spotted twice more in a similar fashion. The brain section was then washed in a cold ethanol solution (75-100%) for 20 seconds in an effort to remove some of the high abundance lipid signal (intense lipid signals tend to overpower weaker peptide signals) and the salts from the reaction buffer. The plate was allowed to dry completely before proceeding. When the plate was dry, alpha cyano buffer was spotted by hand over the area in a low density droplet array, and allowed to dry. Analysis was completed using MALDI-IM-TOFMS (Waters Synapt G2). A typical spectrum of proteolyzed mouse brain is shown in Figure B2.

Labeling reactions were attempted using two different shift reagent labels. The first label was a thiol specific shift reagent for the selective modification the free thiol of cysteine residues via a maleimide moiety. This label was discussed in Chapter III. Briefly, labeling reactions proceeded by chelating a lanthanide metal into the macrocycle (in a 100:1 ratio and pH 5.7) as described previously.[12] The loaded label (100mM, still in the pH 5.7 buffered reaction solution) was spotted onto a brain section, both before and after proteolysis. The spot was allowed to react at room temperature until dry. An additional two aliquots were spotted on the brain section in this manner. Subsequent to the labeling reaction, the brain section was washed in cold ethanol (75-100%) in order to remove some of the lipids and salts from the reaction mixture.[13] Alpha cyano matrix was spotted and analysis was completed using MALDI-IM-TOFMS as illustrated in Figure B3.

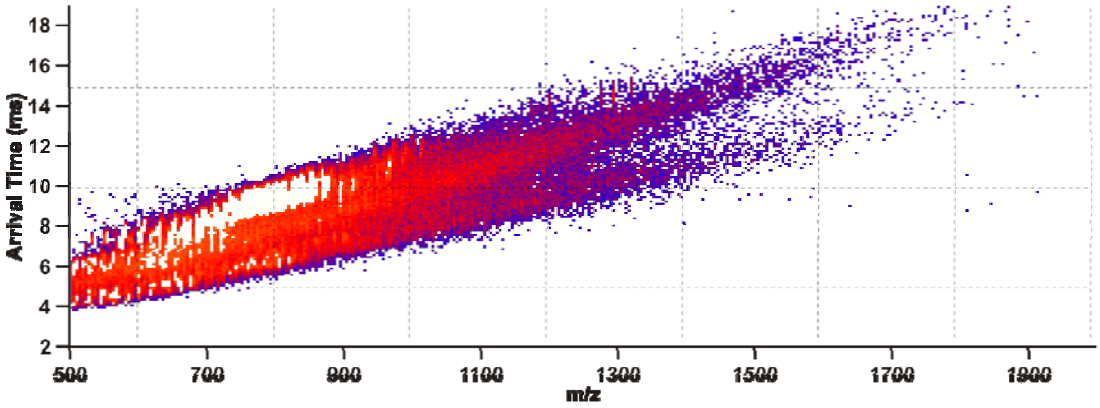


Figure B1. A two dimensional representation of IM-MS conformation space data for an unproteolyzed, unlabeled mouse brain tissue section. The brain was flash frozen and a 12 micron thick section was thaw mounted onto a stainless steel MALDI target plate. Alpha cyano matrix was applied by hand in aliquots of 0.6 microliters in a low density droplet array. Analysis was performed on a Waters Synapt G2 instrument.

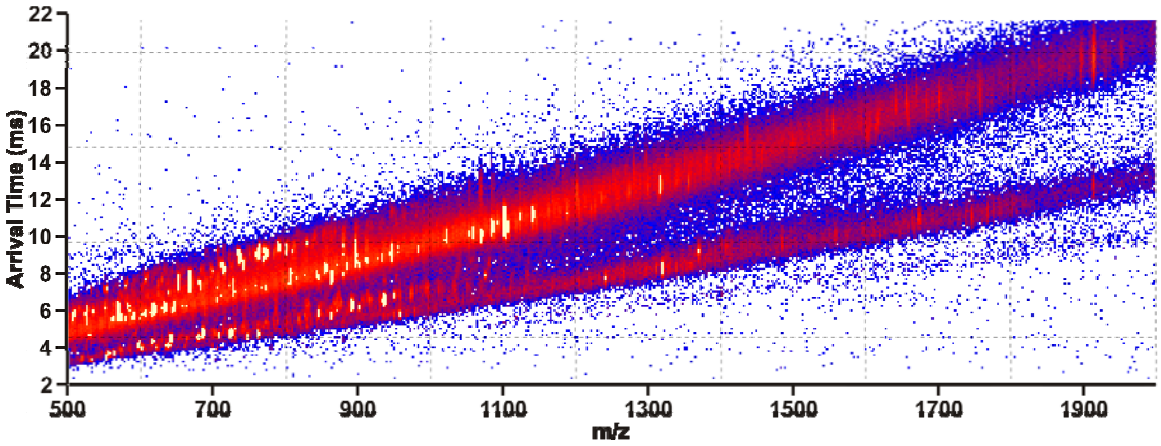


Figure B2. A representation of two dimensional IM-MS conformation space for analysis of an *in situ* trypsin proteolyzed, unlabeled mouse brain. The brain was flash frozen and a 12 micron thick slice was thaw mounted directly onto a stainless steel MALDI target plate. Trypin was applied in a low density droplet array by hand in 0.6 microliter aliquots. After digestion, alpha cyano matrix was applied in the same manner. Analysis was performed on a Waters Synapt G2 instrument.

Primary amine specific on tissue labeling reaction was explored using a shift reagent containing an isothiocyanate moiety. Labeling reactions proceeded by chelating a lanthanide metal to the macrocycle at pH 5.7 (a 100:1 ratio)[12] and then adding ammonium bicarbonate buffer to adjust the pH to 8. This loaded label solution was spotted onto a brain section both before and after proteolysis. The plate was placed on a heating block and heated to 55˚C under an atmosphere of methanol and water. This atmosphere was created by placing the MALDI plate on the heat block next to a tube filled with 50% methanol and inverting a glass petri dish over both. The reaction was allowed to proceed until the spot was dry. The reaction mixture was spotted twice more in the same manner. Subsequent to labeling, the brain section was washed in cold ethanol (75-100%) for 20 seconds to remove some of the lipids and salts from the reaction buffers.[13] Alpha cyano matrix was spotted and analysis was completed using MALDI-IM-TOFMS.

Tandem MS analysis of high intensity signals in the peptide correlation band was done to try to identify several of the peptides, and thus proteins, available in the mouse brain samples. Signals within the peptide correlation band were selected for CID fragmentation in the Waters Synapt G2 instrument. After fragmentation, the resulting MS/MS spectra were analyzed for sequence information. Mascot was used to aid in the identification of the sequence of peptides analyzed by MS/MS.[14]

The mascot utility was initially allowed to search all categories to try to get the best match for some of the short peptides. A second round of analysis was done by restricting the search to known mouse proteins. The results of these searches yielded several tentative peptide sequence matches, however the scores for these matches was low. At a p value of 0.5, a mascot score of 45 would be required for a significant hit. Many of the scores analyzed so far have been in the 20s and 30s, and are insignificant.

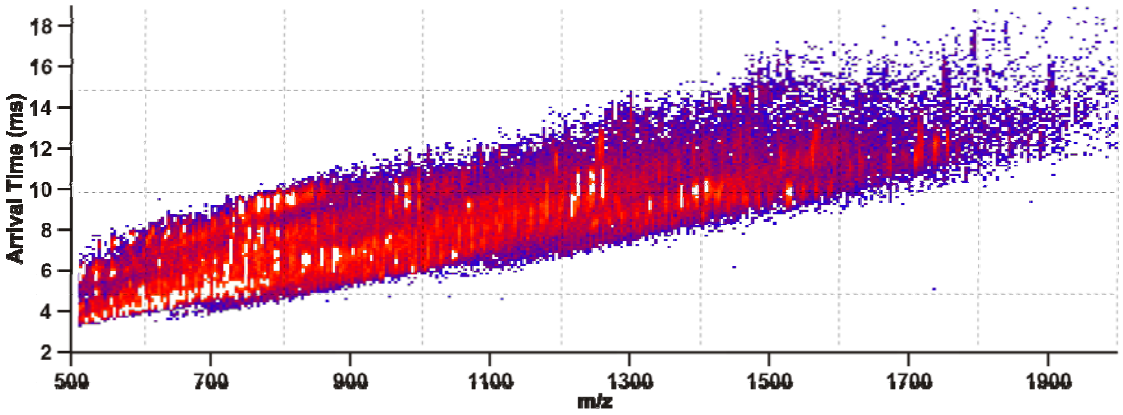


Figure B3. A two dimensional representation of IM-MS conformation space data for a section of labeled mouse brain. The brain was flash frozen and a 12 micron thick slice was thaw mounted directly onto a stainless steel MALDI target plate. Alpha cyano matrix was applied in a low density droplet array by hand in 0.6 microliter aliquots. The labeling reaction consisted of a maleimide based shift reagent coordinated to terbium metal. Cysteine containing peptides are believed to be labeled and shifted to an area around 1300-1500 Daltons and about 10ms drift time. Analysis was performed on a Waters Synapt G2 instrument.

Further analysis and confident identification would require more signal in the MS/MS spectrum. More signal could be acquired through optimization of instrument settings and wash steps prior to analysis.

Further work in this area would involve the optimization of the *in situ* labeling reactions for both the cysteine specific label and the primary amine specific label. Instrument settings could possibly be optimized further, however most improvements are expected to come from changing reaction conditions. Further MS/MS analysis would be completed, including the use of other bioinformatics platforms to aid in the identification of peptide signals. The sequence identification of several peptides that are labeled would be a critical next piece of data.

We are grateful for financial support provided by the Defense Threat Reduction Agency (HDTRA1-09-1-0013), the National Institute on Drug Abuse (RC2DA028981), the Vanderbilt University College of Arts and Sciences, the Vanderbilt Institutes for Chemical Biology, the Vanderbilt Institute for Integrative Biosystems Research, and Ionwerks, Inc (Houston, TX).

References

1. Stoeckli, M.; Farmer, T.B.; Caprioli, R.M. “Automated Mass Spectrometry Imaging with a Matrix-assisted Laser Desorptio n Ionization Time-of-flight Instrument.” *J. Am. Soc. Mass Spectrom.* **1999**, 10, 67-71.

2. McDonnell, L.A.; Heeren, R.M.A. “Imaging Mass Spectrometry.” *Mass Spectrom. Rev.* **2007**, 26, 606-643.

3. Burnum, K.E.; Cornett, D.S.; Puolitaival, S.M.; Milne, S.B.; Meyers, D.S.; Tranguch, S.; Brown, H.A.; Dey, S.K.; Caprioli, R.M. “Spatial and Temporal Alterations of Phospholipids Determined by Mass Spectrometry During Mouse Embryo Implantation.” *J. Lipid Res.* **2009**, 50, 2290-2298.

4. Ridenour, W.B.; Kliman, M.; McLean, J.A.; Caprioli, R.M. “Structural Characterization of Phospholipids and Peptides Directly from Tissue Sections by MALDI Traveling-wave Ion Mobility-mass Spectrometry.” *Anal. Chem.* **2010**, 82, 1881-1889.

5. Cornett, D.S.; Frappier, S.L.; Caprioli, R.M. “MALDI-FTICR Imaging Mass Spectrometry of Drugs and Metabolites in Tissue.” *Anal. Chem.* **2008**, 80, 5648-5653.

6. Chaurand, P.; Schriver, K.E.; Caprioli, R.M. “Instrument Design and Characterization for High Resolution MALDI-MS Imaging of Tissue Sections.” *J. Mass Spectrom.* **2007**, 42, 476-489.

7. Möhler, H.; Battersby, M.K.; Richards, J.G. “Benzodiazepine Receptor Protein Identified and Visualized in Brain Tissues by a Photoaffinity Label.” *Proc. Natl. Acad. Sci. USA* **1980**, 77, 1666-1670.

8. Gavrieli, Y.; Sherman, Y.; Ben-Sasson, S.A. “Identification of Programmed Cell Death In Situ via Labeling of Nuclear DNA Fragmentation.” *J. Cell. Biol.* **1992**, 119, 493-501.

9. Godement, P.; Vanselow, J.; Thanos, S.; Bonhoeffer, F. “A Study in Developing Visual Systems with a New Method of Staining Nurones and their Processes in Fixed Tissue.” *Development.* **1987**, 101, 697-713.

10. Cornett, D.S.; Reyzer, M.L.; Chaurand, P.; Caprioli, R.M. “MALDI Imaging Mass Spectrometry: Molecular Snapshots of Biochemical Systems.” *Nature Methods.* **2007**, 828-833.

11. McLean, J.A.; Ruotolo, B.T.; Gillig, K.J.; Russell, D.H. “Ion Mobility-mass Spectrometry: A New Paradigm for Proteomics.” *Int. J. Mass Spectrom.* **2005**, 240, 301-315.

12. Zhu, X.; Lever, S.Z. “Formation Kinetics and Stability Studies on the Lanthanide Complexes of 1,4,7,10-tetra-azacyclododecane-*N,N’,N’’,N’’’*-tetraacetic acid by Capillary Electrophoresis.” *Electrophoresis.* **2002**, 23, 1348-1356.

13. McLean, J.A.; Ridenour, W.B.; Caprioli, R.M. “Profiling and Imaging of Tissues by Imaging Ion Mobility-mass Spectrometry.” *J. Mass Spectrom.* **2007**, 42, 1099-1105.

14. Perkins, D.N.; Pappin, D.J.C.; Creasy, D.M.; Cottrell, J.S. “Probability-based Protein Identification by Searching Sequence Databases Using Mass Spectrometry Data.” *Electrophoresis.* **1999**, 20, 3551-3567.

APPENDIX C

REFERENCES OF ADAPTATION FOR CHAPTERS

Chapter I: Sections adapted from Randi L. Gant-Branum, Thomas J. Kerr, and John A. McLean, “Labeling Strategies in Mass Spectrometry-Based Protein Quantitation”, *The Analyst* **2009**, 134, 1525-1530.

Chapter II: Adapted from Thomas J. Kerr and John A. McLean, “Peptide Quantitation Using Primary Amine Selective Metal Chelation Labels for Mass Spectrometry”, *Chemical Communications,* **2010**, 46, 5479-5481.

Chapter III: Adapted from Thomas J. Kerr and John A. McLean, “High Density Shift Reagents for Multiplexed Peptide Characterization by Ion Mobility-Mass Spectrometry”, In preparation for submission to *Analytical Chemistry* 2011.

Chapter IV: Adapted from Thomas J. Kerr, Randi L. Gant-Branum, and John A. McLean, “Multiplexed Analysis of Peptide Functionality Using Lanthanide-Based Structural Shift Reagents”, *International Journal of Mass Spectrometry*, **2011**, in press.

Thomas J. Kerr

806 Erin Lane *Email: Josh.Kerr.11@gmail.com*

Nashville, TN, 37221 *Tel: (706) 830.2883*

Education[1]

Vanderbilt University, Nashville, TN

Attended: June 2006 to present

Advisor: Prof. John A. McLean

Degree obtained: Ph.D. in Analytical Chemistry, *anticipated in March 2011.*

Dissertation title: Development of Novel Ion Mobility-Mass Spectrometry Shift Reagents for Proteomic Applications

GPA: 3.58

Wofford College, Spartanburg, SC

Attended: September 2002 - May 2006

Degree obtained: Bachelor of Science in Chemistry

GPA: 3.34 *cum laude*

Teaching Experience

Teaching Assistant – General Chemistry Lab

August-December 2006-2010 – Head TA 2007, 2008, 2009

Teaching Assistant – Analytical Chemistry Lab

January-May 2007- 2008.

Teaching Assistant – Forensic Chemistry Lab

January-May 2010

[1] CV completed February 17, 2011.

Research Experience

Vanderbilt University Graduate School – McLean Lab, January 2007 to present.

Specialized in ion mobility-mass spectrometry instrumentation (research and commercial grade, Water Synapt G2); matrix-assisted laser desorption/ionization and electrospray ionization; concentrated on development of IM-MS shift reagents for proteomic applications.

Publications – peer reviewed

1. Gant-Branum, R.L.; **Kerr, T.J.**; McLean, J.A. "Labeling Strategies in Mass Spectrometry-based Protein Quantitation." *The Analyst.* **2009**, 134, 1525-1530.

2. **Kerr, T.J.**; McLean, J.A. "Peptide Quantitation Using Primary Amine Selective Metal Chelation Labels for Mass Spectrometry." *Chemical Communications.* **2010**, 46, 5479-5481.

3. **Kerr, T.J.**; Gant-Branum, R.L.; McLean, J.A. "Multiplexed Analysis of Peptide Functionality Using Lanthanide-based Structural Shift Reagents." *International Journal of Mass Spectrometry.* **2011**, *in press.*

4. **Kerr, T.J.**; McLean, J.A. "High Density Shift Reagents for Multiplexed Peptide Characterization by Ion Mobility-mass Spectrometry." *Analytical Chemistry.* **2011**, *in preparation for submission.*

5. Gant-Branum, R.L.; **Kerr, T.J.**; McLean, J.A. "Relative Quantitation of Phosphorylated Peptides and Proteins Using Phosphopeptide Element-coded Affinity Tagging." *Analytical Chemistry.* **2011**, *in preparation for submission.*

Presentations

1. Fenn, L.S.; **Kerr, T.J.**; McLean, J.A., *Advanced structural mass spectrometry approaches for proteomics and biophysics.* Vanderbilt Institute of Chemical Biology/ St. Jude yearly retreat, Florence, AL (March 2007).

2. **Kerr, T.J.**; McLean, J.A., *Functionally selective ion mobility shift reagents for proteomic applications.* Proceedings of the 55th annual ASMS conference on mass spectrometry and allied topics, Indianapolis, IN (June 2007).

3. McLean, J.A.; Fenn, L.S.; **Kerr, T.J.**, *Structural mass spectrometry separation strategies for proteomics, glycomics, and lipidomics: Ion mobility-MS and imaging ion mobility-MS.* Protein Society 21st annual symposium, Boston, MA (July 2007).

4. **Kerr, T.J.**; McLean, J.A., *Ion mobility-mass spectrometry shift reagents for multiplexed peptide and protein characterization.* Proceedings of the annual meeting of the Federation of Analytical Chemistry and Spectroscopy Societies, Memphis, TN (October 2007).

5. **Kerr, T.J.**; McLean, J.A., *Structural characterization and relative quantitation of peptides by ion mobility MS.* Vanderbilt Institute of Chemical Biology/ St. Jude yearly retreat, Florence, AL (April 2008).

6. **Kerr,T.J.**; McLean, J.A., *Ion mobility shift reagents for identification of primary amines.* Proceedings of the 56th annual ASMS conference on mass spectrometry and allied topics, Denver, CO (June 2008).

7. Gant-Branum, R.L.; **Kerr, T.J.**; McLean, J.A., *Phosphoproteomics using selective derivatization and structural separations by ion mobility-mass spectrometry.* Proceedings of the 56th annual ASMS conference on mass spectrometry and allied topics, Denver, CO (June 2008).

8. **Kerr, T.J.**; McLean, J.A., *Relative quantitation of peptides and proteins using lanthanide chelation tags as ion mobility shift reagents.* Proceedings of the 60th Southeastern Regional Meeting of the American Chemical Society, Nashville, TN (November 2008).

9. Fenn, L.S.; Kliman, M.; Mahsut, A.; Zhao, S.; **Kerr, T.J.**; Gant, R.L.; McLean, J.A., *Integrated 'omics' using structural mass spectrometry strategies*. Proceedings of the 60th Southeastern Regional Meeting of the American Chemical Society, Nashville, TN (November 2008).

10. **Kerr, T.J.**; McLean, J.A., *Collision cross section and ionization efficiency of IM-MS shift reagents.* Vanderbilt Institute of Chemical Biology/ St. Jude yearly retreat, Nashville, TN (April 2009).

11. Fenn, L.S.; Kliman, M.; Mahsut, A.; Zhao, S.; **Kerr, T.J.**; Gant, R.L.; McLean, J.A., *Integrated 'omics' on the basis of structural separations by ion mobility-mass spectrometry.* Proceedings of the 57th annual ASMS conference on mass spectrometry and allied topics, Philadelphia, PA (June 2009).

12. Fenn, L.S.; Kliman, M.; Gant, R.L.; **Kerr, T.J.**; McLean, J.A., *Characterization of complex biological samples on the basis of structural separations by ion mobility-mass spectrometry.* 18th International mass spectrometry conference, Bremen, Germany (September 2009).

13. **Kerr, T.J.**; McLean, J.A., *Functionally selective lanthanide-based shift reagents for biomolecular mass spectrometry applications.* Proceedings of the annual meeting of the Federation of Analytical Chemistry and Spectroscopy Societies, Louisville, KY (October 2009).

14. **Kerr, T.J.**; McLean, J.A., *Enhanced ion mobility shift reagents for proteomic applications.* Proceedings of the 58th annual ASMS conference on mass spectrometry and allied topics, Salt Lake City, UT (May 2010).

15. **Kerr, T.J.**; McLean, J.A., *Enhanced ion mobility shift reagents for peptide labeling.* Proceedings of the annual meeting of the Federation of Analytical Chemistry and Spectroscopy Societies, Raleigh, NC (October 2010).